Technische Universität München

Fakultät für Chemie

Bayerisches NMR Zentrum

Lehrstuhl für Biomolekulare NMR-Spektroskopie

Solution NMR studies of the splicing factor RBM10 and the molecular chaperone Hsp90

Martin Rübbelke

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Bernd Reif
Prüfer der Dissertation:	1. Prof. Dr. Michael Sattler 2. Prof. Dr. Johannes Buchner 3. Prof. Dr. Dierk Niessing

Die Dissertation wurde am 17.08.2017 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 18.10.2017 angenommen.

Declaration

I hereby declare that parts of this thesis have already been published in the following scientific journals:

Zierer, Bettina K., Matthias Weiwad, Martin Rübbelke, Lee Freiburger, Gunter Fischer, Oliver R. Lorenz, Michael Sattler, Klaus Richter, and Johannes Buchner. "Artificial Accelerators of the Molecular Chaperone Hsp90 Facilitate Rate-Limiting Conformational Transitions." *Angewandte Chemie International Edition* 53, no. 45 (November 3, 2014): 12257–62. doi:10.1002/anie.201406578

Zierer, Bettina K., Martin Rübbelke, Franziska Tippel, Tobias Madl, Florian H. Schopf, Daniel A. Rutz, Klaus Richter, Michael Sattler, and Johannes Buchner. "Importance of Cycle Timing for the Function of the Molecular Chaperone Hsp90." *Nature Structural & Molecular Biology* 23, no. 11 (November 2016): 1020–28. doi:10.1038/nsmb.3305.

Contents

1 Abstract ... 9

2 Zusammenfassung ... 13

3 RBM10 Introduction ... 19

3.1 Splicing ... 19

3.2 Alternative Splicing and the Splicing Code ... 21

3.3 The Splicing Factor RBM10 ... 22

3.4 Scope of the thesis ... 26

4 Background of Applied Methods ... 29

4.1 Basics of NMR ... 29

4.2 Relaxation ... 31

4.3 The 1D NMR experiment ... 32

4.4 Multidimensional NMR ... 33

4.5 NMR resonance assignment ... 34

4.5.1 Backbone assignment ... 34

4.5.2 Side chain assignment ... 35

4.6 Ligand binding and exchange regimes ... 37

4.7 Nuclear Overhauser Effect ... 38

4.8 Structure Calculation ... 41

4.8.1 nOe based structure calculation of small protein domains ... 41

4.8.2 Using RDC, PRE and SAXS data for structure Calculation of Multidomain Proteins ... 42

4.9 Paramagnetic Relaxation Enhancement ... 42

4.10 Residual Dipolar Couplings ... 43

4.11 Small Angle X-ray Scattering ... 44

5 RBM10 Material and Methods ... 47

5.1 Cloning ... 47

5.1.1 Sub-cloning of RBM10 constructs from full length protein ... 47

5.1.2 Quick-Change-PCR for Side directed mutagenesis ... 48

5.1.3 Transformation and Mini-Prep 49

5.2 Protein Expression 49

5.3 Protein Purification 49

5.4 NMR Backbone Assignment 50

5.5 NMR Titrations 51

5.6 NMR Side Chain Assignment and Structure Calculation 51

5.7 RNA TOCSY spectra and ω1-filtered 2D NOESY spectra 52

5.8 NMR Relaxation experiments 53

5.9 Paramagnetic Relaxation Enhancement Measurements 54

5.9.1 Attachment of 3-(2-Iodoacetamido)-PROXYL (IPSL) spinlabel to the protein. 54

5.9.2 Paramagnetic Relaxation Enhancement NMR measurements 55

5.10 Isothermal Titration Calorimetry (ITC) 55

5.11 Small angle X-ray scattering (SAXS) 56

5.12 Static light scattering (SLS) 56

6 RBM10 Results 57

6.1 RRM1-ZF1 and RRM2 57

6.1.1 RRM1-ZF1 apo-protein adopts a flexible domain orientation in solution 57

6.1.2 Both domains are involved in RNA binding in a 1:1 complex 61

6.1.3 NUMB-derived RNA oligo shows strong line broadening in complex with RBM10 67

6.1.4 RRM2 structure 69

6.1.5 RNA binding of RRM2 72

6.2 Domain organization and RNA recognition by RBM10 RRM1-ZF1-RRM2 74

6.2.1 RNA binding domains of RBM10 are independent in the apo-protein 74

6.2.2 All three domains are involved in RNA binding 78

6.2.3 RNA TOCSY spectra when bound to the RRM1-Zn-RRM2 83

6.2.4 SAXS provides information about the shape of the three RNA binding domains 85

6.2.5 Paramagnetic Relaxation enhancement provides inter domain distance information 86

6.3 Analysis of disease-linked mutants and isoforms 92

6.3.1 The RBM10 splicing isoform +V354 and the cancer-linked mutant V354E 92

6.3.2 RNA binding of the +V354 isoform and the V354E mutant 96

6.3.3 Domain-Domain interaction are also not influenced by RNA binding 98

7 RBM10 Discussion 101

8 Hsp90 Introduction 107

8.1 The Hsp90 chaperone 107

8.2 Scope of the project 109

9 Hsp90 Material and Methods 111

9.1 NMR analysis of Hsp90 modulator 4 111

9.2 NMR analysis of Hsp90-Aha1 modulator 1. 111

9.3 NMR analysis of the NTD mutations 111

10 Hsp90 Results 113

10.1 Hsp90 Modulators 113

10.1.1 Modulators where identified by a FRET based screening 113

10.1.2 NMR titration reveal the binding site of an accelerator of the Hsp90 chaperone cycle 114

10.2 Hsp90-Aha1 Modulators 116

10.2.1 Identification of Hsp90-Aha1 modulators by a FRET based screen 116

10.2.2 NMR Analysis reveals that Modulator 1 can release ATP from the NTD of Hsp90 117

10.3 Cycling timing of Hsp90 Chaperone Cycle 120

10.3.1 ATPase activity of Hsp90 mutants does not correlate with viability in yeast 120

10.3.2 NMR analysis of the mutants in the N-terminal domain of Hsp90 120

11 Hsp90 Discussion 125

11.1 Hsp90 and Hsp90-Aha1 modulators 125

11.2 Importance of cycling time of the molecular chaperone Hsp90 125

12 Concluding Remarks 127

13 Literature 129

14 Acknowledgements 135

16 Appendix 137
16.1 Buffers and Media 137
16.2 RBM 10 Constructs 139
16.3 Abbreviations: 139
16.4 List of Figures 141
16.5 List of Tables 143
17 Lebenslauf 145

1 Abstract

NMR spectroscopy has been proven to be useful to study the structure, dynamics and ligand binding of multidomain proteins in solution. Here, two distinct systems are studied: The splicing factor RBM10 that is involved in the posttranscriptional maturation of pre-RNA by pre-mRNA splicing and the molecular chaperone Hsp90 that ensures the correct folding of client proteins.

Splicing is an important processing step in the biogenesis and posttranscriptional processing of mRNA, whereby non-coding introns are removed and coding exons are joined together. Through alternative splicing exons can be skipped, or there can be alternative exon-intron boundaries. Thereby one gene can produce several different proteins. The process of alternative splicing is controlled through *cis*-elements coded in the pre-mRNA and through splicing-factors, which act *in trans* by binding to the *cis*-elements on the pre-mRNA. This regulation allows tissue and developmental stage specific expression of isoforms and thereby increases the coding capacity of the genome.

In this work the RNA binding of the multidomain splicing factor RBM10 is characterized by a combination of biochemical, biophysical methods and structural analysis. CLIP data have shown that RBM10 binding is enriched upstream of 5' splice sites and downstream of 3' splice sites on pre-mRNAs and that RBM10 is involved in the regulation of alternative splicing. As a concrete example RBM10 has been shown to promote exon skipping in the *NUMB* gene, leading to a shorter gene product, which can inhibit Notch signaling and thereby regulate cell proliferation. RBM10 comprises three N-terminal RNA binding domains: two RNA recognition motifs (RRM) and a Zinc finger (ZF1). Results presented show that all three domains of RBM10 interact with a single-stranded 12-mer RNA sequence motif derived from the *NUMB* gene in a high affinity 1:1 complex. This is of high importance, since it shows that all three domains are contributing to the effect of RBM10 in alternative splicing of the *NUMB* gene. The first RRM (RRM1) is connected to the ZF (ZF1) by a short 4-residue linker. Despite this short linker the domains tumble independently in the apostate. The domain orientation becomes rigidified upon binding of the NUMB derived 12-mer RNA. The second RRM domain (RRM2) is connected to ZF1 by a long 57 residue linker. As expected it is also flexible in respect to RRM1-ZF1 in the apo-protein. Also, in the three domain construct the domain orientation becomes more rigid upon RNA binding. Furthermore, according to small angle X-ray scattering (SAXS) the domains become more compact in the RNA bound form. All three domains contribute to the RNA binding affinity, indicating that the domains cooperate in the recognition of the biologically important NUMB RNA. To further characterize the domain arrangement and the domain rearrangement upon RNA binding, paramagnetic relaxation enhancement (PRE) data are presented. The PRE-data show that the domains come into close contact already in the apostate, but considering the dynamic domain orientation the contacts have to be transient. The PRE data on the protein:RNA complex indicate a domain rearrangement upon RNA binding. In the future the PRE and SAXS data, together with RDC data, shall be used to generate a structure model of the three RNA binding domains in complex with the biological relevant NUMB derived 12-mer RNA.

RBM10 exits in two splicing isoforms, which differ in the absence (-V354) and presence (+V354) of valine 354, which is positioned at the N-terminus of α-helix 2 in RRM2. In this work the solution NMR structure of the -V354 RRM2 and the +V354 RRM2 are solved. The domains show minor differences in the orientation of β-strand 4 and α-helix 2. Interestingly the mutation of the extra valine to glutamate (V354E), which was identified in lung cancer cells, was shown to abolish the effects of RBM10 on the alternative splicing of the NUMB pre-mRNA. In this work it is shown that the mutation does not change the recognition of the NUMB derived 12-mer RNA oligo by RRM2. It is therefore likely that the mutation effects the interaction with another component downstream of RNA recognition.

The structural analysis of RBM10 revealed two important points. First that all three RNA binding domains cooperate to recognize RNA sequences from the *NUMB* gene, which is one of RBM10's natural targets, and second that the known V354E mutation found in lung cancer cells does not influence RNA recognition of RBM10. Therefore the known effect of this mutation in splicing assays must depend on interactions with other factors. In this way this study contributed to the understanding of RBM10's role in alternative splicing and in disease.

All the information for the three dimensional fold of a protein is coded in its primary sequence. During protein folding it can happen that intermediates aggregate, because folding intermediates often present hydrophobic surfaces, which are hidden in the inside of the final fold. Molecular chaperones can prevent this aggregation. Hsp90 is an ATP-dependent molecular chaperone. Hsp90 forms a homodimer via the CTD. Upon ATP binding it undergoes large conformational changes leading to an N-terminal closed conformation that can hydrolyze ATP. The Hsp90 chaperone is controlled by a number of co-chaperones, which influence the ATPase activity or recruit client proteins.

In this work the binding site of an activator of the Hsp90 ATPase, which was identified in a FRET based assay, is mapped by NMR spectroscopy. The activator binds close to the ATP binding site and the affected residues do not overlap with the known binding site of the ATPase activating co-chaperone Aha1. In agreement with this the activator can further stimulate the Hsp90:Aha1 complex.

Furthermore, the binding site of an inhibitor specific to the Hsp90:Aha1 complex, which was identified in a FRET based assay, is mapped by NMR-spectroscopy to the NTD of Hsp90. Interestingly this inhibitor is able to release ATP from the ATP-binding pocket of Hsp90. The targeting of the Hsp90:Aha1 complex is of interest, since it is important for the maturation of mutants of cystic fibrosis transmembrane conductance regulator (CFTR), which is associated cystic fibrosis. An effect of the studied inhibitor on the stability of the CFTR mutant ΔF508 could be shown.

The effects of mutants with altered ATPase activity on the function of Hsp90 was studied. There was no correlation between ATPase activity and the ability of mutants to sustain viability in yeast. In this thesis, the effect of the mutants on the NTD of Hsp90 is checked by NMR. Removing the first 8 residues (Δ8) leads to changes in chemical shift spread over the entire domain, indicating a structural rearrangement. This fits to SAXS data were the full length Hsp90 lacking the first 8 residues shows a more compact state in the absence of ATP than the wildtype and is further compacted upon nucleotide binding. This mutant cannot support viability probably because the fully open state is not or insufficiently populated. The mutant E33A, which is ATPase deficient, surprisingly can support viability. By NMR spectroscopy and SAXS it could be shown that the E33A variant can react to ATP binding. The mutant D79N, in contrast cannot bind ATP, and therefore cannot change its conformation in an ATP dependent manner and cannot support viability in yeast. The presented study shows that not the speed of the ATPase reaction determines if an Hsp90 variant is functional, but the ability to adopt the different conformations in the chaperone cycle.

Both parts of the thesis, the investigation of the splicing factor RBM10 and the molecular chaperone Hsp90, made use of the ability of NMR spectroscopy to study proteins in solution. The complex interaction of the three dynamic RNA binding domains of RBM10 was studied using NMR relaxation and PRE experiments, combined with SAXS. NMR spectroscopy was also the method of choice to map the binding site of the modulators of the molecular chaperone Hsp90, which bind Hsp90 only with moderate affinities, and to study the effects of mutations in the N-terminal domain of Hsp90, which undergoes conformational changes throughout Hsp90s chaperone cycle and therefore shows dynamics.

2 Zusammenfassung

NMR-Spektroskopie ist erwiesenermaßen nützlich um die Struktur, Dynamik und Ligandenbindung von Multidomänenproteinen zu untersuchen. Hier werden zwei unterschiedliche Systeme untersucht: Der Spleißfaktor RBM10, der an der posttranskriptionellen Reifung der pre-mRNA durch pre-mRNA Spleißen beteiligt ist und das molekulare Chaperon Hsp90, das für die korrekte Faltung seiner Klientenproteine sorgt.

Spleißen ist ein wichtiger Prozessierungsschritt in der Biogenese von mRNA, bei dem nicht kodierende Introns entfernt und kodierende Exons verbunden werden. Durch alternatives Spleißen können Exons übersprungen werden, oder es können sich alternative Exon-Intron-Grenzen ergeben. Dadurch kann ein Gen mehrere verschiedene Proteine kodieren. Der Prozess des alternativen Spleißens wird durch auf der pre-mRNA kodierte *cis*-Elemente und durch *in trans* agierende Spleißfaktoren, die an die *cis*-Elemente auf der pre-RNA binden, kontrolliert. Diese Regulierung erlaubt Gewebe und Entwicklungsstadien spezifische Expression von Isoformen und vergrößert dadurch die Speicherkapazität des Genoms.

In dieser Arbeit wird durch eine Kombination von biochemischen und biophysikalischen Methoden und strukturelle Analyse die RNA Bindung des Multidomänenspleißfaktors RBM10 untersucht. CLIP Daten haben gezeigt, dass die RBM10 Bindungsstellen in Regionen vor der 5‘-Spleißstelle und nach der 3’-Spleißstelle auf der pre-mRNA angereichert sind und das RBM10 an der Regulierung von alternativem Spleißen beteiligt ist. Als konkretes Beispiel, konnte gezeigt werden, dass RBM10 das Überspringen eines Exons im *NUMB*-Gen befördert, was zu einem kürzerem Genprodukt führt, das den Notch-Signalweg unterdrücken und somit das Proliferieren von Zellen kontrollieren kann. RBM10 beinhaltet drei N-terminale RNA-Bindedomänen: Zwei RNA recognition motifs (RBM) und einen Zinkfinger (ZF1). In dieser Arbeit präsentierte Ergebnisse zeigen, dass alle drei RNA-Bindedomänen mit einem aus dem *NUMB*-Gen abgeleiteten 12-mer RNA-Oligo einem hochaffinen 1:1 Komplex bilden. Dies ist von hoher Wichtigkeit, da es zeigt, dass alle drei Domänen zu dem Effekt von RBM10 auf das alternative Spleißen der NUMB pre-mRNA beitragen. Die erste RRM (RRM1) ist mit dem ZF (ZF1) über eine kurzen 4 Aminosäuren langen Linker verbunden. Trotz der Kürze des Linkers orientieren sich die beiden Domänen unabhängig voneinander im Apoprotein. Die Domänenorientierung wird weniger flexibel im RNA gebundenen Zustand. Die zweite RRM Domäne (RRM2) ist mit ZF1 über einen 57 Aminosäuren langen Linker verbunden und ist wie erwartet im Apoprotein flexibel gegenüber RRM1-ZF1. Auch im Dreidomänenkonstrukt wird die Domänenorientierung weniger flexibel im Protein-RNA-Komplex. Darüber hinaus zeigen SAXS-Daten, dass die Domänen im Protein-RNA komplex kompakter angeordnet sind. Dass alle drei Domänen zur Affinität zur NUMB RNA beitragen, zeigt dass die Domänen in der

Erkennung der NUMB RNA kooperieren. Um das Arrangement der Domänen weiter zu charakterisieren werden PRE-Daten präsentiert. Die PRE-Daten zeigen, dass die Domänen bereits im Apozustand in engen Kontakt kommen, der aber wegen dem dynamischen Verhaltend der Domänen transient sein muss. Die PRE-Daten des Protein:RNA Komplexes zeigen eine Neuanordnung der Domänen ausgelöst durch die RNA-Bindung. In der Zukunft sollen die PRE- und SAXS-Daten zusammen mit RDC-Daten dazu eingesetzt werden ein Strukturmodell der RNA-Bindedomänen im Komplex mit der biologisch relevanten vom NUMB-Gen abgeleiteten 12-mer RNA zu erstellen.

RBM10 existiert in zwei Spleißisoformen, welche sich in dem Fehlen (-V354) oder der Präsenz (+V354) von Valin 354, das am N-Terminus der α-Helix 2 in der RRM2 positioniert ist, unterscheiden. In dieser Arbeit werden die Strukturen der -V354 RRM2 und der +V354 RRM2 präsentiert. Die Domänen zeigen kleinere Unterschiede in der Orientierung des β-Strangs 4 und der α-Helix 2. Interessanterweise hebt die Mutation des extra Valins zu Glutamat, die in Krebszellen präsent ist, den Effekt von RBM10 auf das alternative Spleißen der NUMB pre-mRNA auf. In dieser Arbeit wird gezeigt, dass die Mutation keinen Einfluss auf die Bindung des vom *NUMB*-Gen abgeleiteten 12-mers hat. Daher müssen die bekannten Effekte der Mutation auf Spleißassays von der Interaktion mit anderen Faktoren abhängen.

Die strukturelle Analyse von RBM10 hat zwei wichtige Punkte offengelegt: Erstens, dass alle drei RNA bindenden Domänen kooperieren um eine RNA-Sequenz des *NUMB*-Gens, dass ein natürlicher Bindepartner von RBM10 ist, zu erkennen, und zweitens, dass die V354E Mutation, die in Lungenkrebszellen präsent ist, die RNA-Erkennung von RBM10 nicht beeinflusst und daher die Effekte auf alternatives Spleißen, die in Spleißassays beobachtet wurden, von anderen Faktoren abhängen müssen. Auf diese Weise hat diese Studie zum Verständnis von RBM10s Rolle in alternativem Spleißen und Krankheiten beigetragen.

Alle Informationen der dreidimensionalen Struktur von Proteinen sind in dessen primären Sequenz enthalten. Während der Proteinfaltung kann es passieren, dass Intermediate aggregieren, da Faltungsintermediate häufig hydrophobe Oberfläche präsentieren, die im korrekt gefalteten Protein im inneren verborgen sind. Molekulare Chaperons können diese Aggregation verhindern. Hsp90 ist ein ATP-abhängiges molekulares Chaperon. Hsp90 bildet ein Homodimer über seine CTD. Nach Bindung von ATP durchläuft Hsp90 große strukturelle Änderungen, die zu einem N-terminal geschlossenen Zustand führen, der ATP hydrolysieren kann. Hsp90 wird durch eine Zahl von co-Chaperons kontrolliert, die die ATPase Aktivität beeinflussen oder Klientenproteine rekrutieren.

In dieser Arbeit wird die Bindestelle eines Aktivators der ATPase Aktivität von Hsp90, der in einem FRET basierten Screen entdeckt wurde, mittels NMR-Spektroskopie bestimmt. Der Aktivator bindet nahe an der ATP-Bindestelle und die durch die Bindung beeinflussten Residuen überlappen nicht mit der bekannten Bindestelle des die ATPase aktivierenden co-Chaperons Aha1. In Übereinstimmung damit kann der Aktivator auch den Hsp90:Aha1-Komplex weiter stimulieren.

Darüber hinaus wird die Bindestelle eines für den Hsp90:Aha1-Komplex spezifischen Inhibitors, der in einem FRET basierten Screen entdeckt wurde, mittels NMR-Spektroskopie bestimmt. Interessanterweise kann dieser Inhibitor ATP aus der ATP-Bindestelle verdrängen. Der Hsp90:Aha1-Komplex ist ein interessantes Ziel für einen Inhibitor, da es wichtig für die Reifung von Mutanten des Cystic Fibrosis Transmembrane Conductance Regulators (CFTR) ist, der mit zystischer Fibrose zusammenhängt. Ein Effekt des untersuchten Inhibitors auf die Stabilität der CFTR Mutante ΔF508 konnte nachgewiesen werden.

Es wurden Effekte von Mutationen mit veränderter ATPase Aktivität auf die Funktion von Hsp90 untersucht. Es gab keine Korrelation zwischen ATPase Aktivität und der Fähigkeit der Mutanten das Überleben von Hefe zu gewährleisten. In dieser Arbeit, wird der Effekt der Mutanten auf die NTD von Hsp90 mittels NMR-Spektroskopie untersucht. Das Entfernen der ersten acht Residuen (Δ8) führt zu Unterschieden in der chemischen Verschiebung verteilt über die gesamte Domäne, was auf eine weitreichende strukturelle Veränderung hindeutet. Das passt du SAXS-Daten des Hsp90 Homodimers, in denen die Δ8 Variante im Vergleich zum Wildtyp bereits ohne ATP kompakter ist und auf Bindung von ATP hin noch kompakter wird. Die Δ8 Mutante kann Lebensfähigkeit in Hefe nicht gewährleisten. Die E33A Mutante, die nicht fähig ist ATP zu hydrolysieren, kann überraschenderweise die Lebensfähigkeit in Hefe gewährleisten. Durch NMR-Spektroskopie und SAXS konnte gezeigt werden, dass E33A auf die Bindung von ATP reagieren kann. Die Mutante D79N kann im Gegensatz dazu kein ATP binden und daher auch nicht seine Struktur ATP abhängig verändern und auch nicht die Lebensfähigkeit in Hefe gewährleisten. Die präsentierte Studie zeigt, dass nicht die Geschwindigkeit der ATPase Reaktion für die Funktionalität von Hsp90 entscheidend ist, sondern die Fähigkeit die Unterschiedlichen Zustände im Chaperonzyklus anzunehmen.

Beide Teile dieser Arbeit, die Untersuchung des Spleißfaktors RBM10 und des molekularen Chaperons Hsp90, nutzen die Fähigkeit der NMR Spektroskopie Proteine in Lösung zu untersuchen. Die komplexen Interaktionen der drei dynamischen RNA Bindedomänen von RBM10 wurde mittels NMR Relaxations- und PRE-Experimenten in Kombination mit SAXS untersucht. NMR Spektroskopie war auch die Methode der Wahl um die Bindestelle der Modulatoren des molekularen Chaperons Hsp90, die Hsp90 nur mit moderater Affinität binden, zu bestimmen und die Effekte der Mutationen in der N-terminalen Domäne von Hsp90 zu untersuchen, die während des Chaperonzykluses strukturelle Änderungen durchläuft und daher dynamisch ist.

Part 1

RNA recognition by the splicing factor RBM10

3 RBM10 Introduction

The regulation of gene expression defines, which proteins are expressed in each of our cells. And thereby defines how our tissues, organs and organism are build. Before the entire human genome was sequenced many scientists expected humans to have 100.000 or more genes (Pennisi 2003). So it is surprising that nowadays it is believed that there are only around 20.000 genes in the human genome (Clamp et al. 2007). How can such a small number of genes generate all the different cell and tissue types in the human organism? A possible explanation is the process of alternative splicing, which allows one gene to encode for different proteins. Thereby allowing the presence of different isoforms of a protein in different cell types or developmental stages (Ben-Dov et al. 2008; Kornblihtt et al. 2013). To understand this process first maturation of pre-mRNA to mature mRNA has to be discussed.

3.1 Splicing

After transcription the pre-mRNA has to be processed to become mature mRNA which is transported from the nucleus to the cytoplasm, where it serves as a template for the protein production through the ribosome. The processing includes the 5’ capping, where the first guanine is methylated, splicing, where non coding introns are removed and coding exons are joined together, and 3’ polyadenylation, where a polyadenyl tail is added to the 3’ end of the mRNA.

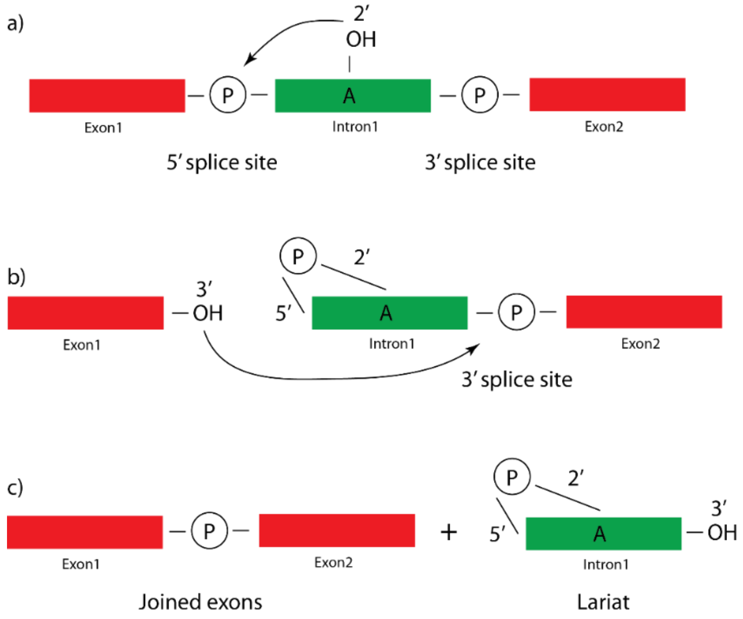


Figure 1 Splicing reaction: *a) During the first transesterification reaction the 2’ hydroxyl group of the branch point adenosine attacks the phosphodiester bond at the 5’ splice site. b) As a result a lariat is formed through a phosphodiester bond between the 5’ end of intron 1 and the branch point adenosine. The now free 3’ hydroxyl group of exon 1 attacks the phosphodiester bond in the 3’ splice site. c) As products the joined exons and the intron 1 with a free 3’ hydroxyl group and the lariat from b is formed.*

During splicing two consecutive transesterification reactions take place. During the first step the branch point adenosine's free 2' hydroxyl group cleaves the phosphodiester at the 5' splice site, which is the 3' end of exon 1 (**Figure 1A**). Thereby the lariat is formed, where the 2' hydroxyl group of the branch point adenosine is bound to 5' end of intron 1. Now the newly formed free 3' hydroxyl group of exon 1 attacks the phosphodiester bond at the 3' splice site (**Figure 1B**). The products of this second transesterification are the two joined exons and the free lariat, where the branch point adenosine is linked to the 5' end of the intron, while the 3' end of the intron has a free hydroxyl group (**Figure 1C**) (Berg, Tymoczko, and Stryer 2010).

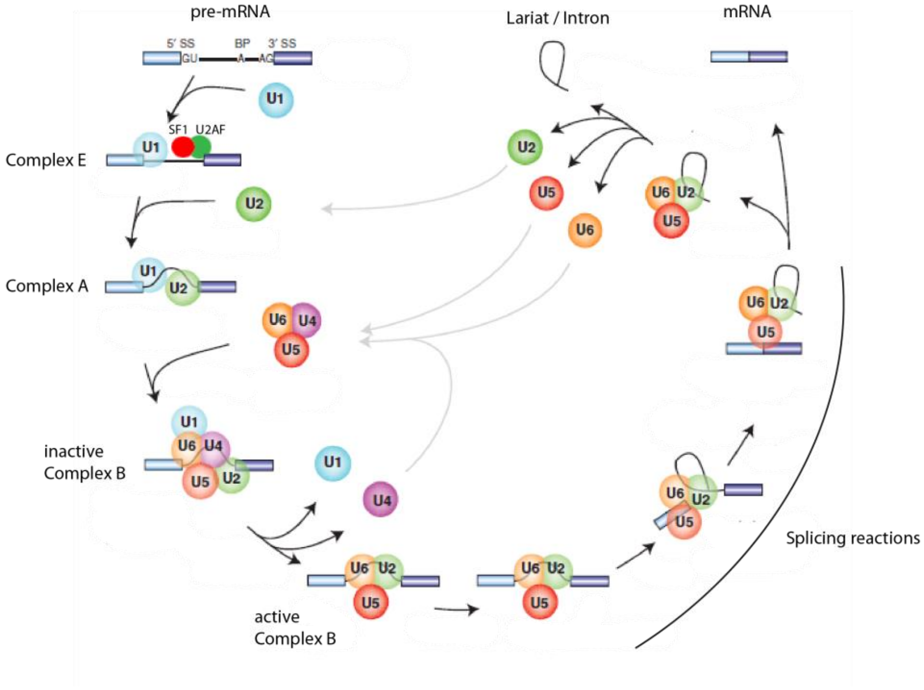


*Figure 2 Spliceosome assembly: U1 binds the 5' splice site of a pre-mRNA, SF1 binds the branch point and U2AF65 binds the polypyrimidine tract close to the 3' splice site forming the complex E. Then U2 interacts stably to the branch site and forms the complex A. The pre-formed tri-snRNP U4/U5/U6 interacts with complex A forming the inactive complex B. After conformational rearrangement U1 and U4 are released forming the active complex B, which performs than the two splicing reactions shown in **Figure 1**. Afterwards the spliceosome falls apart and releases the joined exons and the lariat. Adapted from (Will and Lührmann 2011).*

There are self-splicing introns, where the transesterification reactions are catalyzed by the RNA itself and there are introns which are dependent on the spliceosome. The spliceosome is a highly dynamic protein-RNA complex. It is build up from 4 snRNPs (U1,U2,U5,U4/U6), which are ribonuclear proteins containing small nuclear RNAs (snRNAs) and several proteins. The spliceosome assembles on the pre-mRNA at the splice sites. The mRNA harbors several cis elements at the 5'- and 3'-splice site, as well as at the branch point, which is usually located 18-30 nucleotides upstream of the 3'-splice site. In higher eukaryotes the branch point is followed by the poly-pyrimidine tract. The assembly starts with the binding of U1 to the 5' splice site, the binding of SF1 to the branch point and of U2AF to the polypyrimidine tract close to the 3’ splice site to form the complex E (**Figure 2**). Then U2 stably associates with the branch site forming the complex A. The preassembled U5/U4/U6 tri snRNP associates with the complex A, forming the pre-catalytic complex B. Major rearrangements of the conformation and release of the U1 and U4 snRNPs give rise to the active complex B. The active complex catalyzes then the two splicing reaction steps described before. Afterwards the spliceosome falls apart again and releases the joined exons and the lariat intron (Will and Lührmann 2011).

3.2 Alternative Splicing and the Splicing Code

Splicing is not only a simple processing step, where always the same sequences (introns) are removed from pre-mRNAs. Splicing is better described as an additional level of gene regulation. Through alternative splicing one gene can lead to different mRNAs and thereby to the expression of different proteins. In the human genome most genes are spliced alternatively (Ben-Dov et al. 2008). Thereby one gene can produce several different proteins, which can fulfill different functions. For example there are different splicing isoforms of a protein expressed in different tissues or at different developmental stages (Stolc et al. 2004). This means that the coding capacity of the human genome is much larger than the ca. 20.000 genes (Clamp et al. 2007) would suggest. Alternative splicing can be seen as a way to store information in the genome more efficiently and to regulate the read out of this information.

If a certain splicing event takes place depends on several factors. First the strength of the splice site is important. As described in Chapter 3.1, the 5’ splice site is recognized by the U1 snRNP and the 3’ splice site by U2AF, which binds to the polypyrimidine tract and U2, which interacts with the branch point. A stronger splice site e.g. would have a more complementary sequence to the U1 snRNA, or a longer polypyrimidine tract and would thereby bind the consecutive splicing machinery better (Hertel 2008). But it is not only the strength of the consecutive splice sites that defines if a splicing event takes place. In addition, there are cis-elements in the RNA, which can act as splicing enhancers or silencers. These elements can be encoded either in introns or exons. These cis elements are recognized by splicing factors (trans-factors), which bind to the RNA. E.g. often serine/arginine rich proteins (SR-proteins) bind to exonic splicing enhancers (ESE) and recruit the spliceosome machinery and thereby promote splicing. Heterogeneous nuclear RNPs (hnRNPs) in contrast bind to splicing silencers. They act by looping out exons or by blocking the binding of snRNPs (Hertel 2008) (**Figure 3**).

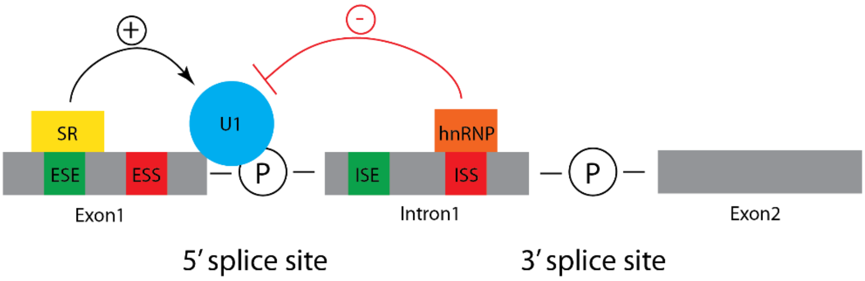


Figure 3 Schematic of regulation of alternative splicing: *pre-mRNA encodes cis-elements which can enhance or inhibit splicing. These sequences can be encoded in exons or in introns. They are called accordingly exonic splicing enhancer (ESE), intronic splicing enhancer (ISE), exonic splicing silencer (ESS) and intronic splicing silencer (ISS). These cis-acting sequencing are bound by splicing factors in trans. An example for activating splicing factors are serine/arginine rich (SR) proteins, which can recruit snRNPs and thereby promote splicing. hnRNPs are an example for inhibiting splicing factors. They are heterogeneous nuclear (hn) RNPs, which can block the binding of snRNPs.*

In practice not one single factor defines if a splicing event takes place. It is the combination of many different factors. These factors include the strength of the consecutive splice sites, the length of the intron, which cis-elements are coded in the mRNA and which trans-acting splicing factors are present. The combination of all these factors can be understand as a splicing code, which determines which final mRNA or mRNAs are generated from a particular gene in particular tissue (Barash et al. 2010). Despite the combinatory effect of many factors, it is important to understand the function of particular splicing factors on a structural level. In the following chapter the splicing factor RBM10 is presented in detail.

3.3 The Splicing Factor RBM10

The RNA binding motif 10 (RBM10) is a multidomain protein that forms together with its closely related homologs RBM 5 and 6 a family of splicing factors (Sutherland et al. 2005). All three proteins share the same domain architecture, with N-terminal RNA interacting and C-terminal protein interacting domains (**Figure 4**). RBM10 also has a nuclear localization signal (Y. Wang et al. 2013). The N-terminal RNA interacting domains comprise two RNA recognition motifs (RRM1 and RRM2) and a zinc finger (ZF1). RRM1 is connected to ZF1 by a short 4-residue linker, while the linker between ZF1 and RRM2 is 57 residues long. Single domain structures of the three RNA binding domains are available in the protein data base (RRM1 PDB: 2LXI, ZF1 PDB: 2MXV, RRM2 PDB: 2M2B), but there are no structures of the domains bound to RNA or structures elucidating the domain arrangement.

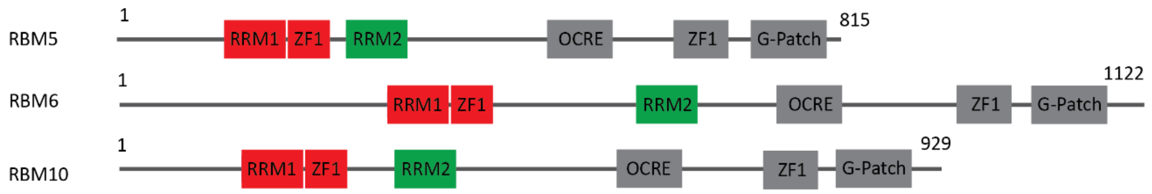

RBM5
1
RRM1
ZF1
RRM2
OCRE
ZF1
G-Patch
815
RBM6
1
RRM1
ZF1
RRM2
OCRE
ZF1
G-Patch
1122
RBM10
1
RRM1
ZF1
RRM2
OCRE
ZF1
G-Patch
929

Figure 4 Domain architecture of the homolog proteins RBM 5, 6 and 10: *The RNA binding domains RRM1, ZF1 and RRM2 are shown in red and green, while the C-terminal protein interacting domains are shown in gray.*

RBM10 exits in two isoforms, which differ in the presence or absence of residue V354, which is positioned in the N-terminus of α-helix 2 in RRM2. The two isoforms are generated through alternative splicing. The exon 10 intron 10 boundary is not sharp, so there are two alternative splicing products. A longer isoform (+V354) where exon 10 is three nucleotides longer, which code for the extra V354, and a shorter isoform (-V354), where these three nucleotides are spliced out with intron 10 (**Figure 5**) (Tessier et al. 2015). It has been speculated that this extra valine disrupts the α-helix 2 of RRM2 (Tessier et al. 2015).

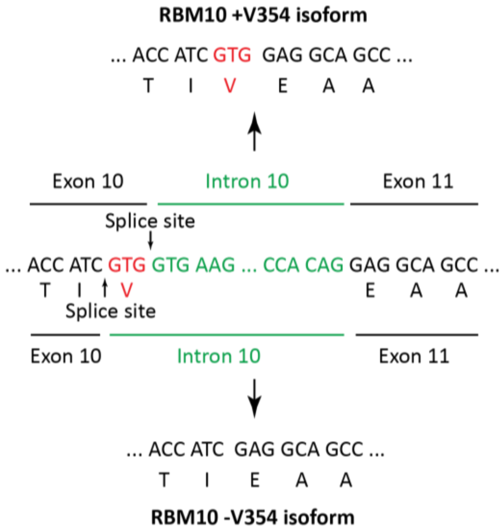

RBM10 +V354 isoform
... ACC ATC GTG GAG GCA GCC ...
T I V E A A
Exon 10
Intron 10
Exon 11
Splice site
... ACC ATC GTG GTG AAG ... CCA CAG GAG GCA GCC ...
T I V E A A
Splice site
Exon 10
Intron 10
Exon 11
... ACC ATC GAG GCA GCC ...
T I E A A
RBM10 -V354 isoform

Figure 5 RMB10 isoforms are generated through splicing: *Alternative splicing of the exon 10 intron 10 splice site leads to two isoforms of RBM10. The splice site is not sharp and can vary by the three nucleotides GTG (red), which are either included in exon 10, which leads to the longer +V354 isoform, or in intron 10, which leads to the shorter -V354 isoform (Tessier et al. 2015).*

The RNA interaction of RBM10 has been studied by crosslinking-immunoprecipitation (CLIP), RNA-Compete and Scaffold independent analysis. According to CLIP-Seq and PAR-CLIP data RBM10 is enriched upstream of the 5'- and downstream of the 3'-splice site and thereby is likely to be associated with splicing regulation(Bechara et al. 2013; Y. Wang et al. 2013). The CLIP-Seq data reveals two pyrimidine rich consensus motifs (CUCUGAACUC and CGAUCCCU) as most significant hits (Bechara et al. 2013). An analysis of the PAR-Seq data, does not find these motifs as statistically significant, instead an exonic purine rich motif (GAAGA) and an intronic pyrimidine rich (UUNCU) motif are identified (Maaskola and Rajewsky 2014).

ZF1 of RBM10 is a RanBP-type-ZF, which are known to interact with GGU motifs (Nguyen et al. 2011). RNA compete data for RBM10 found purine rich heptamers containing a GGA motif for RBM10, which is most likely recognized by ZF1. By isothermal titration calorimetry it was verified that RBM10 has a 3-fold higher affinity for GGA than for GGU (Collins et al. 2017). RRM1 and ZF1 cooperatively recognize a longer CUGUGGA RNA oligo. The main specificity for this oligo is provided by ZF1 (Collins et al. 2017). This heptamer was also discovered in the CLIP-Seq data, but only as the 4th hit (Bechara et al. 2013). Collins et al. also showed with scaffold independent analysis that the RRM2 of RBM10 recognizes a CCNC motif and by ITC that the RRM2 binds a CCCAC oligo with a K_D of 7 µM.

Multi-domain RNA binding proteins often recognize their targets through multiple weak interactions. The cooperativity of this interactions leads to high affinity and specificity. The domains are often flexible in respect to each other and rearrange upon RNA binding. The dynamics and solution domain organization is therefore of key importance to understand the mechanism of RNA recognition by multidomain RNA binding proteins (Mackereth and Sattler 2012; Mackereth et al. 2011). How and if RRM-ZF1 and RRM2 recognize RNA cooperatively remains so far unknown.

The biological importance of the splicing regulation of RBM10 has been shown by splicing assays (Bechara et al. 2013). RBM10 is involved in the alternative splicing of the NUMB gene. It promotes skipping of exon 9 in the NUMB gene which leads to a shorter gene product. Interestingly, the shorter and the longer NUMB gene product have opposite effects on Notch signaling. While the shorter product inhibits Notch signaling, the longer product leads to an overall downregulation of NUMB protein levels and thereby to activation of Notch signaling (Misquitta-Ali et al. 2011). NUMB mediated alterations of Notch signaling have been linked to cancer (Westhoff et al. 2009) (**Figure 6**).

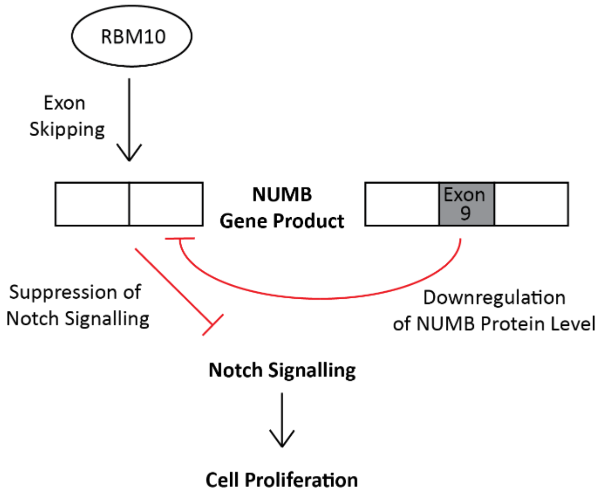


Figure 6 Effects of RBM10 on the alternative splicing of the NUMB gene and Notch signaling: *RBM10 induces skipping of exon 9 of the NUMB gene. The resulting shorter NUMB gene product suppresses Notch signaling and thereby cell proliferation. The longer NUMB gene product in contrast, downregulates the overall NUMB protein level and does not suppress Notch signaling.*

The involvement in the regulation of cell proliferation fits well to the proposed role of RBM10 in lung cancer. Several truncation mutations and the I316F point mutation have been identified in human lung cancer adenocarcinomas (Imielinski et al. 2012). In addition, the adenocarcinoma derived A549 cell line harbors the V354E mutation, which effects the residue that is present additionally in the described splicing isoform of RBM10. This mutation disrupts the function of RBM10 in exon skipping of the NUMB gene in splicing assays (Bechara et al. 2013; Hernández et al. 2016).

In addition to the role in cancer, RBM10 is linked to the developmental disease Talipes equinovarus, Atrial septal defect, Robin sequence, and Persistent left superior vena cava (TARP) syndrome. The TARP syndrome is a pre- or postnatal lethal syndromic form of cleft palate. It has been linked to mutations in RBM10 downstream of the RNA binding domains (Johnston et al. 2010). In a milder form of the TARP syndrome, a C-terminal truncated form of RBM10 is expressed, which lacks the C-terminal nuclear localization site and shows less nuclear localization (Y. Wang et al. 2013).

3.4 Scope of the RBM10 project

The goal the thesis is to understand how RBM10 recognizes pre-mRNA and to study if there is cooperativity between RRM1-ZF1 and RRM2, which are separated by a long linker. As a model system the alternative splicing of exon 9 of the NUMB gene is used. RNA binding of a 12-mer RNA oligo derived from the sequence used in splicing assays (Bechara et al. 2013) will be used. For this goal the two parts, RRM-ZF1 and RRM2, will be first characterized independently. The structure of the two parts will be solved using a classical nOe-based NMR structure calculation approach. RNA binding will be characterized with NMR titrations, isothermal titration calorimetry and small angle light scattering. It will be attempted to gain structural information of the protein RNA complex through intermolecular nOes between the RNA binding domains and the RNA.

The interaction of the three-domain construct with the same NUMB derived 12-mer RNA-oligo will be characterized with the same set of methods. In addition the domain arrangement and dynamics of all RNA binding domains will be studied by ^{15}N-NMR-relaxation experiments, paramagnetic relaxation enhancement (PRE) NMR measurement, residual dipolar coupling (RDC) NMR experiments and small angle X-ray scattering (SAXS). With the same set of experiments the changes in dynamics and the domain rearrangement upon RNA binding will be characterized.

For the generation of a structure model of the RBM10-RNA complex the structures of RRM1, ZF1 and RRM2 shall be used alongside with orientation information derived from RDCs, long range distance information from PREs and shape information from SAXS to build a structural model (**Figure 7**).

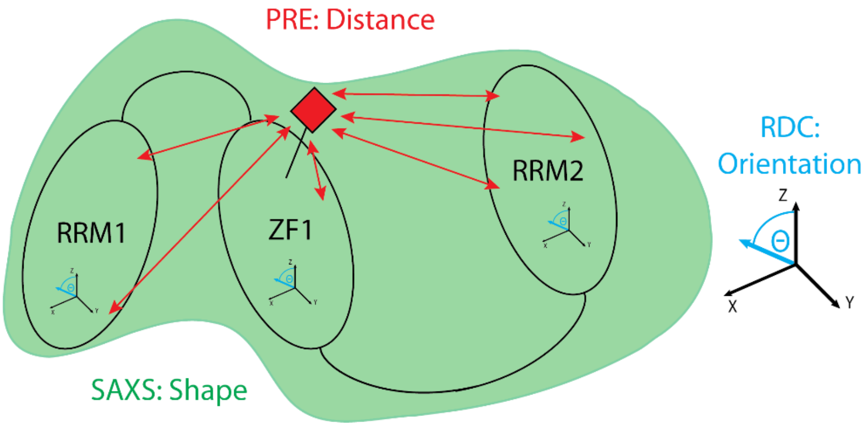


Figure 7: Restraints for structure calculation for RBM10 RRM1-ZF-RRM2: *The structure of the all three RNA binding domains shall be determined using distance restraints from PRE data (red), orientation restraints from RDC data (blue) and shape information from SAXS (green).*

In addition, the structural and functional difference of the RBM10 +V354 isoform as well as the impact of known disease linked point mutations in the RNA binding domains (e.g. V354E), will be characterized.

These findings about the function of the RNA binding domains and the effect of known disease mutations will help to better understand RBM10s function in the regulation of alternative splicing and its implication in diseases like lung cancer and the TARP syndrome.

4 Background of Applied Methods

4.1 Basics of NMR

NMR spectroscopy is based on the energy differences of nuclear spins in a static magnetic field (B_0). In the static magnetic field there is an energy differences depending on the orientation. The energy differences (E) is given by the magnetic moment of the spins in the direction of the static magnetic field along the z-axis (μ_z).

$$E = -\mu_z \cdot B_0$$

The magnetic moment of the spins are given by the spin quantum number in the z-direction (m_z), the gyromagnetic ration (γ), which is a constant for the used element, and the reduced Planck constant (ħ).

$$\mu_z = m_z \cdot \hbar \cdot \gamma$$

Usually for NMR studies nuclei with a spin quantum number I of ½, such as ^{1}H, ^{13}C and ^{15}N, are used. These have to states α (s=+½) and β (s=-½). The energies of these states are:

$$E_\alpha = -1/2 \cdot \hbar \cdot \gamma \cdot B_0$$

$$E_ß = 1/2 \cdot \hbar \cdot \gamma \cdot B_0$$

The energy difference between the α and β state is ΔE.

$$\Delta E = \hbar \cdot \gamma \cdot B_0$$

The population ratio of the α- and the β-state (N_β/N_α) is given by the Boltzmann distribution.

$$\frac{N_\beta}{N_\alpha} = e^{-\frac{\Delta E}{k_B T}}$$

Only the number of spins which make the difference in population participate in the NMR experiment, therefore the higher the energy difference the more sensitive the NMR experiment becomes because more nuclear spins participate in the experiment. Since the energy differences depents on the strength of the magnetic field, the strength of the magnetic field influences the sensitivity. The energy difference is also influenced by the gyramagnetic ratio. This explains why ^{1}H nuclei are more sensitive than ^{13}C (γ_{13C}≈1/4 γ_{1H}) and ^{15}N (γ_{15N}≈1/10 γ_{1H}) nuclei.

To be able to change spin states the sample hast to be irradiated with a wavelength which is in resonance with the spin precision and therefore has the same energy as the energy difference between the two states. This frequency is called the Larmor frequency (ν_L).

$$\Delta E = h \cdot \nu_L$$

$$\hbar \cdot \gamma \cdot B_0 = h \cdot \nu_L$$

$$\frac{\gamma}{2\pi} \cdot B_0 = \nu_L$$

The Larmor frequency does not only depend on the kind of atom, but also on the chemical environment of the atoms. Electrons interact with the static magnetic field and shield the nuclei. Basically the static magnetic field induces an electric current in the valence electrons, which induce a weaker magnetic field $(B_0 \cdot \sigma)$ which works against the static magnetic field. σ is the shielding constant, which depends on the chemical environment and is independent of the magnetic field. In conclusion, the nuclei experience a weaker effective magnetic field (B_{eff}) and thereby have a different Larmor frequency.

$$B_{eff} = B_0 - B_0 \cdot \sigma$$

Instead of reporting the different frequencies, which are field dependent, chemical shift values (δ) in parts per million (ppm) are used, which are normalized to a standard (ν_{ref}), e.g. for ^{1}H and ^{13}C nuclei the frequencies of tetramethylsilane.

$$\delta = \frac{\nu - \nu_{\mathrm{ref}}}{\nu_{ref}} \cdot 10^6$$

Nuclei experience spin-spin or J-couplings through covalent bonds. These couplings can be usually observed for nuclei which are three or less bonds apart from each other. The coupling is mediated by the binding electrons, which form the bonds. The coupling splits up the NMR signal in multiplets. If a nuclear spin is coupled to one other spin, the other spin can be either in the α- or in the β-state. Therefore, the corresponding peak gets split up in two peaks, which are separated by the J-coupling in hertz. The J-coupling is field independent. If a spin is coupled to two other equivalent spins, the two other spins can either both be in the α-state (αα), spin 1 can be in the α- and spin 2 in the β-state (αβ), or vice versa (βα), or both spins can be in the β-state (ββ). Since the αβ and the βα states are energetically equivalent, the peak will be split in a triplet with the intensity ration 1:2:1. If there are n spins, there will be n+1 peaks and the intensity will be distributed according to the binomial series. J-couplings can be used to transfer magnetization between nuclei, which is used in multidimensional NMR experiments, such as the COSY and the HSQC (Keeler 2010).

4.2 Relaxation

The process of relaxation describes the return of the system to its initial state. Relaxation can be divided in two processes. T1 relaxation describes the recovery of the initial magnetization along the z-axis. This type of relaxation is driven by local magnetic fields that have transverse components which fit the Larmor frequency and thereby can change the spin state between α and β. Through this exchange between states the system is driven back to equilibrium.

T2 relaxation describes the loss of coherence in the transverse plane and thereby the loss of signal. This can be caused by the same local fields in the transverse plane which fit the Larmor frequency and thereby locally disturb the coherence. This is called the non-secular contribution to transverse relaxation. In addition, local fields from the z-direction can make the static magnetic field locally a bit stronger or weaker and thereby influence the precision frequency in the transverse plane. Because this effects are local, they effect the sample in different locations of the sample to different extend and thereby lead to loss of coherence in the transverse plane.

For nuclei with spin quantum number of ½, like ^1H, ^{13}C and ^{15}N, there is two main origins for these local magnetic fields. The first are dipolar couplings. Nuclear spins have a dipolar moment which interact with each other through space. This dipolar coupling is usually not observable in solution NMR, because it gets averaged out through the isotropic tumbling, but it still contributes to relaxation. Dipolar couplings will be discussed in more detail in chapter 4.10. The second cause for the local magnetic fields is chemical shift anisotropy. As described before electrons interact with the static magnetic field and shield the nuclei against it. The shielding magnetic field of the electrons depends on the orientation of the molecule in the static magnetic field. This anisotropy of the chemical shifts gives rise to local magnetic fields and thereby contributes to relaxation.

Relaxation correlates with the molecular motion of the molecule in solution. The molecular tumbling can be described by an autocorrelation function.

$$G(\tau) = \overline{B_{loc}{}^2} exp\,(\frac{-|\tau|}{\tau_C})$$

The rotation of an spherical molecule in solution can be described by the Stokes law with the correlation time (τ_C), where η is the viscosity and r_H is the hydrodynmaic radius, k_B is the Boltzmann constant and T is the temperature.

$$\tau_c = \frac{4\pi\eta {r_H}^3}{3k_B T}$$

The Fourier transformation of the autocorrelation functions results in the spectral density function, which quantifies the presence of local magnetic fields of certain frequencies, which depends on the motion of the molecule.

$$J(\omega) = \overline{B_{loc}{}^2} = \frac{2\tau_C}{1 + \omega^2 \tau_C{}^2}$$

For T1 relaxation and the non-secular part of T2 relaxation mainly frequencies around the Larmor frequency are relevant, because they lead to a transition between the α- and the β-states, which requires radiofrequencies in resonance with the spin precision. The secular contribution to T2 relaxation in contrast, is dependent on frequency contributions around zero frequency. While large molecules tumble slowly and thereby present mainly slow frequencies, smaller molecules have a more even distribution of frequencies, with faster frequencies of rotational motion present. In this way relaxation is directly coupled to the molecular motion. Therefore, it is possible to study the molecular motion with relaxation experiments.

The T1 relaxation time can be measured in an inversion recovery experiment. The equilibrium magnetization along +z is altered by a 180° pulse to magnetization along -z. In a delay t, which is varied in length, the magnetization relaxes back to equilibrium magnetization. After the delay the magnetization is turned to the transverse plane by a 90°. The magnetization along the -z axis decays exponentially and the time constant can be detected by plotting the peak intensity against the varied delay t (Keeler 2010).

The T2 relaxation time can be measured by a Hahn-Spin-Echo. The magnetization is turned to the transverse plane by a 90° pulse. Then there is a delay t with a 180° pulse in the middle. After the delay the transverse magnetization is detected. The 180° pulse in the middle refocused chemical shift evolution on the nuclei, but not the loss of coherence through T2 relaxation. By plotting the delay length t against the peak intensity and fitting the graph with an exponential decay function the T2 relaxation time can be determined. In relaxation experiments instead of a single 180° pulse often an CPMG experiment is used (Cavanagh et al. 2006)

4.3 The 1D NMR experiment

With on resonance pulses the magnetization can be manipulated. With an 90° pulse along x or y the magnetization can be transferred from the equilibrium magnetization along the z-axis, parallel to the static magnetic field, in the transverse plane (xy-plane), where the magnetization can be detected. In modern Fourier-transform-NMR brought band pulses are used to excite a wide range of frequencies at the same time, which are than detected simultaneously. Recorded is a free induction decay (FID), which is an overlay of cosine and sine functions of different frequencies with an exponential decay. The FID has to be Fourier transformed to the frequency domain. The Fourier-transformation analyses which frequencies are present in the FID and as a result gives a spectrum, where at each frequency which is present in the FID appears a peak. Due to the exponential decay this peak has a lorentzian line shape (Keeler 2010).

4.4 Multidimensional NMR

In biomolecular NMR 1D experiments usually show a lot of peak overlap, which makes it impossible to analyze spectra of even small proteins. 2D NMR experiments can help to solve this peak overlap problem. A common 2D NMR experiments is the $^{1}H,^{15}N$ heteronuclear single quantum coherence ($^{1}H,^{15}N$-HSQC) experiment. This experiment gives a cross peak for every ^{15}N bound proton and thereby yields to a fingerprint spectrum of proteins, where there is one peak for every backbone amide, except for prolines. In addition, side chain amides of glutamine, asparagine, tryptophan and arginine appear also in the $^{1}H,^{15}N$-HSQC. Because the natural predominant nitrogen isotope is ^{14}N, samples have to be prepared in minimal medium with ^{15}N-isotopes as only nitrogen source (See also chapter 5.2). The chemical shifts observed in a $^{1}H,^{15}N$-HSQC are very sensitive to the changes in the chemical environment, therefore the $^{1}H,^{15}N$-HSQC is an excellent tool to observe ligand binding in NMR titrations.

During the pulse sequence of an HSQC, proton magnetization is exited with a 90° pulse, then the magnetization is transferred via an INEPT sequence to the hetero nucleus. During the INEPT the heteronuclear J-couplings are allowed to evolve, which leads to antiphase magnetization which is still transverse on the protons. Two 90° pulses on proton and nitrogen each, than convert this antiphase magnetization transverse on protons, to antiphase magnetization transverse on nitrogen. The generated transverse magnetization is than allowed to evolve chemical shift during the delay t1. After the delay the magnetization is transferred back to protons by a second INEPT sequence. Finally protons are directly detected like in a 1D experiment in the time dimension t2.

The t1 delay is increased in length by a fixed increment for each point in the second dimension, which leads to a set of 1D experiments with different t1 delay. The different t1 delays are the second time axis in the 2D experiment. The spectrum then arises from Fourier transforming the two time dimension, the direct of the 1Ds (t2) and the indirect from the series of experiments with increasing t1 delay. (Keeler 2010)

4.5 NMR resonance assignment

For the analysis of NMR spectra it is essential to assign each atom to its resonance frequency. The assignment is usual divided in backbone assignment, where the backbone amide nitrogen residues in the $^{1}H,^{15}N$ correlation spectrum or $^{1}H,^{15}N$-HSQC spectrum are assigned by triple resonance 3D-NMR experiments, and side chain experiments, where the rest of the protons are assigned.

4.5.1 Backbone assignment

The basis of backbone assignment is the $^{1}H,^{15}N$-HSQC (see also chapter 4.4). In the $^{1}H,^{15}N$-HSQC there is one peak for every residue, corresponding to the frequencies of the amide hydrogen and nitrogen in the backbone that lead to a two-dimensional spectrum (**Figure 8A**). Only proline residues do not appear in the $^{1}H,^{15}N$-HSQC, since there is no amide hydrogen present. The side chain of tryptophan gives one additional peak and the side chain amides of asparagine and glutamine give two additional peaks with identical nitrogen, but different proton frequencies. Some arginine side chains appear as negative peaks in the spectrum, but many of them are too much broadened by exchange.

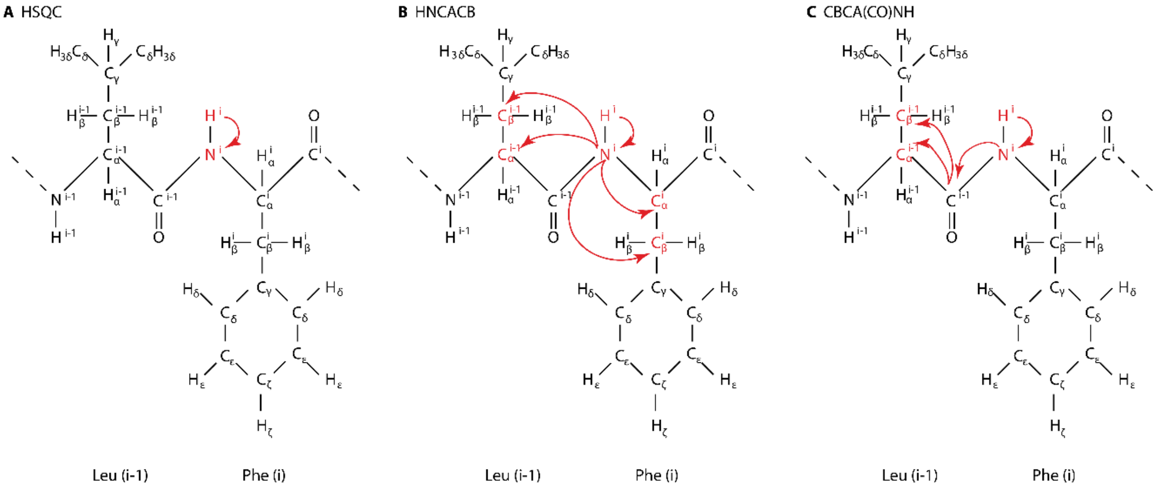


Figure 8 Backbone assignment experiments: *Shown are structural formula of a dipeptide of leucine and phenylalanine. A) The hydrogen and nitrogen of one backbone amide are highlighted in red. A red arrow indicates the transfer of magnetization between the atoms during a* $^{1}H,^{15}N$*-HSQC experiment. B) In addition to the amid atoms from A, the 4 carbon atoms in the same (i) and the previous residue (i-1), which give rise to peaks with the same nitrogen and hydrogen frequencies in an HNCACB spectrum, are highlighted in red. C) Only the two carbon atoms of the previous residue, which give rise to peaks in an CBCA(CO)NH spectrum are highlighted in red in addition to the amide.*

The sequential assignment of the backbone amides is done with two triple resonance experiments (Sattler, Schleucher, and Griesinger 1999). In the HNCACB experiment the magnetization is transferred from the amide proton to the nitrogen and then to the Cα and the Cβ of the same residue (i) and the previous residue (i-1). For each peak in the $^{1}H,^{15}N$-HSQC there are 4 peaks in the HNCACB, with the hydrogen frequency of the amide hydrogen in one dimension, the amide nitrogen frequency in the second dimension and the carbon frequency of either the Cα or the Cβ of the same residue or the Cα or the Cβ of the previous residue in the third dimension (**Figure 8B**). The sequential assignment is done now by making connections between the Cα and Cβ frequencies, which appear twice. Once in the carbon strip in the position in the ^{1}H,^{15}N-HSQC of the residue itself and a second time in the position of the following residue. In contrast, the CBCA(CO)NH experiments transfers the magnetization from the amide nitrogen via the C' only to the Cα and the Cβ of the previous residue (i-1) and therefore leads only to two peaks which have the hydrogen and the nitrogen frequencies of residue i and the carbon frequency of either the Cα or the Cβ of the previous residue i-1 (**Figure 8C**). Superimposing the CBCA(CO)NH with the HNCACB helps to distinguish the peaks coming from the own and the previous residue and in case of overlap allow a more accurate determination of the Cα and the Cβ frequencies. The higher sensitivity of the CBCA(CO)NH can also help to assign weak peaks. To map this connections to the actual sequence, pairs of amino acid types with unambiguous chemical shifts are helpful. E.g. glycine residues have a Cα chemical shift around 45 ppm and alanine has a Cβ chemical shift around 20 ppm.

4.5.2 Side chain assignment

For structural calculation an almost complete assignment of all hydrogen atoms in protein is necessary, therefore the side chain hydrogens have to be assigned. To assign the aliphatic carbon bound hydrogens in a ^{1}H,^{13}C-correlation experiment, a combination of an HBHA(CO)NH and two HCCH-TOCSY spectra is used. In the HBHA(CO)NH spectrum, similar to the CBCA(CO)NH spectrum, two peaks per peak in the ^{1}H,^{15}N-HSQC are visible. Each peak has the hydrogen and nitrogen frequencies of the amide, like in the ^{1}H,^{15}N-HSQC, and either the hydrogen frequency of Hα or Hβ of the previous residue in the third dimension (**Figure 9A**). Combined with the Cα and the Cβ frequencies from the CBCA(CO)NH, the peaks of HαCα- and HβCβ-correlations can be localized in the ^{1}H,^{13}C-HSQC. The other side chain atoms can be assigned from this starting points with the help of two HCCH-TOCSYs (Sattler, Schleucher, and Griesinger 1999). Both TOCSY experiments are based on the ^{1}H,^{13}C-HSQC and show in the third dimensions either, in case of the hCCH-TOCSY, all the carbon frequencies of the other proton bound carbons in the same residue (**Figure 9B**), or, in case of the HcCH-TOSY, all the hydrogen frequencies bound to a carbon in the same residue (**Figure 9C**). The magnetization transfer in a TOCSY is induced by a spinlock and therefore is more wide range than the INEPT transfer from the previous described experiments.

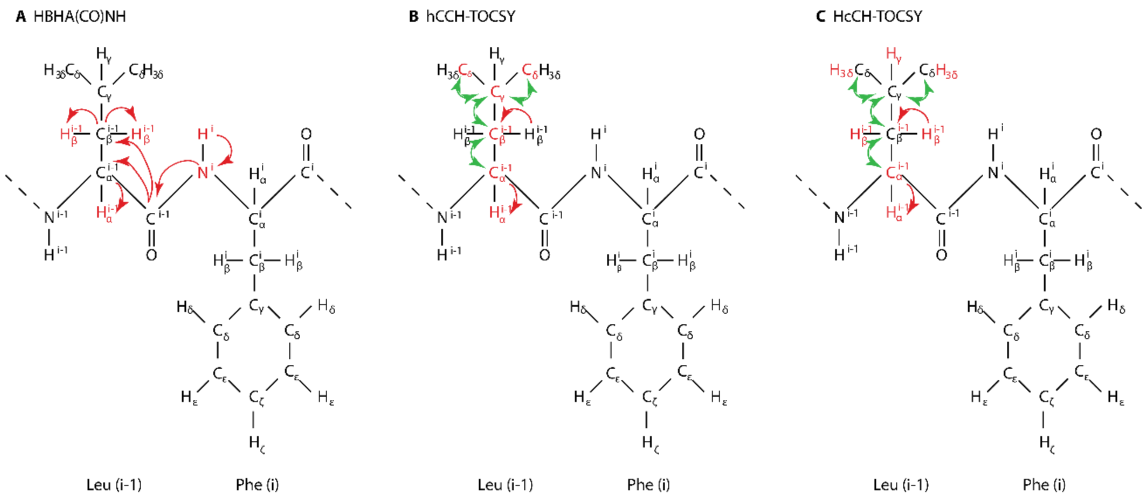


Figure 9 Experiments for aliphatic side chain assignment: *A) Magnetization transfer in the HBHA(CO)NH is indicated by red arrows. The amide hydrogen and nitrogen and the Hα and Hβ of the previous residue are highlighted in red. B) Magnetization transfer for one single peak in an hCCH TOCSY experiment are highlighted by red arrows. In addition the magnetization transfer during the TOCSY spin lock is indicated by green two sided arrows. The atoms which contribute to the correlations are highlighted in red. C) The HcCH-TOCSY is presented as in B.*

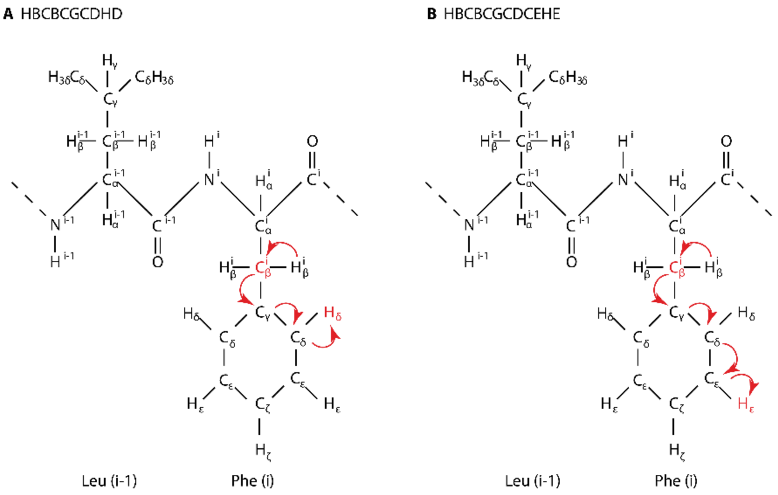


Figure 10 Experiments for aromatic side chain assignment: *Magnetization transfer pathways and atoms whose frequencies contribute to the in a HBCBCGCDHD (A) and an HBCBCGCDE (B) spectra are indicate by red arrows or red written atoms, respectively.*

For the assignment of Hδ and Hε of tyrosine and phenylalanine residues two 2D-experiments are very helpful. The HBCBCGCDHD and the HBCBCGCDCEHE experiments correlates the Cβ frequency of an aromatic residue with its own Hδ and Hε frequency respectively and therefore allows the assignment of the Hδ and Hε (Yamazaki, Forman-Kay, and Kay 1993). These assignments are confirmed with the help of a ^{13}C-edited aromatic NOESY spectrum. The ^{13}C-edited NOESY spectrum is again based on the ^{1}H,^{13}C-HSQC and therefore shows the frequencies hydrogen and carbon correlation in two dimensions and in the third dimension all hydrogen residues in a radius of approximately 5 Å around the carbon bound hydrogen (The nOe is discussed in Chapter 4.7). All other aromatic side chain residues have to be assigned by the ^{13}C-edited aromatic NOESY spectrum.

4.6 Ligand binding and exchange regimes

NMR is widely used to study protein-ligand-interactions. The interaction of a protein with a ligand can be described by the following reaction.

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} PL$$

The rate of the complex association is k_{on} and the rate of the dissociation is k_{off}. The total exchange rate (k_{ex}) is the sum of k_{on} multiplied with the ligand concentration ([L]) and k_{off}.

$$k_{ex} = k_{on} \cdot [L\,] + k_{off}$$

In solution state NMR there are three exchange regimes, which depend on kex and the frequency difference (Δω) between the apo-protein (ω_P) and the protein-ligand-complex (ω_{PL}).

$$\Delta\omega = \omega_P - \omega_{PL}$$

If k_{ex} is much smaller than Δω the exchange process is in the slow regime ($k_{ex} << \Delta\omega$), which means that both states can be observed separately in the NMR spectrum. In case half of the protein would be saturated with the ligand, one would see two sets of peaks with equal peak volume. In the opposite case, when k_{ex} is much larger than Δω the two states interconvert to fast to be distinguished by the NMR experiment and the exchange process is in the fast regime ($k_{ex} >>\Delta\omega$). In this case only one set of peaks appears in the spectrum. The position of this averaged peak is between the frequency position of the apo-protein and the protein-ligand-complex. The position is proportional to the amount of bound ligand. Therefore, if half of the protein is saturated the peak is exactly in the middle of the frequency position of the apo-protein and the protein-ligand-complex. If k_{ex} is compareable to Δω the exchange process is in the intermediate regime ($k_{ex} \approx \Delta\omega$). Here the signal of the protein broadens upon addition of ligand until it disappears completely when half of the protein is saturated with ligand. When adding more protein the signal reappears in the frequency position of the complex.

Changes in chemical shift in an ^{1}H,^{15}N correlation spectrum can be used to identify binding sides of the ligand on the protein. A series of ^{1}H,^{15}N correlation spectra of an ^{15}N-isotope-labelled protein with increasing ligand concentration is recorded. When the ligand is in fast exchange the complex can be often directly assigned by tracing the peaks of the titration series, otherwise the complex has to be assigned again by backbone assignment experiments. From the titration series the chemical shift perturbation (CSP) of each backbone amide can be calculated from the peak positions of the apo-protein and the peak positions of the ligand saturated protein. The CSP is calculated from the difference in chemical shift in the proton and the nitrogen dimension. Since protons have a 10 times higher gyromagnetic ratio than nitrogen, the differences in nitrogen chemical shift has to be divided by 10 to match the differences in proton chemical shift. To avoid that differences with opposite sign chancel, the differences are squared before adding them and then the square root of the sum is taken.

$$CSP = \sqrt{\left(\delta_{H_P} - \delta_{H_{PL}}\right)^2 + \left(\frac{\delta_{N_P} - \delta_{N_{PL}}}{10}\right)^2}$$

The binding site of the ligand can be determined by mapping the CSP on the structure and looking for affected regions. One has to keep in mind that the CSP just reflects a change in the chemical environment of the backbone amide, which can either be caused directly by the ligand or indirectly through conformational changes of the protein. Therefore one has to be careful when using CSP for determining ligand binding sites or as restraints for computational docking of the ligand (Cavanagh et al. 2006).

4.7 Nuclear Overhauser Effect

The nuclear Overhauser effect (nOe) originates from the crossrelaxation between two spins, which depends on dipolar coupling between the two spins. In a two spin system in addition to the auto relaxation, where one spin changes the state from α to β or vice versa, either both spins can change spin states from αα to ββ (W2) or they exchange states with another (αβ to βα or vice versa) (W1). The change of two spin states needs frequency contribution of 2ω ($\omega_1+\omega_2 \approx 2\omega$), while the interchange for 2 nuclei of the same kind depends on frequency contributions around zero frequency ($\omega_1-\omega_2 \approx 0$). The transfer rate from spin 1 to spin 2 is the crossrelaxation rate σ_{12} and depends on the difference of the rates W2 and W0. μ_0 is the permeability of vacuum.

$$W_2 = \frac{1}{20}\left(\frac{\mu_0 \gamma_1 \gamma_2 \hbar}{4\pi r^3}\right)^2 J(2\omega)$$

$$W_0 = \frac{3}{10}\left(\frac{\mu_0 \gamma_1 \gamma_2 \hbar}{4\pi r^3}\right)^2 J(0)$$

$$\sigma_{12} = W_2 - W_0$$

Since the transfer rate (σ_{12}) is the difference of the rates W1 and W2 it can be positive or negative. As discussed before the contributions in the spectral density function depend on the correlation time of the molecule. Large molecules have mainly contributions at zero frequency, which means because of the higher factors that W0 is larger than W2 and therefore the rate σ_{12} is negative. For small molecules the frequency distribution is more equal and W2 is larger than W0 and therefore the rate σ_{12} is positive. For molecules of medium size, where W2 is equal to W0 the rate becomes zero and there is no nOe observable. The cross relaxation rate depends on the distance and is proportional to $1/r^6$. That means that the nOe effects falls of quickly with rising distance between two nuclei. For two protons it is usually only observable up to a distance of 5 Å.

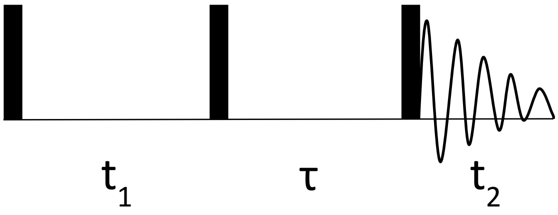


Figure 11 Pulse sequence for 2D-NOESY experiment: *Filled rectangles represent 90° pulses.*

NOESY cross peaks between protons can be measured in a 2D NOESY experiment (**Figure 11**). The first 90° pulse creates transverse magnetization, which gets frequency labelled during the t1 period. The second 90° pulse than creates z-magnetization. Now during the mixing time τ magnetization gets transferred proportional to σ_{12} from spin 1 to spin 2. The third 90° pulse than flips the magnetization to the transverse magnetic plane, where chemical shift evolves again. Magnetization, which did not get transferred from spin 1 to spin 2 during the mixing time τ, remains on the same nucleus during t1 and t2 and therefore has both dimensions the frequency of spin 1. This magnetization gives rise to the diagonal peak of spin 1. Accordingly, there is also a diagonal peak for spin 2. In contrast, the magnetization which was transferred from spin 1 to spin 2 gives rise to cross peaks. In t1 it gets coded with the frequency of spin 1, than magnetization is transferred to spin 2 during the mixing time and during t2 the chemical shift evolves with the frequency of spin 2.The opposite transfer from spin 2 to spin 1 gives rise to a second cross peak, which is mirrored through the diagonal to the other cross peak (**Figure 12**). Since the amount of transferred magnetization during the mixing time depends on σ_{12}, which is distance dependent, the intensity of the cross peaks contain information about the distance between the two nuclei (Keeler 2010).

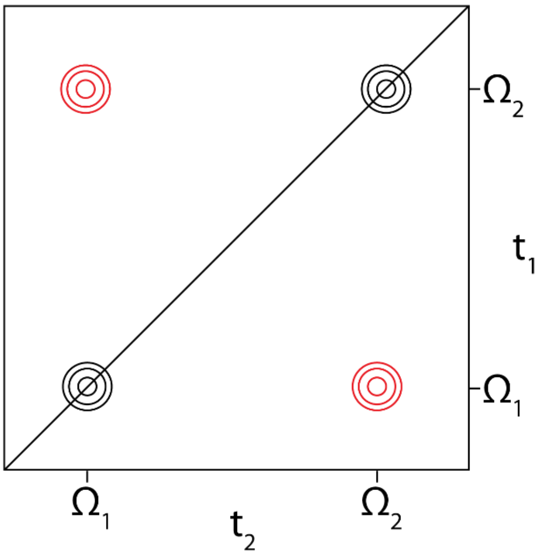


Figure 12 Schematic for 2D-NOESY spectrum: *Diagonal peaks of spin 1 and 2 are plotted in black, while cross peaks between spin 1 and 2 are plotted in red.*

To simplify the spectra and two reduce spectral overlap for proteins usually 3D ^{15}N- or ^{13}C-edited NOESY spectra are recorded (Sattler, Schleucher, and Griesinger 1999). In this case the basis for the NOESY spectrum is either the ^{1}H,^{15}N- or the ^{1}H,^{13}C-correlation spectrum. For each peak in the 2D heteronuclear correlation spectrum there is a strip of NOESY cross peaks in the third dimension with nuclei less than 5 Å away. In the edited proton dimension only protons which are bound to the corresponding hetero nucleus give signals, while in the non-edited proton dimension all protons, which are close in space to these, give rise to cross peaks.

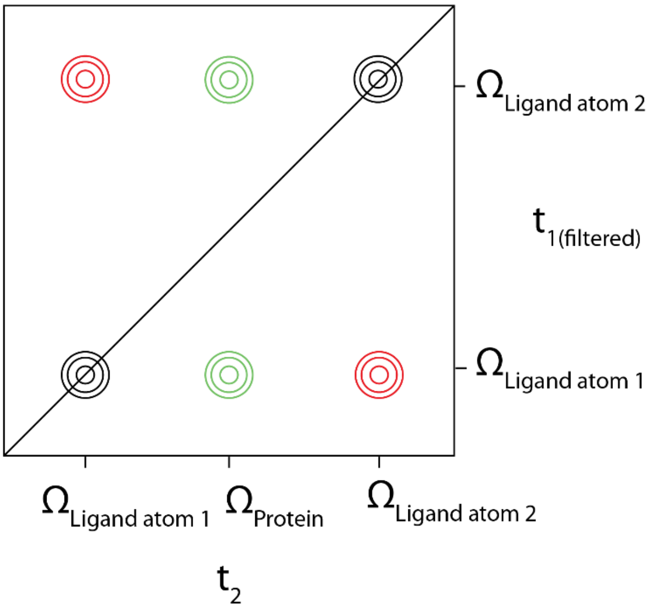


Figure 13 Schematic of an isotope filtered NOESY experiment: *In the example shown here the t1 dimension has a ^{13}C- and ^{15}N-isotope filter, therefore all signals from the ^{13}C and ^{15}N labelled protein get filtered out from this dimension. There are no intramolecular cross peaks of the protein visible. The signals of the unlabeled ligand in contrast appear in both dimensions and show diagonal peaks (black) and intramolecular cross peaks axis symmetric on both sides of the diagonal (red). Since*

protein signals show up only in t2 dimension, intermolecular cross peaks between the protein and the ligand show up only on one side of the diagonal (green). In this example all three nuclei (ligand atom 1 and 2 and protein) are in close enough distance to each other to give rise to cross peaks.

To identify intermolecular nOes between a protein and a ligand it is useful to record isotope filtered NOESY experiment (Zwahlen et al. 1997). An isotope filter basically has the opposite effect as isotope editing. In a filtered dimension of an experiment all nuclei bound to the filtered heteronucleus do not give rise to NMR signals. In a 2D NOESY with an ^{13}C- and ^{15}N-isotope filter in the t1 dimension of an isotope labelled protein and an unlabeled ligand, the intramolecular NOE crosspeaks of the protein do not appear in the spectrum, the intramolecular NOE crosspeaks of the ligand appear on both sides of the diagonal and the intermolecular NOE cross peaks between the protein and the ligand appear only on one side of the diagonal. Basically protein nuclei can appear only in the t2 dimension, while ligand nuclei can appear in both dimensions. Isotope filters can also be introduced to 3D edited NOESY experiments.

4.8 Structure Calculation

4.8.1 nOe based structure calculation of small protein domains

Classical NMR structure calculation is mainly based on nOe distance restraints. It is necessary to assign back-bone and side-chain protons as complete as possible. The assignment of nOe cross peaks can be done automated together with the structure calculation. In this work the program Cyana was used to do so. Cyana needs the sequence of the protein, a shift list containing all the assignments, peak lists of the NOESY spectra with the positions and the volume of the peaks. Torsion angles can be used as additional restraints. Torsion angles can be calculated from chemical shifts using the program TALOS (Shen et al. 2009). In addition, upper and lower distance restraints can be used to indicate known structural elements, like H-bonds or Zn^{2+} coordination e.g. in ZF1 of RBM10.

Cyana runs then in seven cycles of nOe assignment and structure calculation, followed by a final cycle of structure calculation where only unambiguous distance restraints are used for the calculation. The assignment of nOe cross peaks depends on three probabilities: The shift, structure and network probability. The product (P_{tot}) of this three probabilities has to be higher than a threshold value (P_{min}) to make a valid assignment.

$$P_{tot} = P_{shifts} \cdot P_{structure} \cdot P_{network} \geq P_{min}$$

The shift probability compares the frequencies of a nOe cross peaks to the chemical shift considering the given tolerance. The structure probability evaluates based on the structure from the previous run how well the made assignment fits to the structure. The network probability finally scores how well the nOe cross peak assignment fits into a network of other nOe cross peak assignments independent of the calculated structure (Peter Güntert and Buchner 2015). The structure calculation is based on the assigned nOe cross peaks and the additional given restraints and consists of an simulated annealing with torsion angle dynamics (P. Güntert, Mumenthaler, and Wüthrich 1997). Cyana does not use explicit water in the calculation, for a water refinement other structure calculation programs like Aria/CNS can be used (Rieping et al. 2007). In the end the structure should be verified by looking at the torsion angles for example with the program MolProbity (Chen et al. 2010).

4.8.2 Using RDC, PRE and SAXS data for structure Calculation of Multidomain Proteins

The structure determination of multidomain RNA binding proteins is difficult to do with nOe distance restraints alone, since these only give short range distance information up to around 5 Å. For the determination of domain-domain interactions, which are connected often by long flexible linkers it is necessary to get additional long distance and orientation restraints.

Beginning from starting structures of the single domains, which can be either from X-ray crystallography or NMR spectroscopy it is then possible to calculate a structure of the multidomain protein using Aria/CNS. Therefore the structures of the free domains get refined with torsion angles calculated from chemical shifts and RDCs from the multidomain protein in a local refinement. Then the refined single domains are linked by randomized linkers of appropriate length. The domains will be treated as rigid bodies now. First the orientation of the two domains can be determined with the help of RDC data. In a second step the translation of the oriented domains can be determined by the PRE distance restraints from spin labels in several positions. SAXS data can be added as an complementary restraints, which contains information about the shape of the molecule (Simon et al. 2010).

4.9 Paramagnetic Relaxation Enhancement

Paramagnetic relaxation enhancement (PRE) is a technique, where an unpaired electron is introduced in a certain position of a protein or protein-ligand complex. The much higher gyromagnetic ratio of the electron makes it a very strong dipole compared to the nuclei. Because of this it has a very strong effect on the relaxation of nuclei around it. Since the strength of dipole-dipole interactions depend on the distance between the two dipoles the effect on relaxation also depends on the distance. The shorter the T2 relaxation time, the larger the linewidth of peaks become and the lower the intensity of the peaks become. Therefore peaks closer to the unpaired electron become more broadened than peaks farther away from the unpaired electron. The effect of the unpaired electron can be isolated by comparing the peak intensity from diamagnetic and paramagnetic spectra (Battiste and Wagner 2000).

$$\frac{I_{ox}}{I_{red}} = \frac{R2\,\exp(-R2^{sp}t)}{R2 + R2^{sp}}$$

The relaxation rate of the protein (R2) can be estimated from the diamagnetic spectrum. And then the contribution of the spinlabel $R2^{sp}$ can bedetermined. This contribution than can be converted to the distance between the nucleus and the electron r. τ_C is the correlation time between the electron and the nucleus, ω_h^2 is the Larmor frequency of the nucleus and K is $1.23 \cdot 10^{-32}\ cm^6 s^{-2}$, which includes the gyromagnetic ratio of protons, the electronic G factor and the Bohr magneton.

$$r = \left[\frac{K}{R2^{sp}}\left(4\tau_C + \frac{3\tau_C}{1 + \omega_h^2 \cdot \tau_C^2}\right)\right]^{1/6}$$

One way of introducing an unpaired electron is to attach a spin label to a free cysteine. In a protein without free surface accessible cysteines, a cysteine can be engineered to the desired position. The cysteine should be in a rigid structured part of the protein. A good parameter to decide if a residue is rigid enough is the heteronuclear nOe of the backbone amide. If the protein has accessible cysteines, these have to be mutated. The protein than can be labelled by attaching 3-(2-Iodoacetamido)-PROXYL (IPSL) to the free cysteine. After measuring the paramagnetic spectrum, IPSL can be reduced with ascorbic acid and then the diamagnetic sample can be recorded.

4.10 Residual Dipolar Couplings

Nuclear spins interact with each other through spice via dipolar couplings. The dipolar coupling is dependent on the length (r) and the orientation of the vector connecting the two spins relative to the static magnetic field (Θ).

$$D_{ij} = -\frac{\mu_0 \gamma_i \gamma_j h}{(2\pi r)^3} \cdot \langle\frac{3\cos^2\theta - 1}{2}\rangle$$

In an isotropic solution the free tumbling of the molecule averages out the dipolar coupling. In contrast, in solid state NMR spectra the full dipolar coupling is present, which leads to severe line broadening. If a molecule is only partially aligned in an otherwise isotropic solution the dipolar coupling can be reintroduced partially as a residual dipolar coupling (RDC). The residual dipolar coupling is scaled by the degree of alignment. For example the dipolar coupling between the amide proton and nitrogen in the peptide backbone of 20.000 Hz gets reduced to 20 Hz if the sample is 0.1% aligned.

Alignment can be achieved in different ways. Molecules, which have an intrinsic magnetic moment, e.g. the heme containing cyanometmyoglobin, align on their own in the static magnetic field (Tolman et al. 1995). Another way is to attach a lanthanide tag to the molecule (Prestegard, Bougault, and Kishore 2004). The third option, is to use a partially aniosotropic medium or alignment medium. Examples for alignment media are the filamentous Pf1 phage particle (Hansen, Mueller, and Pardi 1998) and the Lα phase of polyethylene glycol-alcohol mixtures (Rückert and Otting 2000).

RDCs can be determined by measuring not decoupled NMR experiments in isotropic (unaligned) and partially anisotropic (aligned) medium. The coupling measured between two peaks in the isotropic case is the J-coupling, while in the anisotropic case it corresponds to the sum of the J-coupling and the RDC. Therefore the RDC can be determined by subtracting the splitting in the spectrum of the isotropic sample from the splitting in the spectrum of the anisotropic sample. The RDC can be positive or negative.

Since measuring a not decoupled HSQC leads to double the amount of peaks, which complicate the spectrum and lead to increased overlap of peaks, in-phase anti-phase experiments (IPAP) have been developed, where through addition and subtraction of in-phase and anti-phase spectra the upfield and downfield component of an doublet can be separated in two spectra (Cordier, Dingley, and Grzesiek 1999).

RDCs contain orientation information and can be used to check domain orientations of crystal structures for crystal packing artifacts, to locally refine NMR structures, or to model domain orientations of rigid complexes, where known domain structures are treated as rigid bodies.

4.11 Small Angle X-ray Scattering

X-rays interact with matter by scattering. In small angle x-ray scattering (SAXS) the scattering is measured at a small angle. From the scattering information about the scattering particle in solution can be drawn. Elastic scattering of X-rays happens through interaction with electrons. After the scattering the X-rays have the same energy but the direction of the X-rays has changed. The scattered X-rays from different atoms interfere with each other constructive or destructive and thereby either add up or get canceled out. The interference depends on the angle of observation and the distance and angle of the scattering atoms to each other. The scattering can be represented as length of the scattering vector q (Schnablegger and Singh 2011).

$$q = \frac{4\pi \sin(\theta)}{\lambda}$$

Where 2θ is the angle of observation and λ is the wave length of the X-rays. When studying liquid samples the background scattering of the solvent has to be subtracted from the sample to gain the scattering originated from the studied molecule. The molecule should be monodisperse, because otherwise the SAXS curve represents the average of all present states.

The SAXS profile has three different regions from which information can be extracted (Boldon, Laliberte, and Liu 2015). From the Guinier region in the beginning of the curve the radius of gyration can be extracted by plotting the scattering intensity against q^2. The radius of gyration is structure independent measure of the size of the particle.

From the second region, the Fourier region, via indirect fourier transformation of the structure factor P(q) the pairwise distance distribution function ρ(r) can be determined, which contains information about the shape of the molecule.

$$\rho(r) = \frac{1}{2\pi^2 r} \int_0^\infty q(P(q)\sin(qr)\, dq$$

From the third region the Porod region information about the surface of the molecule can be extracted.

5 RBM10 Material and Methods

5.1 Cloning

5.1.1 Sub-cloning of RBM10 constructs from full length protein

The different constructs (Table 23) for this study were generated from a full length RBM10 construct by subcloning them to a pET-M11 vector. PCR primers flanking the desired constructs with NcoI in the N-terminus and XhoI in the C-terminus. For constructs including RRM2 BsmB1 had to be used because the DNA-sequence coding the RRM2 domain includes an NcoI restriction site. The BsmB1 sites where designed so they leave the same overhang as NcoI, so the inserts can be ligated to NcoI cut vector. The primers and the vector comprising full length RBM10 are used in a PCR reaction. The PCR samples contain the ingredients listed in Table 1.

Table 1: Ingredients used for PCR

Ingrediant	V (µL)
10x PFU polymerase buffer with $MgSO_4$	5
10 mM dNTP mix	1
Forward primer (100 pmol/µL)	0.5
Reverse primer (100 pmol/µL)	0.5
Template (25 ng/µL)	1
PFU DNA Polymerase (Thermo Scientific)	1
Water	41

The PCR was conducted in a BioRad Thermocycler using the program indicated in Table 1.

Table 2: Temperature program used for PCR

	Temperature	Time	
Initial Melting	95 °C	2 min	
Melting	95 °C	30 s	Loop 30 times
Annealing	56 °C	30 s	
Extension	72 °C	3 min	
Final Extension	72 °C	10 min	

The PCR product and the pET-M11 vector were digested either with NcoI and XhoI from New England Biolabs (NEB) simultaneously in a double digested, or in case of PCR products including RRM2 sequentially first with BsmB1 and then with XhoI. For the digestion reaction the NE buffer 3.1 was used.

After digestion the vector is treated for 1 h with Antarctic phosphatase (NEB, 5000 U/mL) to prevent relegation of the empty vector.

Than the buffers were ligated with T4 ligase. After one our half of the ligation reaction (10 µL) was used for transformation of DH5α cells. (See chapter 5.1.3). An addition one transformation was done from a ligation reaction without insert as a control to see the background of false positive clones. On the next day several colonies were picked and grown in 3 mL LB-medium with kanamycin at 37 °C over night in a 24-well block. In parallel a colony PCR with the same colonies was done to check for positive clones. The colony PCR was done with the same mix as the normal PCR, but instead of PFU- taq-DNA-polymerase was used and instead of the template a little fraction of the picked colony was added. The PCR reaction volume was reduced from 50 µL to 20 µL. The colony PCR was than checked on a 1% agarose gel stained with ethidium bromide. For the positive colonies then the plasmid was isolated from the 3-mL cultures by mini-prep and the sequence of the isolated plasmid was checked by the company GATC. Purified plasmids were stored at 4°C.

5.1.2 Quick-Change-PCR for Side directed mutagenesis

Quick change PCR was used to generate point mutations in the different constructs of the RNA binding domains to analyze the effect of isoforms and mutants and to generate Cys-Mutants for PRE measurements. The primers were designed according to the Agilent “QuikChange II Site-Directed Mutagenesis Kit” manual. The PCR reactions were composed according to Table 1 and run with the temperature program described in Table 3.

Table 3: Temperature program used for quick change PCR

	Temperature	Time	
Initial Melting	95 °C	2 min	
Melting	95 °C	30 s	
Annealing	56 °C	30 s	Loop 20 times
Extension	72 °C	7 min	
Final Extension	72 °C	10 min	

1 µL DpnI (20000 units/mL) was added to the PCR reactions and incubated for 1 hour at 37 °C to degrade the template plasmid. With 5 µL and 45 µL of the reaction mixtures transformations were done into DH5α cells.

5.1.3 Transformation and Mini-Prep

For transformation to 50 µL competent cells of the desired strain, usually BL21(DE3) for protein expression, or DH5α for plasmid amplification, 1 µL of approximately 50 ng/µL of plasmide was added. The cells then were incubated for 15 min on ice. Afterwards cells were heat shocked at 42 °C for 45 s. Then, 900 µL of LB-medium was added and the cells were incubated for 45 min at 37 °C. The cells were spun down and most of the supernatant was discarded. The cells were resuspended in the remaining supernatant (approximately 50 µL). The cells were plated to an agar-plate with 50 µg/mL kanamycin and incubated at 37 °C over night. On the next morning cells were sealed with parafilm and stored at 4 °C. For plasmid isolation Mini Prep a kit from Machinery Nagel was used.

5.2 Protein Expression

A 20 mL pre-culture with 30 µg/ml kanamycin was prepared either in LB- or in M9-medium (for composition of M9 medium see appendix Table 12) for natural abundance or isotope-enriched cultures, respectively. The pre-culture was incubated over night at 37°C. On the next morning, 1 L of medium was inoculated with the pre-culture. The main culture was grown at 37°C until it reached an $OD_{600\ nm}$ between 0.6 and 0.8. Then, the culture was induced with a final concentration of 1 mM IPTG. Afterwards the culture was grown over night at 20°C. On the next morning, the cultures were harvested by centrifuging them at 6000 rpm for 20 min, resuspending the cells in lysis buffer and again spinning them down at 4000 rpm for 20 min in falcons. After harvesting, the cells were either directly used for protein purification or stored at -20 °C.

5.3 Protein Purification

Prior to protein purification cells were lysed by sonication. Frozen cell pellets were thawed in a warm water bath and then resuspended in 30 mL lysis buffer, containing additionally 500 µL of 40 mM proteases inhibitor AEBSF and 100 µL of 1 mg/mL DNAse I, and then lysed on ice by three cycles of 3 min at strength of approximately 50% with the syndicator. Afterwards the lysate is centrifuged at 20,000 rpm at 4°C for 45 min.

For the IMAC purification, the Ni-NTA column has to be stripped and reloaded with Zn^{2+} ions, to avoid exchange of the Zn^{2+} coordinated by the Zinc finger 1 (ZF1) of RBM10 with Ni^{2+}. When purifying constructs lacking ZF1 this step was skipped. For stripping the 4 mL of Ni-NTA beads were first washed with 5 column volumes (cv) of water, then with 5 cv of 0.5 M of EDTA, then again with 5 cv of water, then with 5 cv of 100 mM zinc chloride, then again with 5 cv water and finally with 5 cv of lysis buffer.

The lysate was loaded to the equilibrated IMAC column and the flow through (FT) was collected. The column was washed first with 5 cv of lysis buffer (W1) and then with 5 cv of wash buffer (W2). Afterwards, the protein was eluted with 10 mL of elution buffer.

1 mL of 1 mg/mL TEV protease was added to the eluted protein and then the protein was dialysed against 1 L of TEV-cleavage buffer overnight. On the next morning, the dialyzed protein was filtered with an 0.2 µm filter and then loaded to a 5 mL Heparin Column, which was equilibrated with Heparin buffer A. After the protein was washed with 4 cv of Heparin buffer A, a gradient from 0 to 50% of Heparin Buffer B with a length of 10 cv was applied to elute the protein from the heparin column. The fractions were checked by SDS-PAGE and the fractions containing the cleaved protein were pooled.

The sodium chloride concentration of the sample was adjusted to 1 M by adding 5 M sodium chloride solution. Then the sample was loaded again to the Zn^{2+}-IMAC column, which was equilibrated again with lysis buffer. The column is washed with 1.5 cv of lysis buffer. The FT and the wash fraction are collected together and concentrated to a volume of 1 mL. In case of the three domain construct of RBM10 the second IMAC column was skipped, since the protein tends to bind also without the His-tag to the IMAC column. In this case the elution from the heparin column was directly concentrated to 1 mL.

As a last step of the purification the concentrated protein was loaded to a Superdex 75 size exclusion column. From the size exclusion column, 2 mL fractions were collected and checked by SDS-PAGE. The fractions of pure protein were pooled and concentrated to the desired concentration. The proteins were either directly used for measurements or snap-frozen in liquid nitrogen and stored at -80°C.

5.4 NMR Backbone Assignment

For the assignment experiments $^{13}C/^{15}N$ double labelled protein was used. For the backbone assignment of RRM1-ZF1 HNCACB and CBCA(CO)NH spectra were measured with a 1.1 mM protein sample in NMR buffer. The assignments for the RRM2 (-V354) were taken from the BMRB (Code: 27055) and then confirmed with an HNCACB experiment with a 311 µM sample. Also the assignment of the RRM2 (+354) isoform and the V354E were done with the help of HNCACB spectra. Assignments of the three domain construct were transferred from the RRM1-ZF1 and the RRM2 construct and then confirmed by an HNCACB and an CBCA(CO)NH, which were recorded on a 1 mM sample.

The assignment of the RRM1-ZF1 RNA complex was transferred from the apo-protein following the shifts in the NMR-titration. The backbone assignment was completed using an HNCACB and an CBCA(CO)NH spectrum. The protein-RNA complex of RRM2 was determined as well by following the CSP during the RNA-titration. Finally, the assignment of the RNA-bound three domain construct was done by a combination of the known positions of RNA bound residues in the RRM1-ZF1 and the RRM 2 constructs and the NMR-titration with the three domain construct.

5.5 NMR Titrations

NMR titrations were performed with various constructs of the RNA binding domains of RBM10 (RRM1-ZF1, RRM1-ZF1-RRM2, RRM2, RRM2 isoforms and mutants) with the NUMB RNA 12-mer. Protein concentrations, protein:RNA ratios dimension in the titration steps, fields and points in the ω1 (nitrogen) are listed in Table 4. All titrations were performed at 298 K in 3 mm NMR tubes.

Table 4: Protein concentrations and Protein:RNA ratios used for NMR Titrations

Construct	**[Protein]**	**Ratio #1**	**Ratio #2**	**Ratio #3**	**Ratio #4**	**Field**	**Points ω1**
RRM1-ZF1	170 µM	1:0.3	1:0.6	1:1	1:1.2	800 MHz	150
RRM1-ZF1-RRM2	232 µM	1:0.25	1:0.5	1:1	1:1.5	800 MHz	256
RRM2 -V354	200 µM	1:0.3	1:0.6	1:1	1:1.5	750 MHz	256
RRM2 +V354	200 µM	1:0.3	1:0.6	1:1	1:1.5	750 MHz	256
RRM2 V354E	200 µM	1:0.3	1:0.6	1:1	1:1.5	750 MHz	256

5.6 NMR Side Chain Assignment and Structure Calculation

NMR side chain assignment was done for the RRM1-ZF1, RRM2 -V354 and the RRM2 + V354 construct. For the side chain assignment samples in NMR buffer were used. HBHACONH and HcCH and hCCH-TOCSY experiments (Sattler, Schleucher, and Griesinger 1999) were recorded in normal H_2O buffer, while an $^1H,^{13}C$ correlation spectra and HBCBCGCDHD and HBCBCGCDCEHE spectra (Yamazaki, Forman-Kay, and Kay 1993) were recorded in D_2O buffer. To change a sample from H_2O buffer to D_2O buffer the sample was lyophilized, redissolved in D_2O, lyophilized again and finally redissolved in very pure (99.96%) D_2O. The total volume of the sample was kept the same and thereby the concentration of all the salts in the NMR buffer stayed the same.

For structure calculation in addition to the side chain assignment, distance restrains derived from NOESY spectra are needed. For the RRM1-ZF1, the RRM2 -V354, and the RRM2 +V354 constructs a ^{15}N-edited NOESY spectrum in H_2O NMR-buffer, a aliphatic ^{13}C-edited NOESY spectrum in D_2O buffer and an aromatic ^{13}C-edited NOESY spectrum in D_2O buffer were recorded. For the RRM1-ZF1 construct in addition, an aliphatic ^{13}C-edited NOESY spectrum in H_2O buffer was recorded.

The side chain assignment was repeated for the RRM-ZF1 construct in complex with RNA (concentration 864 µM, protein:RNA ratio 1:1.2) using the same side chain assignment experiments as for the apo-protein. In addition to the ^{15}N-edited NOESY spectrum in H_2O NMR-buffer, the aliphatic ^{13}C-edited NOESY spectrum in D_2O buffer and the aromatic ^{13}C-edited NOESY spectrum in D_2O buffer, an aliphatic and an aromatic ω1-filtered ^{13}C-edited NOESY spectra were recorded to identify intermolecular NOEs.

For the automated assignment of the NOESY spectra and the structure calculation the software Cyana was used (Peter Güntert and Buchner 2015). In addition to the NOESY peak lists, predicted dihedral restraints from the program TALOS (Shen et al. 2009) were used as input. The unambiguous distance restraints and dihedral restraints were then transformed to the CNS file format and used in water refinement run with Aria2/CNS (Rieping et al. 2007). With Aria2/CNS 100 structures were generated, from which the 10 lowest energy structures were water refined and reported as a structural bundle.

5.7 RNA TOCSY spectra and ω1-filtered 2D NOESY spectra

The RNA TOCSY spectra were recorded on free RNA in D_2O-NMR buffer and on saturated RNA with either the RRM1-ZF1 construct or the RRM1-ZF1-RRM2 construct of RBM10. All experiments were performed at 298 K with a TOCSY spin lock time of 60 ms. Proteins in normal NMR buffer were lyophilized, the resulting powder, which contained the protein and the salts of the buffer, was resuspended in the same volume of D_2O as sample was present before the lyophilization. Then the resulting protein solution in D_2O buffer was lyophilized again and then again redissolved in D_2O. The two step process is necessary to remove as much H_2O as possible, especially the hydration water, which is still present in the powder after the first lyophilization. Further parameters of the different measurements are listed in Table 5.

Table 5: Experimental parameters of RNA TOCSY experiments

#	RNA	Protein	number of scans	Points ω1	Field
1	500 µM NUMB 12-mer	-/-	40	248	600 MHz
2	500 µM NUMB 12-mer	1 mM RRM1-ZF1	40	248	600 MHz
3	300 µM NUMB 12-mer	450 µM RRM1-ZF1	8	150	800 MHz
4	270 µM NUMB 12-mer	360 µM RRM1-ZF1-RRM2	8	150	800 MHz

Then ω1-filtered 2D NOESY spectra were recorded to test for intermolecular nOes between the RNA and the protein. The RNA samples were saturated with protein to avoid to observe the free and the bound form of RNA. The experiments were, as well as the RNA-TOCSY experiments, recorded in D_2O NMR buffer at 298 K. The NOESY mixing time was 200 ms. The rest of the measurement paramteters can be found in Table 6.

Table 6: Experimental parameters of ω1-filtered 2D-NOESY experiments

#	RNA	Protein	number of scans	Points ω1	Field
1	300 µM NUMB 12-mer	450 µM RRM1-ZF1	256	200	800 MHz
2	270 µM NUMB 12-mer	360 µM RRM1-ZF1-RRM2	368	256	800 MHz

5.8 NMR Relaxation experiments

To determine T1 and T2 and τ_C, ^{15}N T1 and T1rho relaxation experiments were measured. In Table 7 the delays used for the measurement for T1 and T1rho are listed.

Table 7: Delays used for ^{15}N relaxation experiments

#	T1 Delays RRM1-ZF1 (ms)	T1rho delays RRM1-ZF1 (ms)	T1 Delays RRM1-ZF1-RRM2 (ms)	T1rho Delays RRM1-ZF1-RRM2 (ms)
1	10.8	5.0	10.8	5.0
2	43.2 (as duplicate)	10.0	21.6	10.0
3	64.8	15.0	32.4	15.0
4	86.4	20.0	43.2 (as duplicate)	20.0
5	172.8	30.0	64.8	30.0
6	345.6 (as duplicate)	40.0	86.4	40.0
7	518.4	60.0	172.8	60.0
8	691.2	80.0 (as duplicate)	345.6 (as duplicate)	80.0 (as duplicate)
9	1036.8	100.0 (as duplicate)	518.4	100.0 (as duplicate)
10	1382.4	120.0	691.2	120.0
11	1728.0		1036.8	
12	2419.2		1382.4	

The order of the delays were randomized and the experiments were recorded scan wise interleaved. For the T1ρ experiments a 2 kHz spin lock was used. All relaxation experiments were recorded at 25 °C. For the free RRM1-ZF1 construct a 250 mM sample in a 3mm-tube was used. For the RRM1-ZF1 complex a 250 mM sample in a 3mm-tubesaturated with 300 mM NUMB 12-mer RNA was used. Spectra were recorded with 128 complex points in the ω1 dimension and 12 scans. For the free RRM1-ZF1-RRM2 construct a 400 mM sample in a 3mm-tube was used. Spectra were recorded with 75 complex points in the indirect dimensions and 8 scans. For the complex of RRM1-ZF1-RRM2 and the NUMB RNA 12-mer a sample with protein concentration of 232 µM and a NUMB RNA concentration of 348µM was used. Spectra were recorded with 100 points in the ω1 dimension and 24 scans. Peaks were fitted with the build in module in CcpNmr analysis (Vranken et al. 2005), the errors represent the fitting error of the exponential decay including the two duplicate points.

In addition, $\{^1H\}$,^{15}N-heteronuclear NOE experiments were recorded on the same samples for RRM1-ZF1 free and RNA bound states. The saturated and the reference experiments were recorded interleaved. Each of them was recorded with 128 complex points in the ω1 dimension and 16 scans. For the free RRM1-ZF1-RRM2 construct a $\{^1H\}$,^{15}N-heteronuclear nOe experiment with the 128 complex points in the ω1 dimension and 20 scans was recorded on a 1 mM protein sample in a shigemi tube.

5.9 Paramagnetic Relaxation Enhancement Measurements

5.9.1 Attachment of 3-(2-Iodoacetamido)-PROXYL (IPSL) spinlabel to the protein.

For the labelling with IPSL the protein with the engineered cysteine residue in the desired position had to be reduced with 10 mM DTT for at least 2 hours. For this purpose 5 µL 1 M DTT was added to 500 µL of the purified protein with a concentration between 300 µM and 1mM in NMR-buffer. After 2 h the protein was dialysed against 2.5 L IPSL-dialysis buffer (**Table 20**) in a 10 kDa cut off dialysis tubing at 4 °C overnight. On the next morning the dialysis tubing was transferred to 2.5 L of fresh IPSL-dialysis buffer and dialyzed for another 3h at 4 °C. Afterwards, the dialysis tubing was transferred to 500 mL of IPSL-reaction buffer (Table 21) and dialysed for 3h at 4°C. The concentration of the protein sample was measured with the nanodrop and the volume was approximated by weighing the sample. A 10-fold molar excess of IPSL was added to the potein and the reaction mixture was incubated at 4 °C over night in the dark. On the next morning the sample was diluted to 15 mL with NMR-buffer without DTT and then concentrated in a concentrator with a molecular weight cut-off of 10 kDa to a final volume of 500 µL. Afterwards the protein was diluted with NMR-buffer without DTT and concentrated two times more to exchange the buffer and remove the excess of spin label.

5.9.2 Paramagnetic Relaxation Enhancement NMR measurements

For the measurements, a 250 µM protein sample of the IPSL labelled protein was measured in a 3 mm tube. For the paramagnetic protein an ^{1}H,^{15}N-HSQC and an HNCO spectrum was measured. The ^{1}H,^{15}N-HSQC spectrum was measured for 3 h (d1 = 5 s, NS = 8, $TD_{(15N)}$ = 256 points) and the HNCO spectrum for 18 h (d1 = 5 s, NS = 4, $TD_{(15N)}$ = 120 points, $TD_{(13C)}$ = 26 points). Afterwards the IPSL was reduced by adding a final concentration of 3 mM ascorbic acid from a 100 mM stock. Then, the ^{1}H,^{15}N-HSQC and the HNCO is measured again on the diamagnetic sample. In case of the PRE measurements of the protein-RNA complex, 250 µM IPSL labelled protein is saturated with 300 µM of RNA. All measurements were done at 298 K on a Bruker Advance III at 800 MHz with a cryogenic triple resonance probe. For the analysis peaks were picked in CcpNmr analysis and the peak heights of the paramagnetic and the diamagnetic spectrum were exported to calculate the ratio (I_{para}/I_{dia}). The error was determined depending on the spectral noise of both spectra.

5.10 Isothermal Titration Calorimetry (ITC)

ITC experiments were conducted in NMR buffer (Table 22) with 5 mM β-mercaptoethanol instead of 1 mM DTT. When using construct without ZF1 (e.g. RRM1 or RRM2) measurements were done without the reducing agent β-mercaptoethanol. The mesurements were done on a Malvern ITC200 machine at 25 °C. The feedback was set to low and the stirring speed to 750 rpm. Data were than fitted with the Malvern PEAQ ITC software with a one binding site model. The protein constructs and concentrations used for titrations with the NUMB RNA are listed in Table 8.

Table 8: Constructs and Concentrations used in ITC experiments

#	Concentration in Syringe		Concentration in Cell	
1	180 µM	NUMB RNA	30.0 µM	RRM1-ZF1
2	299 µM	RRM1	30.1 µM	NUMB RNA
3	211 µM	NUMB RNA	20.0 µM	RRM1-ZF1-RRM2 (-V354)
4	203 µM	RRM1-ZF1-RRM2 (-V354)	19.9 µM	NUMB RNA
5	197 µM	RRM1-ZF1-RRM2 (+V354)	20.2 µM	NUMB RNA
6	199 µM	RRM1-ZF1-RRM2 (V354E)	20.3 µM	NUMB RNA
7	337 µM	RRM2 (-V354)	27.6 µM	NUMB RNA

5.11 Small angle X-ray scattering (SAXS)

For SAXS measurements the three domain constructs (-V354, V354 and V354E) were measured in NMR buffer (Table 19). The measurements were done at three concentrations (10 mg/mL, 5 mg/mL and 1 mg/mL) at 25 °C. For SAXS measurements of the complex with NUMB RNA the complex was preformed by mixing protein and RNA in a 1:1.2 ratio. The complex was then purified by SEC on a superdex S75 size exclusion column to separate the complex from an excess of RNA. The collected fractions containing the complex are than concentrated to 150 µL, which should correspond to a protein concentration of 10 mg/mL. From this sample the dilutions are made (5mg/mL and 3 mg/mL). Due to losses during the SEC the real concentration is probably a bit smaller.

5.12 Static light scattering (SLS)

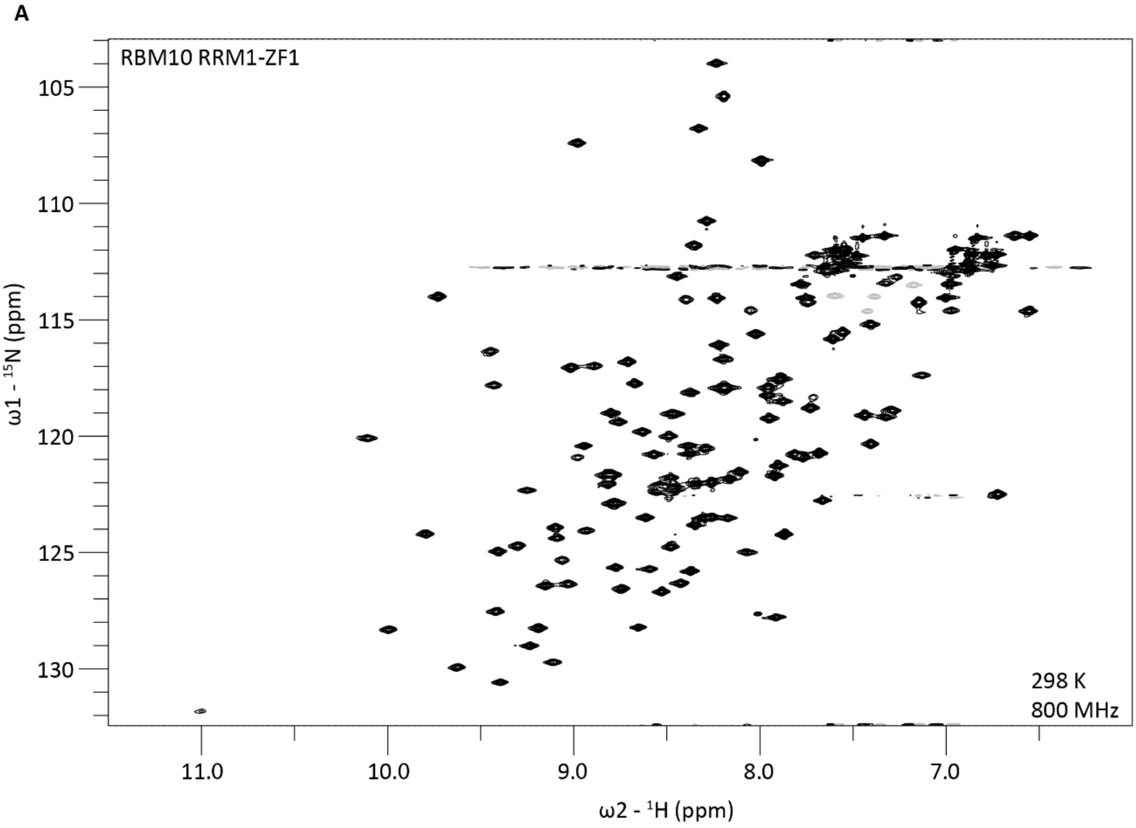
Static light scattering was performed for RBM10 RRM1-ZF1 and RRM1-ZF1-RRM2 with and without the NUMB RNA to confirm the binding stoichiometry. For the measurement a Viscotek TDA 305 machine and a analytical superdex 75 column was used. For the measurements, NMR-buffer (Table 19) was used. The detector was calibrated with 4 mg/mL bovine serum albuminum (BSA). Than the scattering of the free protein (4 mg/mL) was recorded as a reference. Finally, the protein-RNA complex (1:1.5 ratio for RRM1-ZF1 and 1:2 ration for RRM1-ZF1-RRM2) was loaded with a protein concentration of 3.5 mg/mL in case of RRM1-ZF1 and 4 mg/mL in case of RRM1-ZF1-RRM2. From the scattering the mass of the apo-proteins and the complexes were calculated.

6 RBM10 Results

6.1 RRM1-ZF1 and RRM2

6.1.1 RRM1-ZF1 apo-protein adopts a flexible domain orientation in solution

A small linker of four residues connects RRM1 to ZF1. The ^{1}H,^{15}N NMR correlation spectrum of the RRM1-ZF1 construct is well dispersed with signals of similar peak intensity, indicating a well-folded protein (**Figure 14A**). The backbone residues could be assigned to 95%, not considering the four proline residues, which do not appear in the ^{1}H,^{15}N correlation spectrum (**Figure 14B**). The assignment was done based on a HNCACB and a CBCA(CO)NH spectrum.



B

```
      125   130   135   140   145   150   155   160   165   170   175
 GAMA EKASN IVMLR MLPQA ATEDD IRGQL QSHGV QAREV RLMRN KSSGQ SRGFA FVEFS
180   185   190   195   200   205   210   215   220   225   230   235
HLQDA TRWME ANQHS LNILG QKVSM HYSDP KPKIN EDWLC NKCGV QNFKR REKCF KCGVP
240
KSE
```

Figure 14 Assignment of RRM1-ZF1: *A) ^{1}H,^{15}N-correlation spectrum of RBM10 RRM1-ZF1. Positive peaks are plotted in black and negative peaks are plotted in gray. B) Primary sequence of RBM10 RRM1-ZF1 construct. The residues printed in italic indicate artificial residues due to cloning. Proline residues are printed in gray since they do not give rise to signals in the ^{1}H,^{15}N-correlation spectrum. Assigned and unassigned residues are colored in green and red, respectively.*

To test if ZF1 was folded as well, EDTA was added to strip the Zn^{2+} out of the zinc finger. As a result signals of the ZF shift to the middle of the spectrum, which is characteristic of unfolded proteins (**Figure 15**). Since the signals corresponding to the unfolded ZF1 are not present in the initial sample, it can be assumed that RRM1-ZF1 was purified fully coordinated with Zn^{2+}. The exchange of Zn^{2+} with Ni^{2+} during IMAC purification could be avoided by using Zn^{2+}-NTA instead of Ni^{2+}-NTA.

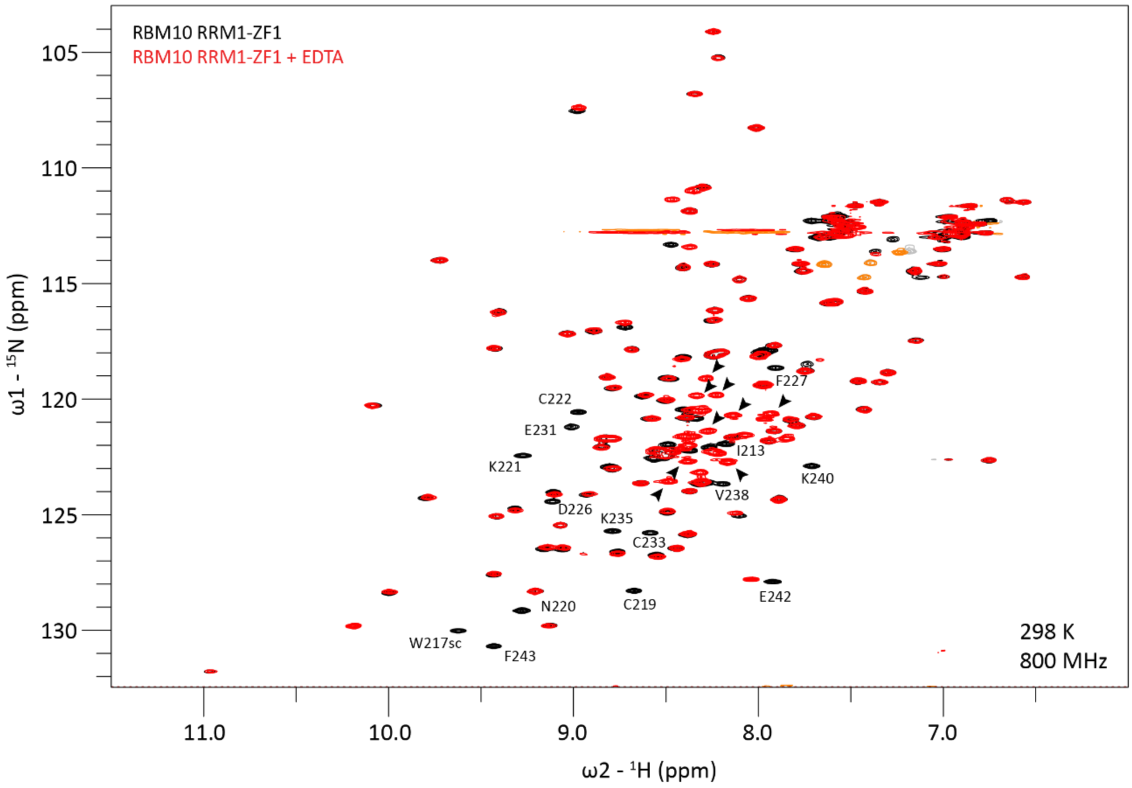
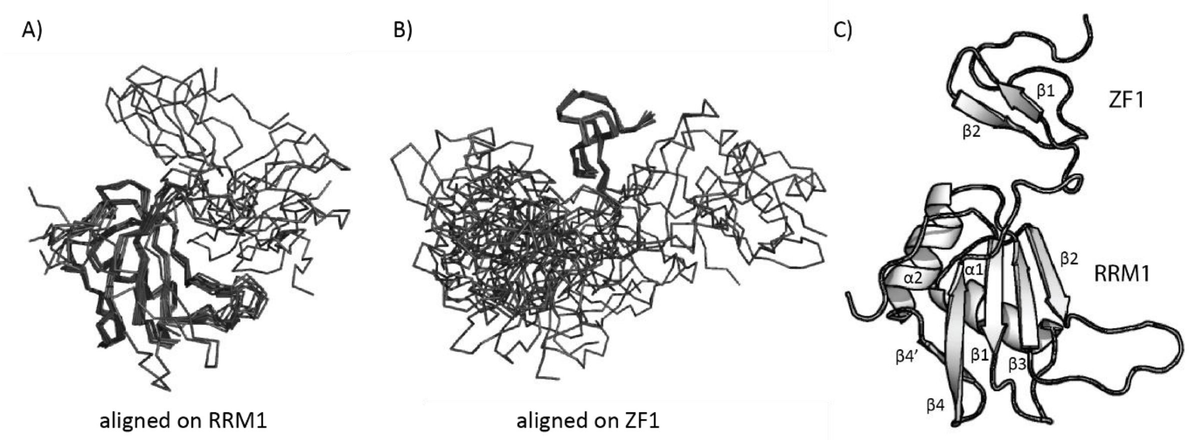


Figure 15 Unfolding of ZF1 with EDTA: *A superimposition of the* ^{1}H,^{15}N*-correlation spectra of RBM10 RRM1-ZF1 in absence (black) and presence (red) of EDTA is shown. Assignment of peaks that disappear upon addition of EDTA are indicated. Peaks that appear in the middle of the spectrum upon addition are indicated by black arrows.*

For structure calculation all protons have to be assigned, including the side chains. It was possible to assign the protons up to 90.8 % using HBHA(CO)NH, HcCH-TOCSY, hCCH-TOCSY, CBCGCDHD, CBCGCDCEHE, ^{15}N edited NOESY, aliphatic ^{13}C edited and aromatic ^{13}C edited NOESY NMR-spectra (Sattler, Schleucher, and Griesinger 1999).

The chemical shift assignment as well as integrated peak lists of the NOESY spectra were used to assign the nOe cross peaks and to calculate the structure using Cyana, followed by water refinement using Aria2/CNS. RRM1 and ZF1 converge both on their own, but the domain orientation between them is not defined by the structure. This becomes clear when aligning the bundle of NMR structures to RRM1 (**Figure 16A**) and ZF1 (**Figure 16B**), respectively. There were no nOe cross peaks between RRM1 and ZF1 identified. This is also reflected in the high rmsd values of the entire ensemble (**Table 9**). Most of the 0.7% outliers in the Ramachandran plot analysis are located in the flexible N-terminus (A122, A124) or the linker between the two domains (I213,N214,E215).



aligned on RRM1

aligned on ZF1

Figure 16: RBM10 RRM1-ZF1 NMR structure: *Ribbon plot of the 10-structure water refined NMR bundle of RBM10 RRM1-ZF1 aligned on RRM1 (A) or ZF1 (B). C) Cartoon plot of the RBM10 RRM1-ZF1 structure. Secondary structure elements are labelled.*

Table 9: Structure calculation statistics for the water refined RBM10 RRM1-ZF1 structure:

Structure calculation restraints	
Distance restraints	
Total NOEs	1877
Sequential (|i- j| ≤ 1)	956
Medium-range (1<|i- j| <5)	263
Long-range (|i- j| ≥ 5)	658
Dihedral restraints (ϕ+ψ)	200
Quality analysis	
Restraints violations (mean ± s.d)	
Distance restraints (Å)	0.022 ± 0.003
Dihedral angle restraints (º)	0.474 ± 0.24
Deviation from idealized geometry	
Bond length (Å)	0.003 ± 0.0001
Bond angles (º)	0.46 ± 0.020
Improper dihedral distribution (º)	1.24 ± 0.088
Average pairwise r.m.s. deviation (Å)	
Heavy	7.96 ± 2.91
Backbone	8.19 ± 2.89
Ramachandran values (%) (MolProbity)	
Most favored regions	93.7 %
Allowed regions	99.3 %

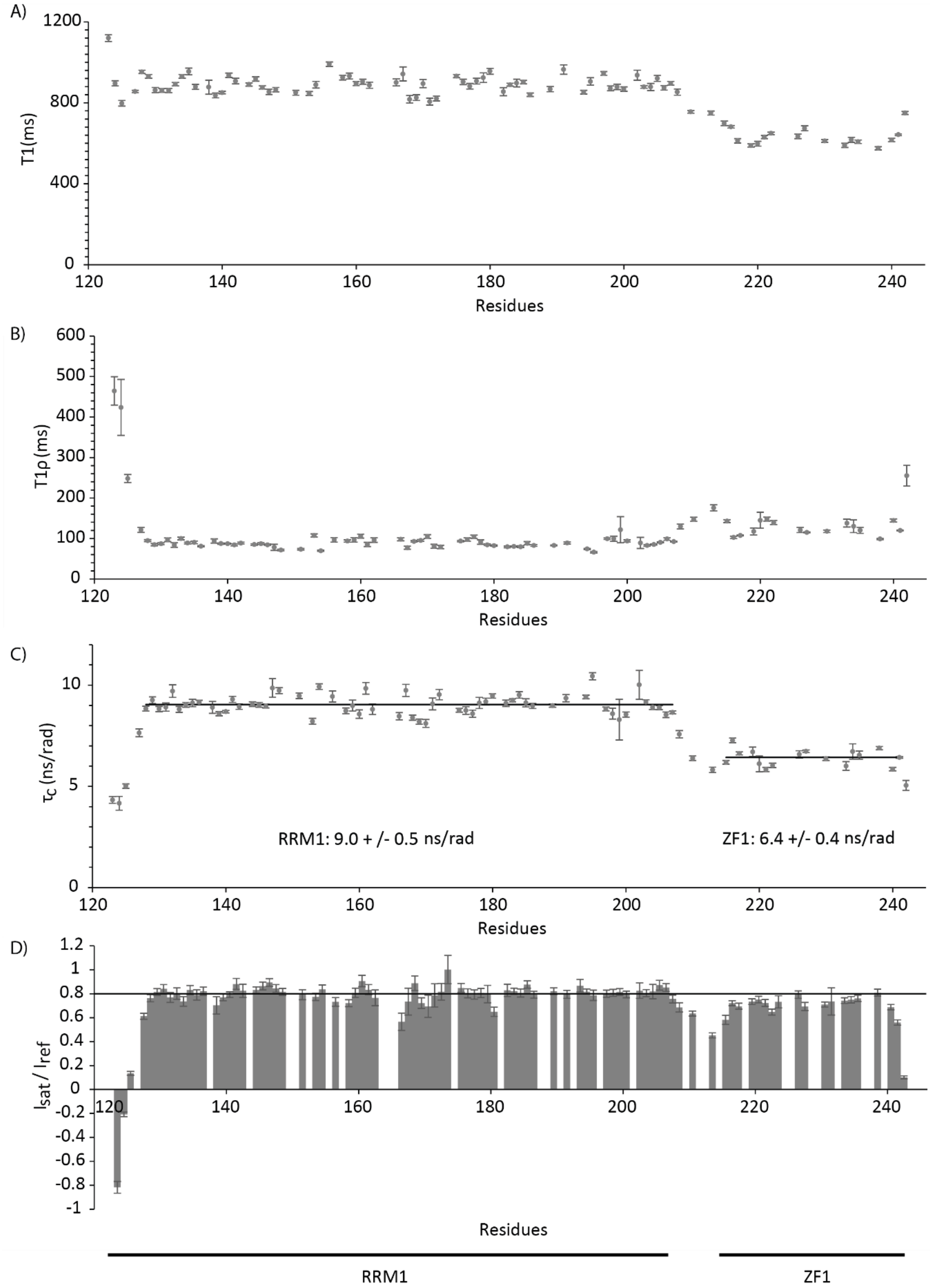


Figure 17: Relaxation data RBM10 RRM1-ZF-1: *A) T1 time constant, B) T1ρ time constant, C) correlation time τ_C and D) {^{1}H},^{15}N-heteronuclear nOe of the RBM10 RRM1-ZF1 apo-protein plotted against the residue number.*

The structure of RRM1 represents a canonical RRM fold. The four β-strands form an anti-parallel β-sheet, which is packed against the two α-helices. There is an additional short β-strand (β4’) present in front of the β4-strand. This short strand and β-strand 4 from a β-hairpin. The topology of the RRM1 domain including β4’ is βαββαββ. The ZF1 domain has two β-strands, which form an antiparallelβ-sheet. The zinc is not depicted here since it was not explicitly present during the structure calculation. Instead upper distance limits between the Cβ of the cysteine residues coordinating Zn^{2+} were used.

The lack of detectable nOes and thus distance restraints between RRM1 and ZF1 hints to a dynamic behavior of the two domains. To test the dynamics of the two domains relative to each other, ^{15}N-relaxation data have been recorded. T1 (**Figure 17A**) and T1ρ (**Figure 17B**) data were recorded and from these the correlation time on a residue resolution was estimated (**Figure 17C**). While RRM1 residues show an average correlation time of 9.0 +/- 0.4 ns/rad, residues belonging to ZF1 show a much lower correlation time of 6.4 +/- 0.4 ns/rad. The big difference in the correlation times indicates that the domains do not have a rigid domain arrangement, but are flexible relative to each other. This dynamic behavior explains the lack of interdomain nOe signals observed during the structure calculation. In addition, a {^{1}H},^{15}N-heteronuclear nOe experiment was recorded. In a {^{1}H},^{15}N-heteronuclear nOe experiment the ratio of a saturated and an unsaturated reference spectrum is made. This ratio is 0.8 for residues with fully rigid backbone, while they are lower or even negative for flexible residues. The RRM-ZF1 is mostly rigid, with the exception of the termini, the linker between the two domains, and a longer loop in RRM1 around residue 165 between β-strands two and three (**Figure 17D**).

6.1.2 Both domains are involved in RNA binding in a 1:1 complex

To answer the question how RRM1 and ZF1 recognize RNA, an NMR titration with NUMB derived RNA 12-mer-oligo (UUGUCUGCUCCC) was done. There are significant changes in chemical shift upon RNA binding in RRM1, ZF1 and the linker connecting the two domains. Shifting peaks are severely broadened in the intermediate titration steps (0.3x and 0.6x molar equivalent of RNA to protein), but they are sharper again at a 1:1 protein:RNA ratio. There is no significant change in peak position upon adding more RNA (1.2x), indicating a high affinity 1:1 complex between protein and RNA (**Figure 18A**). Many assignments could be transferred by following the peaks throughout the titration. The remaining residues in the complex were then assigned using HNCACB and CBCA(CO)NH spectra. Thereby a nearly complete assignment of the complex could be gained. Few residues of the protein-RNA complex could not be identified in the ^{1}H,^{15}N-correlation spectrum (M123, F234, K228). Likely, the signal broadening beyond detection in the ^{1}H,^{15}N-correlation spectrum is due to exchange of the binding RNA.

A closer look at the affected regions in RRM1 shows that the ß-sheet of RRM1, especially the β-strands 2 and 3 and the loop connecting them, and the N-terminus of the domain experience chemical shift perturbation (CSP) upon RNA binding. The canonical RNA binding interface of an RRM is the β-sheet and includes two RNP motifs. The RNP1 motif in β-strand 3 has a consensus sequence of [RK]-G-[FY]-[GA]-[FY]-[ILV]-X-[FY] and the RNP2 motif in β-strand 1 has a [ILV]-[FY]-[ILV]-X-N-L consensus sequence (Cléry, Blatter, and Allain 2008). While the RRM1 of RBM10 has a fully conserved RNP1 motif, the RNP 2 motif has instead of the [FY] in position 2 and the N in position 5 two methionine residues. The CSP, which is stronger for β-strand 2 and 3, and weaker for β-strands 1 and 4, suggests that the RBM10 RRM1 utilizes only the canonical RNP1 motif, but not the non-canonical RNP2 motif for RNA recognition.

The linker between RRM-ZF1 shows strong chemical shift perturbation as well. These could be caused by protein-RNA contacts directly or indirectly through the rearrangement of the domains. The zinc finger itself shows moderate chemical shift perturbation. Additionally the peaks K228 and F234 in the ZF could not be detected after binding.

In conclusion the NMR-titration suggests that RRM1-ZF1 binds the NUMB-derived RNA oligo in a 1:1 complex involving both domains. To further validate and characterize this interaction isothermal titration calorimetry experiments were conducted with the RRM1 alone and the RRM1-ZF1 construct. RRM1 alone has a dissociation constant (K_D) of 6 µM (**Figure 19A**), RRM1-ZF1 in contrast has with a K_D of 0.7 µM a one order of magnitude higher affinity (**Figure 19B**). The high gain of affinity confirms that both domains are important for the recognition of the RNA. The ITC curve of RRM1-ZF1 gives a stoichiometry of approximately 1:1, which is also in agreement with the NMR-titration, which was saturated at a molar ratio of 1:1.

Figure 18 NMR titration of RBM10 RRM1-ZF1 with NUMB derived 12-mer-RNA oligo: *A) Superimposed ^1H,^{15}N-correlation spectra of RBM10 RRM1-ZF1 apo-protein (black) and titration steps with 0.3x (red), 0.6x (green), 1.0x (cyan) and 1.2x (blue) molar equivalents of RNA present. Signals in the black and the blue spectrum of residues that experience high chemical shift perturbation are connected by arrows and labeled. Residue F234, which could not be detected in the complex spectrum, is labeled as well. B) Chemical shift perturbations between RBM10 RRM1-ZF1 free and in complex with the NUMB derived 12-mer RNA oligo are plotted against the residue number. C). CSP from B that are higher than 0.05 ppm (salmon dotted line)or 0.10 ppm (red dotted line) are plotted on the structure of RRM1-ZF1. The name of the secondary structure elements is indicated. Additionally the sequence of the two RNP motifs is indicated and compared to the consensus motif. Red marked residues do not match. (Figure on next page)*

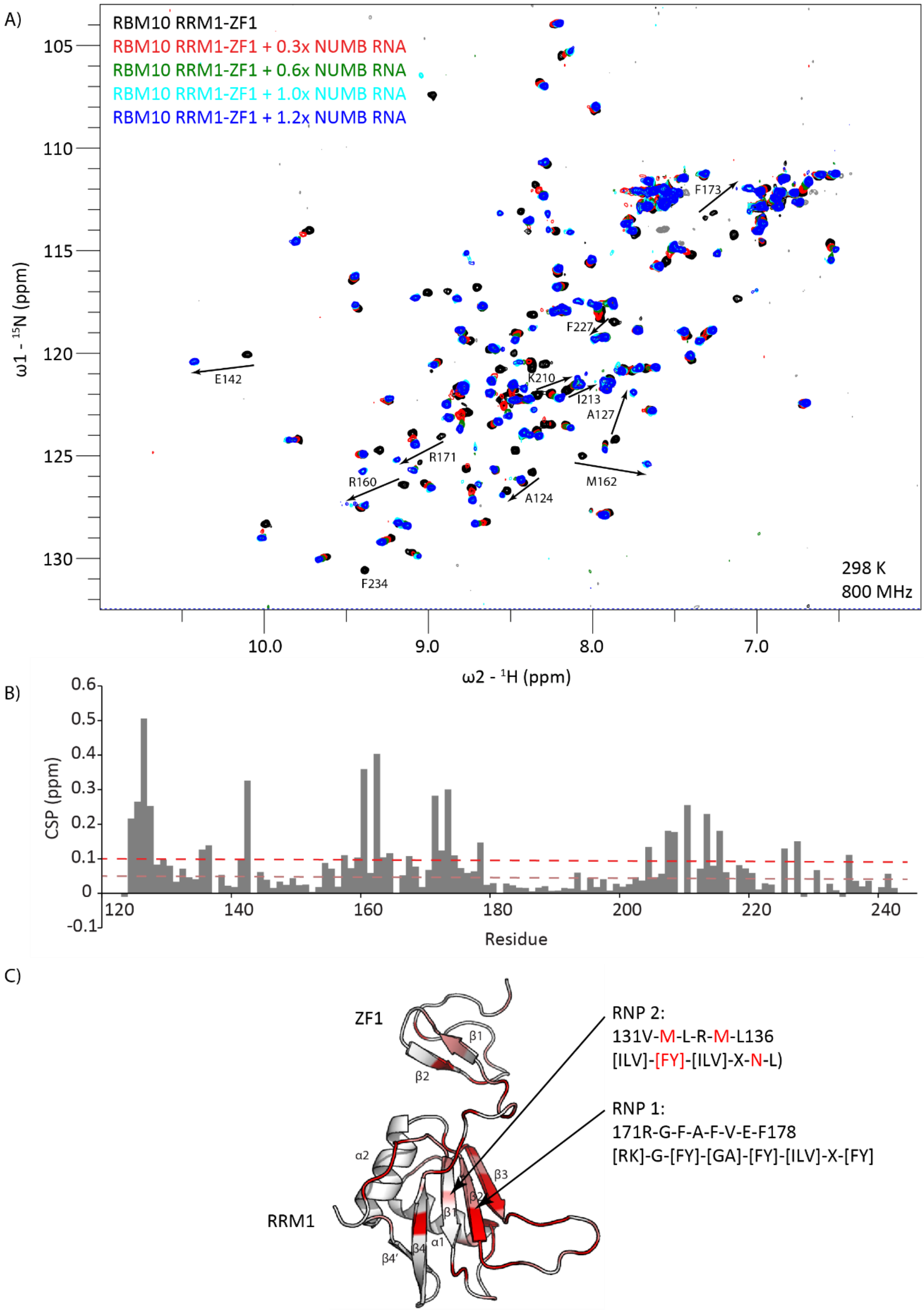
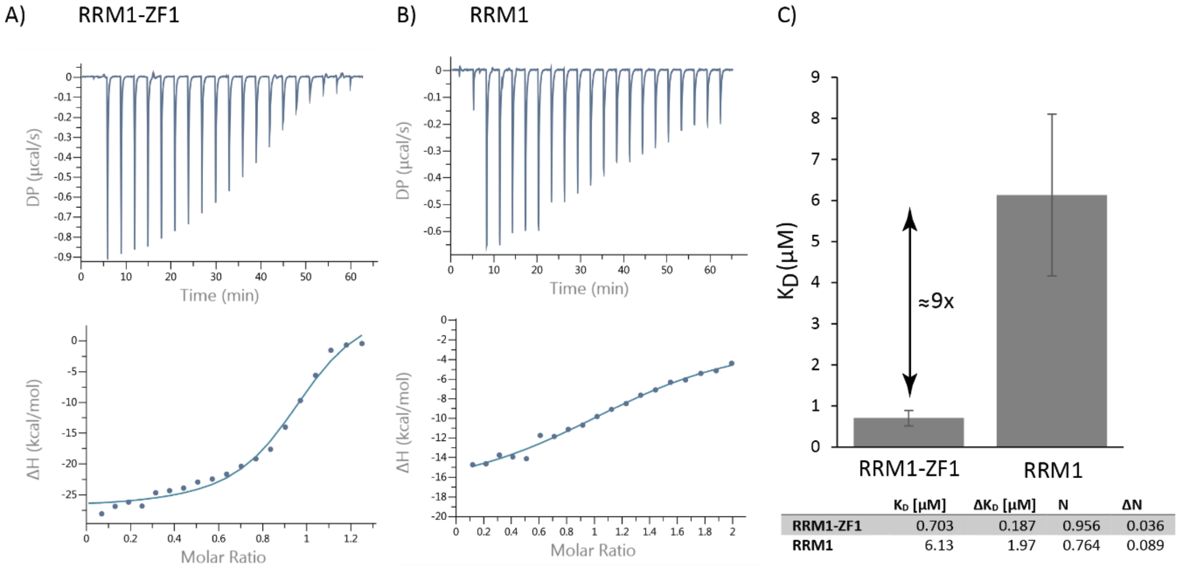


Figure 18 NMR titration of RBM10 RRM1-ZF1 with NUMB derived 12-mer-RNA oligo: *See previous page for the figure note.*



	K_D [µM]	ΔK_D [µM]	N	ΔN
RRM1-ZF1	0.703	0.187	0.956	0.036
RRM1	6.13	1.97	0.764	0.089

Figure 19 ITC-data for RRM1-ZF1 and RRM1 with NUMB derived 12-mer RNA oligo: *ITC data for RRM1-ZF1 construct (A) and RRM1 construct (B) titrated with the NUMB derived 12-mer RNA oligo. C) Plot of dissociation constant for RRM1-ZF1 and RRM2 derived from A and B. Values are listed in the table below. Additionally the binding stoichiometry N is provided.*

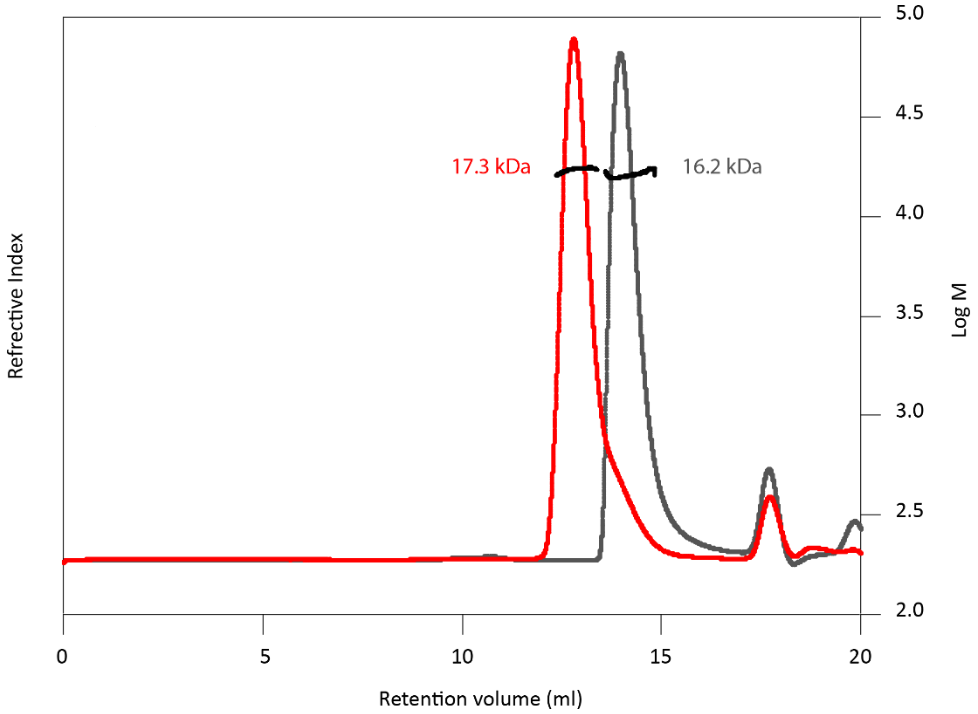


Figure 20 Small angle light scattering data for RBM10 RRM-ZF1 RNA complex: *The refractive index in arbitrary units is plotted for RBM10 RRM1-ZF1 in absence (grey) and in presence (red) of the NUMB derived 12-mer RNA oligo (red). In black the logarithm of the molecular weight is plotted for the main peak of both data sets. The molecular weight extracted is indicated next to the peaks in matching colors.*

To further validate the binding stoichiometry of the complex between RBM10 RRM1-ZF1 and the NUMB derived 12-mer RNA oligo static light scattering (SLS) has been measured by Arie Geerlof. The RRM1-ZF1 construct has an expected mass of 13.9 kDa and the NUMB derived 12-mer RNA oligo has an expected mass of 3.7 kDa. The SLS measurement yielded a mass of 16.2 kDa for the apo-protein, which is around 15% heavier than expected. For the complex of RRM1-ZF1 with the NUMB derived 12-mer RNA oligo a mass of 17.3 kDa was measured, which is quite similar to the calculated mass of 1:1 complex of protein and RNA (13.9 kDa + 3.7 kDa = 17.6 kDa) and far away from theoretical molecular weights of 1:2 (13.9 kDa + 2·3.7 kDa = 21.3 kDa) or 2:1 (2·13.9 kDa + 3.7 kDa 31.5 kDa) complex (**Figure 20**). The 1:1 stoichiometry determined by SLS fits well to the NMR titrations and to the ITC data.

Since the 12-mer-NUMB RNA induces CSP in RRM1 and ZF1 and the domains are flexible to each other in the apo-protein, the next question to answer is if the RNA makes the domain orientation of the two domains rigid. For this purpose the same set of relaxation experiments, which was recorded for the apo-protein, was recorded for the RNA saturated RRM1-ZF1 construct.

T1 (**Figure 21A**) and T1ρ (**Figure 21B**) relaxation data were used to estimate local correlation times τ_C (**Figure 21C**). The average τ_C value for RRM1 in the RNA bound form is 12.1 +/- 0.8 ns/rad, while the average τ_C for ZF1 is 10.7 +/-0.6 ns/rad. Both domains tumble slower when bound to RNA, but the correlation time of ZF1 in the RNA bound form is still smaller than the correlation time of RRM1. The difference in correlation time between the two domains in the RNA bound form is 1.4 ns/rad, which is less than in the apostate (2.6 ns/rad) (**Figure 22**). This data indicates that the domain orientation becomes more rigid upon RNA binding, but not completely rigid, even so the errors for the correlation times for the RNA bound form are quite high.

The $\{^1H\},^{15}N$-heteronuclear nOe data of the RNA bound form or RRM1-ZF1 looks very similar to the $\{^1H\},^{15}N$-heteronuclear nOe data of the free form (**Figure 21D**). The only significant difference is that the N-terminus of the RRM1-ZF1 construct is less flexible in the RNA bound form compared to the free form, since the $\{^1H\},^{15}N$-heteronuclear nOe data show small positive instead of negative values. The N-terminus also showed strong CSP upon RNA binding. These observations might be explained by a transient interaction of the N-terminus with the RNA. If the linker between RRM1 and ZF1 becomes rigid cannot be answered because the corresponding peaks were overlapped with others in the $^1H,^{15}N$-correlation spectrum of either the free or the RNA bound form. For the structure determination of the protein RNA complex it is necessary to obtain intermolecular nOes, to use as distance restraints between the RNA and the protein.

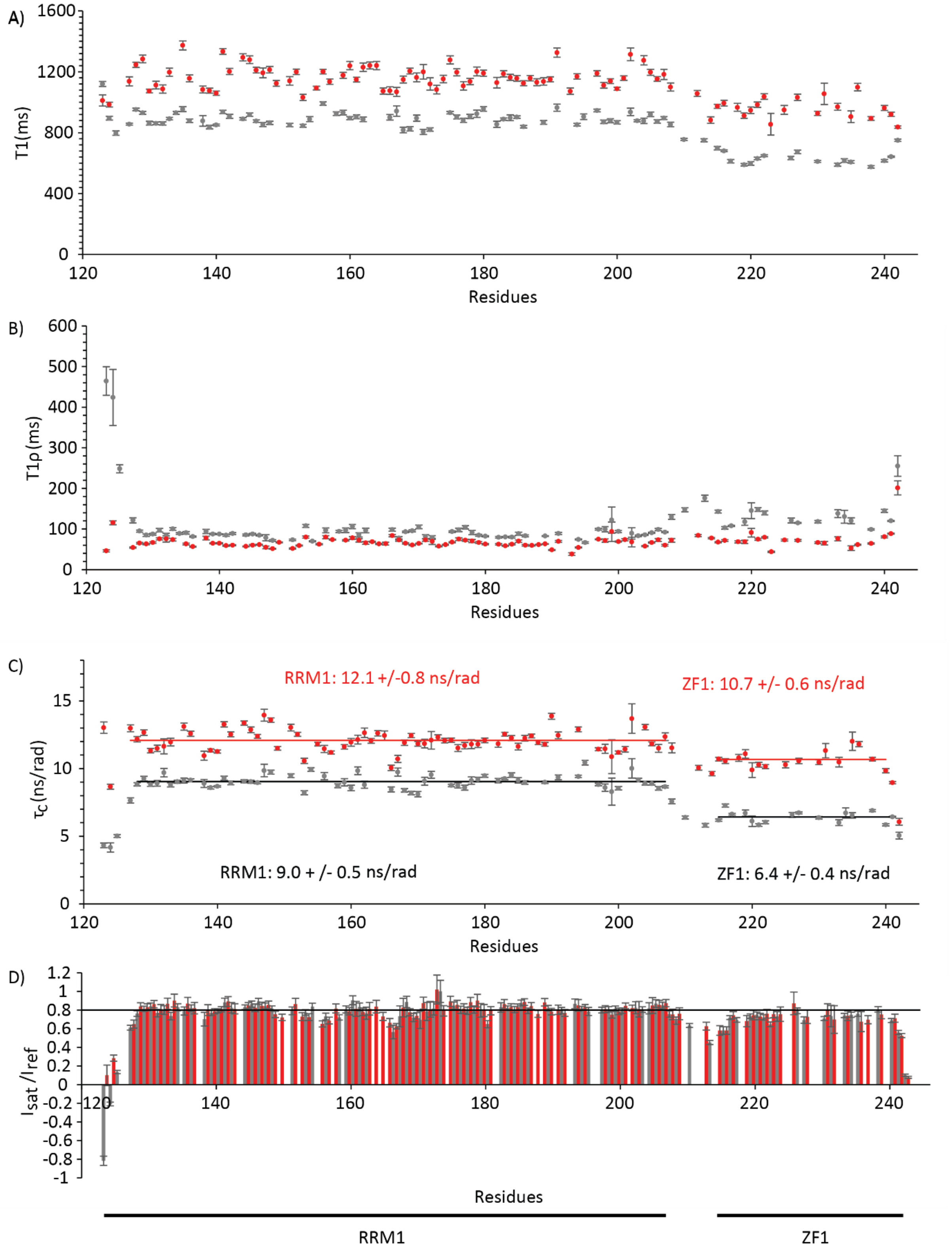


Figure 21: Relaxation data forRBM10 RRM1-ZF1 NUMB RNA complex: *A) T1 time constant, B) T1ρ time constant, C) correlation time* τ_C *and D)* $\{^1H\},^{15}N$*-heteronuclear nOe of the RBM10 RRM1-ZF1 NUMB RNA complex (red) plotted against the residue number. For comparison the same data of the apo-protein is plotted in grey.*

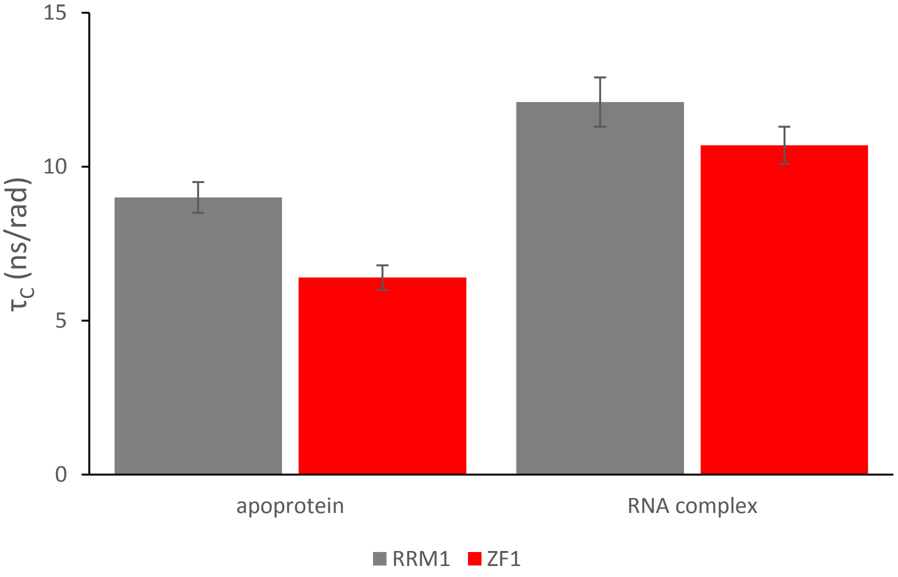


Figure 22 Correlation times of RRM1-ZF1 domains free and bound to RNA: *The correlation time of RRM1 (grey) and ZF1 (red) within the RRM1-ZF1 construct are shown in the free and in the RNA bound form. The correlation time have been averaged over the domains. The error represents the standard derivation.*

6.1.3 NUMB-derived RNA oligo shows strong line broadening in complex with RBM10

So far the interaction between RRM1-ZF1 and the NUMB derived 12-mer RNA oligo has been characterized only from the protein side. In the next step TOCSY spectra of the RNA free and bound to RBM10 RRM1-ZF1 will be presented. The H5,H6-correlations of the pyrimidine bases in RNA give a good fingerprint spectrum, since for each pyrimidine one peak appears. The free NUMB derived 12-mer RNA oligo (UUGUCUGCUCCC) comprises 5 uracil and 5 cytosine nucleotides, therefore ten H5,H6-correlation are expected in the spectrum of this oligo. The TOCSY spectrum of the free NUMB derived 12-mer oligo indeed shows ten H5,H6-correlations (numbered in **Figure 23A**). Seven of these H5,H6 correlations (1-7) give rise to strong peaks of similar intensity, while the last three (8,9 and 10) give rise to weaker peaks, with decreasing intensity (8>9>10). The peaks 1 to 9 are so sharp that the homonuclear coupling can be observed and that the peaks appear as duplets.

When the oligo is saturated with RBM10 RRM1-ZF1 the peaks become strongly broadened and overlap severely. Overall no strong changes in chemical shift can be resolved (**Figure 23B**). Changing the field from 600 MHz to 800 MHz leads to slightly sharper peaks, but the peaks are still overlapped and no chemical shift changes can be resolved (**Figure 23C**). The strong line broadening can be caused by exchange and indicates that the protein-RNA complex is not completely rigid. The RNA does not seem to bind in one well defined binding register, which is crucial for structure determination.

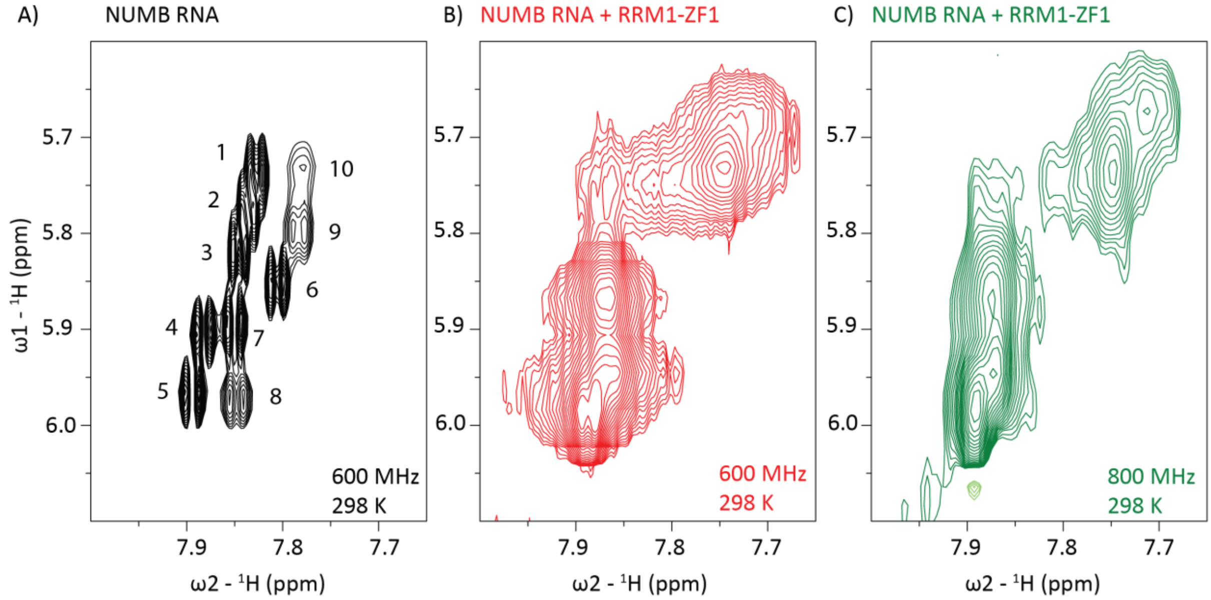


Figure 23 NUMB-RNA-TOCSY spectra with and without RBM10 RRM1-ZF1: *Shown are homonuclear 2D TOCSY spectra of the NUMB derived 12-mer RNA oligo (UUGUCUGCUCCC) free measured at 600 MHz (A), saturated with RBM10 RRM1-ZF1 at 600 MHz (B) and at 800 MHz (C). The shown part of the spectrum comprises the H5,H6-correlations of the pyrimidine nucleotides in RNA. The pyrimidine nucleotides in the NUMB derived 12-mer RNA oligo are highlighted in red.*

Nevertheless, with the NUMB derived 12-mer RNA oligo sample, which was saturated with ^{13}C- and ^{15}N-isotope labeled RBM10 RRM1-ZF1, a 2D ω1-filtered NOESY experiment was recorded. The spectrum shows the intramolecular cross peaks of the unlabeled RNA (green boxes in **Figure 24**). The RNA signals cluster in three areas, between 4 and 5 ppm, where the peaks corresponding to sugar atoms appear, around 6 ppm, where peaks corresponding to the H1' sugar atoms and the H5 of uracil and cytosine appear, and around 8 ppm, where the H2 of adenine, the H6 of uracil and cytosine and the H8 of adenine and guanine appear. The intramolecular cross peaks between the described regions appear twice symmetric to the diagonal (green boxes in **Figure 24**). The signals are broad, similar to the H5,H6-correlation in the TOCSY spectra. In addition, intermolecular cross peaks between all three regions of the RNA and the protein can be observed. There is intermolecular cross peaks from the RNA to aromatic and to aliphatic protons of the protein. The intermolecular cross peaks appear only once with the frequency of the RNA proton on the ω1 axis and the frequency of the protein protons on the ω2 axis (red boxes in **Figure 24**). Most of the intermolecular cross peaks are quite weak. The spectrum also shows some intramolecular cross peaks of the protein, which have passed the filter, but these again appear twice on both side of the diagonal. The intramolecular protein cross peaks probably originate from residual ^{12}C-isotopes in the protein due to preparation of the pre-culture in not ^{13}C-isotope labelled medium. Because of the strong line broadening in the RNA signals upon addition of the protein, the structure determination of the protein:RNA complex was not further perused.

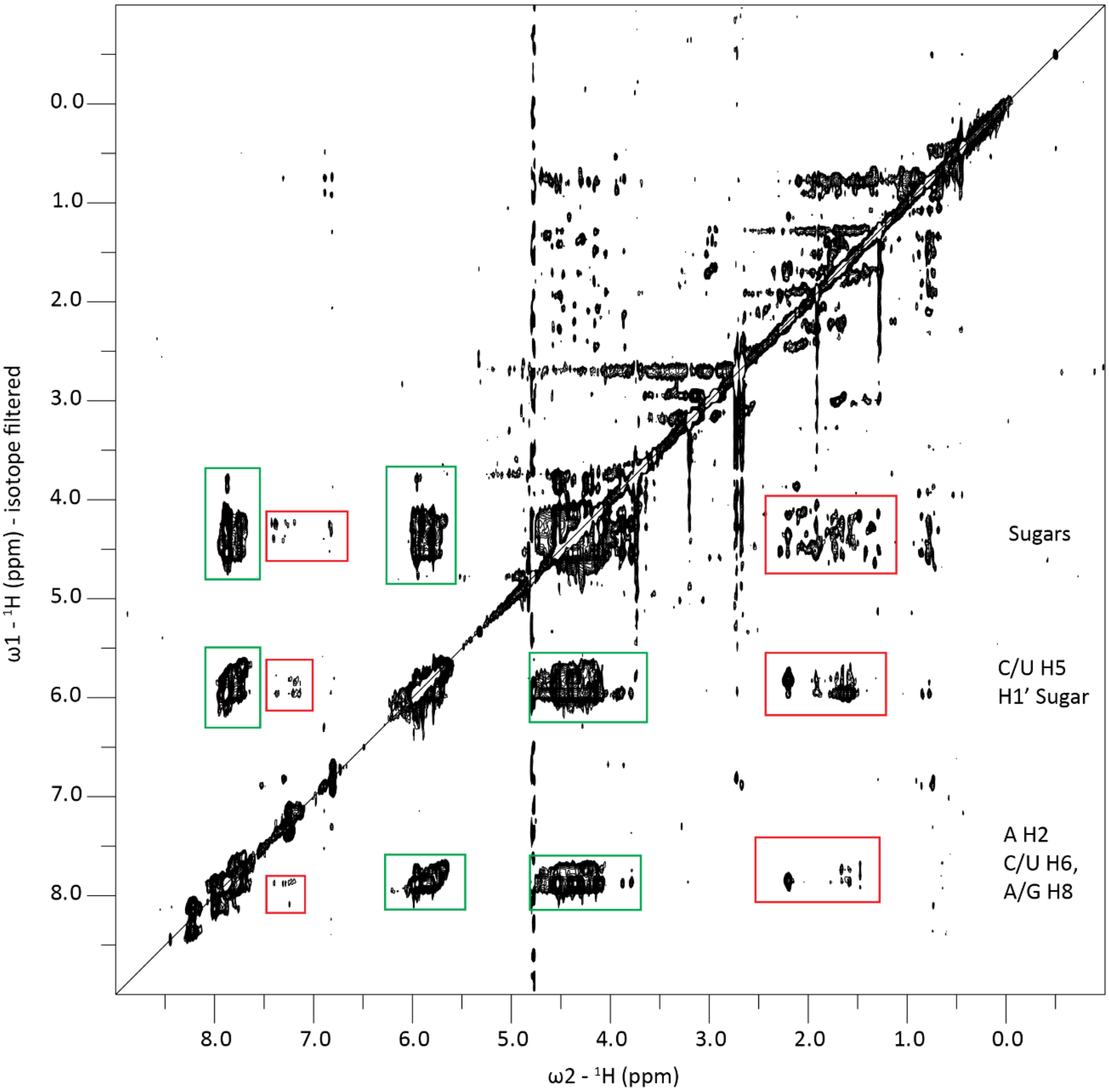
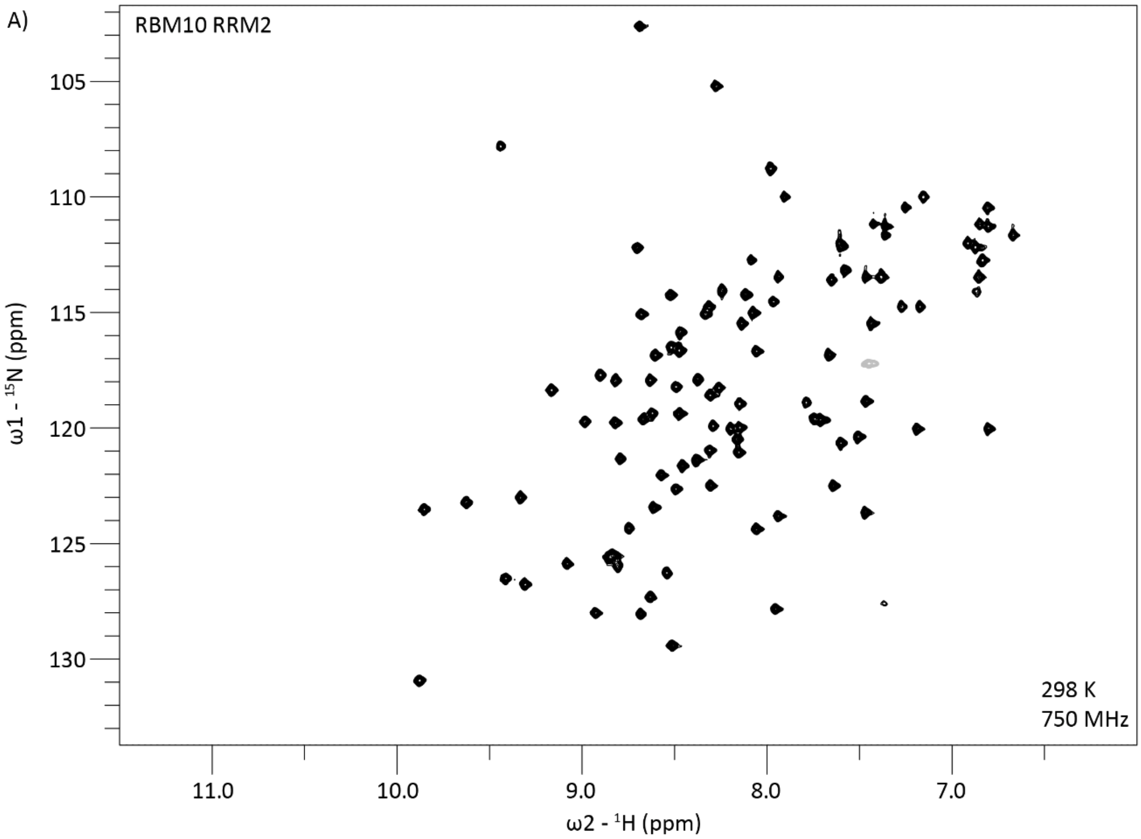


Figure 24 ω1-filtered NOESY spectrum of the NUMB derived RNA 12-mer oligo saturated with RBM10 RRM1-ZF1: *Shown is a ω1-filtered NOESY spectrum of the unlabeled NUMB derived RNA 12-mer oligo, which is saturated with ^{13}C- and ^{15}N-isotope labelled RBM10 RRM1-ZF1. Areas with intramolecular RNA cross peak are marked with green boxes. The intramolecular cross peaks appear twice axis symmetric to the diagonal. Areas with intermolecular cross peaks between the RNA and the protein are marked by red boxes. Labels on the right indicate which RNA protons give rise to signal in that area of the spectrum.*

6.1.4 RRM2 structure

The third RNA binding domain in RBM10 is RRM2, which is C-terminal of ZF1 and connected by a 57 amino acid linker. In the upcoming chapters only the -V354 isoform, will be discussed. The differences to the +V354 isoform will be discussed in chapter 6.3.1 . The ^{1}H,^{15}N-correlation spectrum of RBM10 RRM2 is well dispersed and has uniform peak intensity indicating a well folded domain without larger flexible parts (**Figure 25A**). The backbone of the domain has been assigned to 98% using an HNCACB spectrum (**Figure 25B**).



B)

```
      300   305   310   315   320   325   330   335   340   345   350
  GAM DTIIL RNLNP HSTMD SILGA LAPYA VLSSS NVRVI KDKQT QLNRG FAFIQ LSTIE
355   360   365   370   375   380
AAQLL QILQA LHPPL TIDGK TINVE FAKG
```

Figure 25 Assignment of RBM10 RRM2: *A) ^{1}H,^{15}N-correlation spectrum of RBM10 RRM2. Positive peaks are plotted in black and negative peaks are plotted in gray. B) Primary sequence of RBM10 RRM2 construct. The residues printed in italic indicate artificial residues due to cloning. Proline residues are printed in gray since they don't give rise to signals in the ^{1}H,^{15}N-correlation spectrum. Assigned and unassigned residues are colored in green and red, respectively.*

For structure calculation also the side chain protons have to be assigned. The assignment of all protons was completed to 98.9% using HBHA(CO)NH, HCCH TOCSY, CBCGCDHD, CBCGCDCEHE, aromatic ^{13}C-edited NOESY experiments. NOE cross peaks of ^{15}N-edited and aromatic and aliphatic ^{13}C-edited NOESY spectra were assigned using the software Cyana. In addition to distance restraints dihedral restraints from the software TALOS were used. The used restraints are listed in **Table 10**.

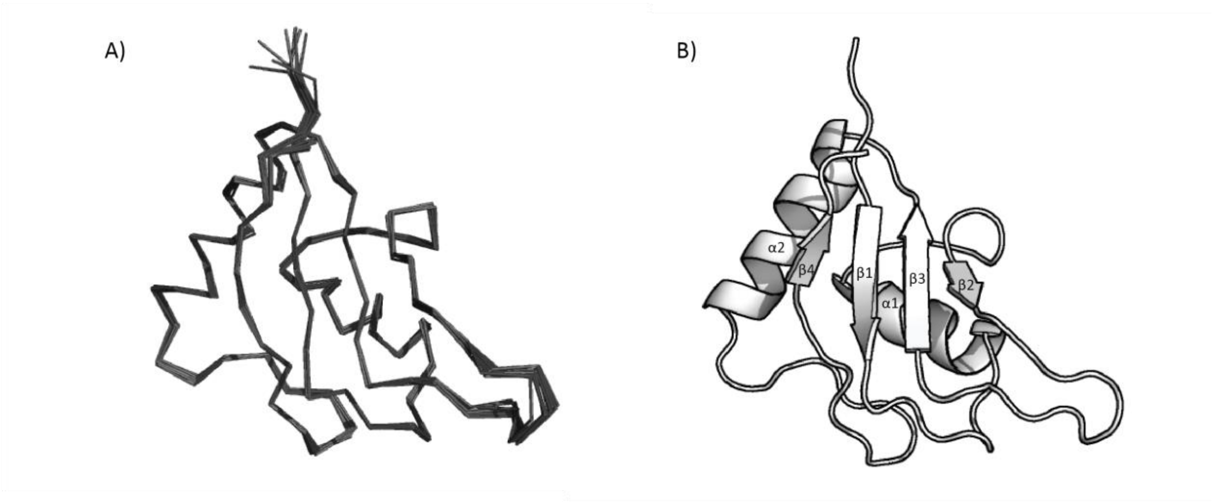


__Figure 26 RBM10 RRM2 NMR structure:__ A) Ribbon plot of the 10-structure water refined NMR bundle of RBM10 RRM2. B) Cartoon plot of the RBM10 RRM2 structure. Secondary structure elements are labelled.

Table 10 Structure calculation statistics for water refined RBM10 RRM2 structure:

Structure calculation restraints	
Distance restraints	
Total NOEs	2074
Sequential (|i- j| ≤ 1)	944
Medium-range (1<|i- j| <5)	434
Long-range (|i- j| ≥ 5)	696
Dihedral restraints (ϕ+ψ)	128
Quality analysis	
Restraints violations (mean ± s.d)	
Distance restraints (Å)	0.021 ± 0.001
Dihedral angle restraints (º)	0.390 ± 0.06
Deviation from idealized geometry	
Bond length (Å)	0.004 ± 0.00001
Bond angles (º)	0.502 ± 0.008
Improper dihedral distribution (º)	1.31 ± 0.075
Average pairwise r.m.s. deviation (Å)	
Heavy	0.69 ± 0.08
Backbone	0.42 ± 0.09
Ramachandran values (%) (MolProbity)	
Most favored regions	94.7 %
Allowed regions	99.8 %

The dihedral and the unambiguous distance restraints generated by Cyana, where then used in ARIA2 to generate 100 structures, from these the ten structures with the lowest energy were water refined. The quality analysis of this water refined structure is also listed in **Table 10** and the structural bundle is shown in **Figure 26A**. The structure of RBM10 is a canonical RRM fold, with a βαββαβ topology (Cléry, Blatter, and Allain 2008). The four β-strands from an antiparallel β-sheet, which is packed against the two α-helices in the back (**Figure 26B**).

6.1.5 RNA binding of RRM2

To characterize the RNA binding of RRM2 the same NUMB derived 12-mer RNA oligo, which was used for the titration of RBM10 RRM-ZF1 was used for an NMR titration of RBM10 RRM2. During the titration most of the peaks are in the fast exchange regime on the NMR time scale. Some peaks (e.g. I334 and F345) show fast-intermediate exchange. The peaks of residues D300, S351, T352 and E379 were not detectable in any of the titration steps, due to line broadening. The titration is not saturated at a 1:1 protein:RNA ratio, since further addition of RNA to 1.5x molar equivalents still leads to stronger CSP (**Figure 27A**). In case of RRM2 the protein:RNA ratios are less accurate since the RRM2 construct does not comprise a tryptophan residue, which makes concentration determination of the RRM2 construct by UV-absorption less accurate.

High CSP upon RNA binding clusters in the β-sheet, involving all four β-strands. The two RNP motifs diverge from the canonical motif. In the RNP 1 (Consensus: [RK]-G-[FY]-[GA]-[FY]-[ILV]-X-[FY]) in position 8, instead of an aromatic amino acid leucine is present and in the RNP 2 (Consensus: [ILV]-[FY]-[ILV]-X-N-L) in position 2 instead of the aromatic amino acid an isoleucine is present. Since both β-strand 1, containing the RNP2, and β-strand 3, containing RNP1, show high CSP upon RNA binding, both RNP motifs of the RRM2 of RBM10 seem to contribute to RNA recognition.

Additionally the N-terminal part of α-helix 2 and the loop connecting it with β-strand 3 show high chemical shift perturbation. This is the same area were in the splicing isoform the residue V354 is inserted. The effect of this insertion on RNA binding will be discussed in chapter 6.3.2.

To determine the K_D of the interaction ITC experiments have been conducted with RBM10 RRM2 and the NUMB derived 12-mer RNA oligo. RRM2 and the RNA have a K_D of 23.8 µM. This means the affinity of RRM2 to the tested RNA is around 4 times lower than the affinity of RRM1 and around 34 times lower than the affinity of RRM1-ZF1. The interplay between the different domains will be discussed later in chapter 6.2.2. The stoichiometry of 1.47 cannot be trusted since determination of the stoichiometry by ITC is strongly dependent on an accurate determination of the protein concentration, which is difficult for RRM2 because of the lack of tryptophan residues.

***Figure 27**: **NMR titration of RBM10 RRM2 with NUMB derived 12-mer-RNA oligo:** A) Superimposed $^{1}H,^{15}N$-correlation spectra of RBM10 RRM2 apo-protein (black) and titration steps with 0.3x (red), 0.6x (green), 1.0x (cyan) and 1.5x (blue) molar equivalents of RNA present. Peaks in the black and the blue spectrum of residues that experience high chemical shift perturbation are connected by arrows and labeled. Residues D300, S351, T352 and E379, which could not be detected in the complex spectrum, are labeled as well. B) Chemical shift perturbations between RBM10 RRM2 free and in complex with the NUMB derived 12-mer RNA oligo are plotted against the residue number. C) CSP from B that are higher than 0.05 ppm (red dotted line) plotted on the structure of RRM2. The name of the secondary structure elements is indicated. Additionally the sequence of the two RNP motifs is indicated and compared to the consensus motif. Red marked residues do not match. Also the position of the extra V354 residue in the +V354 isoform is indicated. (Figure on next page)*

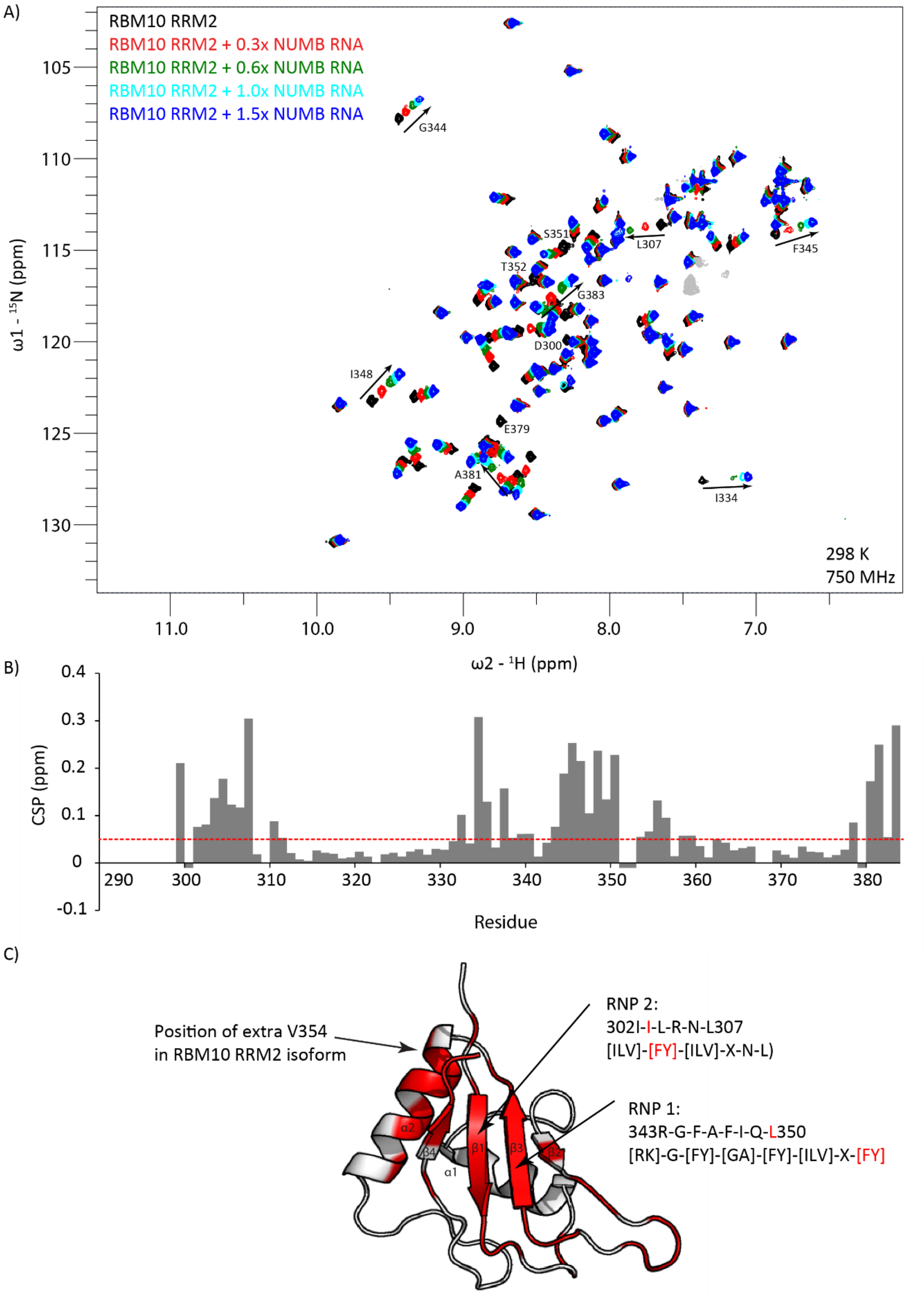
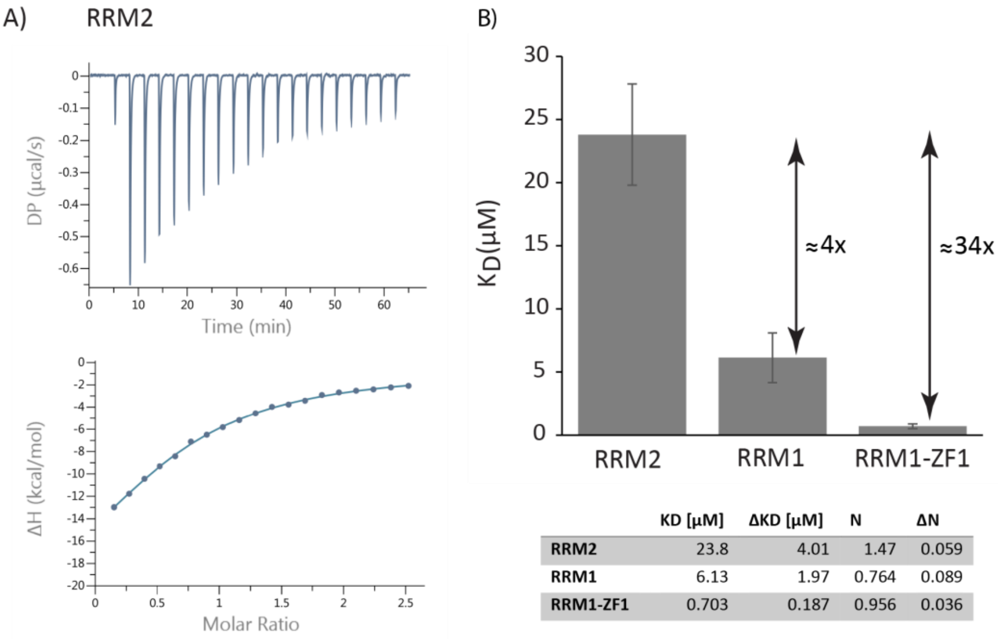


Figure 27 NMR titration of RBM10 RRM2 with NUMB derived 12-mer-RNA oligo: *See previous page for the figure note.*



	KD [µM]	ΔKD [µM]	N	ΔN
RRM2	23.8	4.01	1.47	0.059
RRM1	6.13	1.97	0.764	0.089
RRM1-ZF1	0.703	0.187	0.956	0.036

Figure 28 ITC-data for RRM2 with NUMB derived 12-mer RNA oligo: *A) ITC data for RRM2 construct titrated with the NUMB derived 12-mer RNA oligo. C) Plot of dissociation constant for RRM2 derived from A, RRM1-ZF1 and RRM2 derived from A and B in* ***Figure 19****. Values are listed in the table below. Additionally the binding stoichiometry N is provided.*

6.2 Domain organization and RNA recognition by RBM10 RRM1-ZF1-RRM2

6.2.1 RNA binding domains of RBM10 are independent in the apo-protein

The question remains if the three RNA binding domains RRM1, ZF1 and RRM2 collaborate to recognize RNA. To test if the domains interact with each other in the apo-protein the ^{1}H,^{15}N NMR-correlation spectra of the three domain-construct comprising RRM1, ZF1 and RRM2 were compared to separate spectra of RRM-ZF1 and of RRM2 (**Figure 29A**). The ^{1}H,^{15}N NMR correlation spectrum of the three domain spectrum shows additional peaks compared to the spectra of RRM1-ZF1 and RRM2 in the middle of the protein frequency range, which correspond to the 57 residue flexible linker between ZF1 and RRM2. While RRM1 residues have nearly identical chemical shifts in the three-domain construct as in the RRM1-ZF1 construct, ZF1 and RRM2 have several residues, which have different chemical shifts in context of the three-domain-construct. In ZF1 especially residues close to the N-terminus of the RRM1-ZF1 construct are strongly effected (**Figure 29B**). Other effected residues are located in the N-terminal β-strand 1 of ZF1 and in loops (**Figure 29C**). Residues with different chemical shift in RRM2 cluster in three different parts of the domain: The N-terminus, in the middle around residue 350 and the C-terminus (**Figure 29B**). These clusters correspond to the β-strand1, a stretch from β-strand 3 up to the N-terminal part of α-helix 2 and β-strand 4. All these elements are close in space to each other (**Figure 29C**).

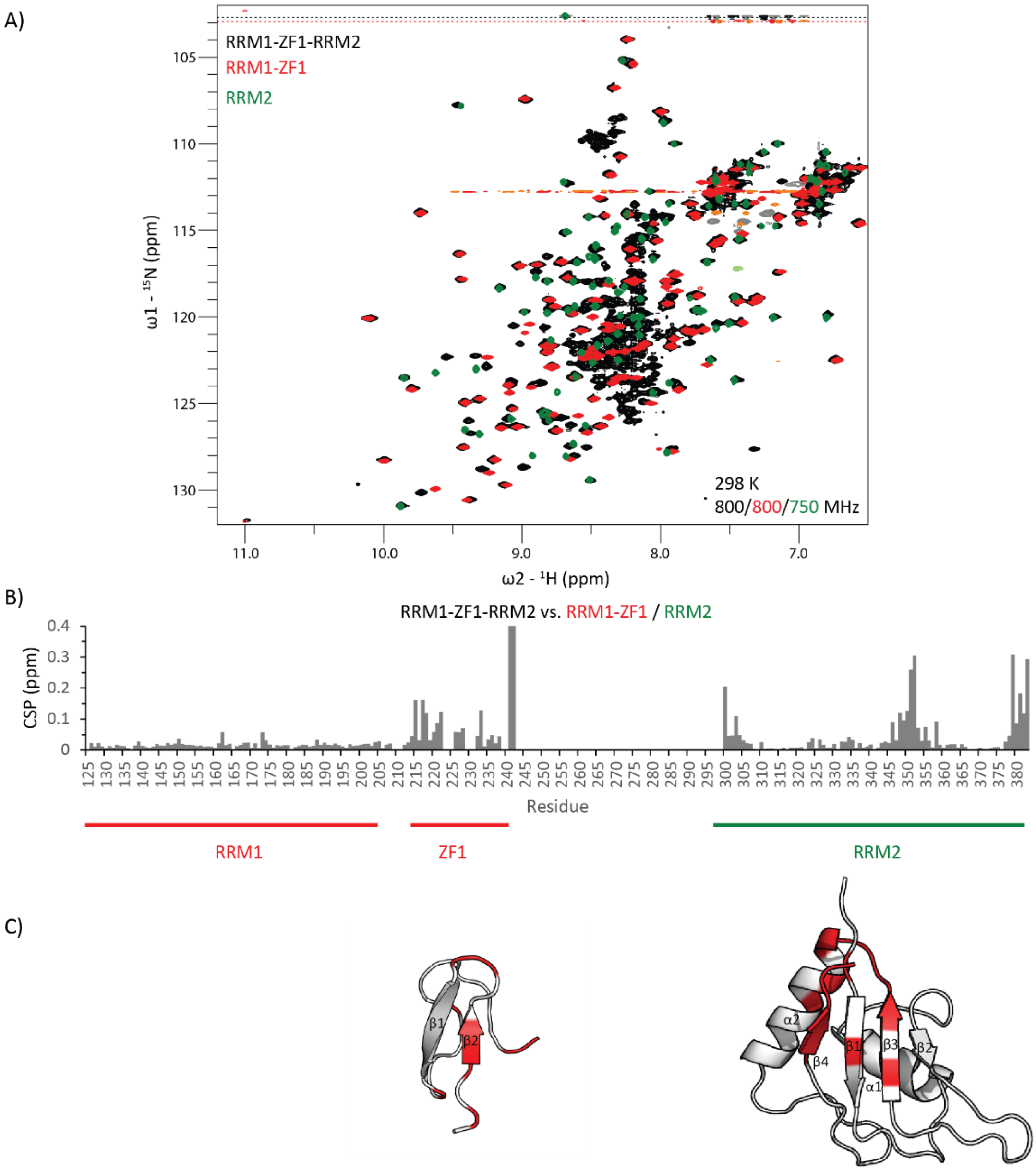


Figure 29 Three-domain construct vs. RRM1-ZF1/RRM2: *(A) ^{1}H,^{15}N-NMR correlation spectra of RBM10 RRM1-ZF1-RRM2 (black) superimposed with spectra of RBM10 RRM1-ZF1 (red) and RRM2 (green). (B) Chemical shift perturbation (CSP) plots of RRM1-ZF1-RRM2 against RRM1-ZF1 and RRM2 in separate constructs. (C) CSP > 0.06 ppm are colored in red on the structures of ZF1 (left) and RRM2 (right).*

To test if the observed differences between the three domain construct and the RRM-ZF1 and RRM2 constructs are caused by an interaction between the two domains or between the two domains and the linker connecting them, two constructs comprising RRM1-ZF1 and the linker (RRM1-ZF1-Linker) and the linker and RRM2 (Linker-RRM2), respectively, were generated.

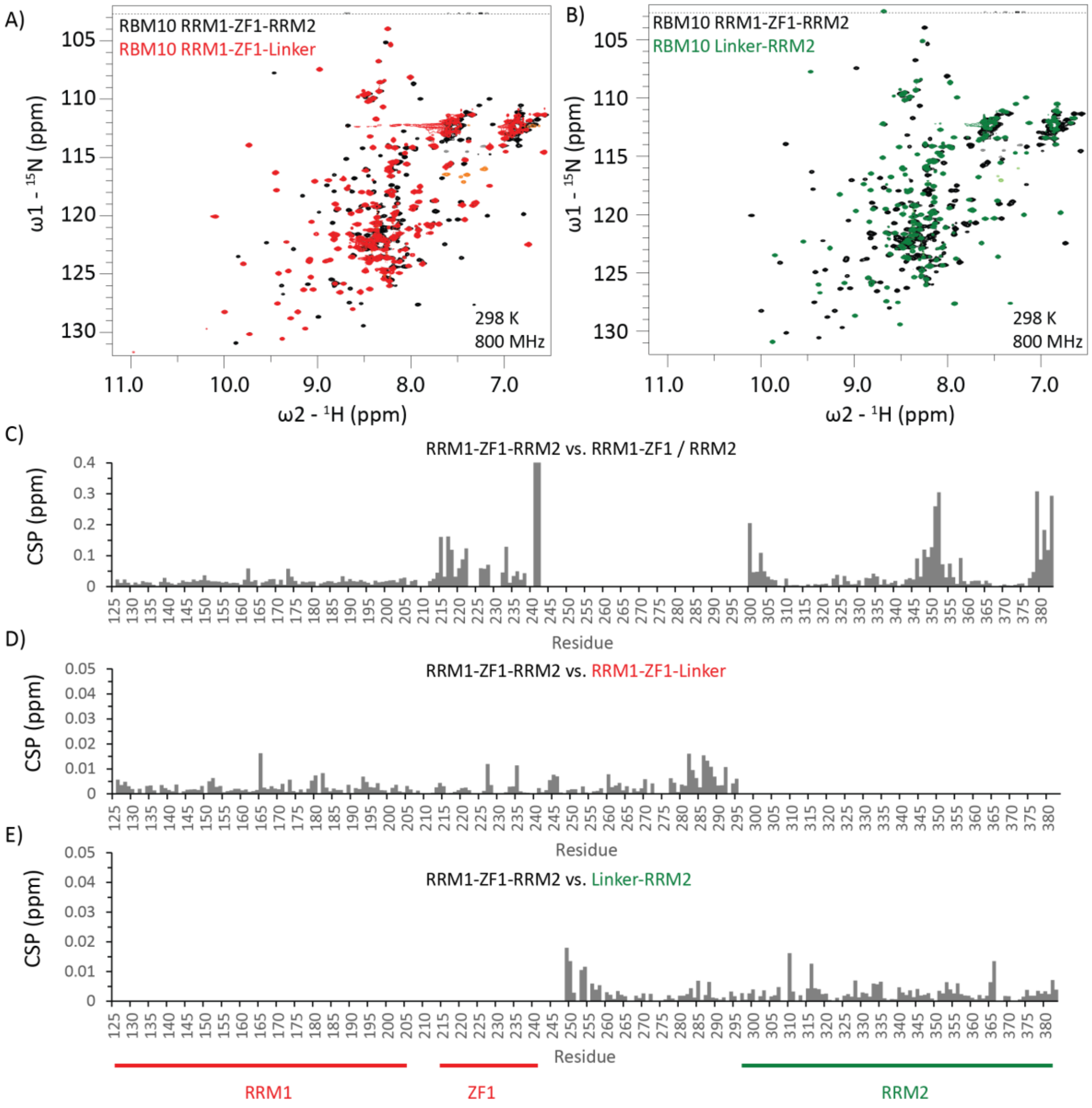


Figure 30 RBM10 RRM-Z1 and RRM2 linker constructs: *^{1}H,^{15}N-correlation spectra of RBM10 RRM1-ZF1-RRM2 (black) superimposed with spectra of RBM10 RRM1-ZF1-Linker (A, red) or Linker-RRM2 (B, green). Chemical shift perturbation (CSP) plots of RRM1-ZF1-RRM2 against RRM1-ZF1 and RRM2 in separate constructs (C), or against the RRM1-ZF1-Linker construct (D), or against the Linker-RRM2 construct (E).*

The ^{1}H,^{15}N NMR correlation spectra of RRM1-ZF1-Linker does not show the differences in chemical shift in ZF1 observed between the three-domain construct and RRM1-ZF1 **(Figure 30A, C and D).** Furthermore, the superimposition of the ^{1}H-^{15}N NMR correlation spectra of the three-domain construct with the Linker-RRM2 construct does not show the differences in chemical shift observed between the three-domain construct and the RRM2 domain **(Figure 30B, C and E)**. Both linker construct show small chemical shift perturbation compared to the three-domain-construct in the free termini of the linker. This is an indication that the CSP in ZF1 observed between the RRM-ZF1 and the three-domain construct are caused by the N-terminus of the linker between ZF1 and RRM2. In a similar way the differences in the β-sheet of RRM2 are probably caused by the C-terminus of the linker between ZF1 and RRM2, which probably effects the β-strand 1, which is directly at the N-terminus of the RRM2 domain.

To get an idea of the dynamic behavior of RRM1, ZF1 and RRM2 within in the three-domain-construct ^{15}N relaxation data were recorded (**Figure 31**).

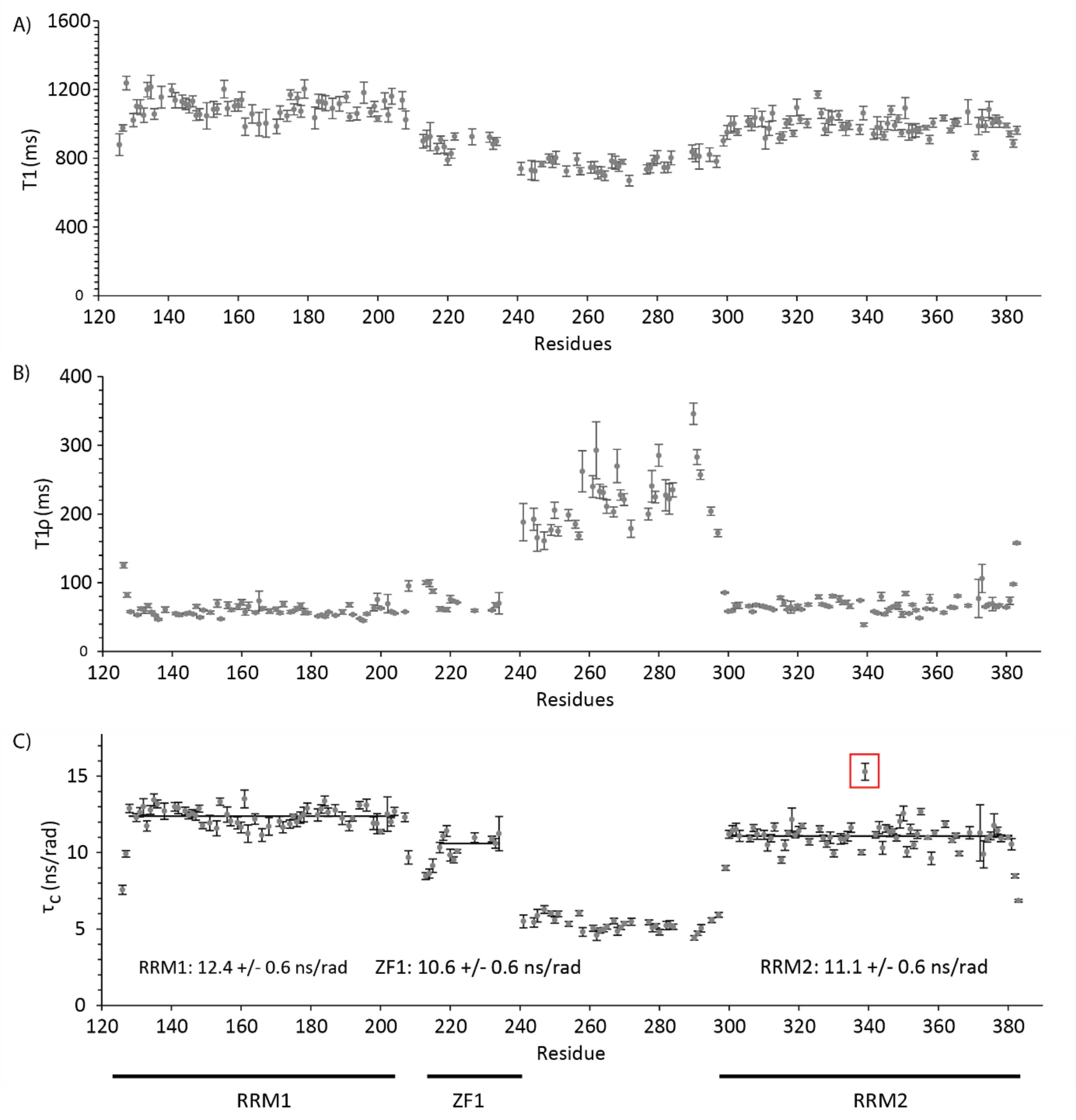


Figure 31 Relaxation data for RBM10 RRM1-ZF1-RRM2: *A) T1 time constant, B) T1ρ time constant and C) correlation time of the RBM10 RRM1-ZF1-RRM2 apo-protein plotted against the residue number. The red box in C indicates an outlier, which was not used for the calculation of the average correlation time of RRM2.*

In context of the three domain RRM1 has a correlation time of 12.4 ns/rad, while ZF1 and RRM2 have lower correlation times of 10.6 ns/rad and 11.1 ns/rad respectively. The long 57 residue between ZF1 and RRM2 is highly flexible with a correlation time around 5 ns/rad. In the three-domain-construct RRM1 and ZF1 both have higher correlation times than in context of the RRM1-ZF1 construct, which can be explained through the additional drag of the flexible linker and the RRM2 (**Figure 32**). Since all three domains have different correlation times and the correlation time increases only slightly compared to the RRM1-ZF1 construct, it can be concluded that in absence of RNA, all three domains are flexible relative to each other.

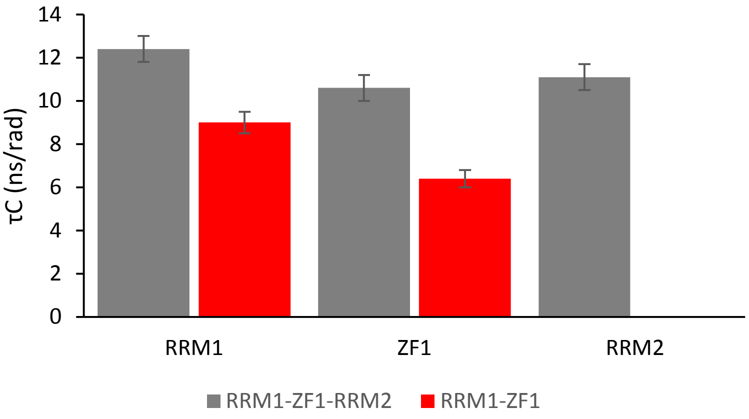


Figure 32 Average domain correlation times per domain for RRM1-ZF1-RRM2: *Plotted are the average domain correlation times for RRM1-ZF1-RRM2 (gray) in comparison with the average domain correlation times of RRM1-ZF1 (red). The error represents the standard derivation.*

6.2.2 All three domains are involved in RNA binding

Knowing that both the RRM1-ZF1 and the RRM2 construct can recognize the NUMB derived 12-mer RNA oligo, the question remains if the two regions cooperate in recognizing RNA. Therefore, an RNA titration of the three-domain construct, comprising all three RNA interacting domains of RBM10, and the NUMB derived 12-mer RNA oligo has been conducted (**Figure 34A**). The titration series is in a fast-intermediate regime on the NMR time scale. It is further in the intermediate regime than the RRM1-ZF1 construct. Peak assignment has been done using the titration series and additionally the known complex chemical shifts of the RRM1-ZF1 construct and the RRM2 construct. Nevertheless the three-domain-construct in complex with the NUMB derived 12-mer RNA oligo is only assigned to 69%. For most signals the three-domain construct is saturated at 1:1, since there are no more changes in chemical shift observable when adding more RNA (1.5x). The observed CSP upon RNA binding for the assigned residues in the three-domain construct for RRM1 and RRM2 is similar to the changes in chemical shift observed for the RRM1-ZF1 and the RRM2 constructs individually upon RNA binding (**Figure 34B and C**). For ZF1 many assignments in the complex of the three-domain-construct with the NUMB derived 12-mer RNA oligo are missing. Therefore no clear answer of the involvement in RNA binding can be given from the titration data. Nevertheless e.g. the residue F234 disappears upon RNA binding for the three-domain-construct like it did for the RRM1-ZF1 construct. The peaks corresponding to the linker between RRM1 and ZF1 that experience strong chemical shift perturbation in the RRM1-ZF1 construct upon RNA binding, are so far not assigned in the three-domain-construct:RNA complex. The long linker between ZF1 and RRM2 does not show changes in chemical shift upon addition on RNA and does therefore not interact with the RNA.

To verify the stoichiometry from the NMR titration SLS data were recorded. RBM10 RRM1-ZF1-RRM2 has an expected molecular weight of 28.9 kDa. With 28.6 kDa the molecular weight determined by SLS matches well. For the complex of RBM10 RRM1-ZF1-RRM2 and the NUMB derived 12-mer RNA oligo a molecular weight of 31.2 kDa was measured (**Figure 33**). This value corresponds best to a 1:1 protein:RNA complex, for which the expected molecular weight is 32.6 kDa. A 1:2 protein:RNA complex would have a molecular weight of 36.3 kDa and is therefore very unlikely. The first half of the peak shows a drop in molecular weight, considering only the second half of the peak a molecular weight of 32.0 kDa is determined. This drop could be caused by an impurity which is lower in molecular weight. The peak of the protein-RNA complex is narrower than the peak for the free protein, indicating a more uniform conformation.

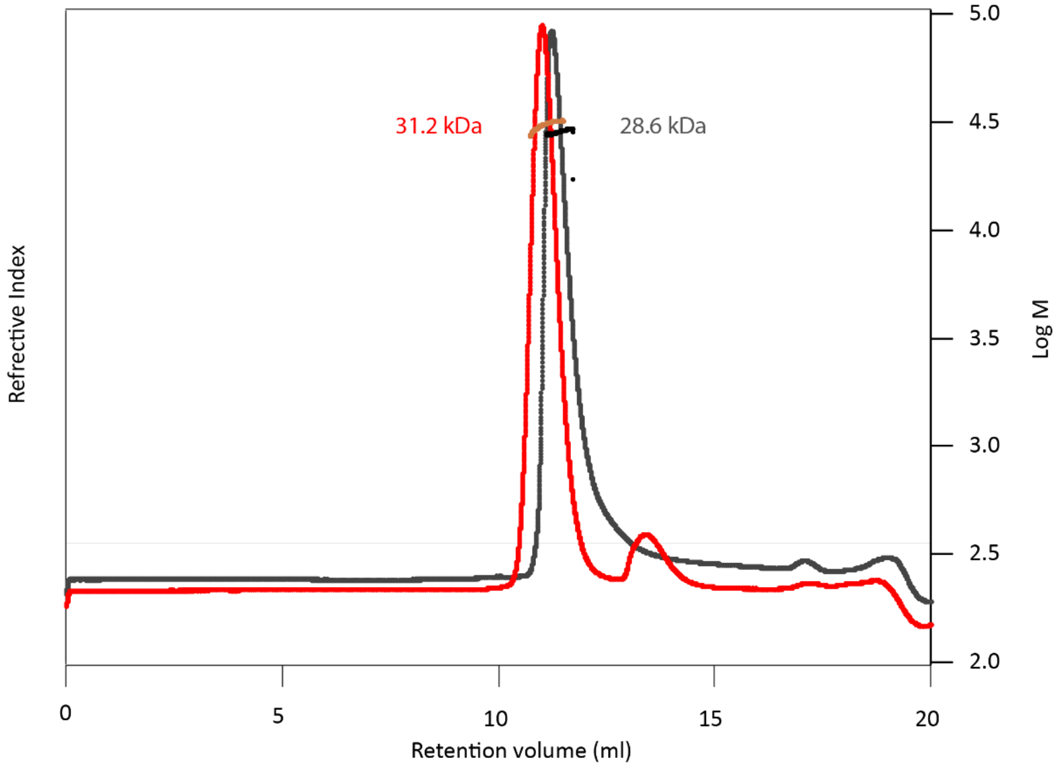


Figure 33 Small angle light scattering of RBM10 RRM-ZF1-RRM2 RNA complex: *The refractive index in arbitrary units is plotted for RBM10 RRM1-ZF1-RRM2 in absence (grey) and in presence (red) of the NUMB derived 12-mer RNA oligo (red). In black and orange the logarithm of the molecular weight is plotted for the main peak of both data sets. The molecular weight extracted is indicated next to the peaks in matching colors.*

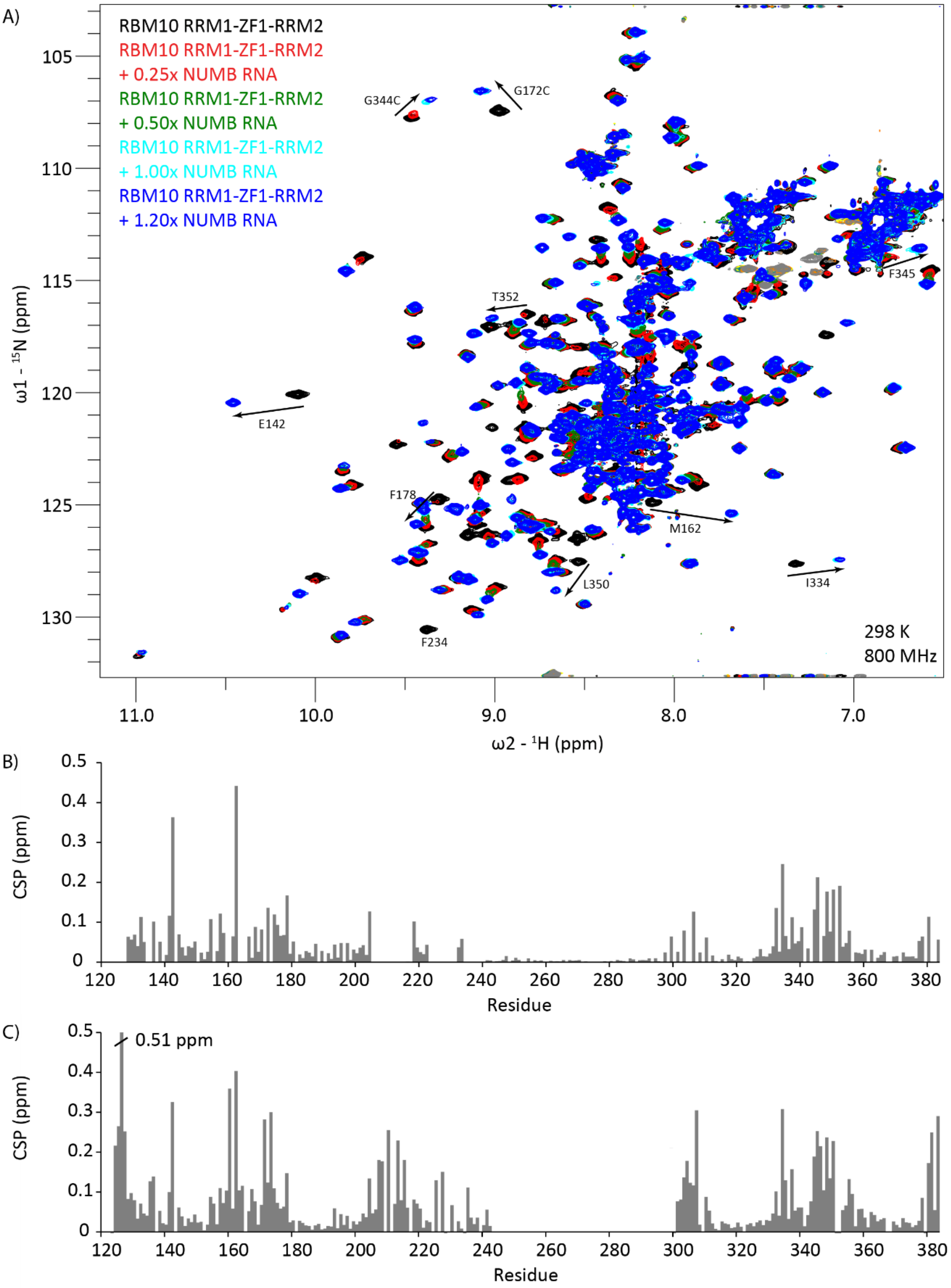
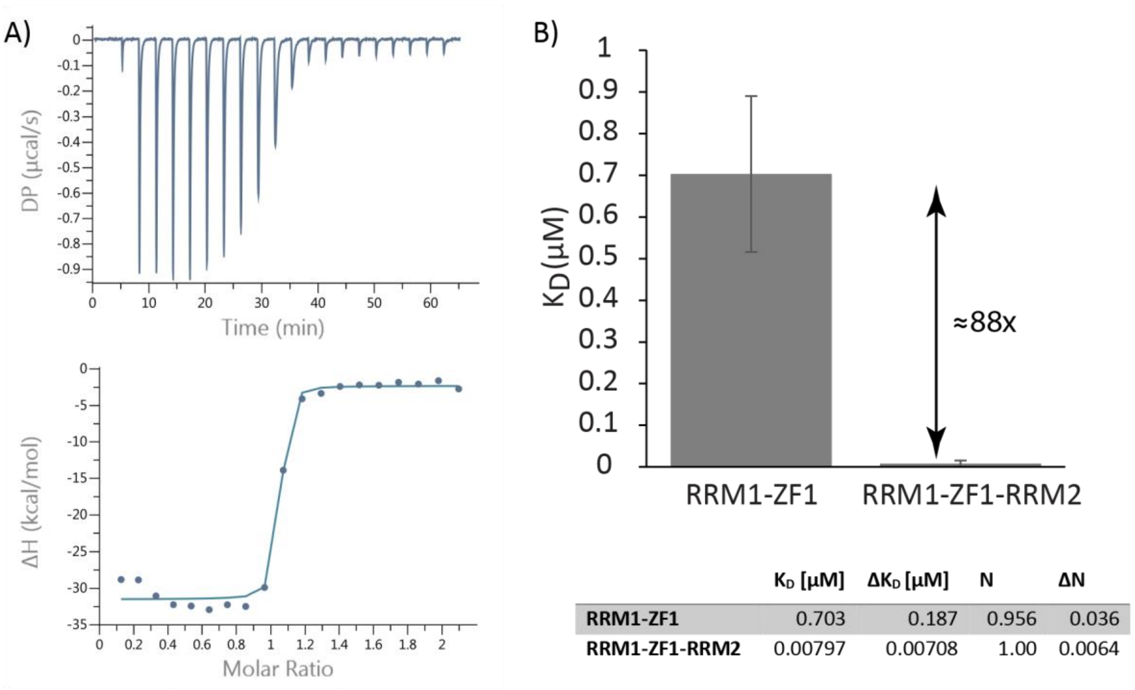


***Figure 34**: **NMR titration of RBM10 RRM1-ZF1-RRM2 with NUMB derived 12-mer-RNA oligo:** A) Superimposed ^{1}H,^{15}N-correlation spectra of RBM10 RRM1-ZF1-RRM2 apo-protein (black) and titration steps with 0.25x (red), 0.50x (green), 1.00x (cyan) and 1.50x (blue) molar equivalents of RNA present. Peaks in the black and the blue spectrum of residues that experience high chemical shift perturbation are connected by arrows and labeled. Residues F234, which could not be detected in the complex spectrum, is labeled as well. B) Chemical shift perturbations between RBM10 RRM1-ZF1-RRM2 free and in complex with the NUMB derived 12-mer RNA oligo are plotted against the residue number. C) For comparison the CSP from*

*titrations with the RRM1-ZF1 (from **Figure 18 NMR titration of RBM10 RRM1-ZF1 with NUMB derived 12-mer-RNA oligo:** See previous page for the figure note.*

*) and the RRM2 (from **Figure 27**) construct are plotted in one graphic.*

The NMR titration and SLS data indicate that both RRM1 and RRM2 and probably also ZF1 interact with the NUMB derived 12-mer RNA oligo in a 1:1 complex. ITC data of the three-domain-construct result in an affinity of 8 nM and a stoichiometry of 1:1 (**Figure 35**). The affinity of the three-domain construct for the NUMB derived 12-mer RNA oligo is therefore around 88x higher than the affinity of RRM1-ZF1 to the same RNA. The RRM2 construct for the same RNA has an affinity of 23.8 µM. From this comparison of binding affinities it becomes clear that despite the long flexible linker between the two RNA binding elements RRM1-ZF1 and RRM2, collaborate to recognize the NUMB derived 12-mer RNA oligo.



	K_D [µM]	ΔK_D [µM]	N	ΔN
RRM1-ZF1	0.703	0.187	0.956	0.036
RRM1-ZF1-RRM2	0.00797	0.00708	1.00	0.0064

***Figure 35 ITC data RRM1-ZF1-RRM2:** A) ITC data for RRM1-ZF1-RRM2 construct titrated with the NUMB derived 12-mer RNA oligo. C) Plot of dissociation constant for RRM1-ZF1-RRM2 derived from A and RRM1-ZF1 derived from **Figure 19**A. Values are listed in the table below. Additionally the binding stoichiometry N is provided.*

Since all three domains seem to be involved in the recognition of a rather small 12-mer RNA oligo it is of interest how the three domains, which seem to be flexible towards each other in the apo-protein, behave dynamically in the RNA bound form. ^{15}N-relaxation data has been recorded on the three-domain-construct in presence of the NUMB derived 12-mer RNA oligo to answer this question (**Figure 36**). The addition of RNA leads to an increase of the correlation time of all three domains. For RRM1 and RRM2 this increase is moderate with around 3.5 ns/rad, while for ZF1 it is bigger with around 5 ns/rad. Because of the line broadening due to RNA binding the spectral quality of the complex spectra is worse and therefore the errors are higher than for the apo-protein. Considering the high error, no significant different in correlation time between the individual domains can be observed for the three-domain construct in complex with RNA (**Figure 37**). The linker in contrast stays as flexible as it as in the free protein, indicating that it is not involved in RNA binding, which is in agreement with the lack of CSP in the linker during the NMR titration with the RNA.

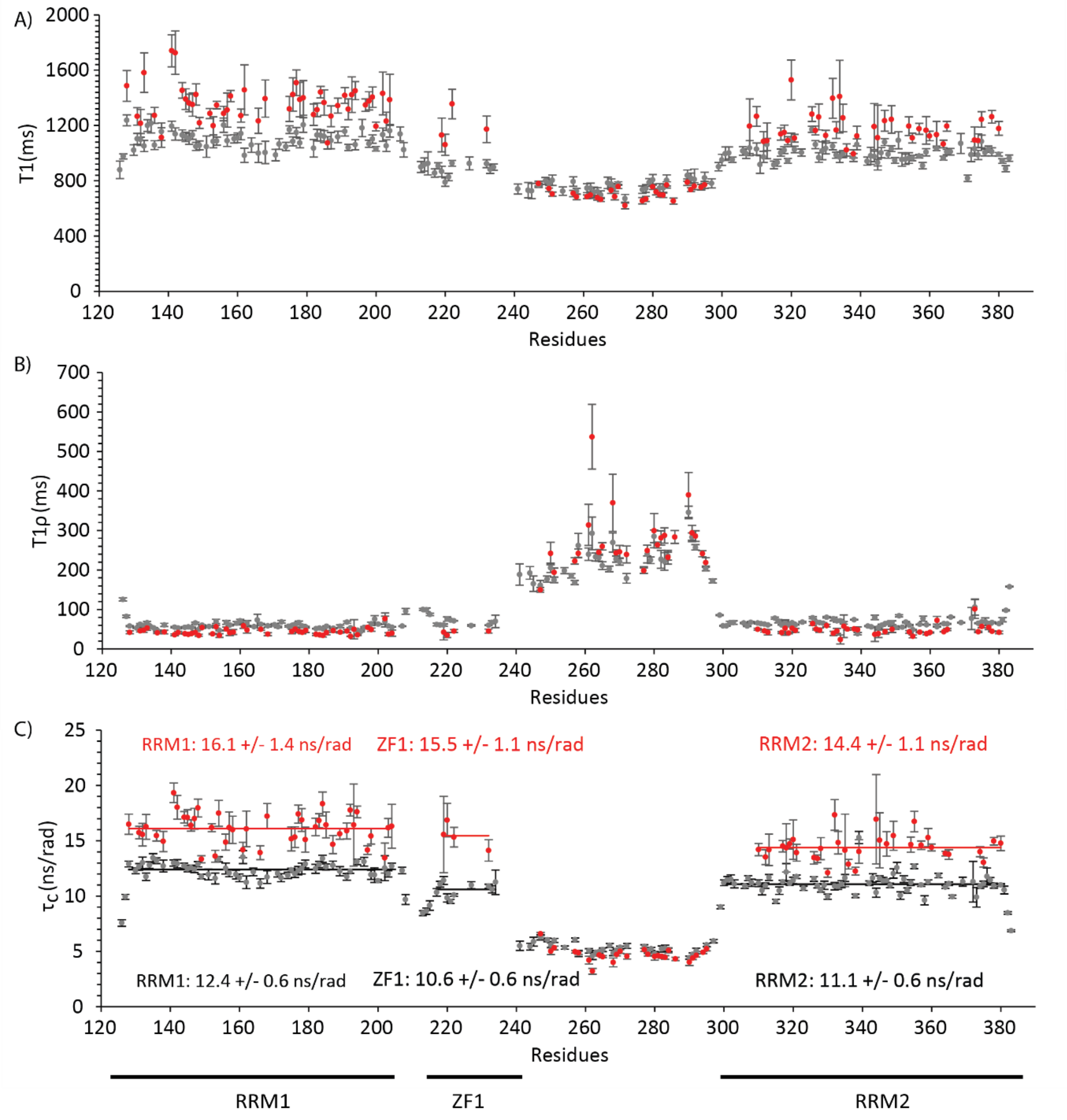


Figure 36 Relaxation data RBM10 RRM1-ZF1-RRM2 NUMB RNA complex: *A) T1 time constant, B) T1ρ time constant and C) correlation time τ_C of the RBM10 RRM1-ZF1 NUMB derived 12-mer RNA oligo complex (red) plotted against the residue number. For comparison the same data of the apo-protein is plotted in grey.*

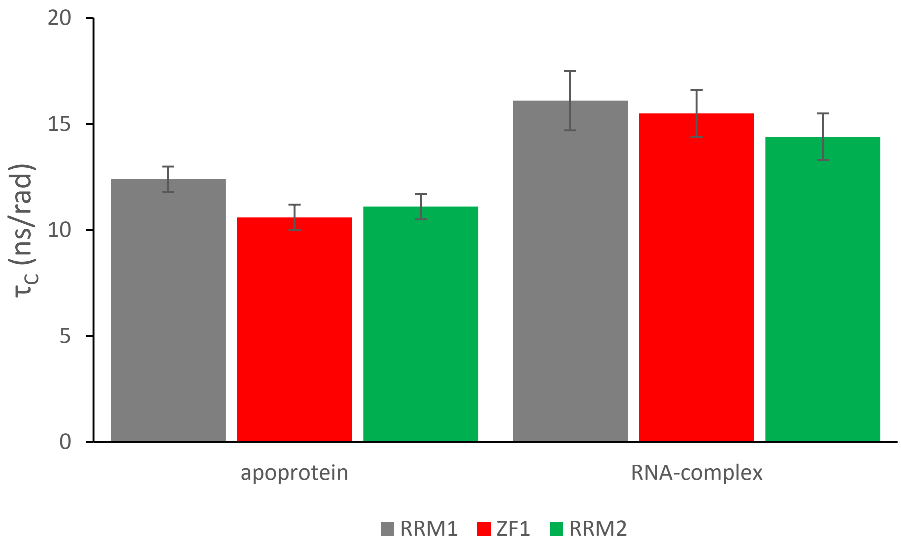


Figure 37 Correlation times of RRM1-ZF1-RRM2 domains free and bound to RNA: *The correlation time of RRM1 (grey), ZF1 (red) and RRM2 (green) within the RRM1-ZF1-RRM2 construct are shown in the free and in the RNA bound form. The correlation times have been averaged over the domains. The indicated error represents the standard deviation.*

6.2.3 RNA TOCSY spectra when bound to the RRM1-Zn-RRM2

The much higher affinity of the three-domain-construct to the NUMB derived 12-mer RNA oligo motivated to measure the TOCSY spectra of the RNA in complex with the three-domain-construct.

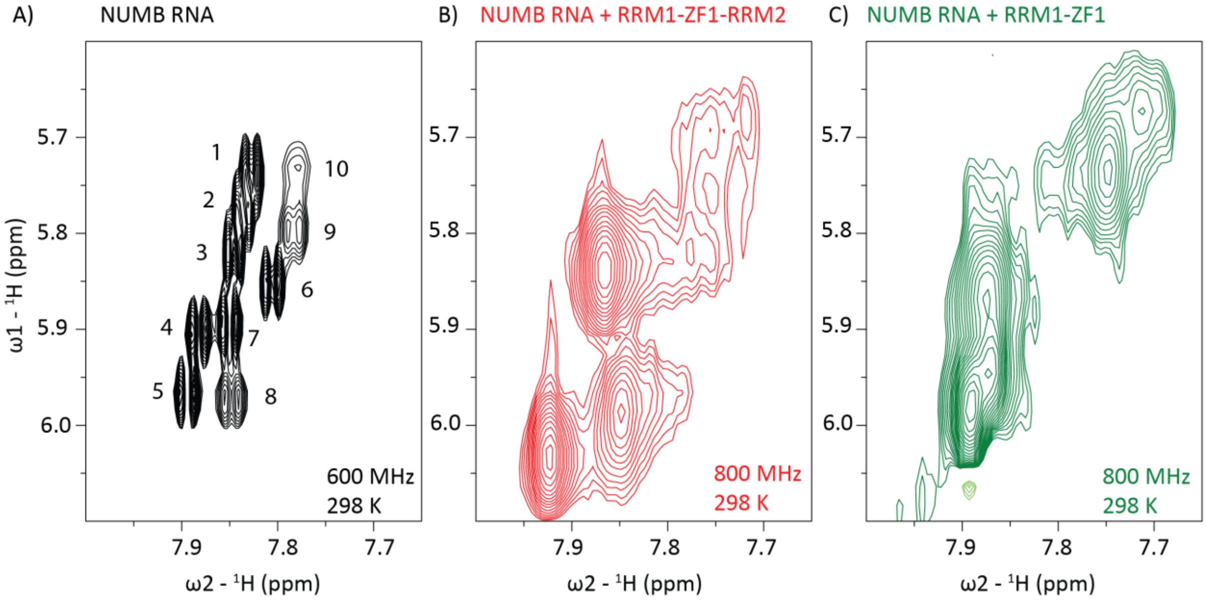


Figure 38 NUMB-RNA-TOCSY spectra with RBM10 RRM1-ZF1-RRM2: *Shown are homonuclear 2D TOCSY spectra of the NUMB derived 12-mer RNA oligo (UUGUCUGCUCCC) free measured at 600 MHz (A), in presence of RBM10 RRM1-ZF1-RRM2 at 800 MHz (B) and for comparison in presence of RRM1-ZF1 at 800 MHz (C). The shown part of the spectrum comprises the H5,H6-correlations of the pyrimidine nucleotides in RNA. The pyrimidine nucleotides in the NUMB derived 12-mer RNA oligo are highlighted in red.*

The TOCSY spectrum of the RNA in complex with the three-domain-construct still shows very broad peaks (**Figure 38B)**, but also shows some differences to the RNA in complex with the RRM1-ZF1 construct (**Figure 38C)**. While three peaks downfield are a bit more disperse, the upfield peaks are broader than in the complex with RRM1-ZF1. To check if the higher affinity leads to more intermolecular nOe cross peaks a 2D ω1-filtered NOESY experiment was recorded with the NUMB RNA and the three-domain construct in excess.

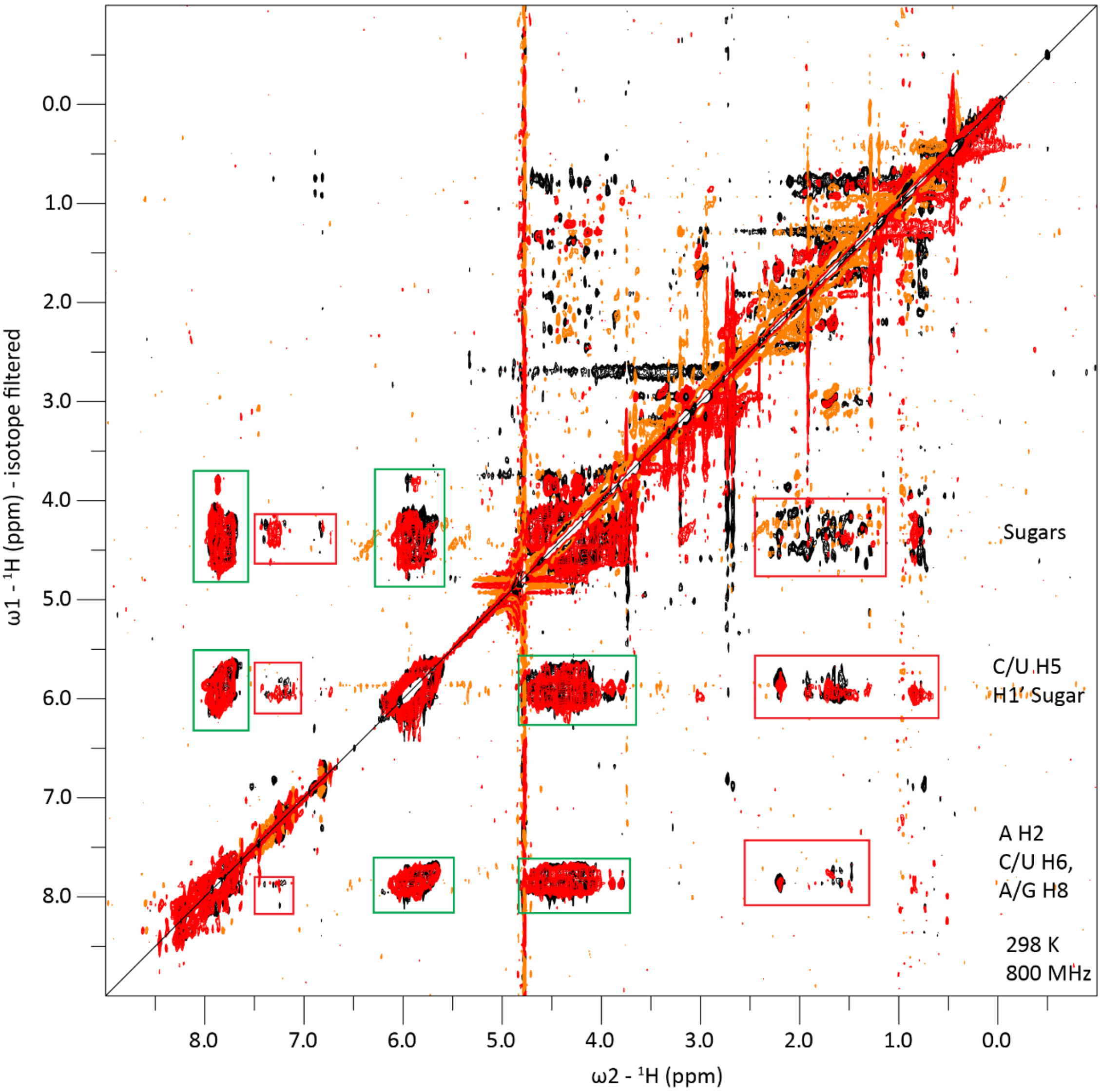


Figure 39 ω1-filtered NOESY spectrum of NUMB derived RNA 12-mer oligo saturated with RBM10 RRM1-ZF1-RRM2: *Shown is a superimposition of ω1-filtered NOESY spectra of the unlabeled NUMB derived 12-mer RNA oligo, which is saturated with* ^{13}*C- and* ^{15}*N-isotope labelled RBM10 RRM1-ZF1 (black) or RRM1-ZF1-RRM2 (red). Areas with intramolecular RNA cross peak are marked with green boxes. The intramolecular cross peaks appear twice axis symmetric to the diagonal. Areas with intermolecular cross peaks between the RNA and the protein are marked by red boxes. Labels on the right indicate which RNA protons give rise to signal in that area of the spectrum.*

The 2D ω1-filtered NOESY experiment of the NUMB derived 12-mer RNA oligo saturated with the RRM1-ZF1-RRM2 construct shows less nOe cross peaks than when saturated with RRM1-ZF1. A possible reason is the higher molecular weight of the complex with the three-domain-construct, which leads to broader lines.

6.2.4 SAXS provides information about the shape of the three RNA binding domains

As an initial information about the domain arrangement in the RBM10 three-domain-construct small angle X-rays scattering curves have been recorded by Ralf Stehle. The extracted P(*r*) curves of the apo-protein (**Figure 40A**) and the protein in complex with the NUMB derived 12-mer RNA oligo (**Figure 40B**) measured at three different concentration show a difference between the lowest (1 mg/mL) and the two higher concentrations (5 and 10 mg/mL).

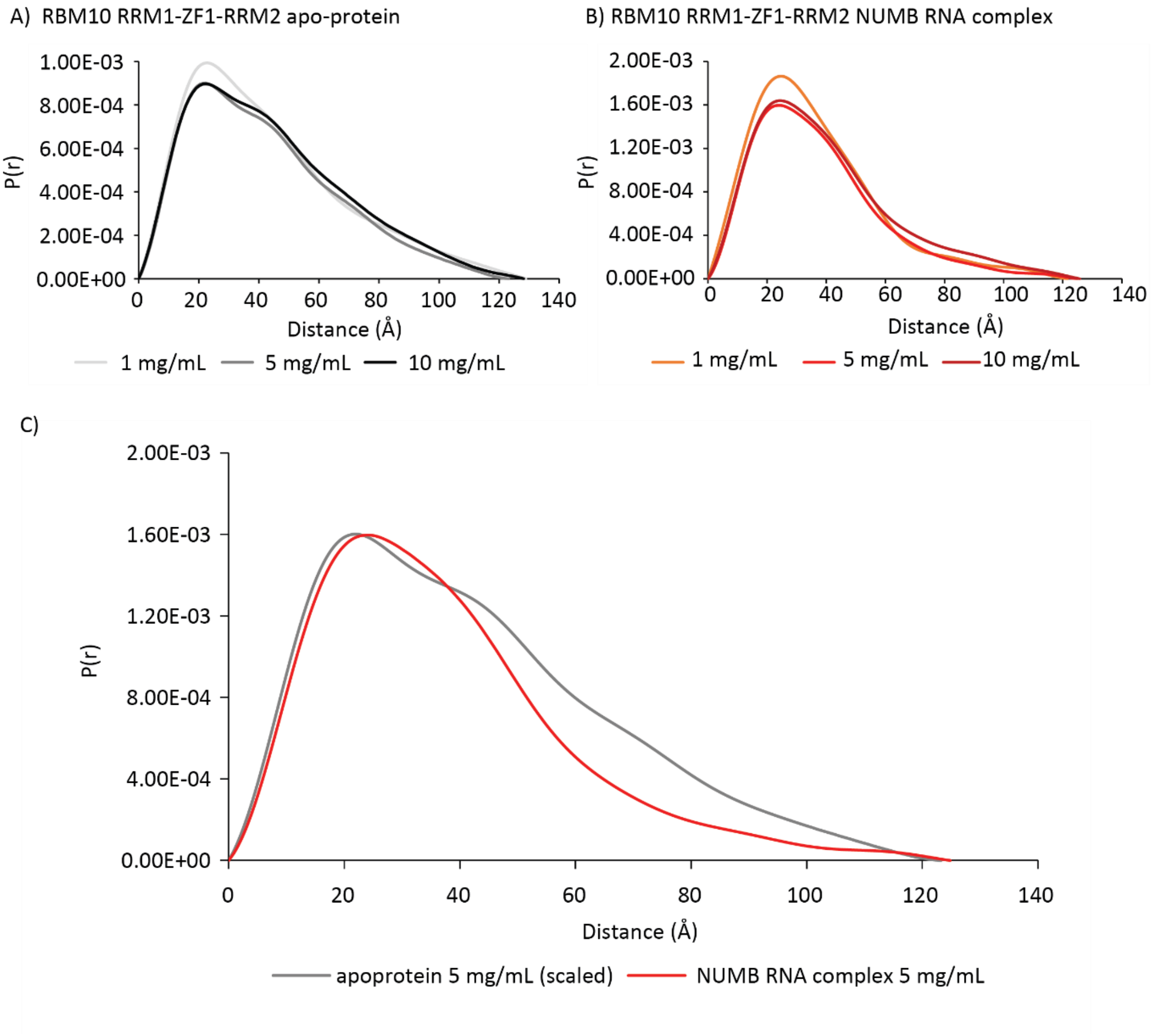


Figure 40 SAXS data for RBM10 RRM1-ZF1-RRM2: *A) SAXS p(r) curves of 1 mg/mL (light grey), 5 mg/mL (grey) and 10 mg/mL (black) RBM10 RRM1-ZF1-RRM2 apo-protein. B) SAXS p(r) curves of of 1 mg/mL (orange), 5 mg/mL (red) and 10 mg/mL (dark red) RBM10 RRM1-ZF1-RRM2 saturated with the NUMB derived 12-mer RNA oligo. C) Superimposition of the 5 mg/L curves from A and B. The curve of the apo-protein is scaled, so that the maximum fits the maximum curve of the protein-RNA complex.*

The intensities of the curves are not directly comparable. To be able to compare the P(r) curves of the apo-protein and the protein-RNA-complex the curve of the apo-protein were scaled so that the maximum fits the curve of the protein-RNA complex. For this comparison the middle concentration of 5 mg/mL is used (**Figure 40C**). The curve of the RNA-complex has higher contributions at shorter distances, indicating a more compact state or ensemble of states. The same effect is also seen when superimposing the curves of all three concentrations (not shown). This compaction is expected, since both RNA binding elements, RRM1-ZF1 and RRM2, have to fit onto the small NUMB RNA derived 12-mer oligo. The compaction also fits to the narrower peaks for the protein:RNA complex compared to the free protein, which was observed in SLS measurements. Surprisingly the D^{max} value does not change upon RNA binding. Probably the highest distance would be a detached state were RRM2 is far away from RRM1-ZF1, which is expected for the apo-protein, where all three domains tumble independently. Possibly in the three domain construct RRM1-ZF1 is tightly bound to the RNA, while RRM2 is bound in the major conformation and is detached in a minor conformation, which is not detectable in the $^{1}H,^{15}N$-correlation spectrum.

6.2.5 Paramagnetic Relaxation enhancement provides inter-domain distance information

Distance restraints derived from NOESY spectra cover only a very short distance range up to 5 Å, which is often not enough to determine the domain orientation in multidomain proteins. Paramagnetic relaxation enhancement (PRE) can provide much longer range distance restraints. To measure PRE data unpaired electrons in defined positions have to be introduced to the protein. For this purpose cysteine mutants are of the protein are designed, which can be labeled with IPSL, which harbors an unpaired electron. The cysteine should be introduced into a rigid position of the protein, but with a solvent exposed side chain, to get a rigid orientation of the spin label and to not disturb the structure of the protein. The spin labels should cover the surface of all domains. For RBM10 16 cysteine mutants have been designed, seven in RRM1, two in ZF1 and seven in RRM2 (**Figure 41**). All mutants have been incorporated in the RBM10 RRM1-ZF1-RRM2 construct.

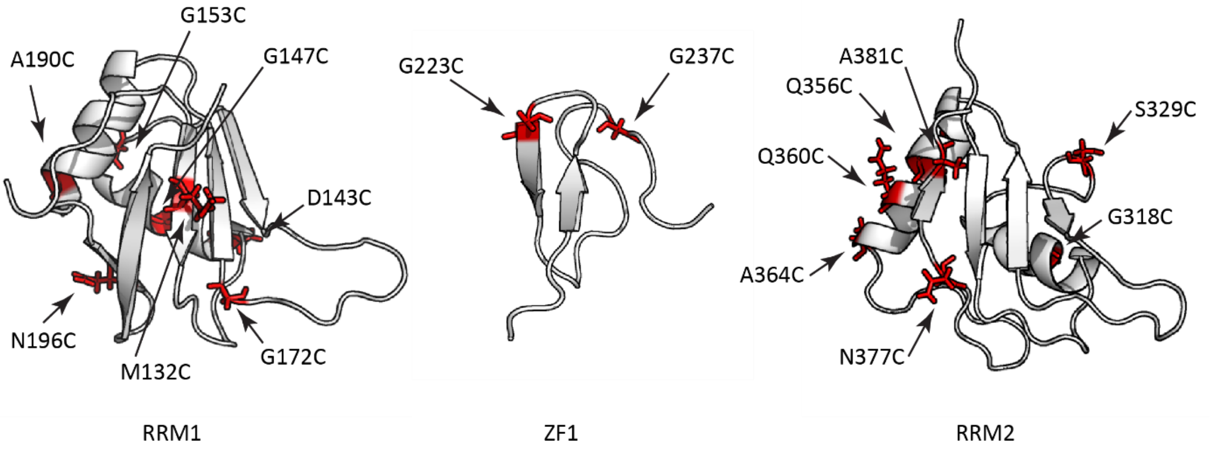


Figure 41 Positions of cysteine mutants for PRE measurements: *Side chains of residues that were mutated to cysteine for PRE measurements are shown as sticks, are colored in red and labeled.*

The seven RRM1 mutants are M132C, D143C, G147C, G153C, G172C, A190C und N196C. The A190C mutant could not be expressed, since the cells transformed with the expression plasmid did not grow in the pre-culture. The G153C mutant showed strongly reduced peak intensities for all residues of RRM1 and ZF1, indicating that this mutation destabilizes both domains (data not shown). Interestingly the G153C is a known cancer linked mutation (J. Zhao et al. 2017). For the remaining five mutants PRE data have been recorded on the apo-protein and the protein in complex with the NUMB derived 12-mer RNA. The ratios of paramagnetic and diamagnetic HNCO spectra per residue are shown in **Figure 42**.

The protein shows for all five spin label positions a PRE effect close to the spin label positions. The spin label effect falls off with increasing distance to the spin label. This can be seen for the spinlabel in the position 143 in RRM1 in **Figure 43A**. The spin labels in position M132C and G172C show pronounced PRE effects in the RRM2 domain, indicating that the RRM2 domain has to come into proximity of the β-sheet of RRM1 for some time. Since the domain orientation in the apo-protein is flexible these contacts have to be transient. It has to be considered that the PRE effect for transient contacts can be very high, if the domains come close together even if only a small population of the protein shows this contacts, because the PRE effect falls off with the distance to the power of six (Battiste and Wagner 2000). This also explains why the flexible linker between ZF1 and RRM2 often experiences a strong PRE effect. This effect on the linker is strongest with the spin label in positions 132 and 172.

Upon RNA binding the PRE pattern for RRM2 observed in M132C, G172C and D143C changes, indicating a difference in domain arrangement and/or flexibility. As an example the PRE data for the apo-protein and the protein:RNA complex with the spin label attached to position D143C is plotted in **Figure 43**. Upon RNA binding the areas at the bottom of the β-sheet, including the loop after β-strand 1

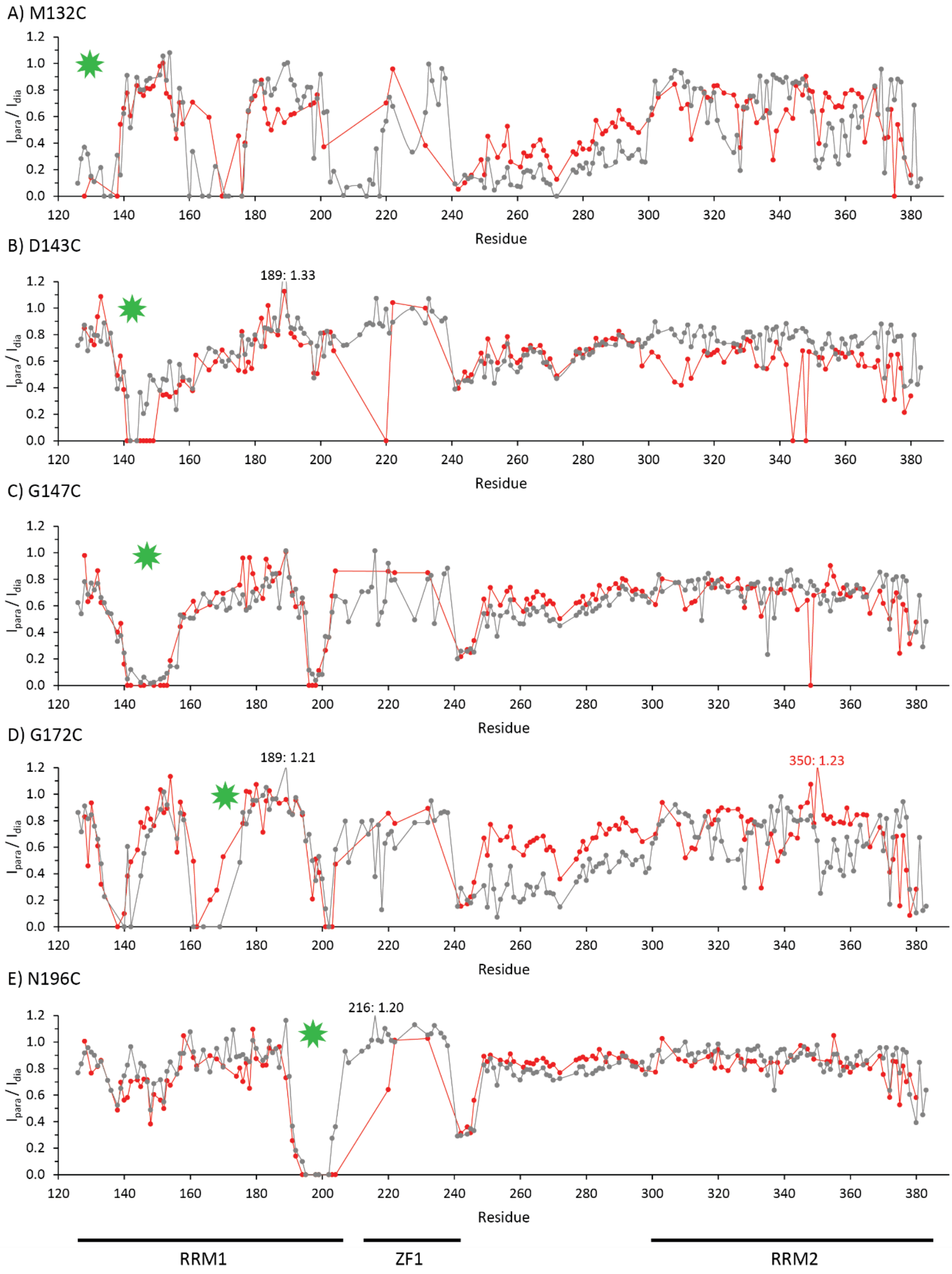


Figure 42 PRE data for RBM10 RRM1-ZF1-RRM2 with spin labels in RRM1: *Plotted is the intensity ratio of HNCO spectra of a paramagnetic and a diamagnetic sample of RBM10 RRM1-ZF1-RRM2 in absence (grey) and presence (red) of the NUMB RNA derived 12-mer RNA oligo. The spin label is linked to the engineered cysteine in position 132 (A), 143 (B), 147 (C), 172 (D) and 196 (E). The green star represents the position of the spin label. Ratios higher than 1.2 are indicated.*

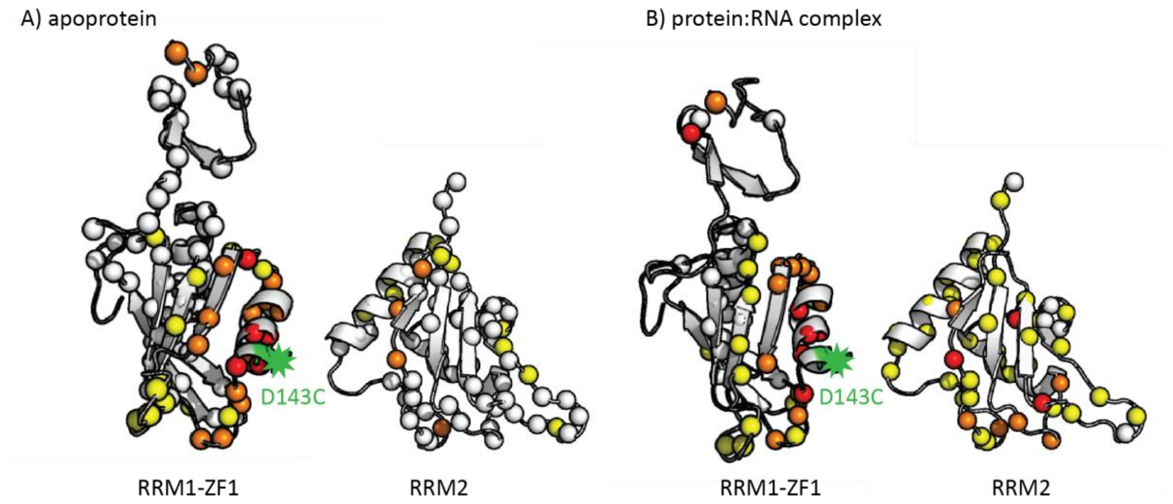


Figure 43 PRE ratios of RBM10 RRM1-ZF1-RRM2 with spin label in position 143: *Spheres indicate residues for which PRE data could be acquired. The spheres are colored in yellow for peak ratios of the diamagnetic and paramagnetic spectra smaller 0.7, orange for a ratio smaller 0.5 and red for a ratio smaller 0.3. The site of the spin label attachment is indicated by a green star. Panel A shows data for the apo-protein, while panel B shows data for the protein in complex with the NUMB derived 12-mer RNA oligo.*

and the loop before β-strand 4 experience a stronger PRE effect compared to the apo-protein. The PRE effect on RRM1-ZF1 in contrast does not change, indicating a difference in domain arrangement. Since the assignment of the protein:RNA complex is less complete than the assignment of the apo-protein, one has to be careful to draw further conclusions. The flexible linker between RRM1-ZF1 experiences a weaker PRE effect in the protein:RNA complex in M132C and G172Cindicating that the linker does not come in as close contact anymore to these positions in the protein:RNA complex.

The two ZF1 mutants are G223C and G237C. The G223C mutant showed peak doubling for many peaks assigned to ZF1 residues. The mutation was designed to this position because the homolog RBM5 has a cysteine in the same position. The cysteine is located directly next to one of the cysteine coordinating the Zn^{2+}. RBM5 is studied by Komal Soni in the Sattler group and she also observes peak doubling for the peaks assigned to ZF1 residues in RBM5. Even more interestingly these peak doubling disappears when this cysteine is mutated out of RBM5 ZF1 (unpublished data). For the other G237C mutant PRE data were recorded on the apo-protein and the protein in complex with the NUMB derived 12-mer RNA. The ratios of paramagnetic and diamagnetic HNCO spectra per residue are shown in **Figure 44**.

A spin label attached in position 237 leads to PRE effects in the RBM10 RRM1-ZF1-RRM2 apo-protein on both RRMs, indicating that the ZF1 comes into close contact to both domains. This fits well to the dynamic domain arrangement shown by ^{15}N-relaxation data. Most signals of the ZF could not be identified in the spin-labelled G237C mutant, since they were broadened also in the diamagnetic spectrum.

Upon RNA binding changes in the PRE pattern of both RRMs can be observed. Interestingly the domains are affected differently. While some residues in RRM1 experience a weaker PRE effect in RNA bound form compared to the apo-protein, some residues in RRM2 experience a stronger PRE effect in the RNA bound form. Again this hints towards a domain rearrangement upon RNA binding.

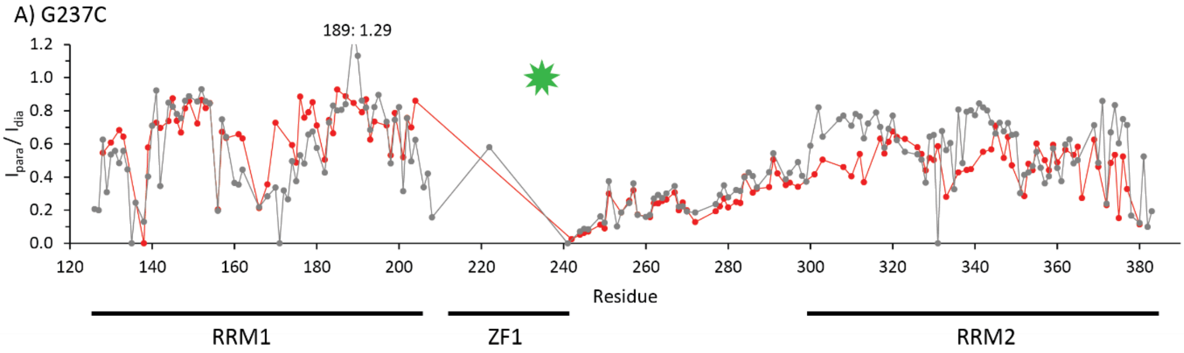


Figure 44 PRE data for RBM10 RRM1-ZF1-RRM2 with the spin label in ZF1: *Plotted is the intensity ratio of HNCO spectra of a paramagnetic and a diamagnetic sample of RBM10 RRM1-ZF1-RRM2 in absence (grey) and presence (red) of the NUMB RNA derived 12-mer RNA oligo. The spin label is linked to the engineered cysteine in position 237. The green star represents the position of the spin label. Ratios higher than 1.2 are indicated.*

Finally the seven designed mutants for the RRM2 domains are G318C, S329C, Q357C, Q360C, A364C, N377C and A381C. The Q360C mutation could not be generated with the ordered primers, without introducing a second mutation close by. Since the neighbored Q357C mutation worked out well, work on the Q360C mutation was not continued. The A381C mutation leads to a destabilization of β-strand 4, which is broadened beyond detection in the ^{1}H,^{15}N-correlation spectrum of the mutant (data not shown). For the other five mutants PRE data were recorded on the apo-protein and the protein in complex with the NUMB derived 12-mer RNA. The ratios of paramagnetic and diamagnetic HNCO spectra per residue are shown in **Figure 45**.

The spin labels in RRM2 also effect the residues in the same domain locally around the spin label attachment side. Additionally, the spin labels attached to residues in the apo-protein of RRM2 also effect residues in RRM1, which is more pronounced in the mutants S329C and Q357C. The spin label in position 329, 357 and 364 also effect ZF1. These three spin label positions also strongly effect the linker between ZF1 and RRM2.

Upon RNA binding the PRE pattern for the positions 357 and 364 of the spin label show clear differences in the PRE pattern, which fits the picture of a domain rearrangement upon RNA binding.

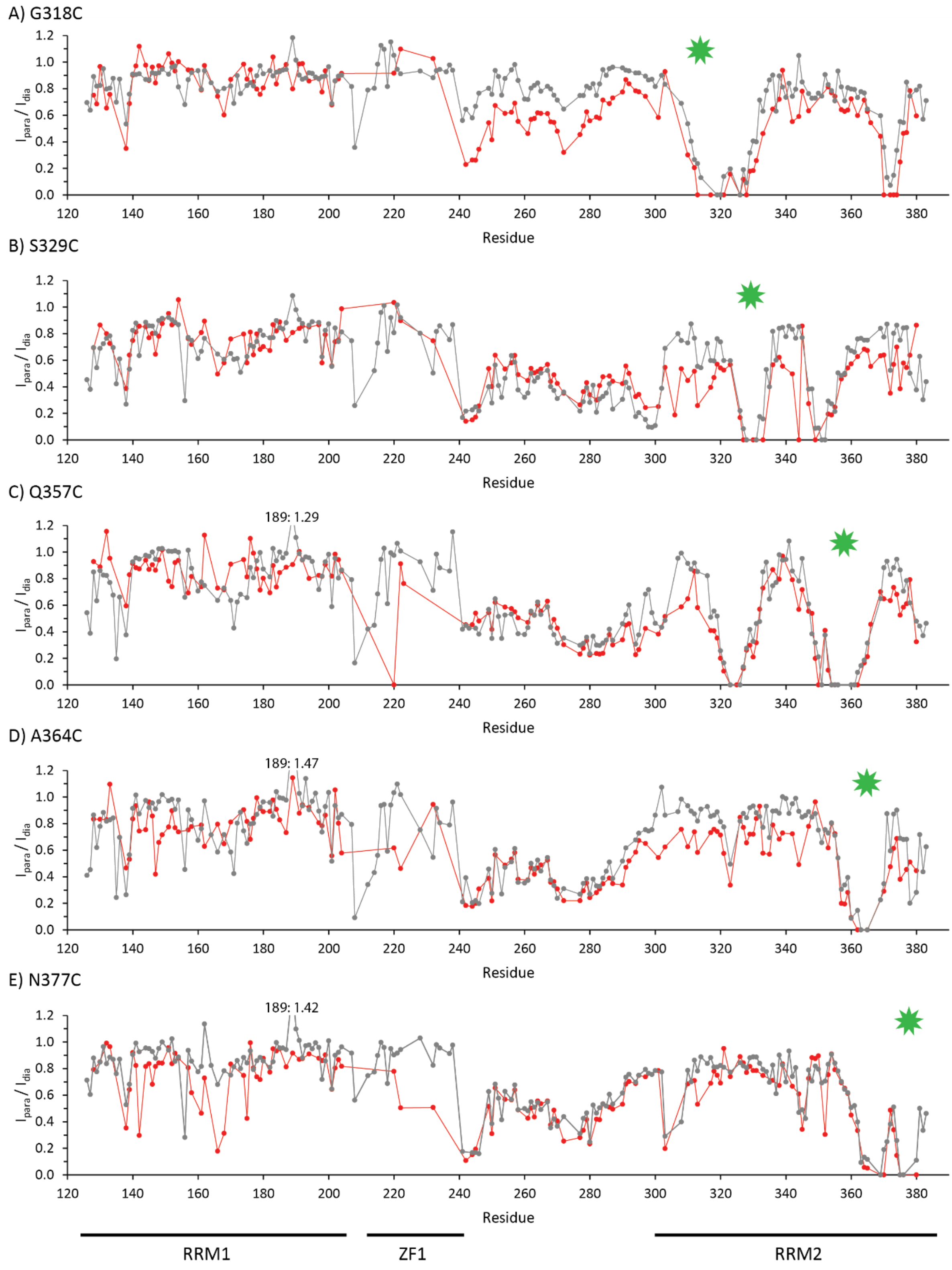


Figure 45 PRE data for RBM10 RRM1-ZF1-RRM2 with spin labels in RRM2: *Plotted is the intensity ratio of HNCO spectra of a paramagnetic and a diamagnetic sample of RBM10 RRM1-ZF1-RRM2 in absence (grey) and presence (red) of the NUMB RNA derived 12-mer RNA oligo. The spin label is linked to the engineered cysteine in position 318 (A), 329 (B), 357 (C), 364 (D) and 377 (E). The green star represents the position of the spin label. Ratios higher than 1.2 are indicated.*

6.3 Analysis of disease-linked mutants and isoforms

6.3.1 The RBM10 splicing isoform +V354 and the cancer-linked mutant V354E

RBM10 exits in two splicing isoforms, which differ by the inclusion or exclusion of amino acid V354. The isoform including the extra valine residue will be called +V354. In this section the residue numbering for the +V354 isoform will be used, also when comparing it to the -V354 isoform, which was described in chapter 6.1.4.

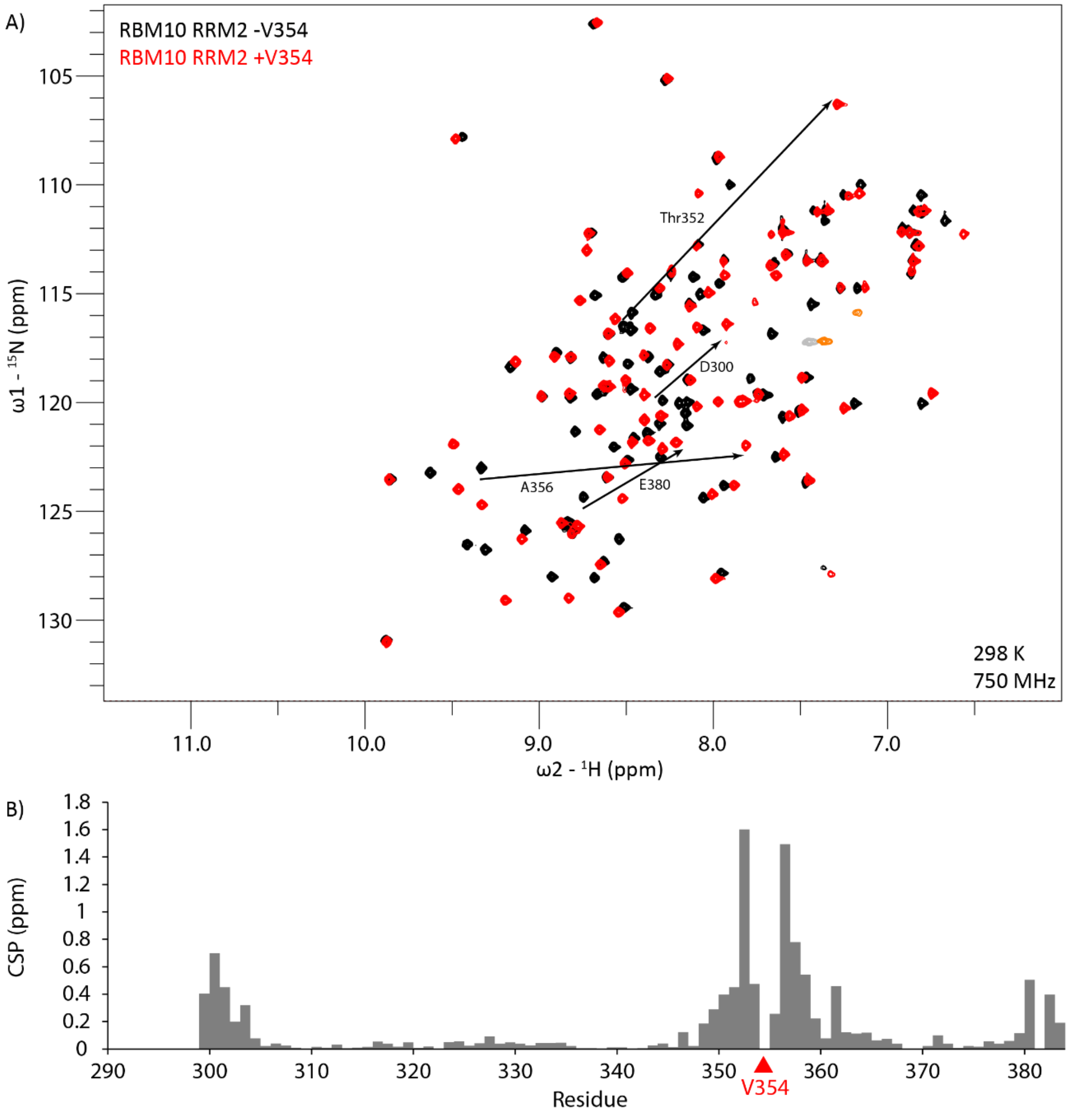


Figure 46 ***^{1}H,^{15}N-correlation spectra of the RRM2 + V354 isoform:*** *A) Superimposition ^{1}H,^{15}N-correlation spectra of the -V354 (black) and the +V354 (red) isoforms of RBM10 RRM2. Residues with strong difference in chemical shift are connected by an arrow and labelled. The numbering follows the numbering of the +V354 isoform, that means that the numbers for the -V354 are +1 from 354 onwards. B) Chemical shift perturbation plot between the RRM2 isoforms -V354 and +V354. The red triangle indicates the position of the extra valine.*

Surprisingly, the extra valine residue causes large changes in chemical shift in the ^{1}H,^{15}N-correlation spectrum of the RRM2 (**Figure 46A**). The backbone of the +V354 variant was therefore assigned using HNCACB spectrum. The chemical shift perturbation plot shows three clusters of high CSP: Around the insertion site of the additional V354 and at the N- and C-terminus of the domain, which correspond to the β-strands 1 and 4.

To find out how one single point mutation can induce such large changes in the ^{1}H,^{15}N-correlation spectrum, side chain assignment and structure calculation has been performed in the same way as for the -V354 variant. 99.1 % of all protons were assigned. The same kind of NOESY spectra were used for structure calculation and water refinement in Cyana and Aria2 respectively as for the -V354 variant.

The +V354 RRM2 structure shows the same fold as the -V354 RRM2 structure. The NMR ensemble converges well except for the linker following α-helix 2 were two structures diverge a little bit from the other eight in the structure bundle **Figure 47**.

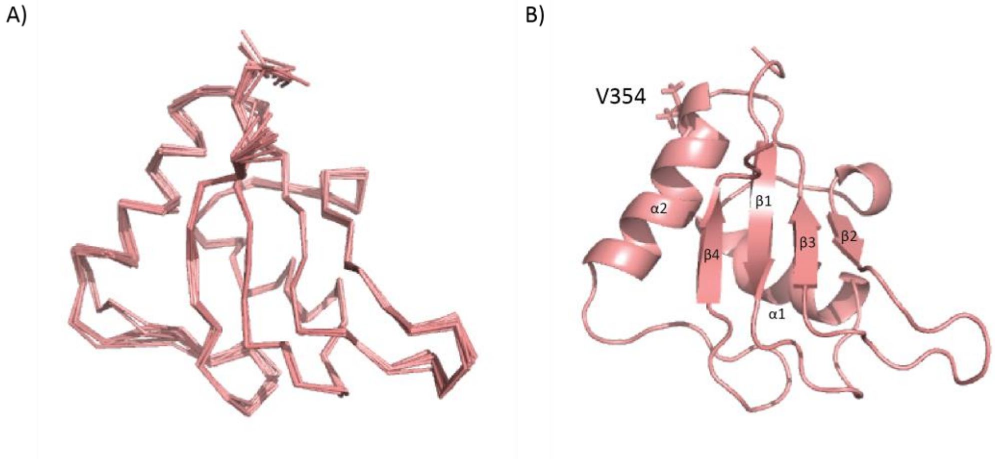


Figure 47 RBM10 RRM2 +V354 NMR structure: *A) Ribbon plot of the 10-structure water refined NMR bundle of RBM10 RRM2. B) Cartoon plot of the RBM10 RRM2 structure. Secondary structure elements are labelled. The additional valine 354 is shown as sticks and labelled.*

The structure of the RRM2 +V354 isoform is overall very similar to the structure of the -V354 isoform. Helix α2 gets slightly bend in the +V354 structure and is slightly differently oriented. The loop directly in front of helix α2 is curved out a bit in the +V354 structure, to make room for the extra residue. This changes the position of threonine 352 drastically, which explains, why this residue experience strong CSP. Also the C-terminus harboring β-strand 4 is oriented differently. The side chain of F381 in the +V354 isoform is twisted compared to the corresponding F380 in -V354. The side chain of this phenylalanine is directly pointing at the helix α2 (**Figure 48**). The twist of F380/381 explains the strong chemical shift of alanine 356, which is in part caused by the change in ring current. While the contribution of ring currents to the chemical shift of the amide nitrogen and proton for alanine 356 in the -V354 is calculated by MOLMOL (Koradi, Billeter, and Wüthrich 1996; Case 1995) to be 0.216 and 0.174 ppm, it is -0.049 and 0.003 ppm in the +V354 isoform.

Table 11: Structure calculation statistics for RBM10 RRM2 +V354:

Structure calculation restraints	
Distance restraints	
Total NOEs	1805
Sequential ($\|i- j\| \leq 1$)	866
Medium-range ($1<\|i- j\| <5$)	364
Long-range ($\|i- j\| \geq 5$)	575
Dihedral restraints ($\phi+\psi$)	130
Quality analysis	
Restraints violations (mean ± s.d)	
Distance restraints (Å)	0.023 ± 0.002
Dihedral angle restraints (º)	0.400 ± 0.06
Deviation from idealized geometry	
Bond length (Å)	0.004 ± 0.0002
Bond angles (º)	0.48 ± 0.010
Improper dihedral distribution (º)	1.26 ± 0.077
Average pairwise r.m.s. deviation (Å)	
Heavy	0.65 ± 0.04
Backbone	0.23 ± 0.05
Ramachandran values (%) (MolProbity)	
Most favored regions	93.0 %
Allowed regions	99.4 %

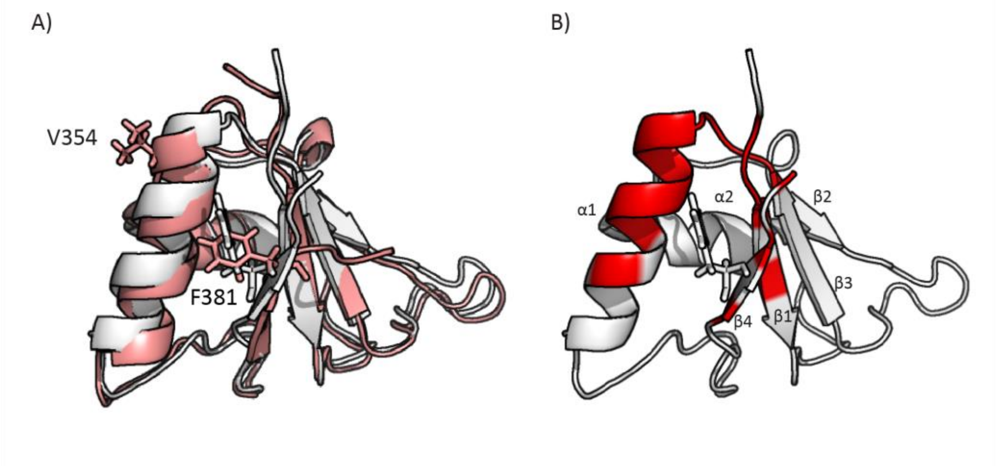


Figure 48 RBM10 RRM2 +/- V354 comparison: *A) Superimposition of RBM10 RRM2 -V354 (white) and +V354 (salmon) NMR structures. Side chains of residues V354 and F381 (numbered according to +V354 numbering) are shown as sticks. B) CSP extracted from* ***Figure 46*** *larger than 0.2 ppm are highlighted in red on the RBM10 RRM2 -V354 structure.*

Interestingly the residue V354, is known to be mutated in lung cancer cells and the mutation has shown to alter the effect of RBM10 on the alternative splicing of the NUMB gene. Comparing the ^{1}H,^{15}N-correlation spectra of the RRM2 +V354 and the RRM2 V354E mutant shows only small changes in chemical shift around the mutation site (**Figure 49**), indicating that the mutation does not change the structure of RRM2 +V354 globally.

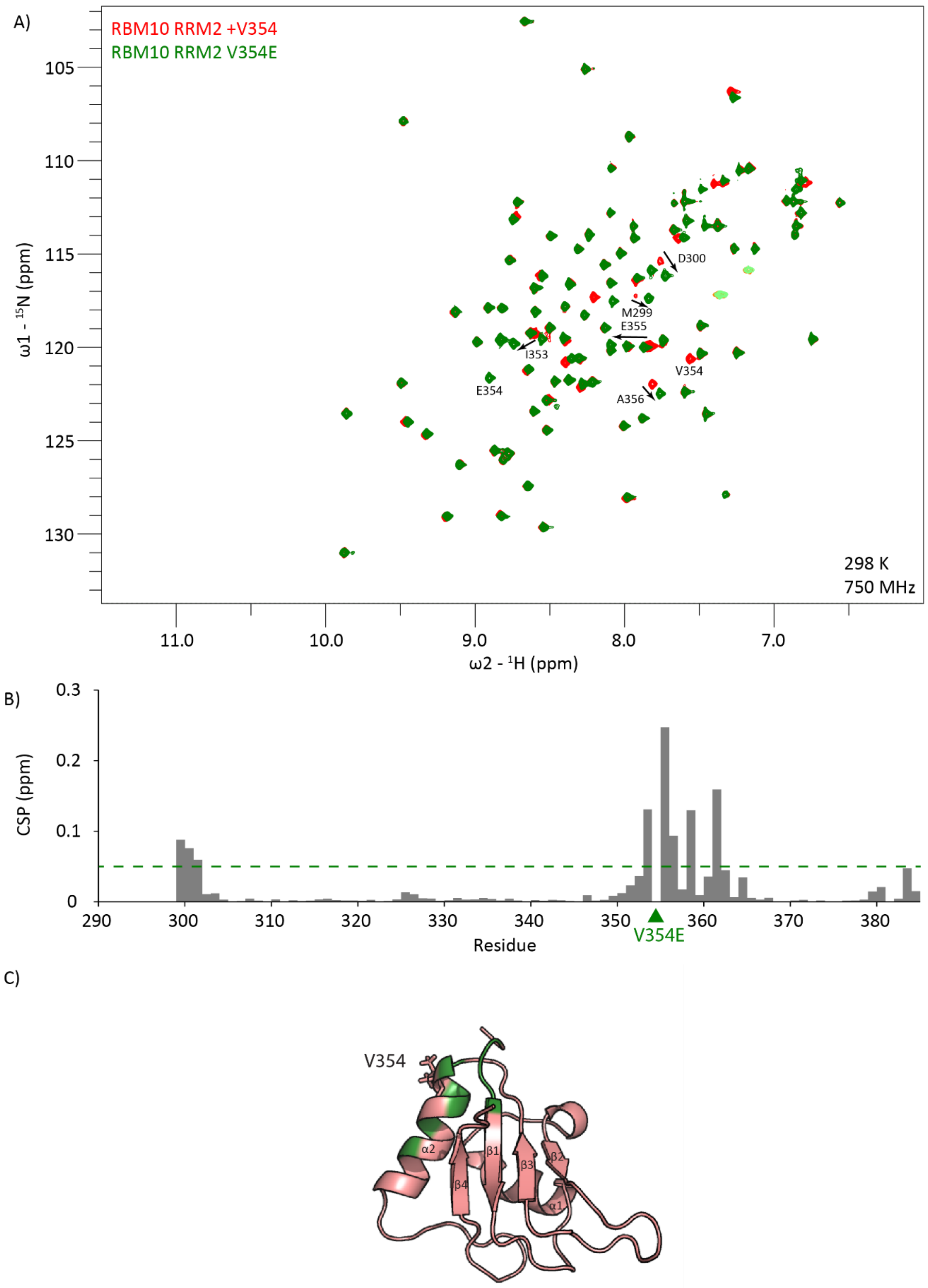


Figure 49 ^{1}H,^{15}N-correlation spectra of the RRM2 + V354 isoform: *See next page for the figure note.*

Figure 49 ^{1}H,^{15}N-correlation spectra of the RRM2 + V354 isoform: *A) Superimposition of ^{1}H,^{15}N-correlation spectra of the +V354 (red) isoform and the V354E mutant (green) of RBM10 RRM2. Residues with strong difference in chemical shift are connected by arrows and labelled B) Chemical shift perturbation plot between the RRM2 +V354 isoform and the V354E mutant. The green triangle indicates the position of the mutation. C) CSP larger 0.05 ppm from B plotted onto the NMR structure of the +V354 isoform. Secondary structure elements are labelled. The residue V354 is shown as sticks and labelled.*

Despite the only small changes the backbone assignment of V354E has been checked using an HNCACB spectrum. The first residues M299 and D300, which appear as very weak peaks in the ^{1}H,^{15}N-correlation spectrum of RRM2 +V354, have a little bit different shift in RRM2 V354E and appear here as strong as most of the peaks. There seems to be some additional exchange present in the +V354 construct which is not present in the V354E construct.

6.3.2 RNA binding of the +V354 isoform and the V354E mutant

To find out if the isoforms and the V354E mutant behave differently in RNA binding, NMR titration experiments have been performed with all three forms. The Chemical shift perturbations from all three titrations are plotted in **Figure 50A**. The overall pattern of the chemical shift perturbation is the same for all three proteins. Only the loop between β-strand 3 and α-helix 2 and the N-terminal part of α-helix 2, show stronger chemical shift perturbation in the -V354 isoform compared to the +V354 isoform (**Figure 50B**). This part of the structure also showed differences in the structure. There is no difference between the +V354 isoform and the V354E mutant. This means that the different behavior of +V354 and V354E in splicing assays cannot be explained with the RNA recognition by RRM2.

To answer if the different chemical shift perturbation pattern also results in a difference in affinity, ITC experiments with the two isoforms and the V354E mutant have been conducted in context of the three-domain-construct (**Figure 51A and B**). The three-domain-construct was chosen since the RRM2 domain lacks tryptophan residues, which makes the determination of the protein concentration more error prone. The ITC curves show small differences in K_D. The -V354 isoform shows the highest affinity with a K_D of 8 +/- 7 nM. The +V354 and V354E have slightly lower affinities with a K_D of 18 +/- 5 nM and 40 +/ 9 nM. Considering the error, there is no relevant difference detectable between the three variants (**Figure 51C**).

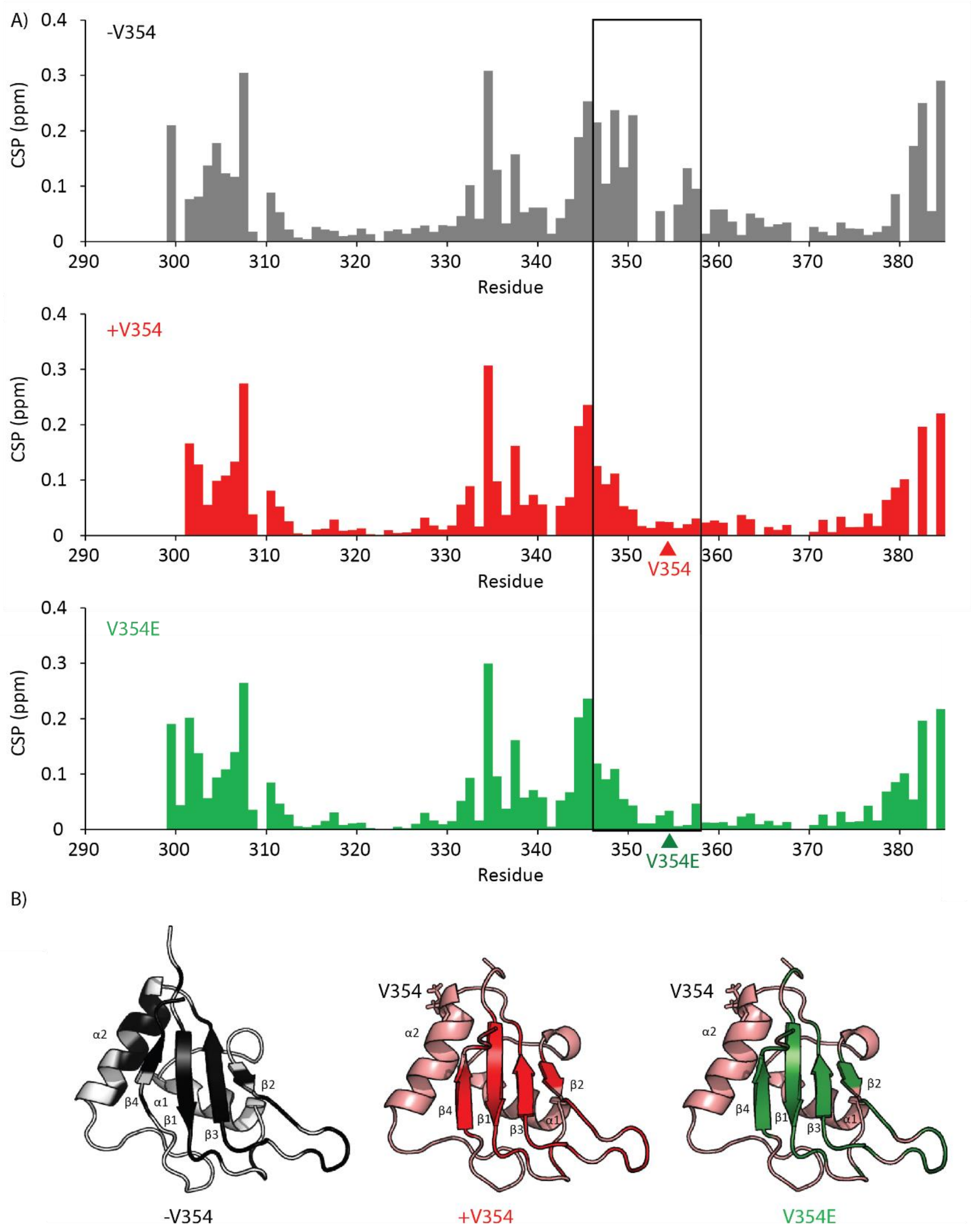
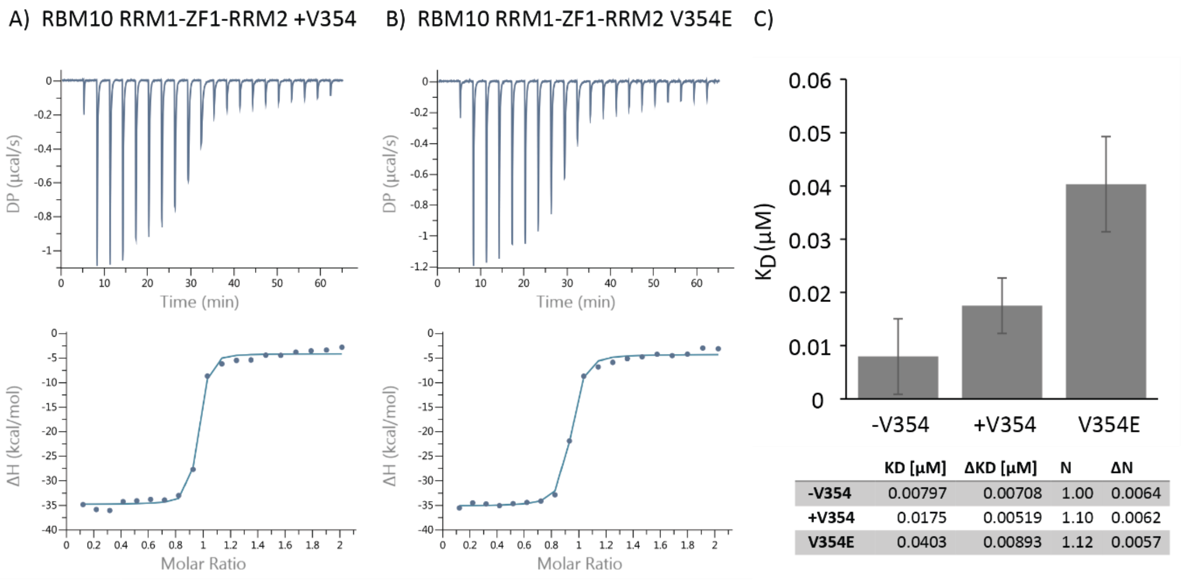


Figure 50 RNA binding of RBM10 RRM2 +V354 isoform and V354E mutant: *A) chemical shift perturbation plots extracted from NMR titrations of the -V354 isoform (grey), the +V354 isoform (red) and the V354E mutant (green) with the NUMB derived 12-mer RNA oligo. Small arrows indicate the additional V354 and the V354E mutation. The box indicates a cluster of residues that behave different in the -V354 isoform in comparison with +V354 and V354E. B) Data from A plotted on the -V354 (white) or +V354 (salmon) RRM2 structure in corresponding colors to A.*



A) RBM10 RRM1-ZF1-RRM2 +V354 B) RBM10 RRM1-ZF1-RRM2 V354E C)

	KD [µM]	ΔKD [µM]	N	ΔN
-V354	0.00797	0.00708	1.00	0.0064
+V354	0.0175	0.00519	1.10	0.0062
V354E	0.0403	0.00893	1.12	0.0057

Figure 51: ITC data RRM1-ZF1-RRM2 +V354 isoform and V354E mutant: *ITC data for RRM1-ZF1-RRM2 +V354 (A) and V354E (B) titrated with the NUMB derived 12-mer RNA oligo. C) Plot of dissociation constant for RRM1-ZF1-RRM2 -V354, +V354 and V354E derived from* ***Figure 35****A, A and B from this figure, respectively. Values are listed in the table below. Additionally the binding stoichiometry N is provided.*

6.3.3 Domain-Domain interaction are also not influenced by RNA binding

Next the effect of the extra residue V354 and the V354E mutation was studied in context of the three-domain construct. This is especially interesting for the V354E mutation. Since there is no big differences in RNA binding induced by the mutation, maybe the differences seen in splicing assays can be explained by differences in domain-domain interaction of the three RNA binding domains.

^{1}H,^{15}N-correlation spectra of RBM10 RRM1-ZF1-RRM2 -V354 and +V354 show only differences in chemical shift in the RRM2 domain, RRM1 and ZF1 do not show significant changes in chemical shift (**Figure 52A and C**). Therefore it can be assumed that the +V354 isoform has no stable domain-domain contacts in the apo-protein, as was shown for the -V354 isoform (**Figure 30**). Not all residues of the RRM2 could be assigned in the spectrum of the +V354 isoform, therefore the CSP plot has some gaps for RRM2.

Comparing the ^{1}H,^{15}N-correlation spectra of RBM10 RRM1-ZF1-RRM2 +V354 and V354E shows also only differences in the RRM2 domain, while chemical shifts of RRM1 and ZF1 residues do not differ in chemical shift. So also the V354E does not introduce stable interactions between the RNA binding domains in the apostate.

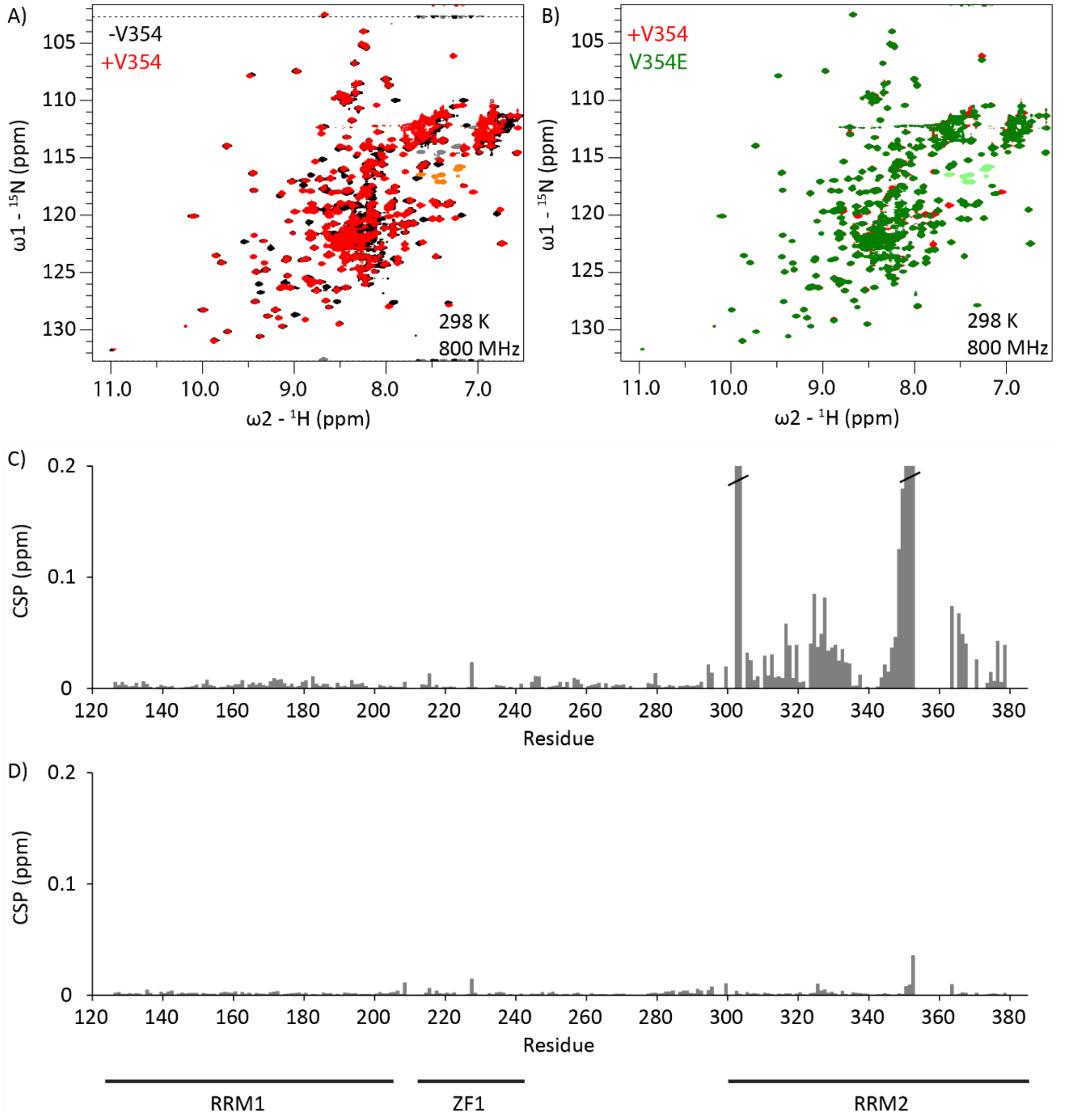


Figure 52 -V354 isoform and V354E mutant in context of RRM1-ZF1-RRM2: *A) Superimposition of* ^{1}*H,*^{15}*N-correlation spectra of RBM10 RRM1-ZF1-RRM2 -V354 (black) with RBM10 RRM1-ZF1-RRM2 +V354 (red). B) Superimposition of* ^{1}*H,*^{15}*N-correlation spectra of RBM10 RRM1-ZF1-RRM2 +V354 (red) with RBM10 RRM1-ZF1-RRM2 +V354 (green). C) Differences in chemical shift between spectra from A plotted against the residue number. Black bars indicate cut-off bars. D) Differences in chemical shift between spectra from B plotted against the residue number.*

7 RBM10 Discussion

RBM10 is a regulator of alternative splicing that is involved in the regulation of apoptosis (Sutherland et al. 2005). It contains N-terminal RNA interacting and C-terminal protein interacting domains. RBM10 has been shown to bind pre-mRNAs upstream of the 5' splice site and downstream of the 3'splice site (Bechara et al. 2013; Y. Wang et al. 2013) and particularly to regulate the alternative splicing of exon 9 of the NUMB gene, leading to two different NUMB isoforms that have opposite effects on the regulation of Notch signaling and thereby RBM10 regulates cell proliferation (Misquitta-Ali et al. 2011). Using a 12-mer RNA oligo derived from the NUMB gene, as it was used in splicing assays (Bechara et al. 2013), it could be shown in this thesis that the three N-terminal RNA interacting domains RRM1, ZF1 and RRM2 interact with this oligo in a high affinity 1:1 complex. With an affinity of 704 nM RRM1-ZF1, which's domains are connected by a short linker, contributes the majority of the affinity. This affinity of RRM1-ZF1 to the NUMB derived 12-mer RNA oligo is around 10 fold lower than the affinity found by Collins et al. for a CUGUGGA 7-mer RNA oligo, which is present in an exonic sequence of the TNR6A pre-mRNA and allows RBM10 to promote exon skipping in this pre-mRNA (Collins et al. 2017). Collins et al. reported that the main sequence specificity is governed by ZF1 which recognizes the GGA motif in the oligo. This is consistent with a report by Nguyen et al., who find that a family of RanBP2 type ZFs, including ZF1 in RBM10, recognize a GGU motif. The absence of two consecutive guanine in the NUMB derived 12-mer RNA oligo used in this study might explain the lower affinity. But still ZF1 contributes to affinity, since the RRM1 alone binds the NUMB derived 12-mer RNA oligo around 9 fold weaker than the RRM1-ZF1 construct. The contribution of RRM1 and ZF1 to recognition of the NUMB derived RNA oligo also gets supported by the chemical shift perturbation extracted from NMR titration data, which show that both domains and the linker between them are affected by RNA binding. Another point supporting this is that while in the apo-protein the domains tumble independently in solution, the domain orientation becomes more rigid when bound to RNA (**Figure 53**).

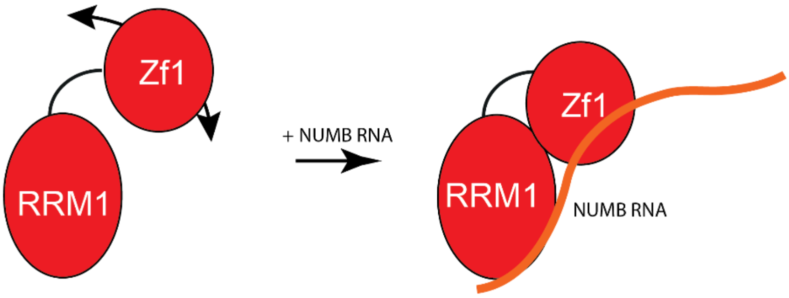


Figure 53 RNA binding of RBM10 RRM1-ZF1*: In the RBM10 RRM1-ZF1 apo-protein the two domain orientation is flexible in respect towards each other. Upon RNA binding the domains become more rigid with respect to each other. Both domains are involved in RNA binding.*

Collins et. al. also found that the same 7mer RNA gets recognized by RBM10 RRM1-ZF1-RRM2 in a consecutive manner. They observe peaks belonging to RRM1-ZF1 changing first in a slow exchange regime and after these are saturated at a molar ratio of 1:1, they observe peaks corresponding to RRM2 to shift in a fast exchange regime (Collins et al. 2017). A titration with the NUMB RNA derived 12-mer RNA oligo in contrast shows that peaks of RRM1-ZF1 and RRM2 change simultaneously in a fast-intermediate exchange regime and that the protein is saturated at a molar ratio of 1:1. This 1:1 stoichiometry is also confirmed by ITC and SLS data. Interestingly the 5' end of the NUMB derived 12-mer RNA oligo is similar to the 5' end of the oligo used by Collins et al. and the 3' end of the NUMB derived 12-mer RNA oligo contains three consecutive cytosine nucleotides. Collins et al. have also shown by scaffold independent analysis and ITC that the RRM2 binds a CCCAC 5mer RNA oligo. Taken all this together one can propose a model, where the 5' end of the 12-mer RNA oligo is recognized by RRM1, the second guanosine residue in the sequence might be recognized by ZF1, and the consecutive cytosine nucleotides in the 3' end by RRM2 (**Figure 54**).

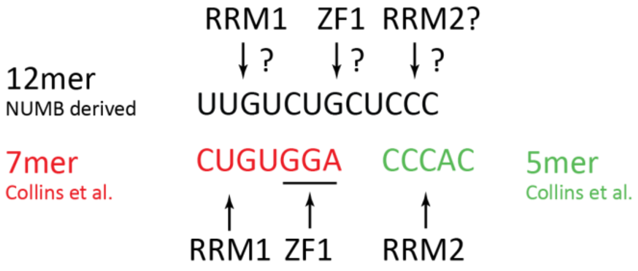


Figure 54 Comparison of RRM1-ZF1-RRM2 RNA binding: *Shown is the NUMB derived RNA sequence (black), compared to the RNA motifs from the literature that are known to be recognized by RRM-ZF1 and RRM2 (Collins et al. 2017).*

The affinity of the RRM2 alone to the NUMB derived 12-mer RNA oligo is with 23.8 µM rather low compared to the 704 nM observed for RRM1-ZF1. Despite that, the three domain construct comprising all three RNA binding domains has an affinity of 8 nM, which indicates that for the recognition of the intronic NUMB RNA sequence, despite the lack of a GGA motif, all three domains cooperate. In SAXS data upon RNA binding a compaction of the three-domain-construct is observed, which fits the expectation that all three domains fit on a 12-mer RNA oligo. Interestingly the D^{max} of the free and bound RBM10 RRM-ZF1 three domain-construct extracted from SAXS data does not change. Since the D^{max} probably corresponds to a detached conformation, this might be explained with the RRM2 being not completely bound but being detached in a minor state. Such a minor detached state would also explain the slightly lower correlation time of RRM2 compared to RRM1 determined by ^{15}N-relxation data in the context of RBM10 RRM1-ZF1-RRM2 bound to the NUMB derived 12-mer RNA oligo.

Looking at the RNA side of the interaction the peaks corresponding to H5-H6-correlations in a TOCSY spectrum get broadened strongly upon binding of RBM10 RRM1-ZF1 or RRM1-ZF1-RRM2. This indicates the presence of exchange. This might be caused by the absence of a GG in the NUMB derived 12-mer RNA oligo, which might lead to a not well defined binding register of ZF1. In addition the possibility of an minor conformation where RRM2 is detached, might add to the dynamics in the RRM1-ZF1-RRM2:RNA complex (**Figure 55**).

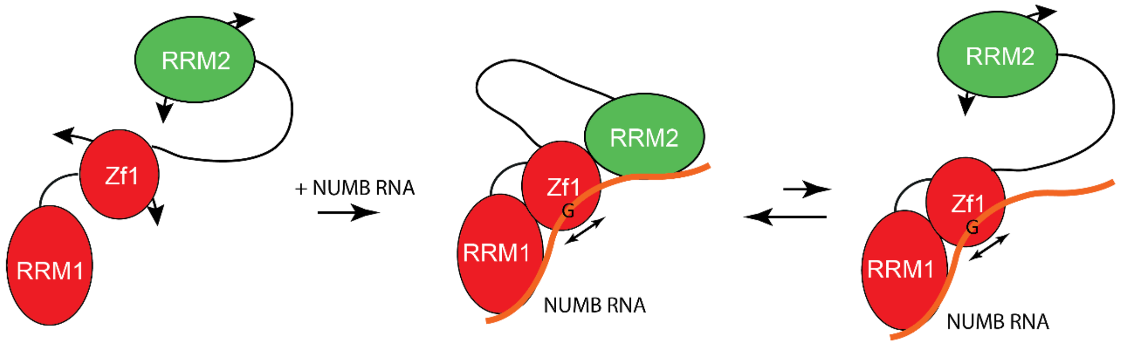


Figure 55 RNA binding of RBM10 RRM1-ZF1-RRM2: *In the RBM10 RRM1-ZF1-RRM2 apo-protein all three domains are flexible in respect to each other. Upon RNA binding the domains become more rigid with respect to each other and the construct is compacted according to SAXS data. Since the* D^{max} *from SAXS does not decrease upon RNA binding, there might be a minor population where the RRM2 domain is detached. Because of the absence of a GG sequence in the NUMB derived 12-mer RNA oligo, the binding register of the ZF1 might not be well defined.*

In conclusion the three RNA binding domains of RBM10 show a complex dynamic behavior. The paramagnetic relaxation data presented here indicate that the conformation of the three domains change upon RNA binding. While the apo-protein is highly dynamic, the protein RNA complex shows reduced dynamics. It remains the question if it is possible to generate a structural model of the protein RNA complex using the PRE and SAXS data presented here and additionally RDC data, which still have to be recorded, in a structure calculation where the domains are treaded as rigid bodies, like it has been presented by Simon et al. (Simon et al. 2010). Alternatively the data can be used to generate a structural ensemble like it has been done by Huang et al. (Huang et al. 2014) with the same data as input. Considering the highly dynamic behavior of the apo-protein, it can only be modelled by an ensemble approach.

RBM10 exhibits an isoform generated through alternative splicing, which differs from the variant discussed so far by an additional valine at the N-terminus of α-helix 2 in RRM2 in position 354. It has been suggested by molecular modelling that this additional valine might disrupt the α-helix 2 (Tessier et al. 2015). In this work the solution NMR structures of the two variants of RRM2 (-V354 and +V354) are compared. The addition of the extra valine does not disrupt the α-helix 2, instead the α-helix 2 becomes a bit bended and the orientation of β-strand 4 changes a little. The extra valine does not show a dramatic effect on RNA binding. In NMR-titrations α-helix 2 and the loop N-terminal of α-helix 2 show lower chemical shift perturbation upon RNA binding in the +V354 variant compared to the -V354 variant, but the affinity does not change significantly.

Interestingly the additional residues V354 is found mutated to glutamate in lung cancer cells. The V354E mutant also does not show increased skipping of exon 9 in the NUMB gene as the -V354 and +V354 isoforms of the protein do (Bechara et al. 2013; Hernández et al. 2016). ^{1}H,^{15}N-correlation spectra of the RBM10 RRM2 V354E show only local changes in chemical shift around the mutation site, indicating no change of the overall protein structure. NMR titrations with the NUMB derived 12-mer RNA oligo show the same chemical shift perturbation pattern for the V354E mutant as for the V354 isoform. Also the affinity determined by ITC does not change significantly, indicating that the mutation does not influence RNA binding, which makes sense since the side chain of V354 is located in an area not involved in RNA binding. The V354E mutation also does not induce changes in chemical shift in the other domains in context of the three-domain-construct, indicating that the mutation does not change domain arrangement of the RNA binding domains in the apo-protein. Therefore it is likely that the mutation effects either interactions with the C-terminal protein interacting domains or the interaction with an external factor.

This thesis characterizes the complex dynamic domain arrangement and RNA binding properties of RBM10. It is shown that despite the long flexible linker between ZF1 and RRM2 all the three RNA binding domains of RBM10 together recognize a 12-mer RNA oligo derived from the *NUMB* gene in a high affinity 1:1 complex. These characterized interaction is of high importance since RBM10 can induce skipping of exon 9 during splicing of the NUMB pre-mRNA and the resulting shorter *NUMB* gene product can inhibit Notch signaling and thereby affect cell proliferation.

Part 2

The molecular chaperone Hsp90

8 Hsp90 Introduction

8.1 The Hsp90 chaperone

The structure of a protein is coded in its primary sequence. It contains all the information about the formation of secondary structure elements like α-helices and β-sheets and about the tertiary three dimensional structure. Proteins fold after or during their translation through the ribosome. Even so the protein sequence determines the structure, proteins can misfold and aggregate. Aggregation can happen through hydrophobic parts of the protein, which are solvent exposed during the folding process and would be buried inside the correctly folded protein (Berg, Tymoczko, and Stryer 2010).

Molecular chaperones can prevent aggregation by binding to unfolded proteins or folding intermediates and preventing them from aggregation. Chaperones are often called heat shock proteins (Hsp), because they are found to be overexpressed at higher temperature, where aggregation is more likely (Alberts, Johnson, and Lewis 2007). The chaperone Hsp90 forms a homodimer. Each monomer comprises three domains. A N-terminal nucleotide binding domain (NTD), a middle domain (MD), which is responsible for client protein binding, and a C-terminal dimerization domain (CTD) (**Figure 56**).

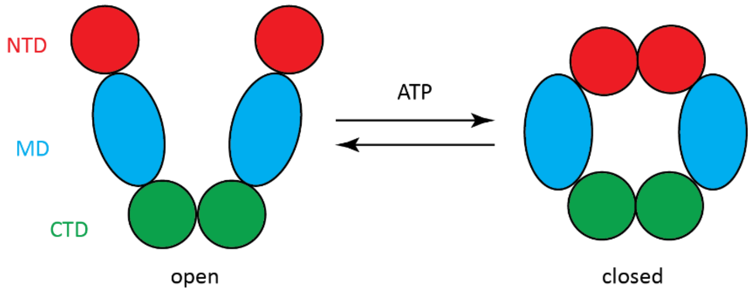


Figure 56 Architecture of the Hsp90 homodimer: *A) Blob representation of the Hsp90 homodimer in the open and closed conformation. The N-terminal domain is depicted in red, the middle domain in blue and the C-terminal domain in green.*

In the nucleotide free form the Hsp90 forms a V-shape, where the protein is dimerized only via the CTD domains. Upon ATP binding the protein undergoes a rearrangement via several intermediates during which it adopts an N-terminal closed state. Upon ATP binding the NTD rearranges, closing the ATP binding site with the α-helical lid (intermediate state), then the NTDs dimerize (closed-1 state) and afterwards associate with the MD (closed-2 state). After ATP hydrolysis, which involves also the MD, ADP is released and Hsp90 opens again to the nucleotide-free-state. The Hsp90 chaperone cycle is controlled by a large set of co-chaperones which can influence the ATPase chaperone cycle or recruit client proteins to Hsp90 (Schopf, Biebl, and Buchner 2017).

The co-chaperone activator of Hsp90 ATPase activity (Aha1) accelerates the Hsp90 ATPase activity (Barry Panaretou et al. 2002). The Aha1 protein comprises two domains, an N-terminal domain (ND) and a C-terminal domain (CD). The interaction of the ND of Aha1 with the MD of Hsp90 has been shown by a crystal structure (Meyer et al. 2004), while the interaction of both domains of Aha1, but mainly the CD, with the NTD of Hsp90 have been shown by NMR (Retzlaff et al. 2010). Aha1 preferably binds the N-terminal closed state of Hsp90. Hsp90 together with Aha1 plays an important role in the maturation of cystic fibrosis transmembrane conductance regulator (CFTR) mutants. Partial silencing of Aha1 by siRNA rescues the most common disease mutant (ΔF508) of CFTR in cystic fibrosis (X. Wang et al. 2006).

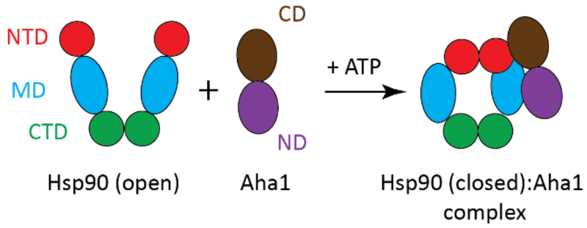


Figure 57 Architecture of the Hsp90:Aha1 complex: *Blobs representation of the Hsp90:Aha1 complex. The N-terminal domain of Hsp90 is depicted in red, the middle domain in blue and the C-terminal domain in green. The C-terminal domain of Aha1 (CD) is shown in brown and the N-terminal domain (ND) in purple.*

Geldanamycin and radicicol are known as competitive inhibitors of the ATPase activity of Hsp90 (Roe et al. 1999). Additionally, the antibiotic novobiocin was found to interact with a novel site in the C-terminal domain of Hsp90 (Marcu et al. 2000). With 4-hydroxytamoxifen and tamoxifen also two compounds have been identified, which activate the ATPase activity of Hsp90 (R. Zhao et al. 2010).

Several mutations in Hsp90 are known from the literature which influence the ATPase activity. The deletion of the first 8 amino acids of Hsp90 (Δ8) leads to higher ATPase activity and promotes dimerization (Richter, Reinstein, and Buchner 2002).The temperature sensitive T22I mutant and the A107N mutant, which was designed to favor lid closure, feature increased ATPase activity (Prodromou et al. 2000). The D79N mutant is deficient of binding ATP, while the E33A mutant can bind but not hydrolyze ATP (B Panaretou et al. 1998).

Scope of the Hsp90 projects

Hsp90 is a molecular chaperone, which is important for correct folding of proteins. It undergoes conformational changes in an ATP dependent manner. Here the binding sites of an Hsp90 activator and an inhibitor of Hsp90 in complex with the activating co-chaperone Aha1 will be mapped by NMR-spectroscopy. In addition, an analysis of the effect of several mutants, with known effect on the ATPase activity of Hsp90 is done by NMR, to get an idea of the importance of cycling timings in the Hsp90 chaperone cycle.

9 Hsp90 Material and Methods

In this part the details only for the presented NMR data are given, experimental details to the other data can be found in the original publications.

9.1 NMR analysis of Hsp90 modulator 4

All NMR spectra were recorded on a Bruker Advance III 600 MHz NMR spectrometer equipped with a cryogenic triple resonance probe. The construct of the NTD of yeast Hsp90 comprises residues 1 to 210 and the construct of the MD of Hsp90 comprises the residues 259 to 529. The proteins were prepared by my collaborator Bettina Zierer. For the NMR measurement the protein samples were dissolved in 20 mM sodium phosphate buffer (pH 6.5) containing 100 mM sodium chloride, 2 mM EDTA, 1 mM DTT and 5% D_2O. For the analysis of the NTD a protein concentration of 100 µM was used and 200 µM of modulator 4 was added. For the MD a protein concentration of 50 µM was used and 50 µM of modulator 4 were added. Modulator 4 was added from a 20 mM stock dissolved in deuterated DMSO. To the reference spectra the same amount of deuterated DMSO was added. All NMR spectra were recorded at 298 K. Processing of the NMR spectra was done using nmrPipe (Delaglio et al. 1995) and analysis of the NMR spectra was done with CCPNMR analysis (Vranken et al. 2005).

9.2 NMR analysis of Hsp90-Aha1 modulator 1.

NMR experiments for the characterization of the Hsp90-Aha1 modulator 1 were recorded on the same NMR equipment and using the same buffer as for the Hsp90 modulator. Additionally constructs for ND (1-156) and CD (156-356) of Aha1 were used. The proteins were expressed and purified by Sandrine Stiegler. For measurements of the Hsp90 MD, the Aha1 ND and the Aha1 CD a protein concentration of 200, 167 and 200 µM, respectively and a concentration of modulator 4 of 200 µM were used. For the NTD of Hsp90 a protein concentration of 95 µM were used and a titration was done with a modulator 4 concentration up to 285 µM. Changes in chemical shift were already observed at lower concentrations of the modulator, but not to saturation (data not shown). For the spectra containing ATP 4 mM ATP and 6 mM magnesium sulfate were added to a 100 µM sample of the Hsp90 NTD. To this than 600 µM of modulator 1 were added. Processing and analysis of NMR spectra was done with the same programs as for the Hsp90 modulators.

9.3 NMR analysis of the NTD mutations

NMR spectra for the mutants were recorded on the same NMR equipment and with the same buffer as for the modulator studies. The NTD comprising residues 1-210 were used in concentrations between 100 and 465 µM. Spectra were recorded for the wildtype as well as the Δ8, T22I, E33A, D79N and

A107N variants. For the variants E33A and D79N the assignment of the mutant was confirmed using ^{15}N-edited NOESY spectra. The wt, T22I, E33A and A107N were also measured in the presence of 4 mM ATP and 6 mM magnesium sulfate. The E33A and D79N variants were measured also in the presence of a 2:1 excess of the non-hydrolysable ATP analog AMP-PNP. The spectra were processed and analyzed using the same programs as for the modulator studies.

10 Hsp90 Results

10.1 Hsp90 Modulators

The data presented in this chapter have been published in:

“Zierer, Bettina K., Matthias Weiwad, Martin Rübbelke, Lee Freiburger, Gunter Fischer, Oliver R. Lorenz, Michael Sattler, Klaus Richter, and Johannes Buchner. “Artificial Accelerators of the Molecular Chaperone Hsp90 Facilitate Rate-Limiting Conformational Transitions.” Angewandte Chemie International Edition 53, no. 45 (November 3, 2014): 12257–62.“

10.1.1 Modulators where identified by a FRET based screening

My collaborator Bettina Zierer in the group of Johannes Buchner characterized modulators of the Hsp90 chaperone cycle, which were identified in a FRET based screen. To monitor the closing reaction a donor and an acceptor fluorophore were introduced to Hsp90 in the NTD and the MD respectively. The fluorophores were positioned that the fluorophores are far apart in the open conformation and close together in the N-terminal closed conformation and thereby lead to low FRET in the open and high FRET in the closed conformation. The closure in presence of compounds from a library was induced by the non-hydrolysable ATP analog AMP-PNP (**Figure 58A**).

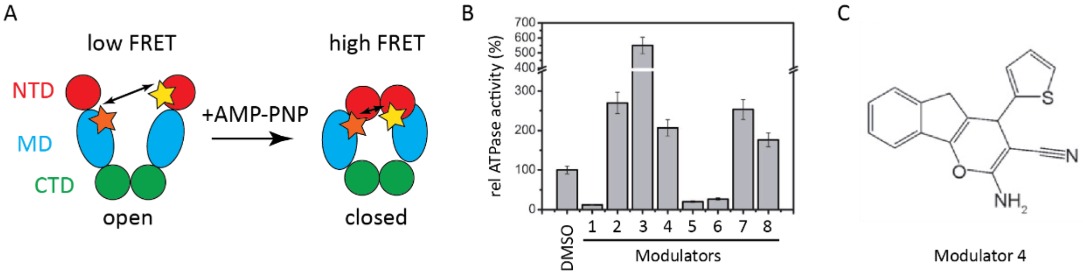


Figure 58: Identification of Hsp90 activator modulator 4: *A) Hsp90 is labelled in the MD and NTD with FRET donor and acceptor fluorophores, respectively. Now in mixed dimers, one monomer carrying a donor fluorophore and the other monomer carrying an acceptor fluorophore one can read out the N-terminal closing of the Hsp90 upon AMP-PNP binding and the influence of a compound library on this closing. B) Relative ATPase activity (*k_{cat}*) in presence of the modulators identified in the FRET based screen. C) Chemical structure of modulator4, which is the activator with the lowest apparent* K_D*.*

The influence on Hsp90 ATPase activity of the hits from the FRET based assay was tested (**Figure 58B**). Modulator 4 (**Figure 58C**) shows a stimulation of the ATPase activity by 2.8 fold and has with a K_D of 6 µM a relatively high affinity. In presence of modulator 4 the K_M value of Hsp90 to ATP is reduced from 0.32 mM to 0.18 mM, which indicates allosteric binding. The compound is soluble enough to be characterized by NMR to identify the binding site (Zierer et al. 2014).

10.1.2 NMR titration reveal the binding site of an accelerator of the Hsp90 chaperone cycle

^{1}H,^{15}N-corralation spectra of the Hsp90 NTD and MD in absence and presence of modulator 4 were recorded (**Figure 59**A).

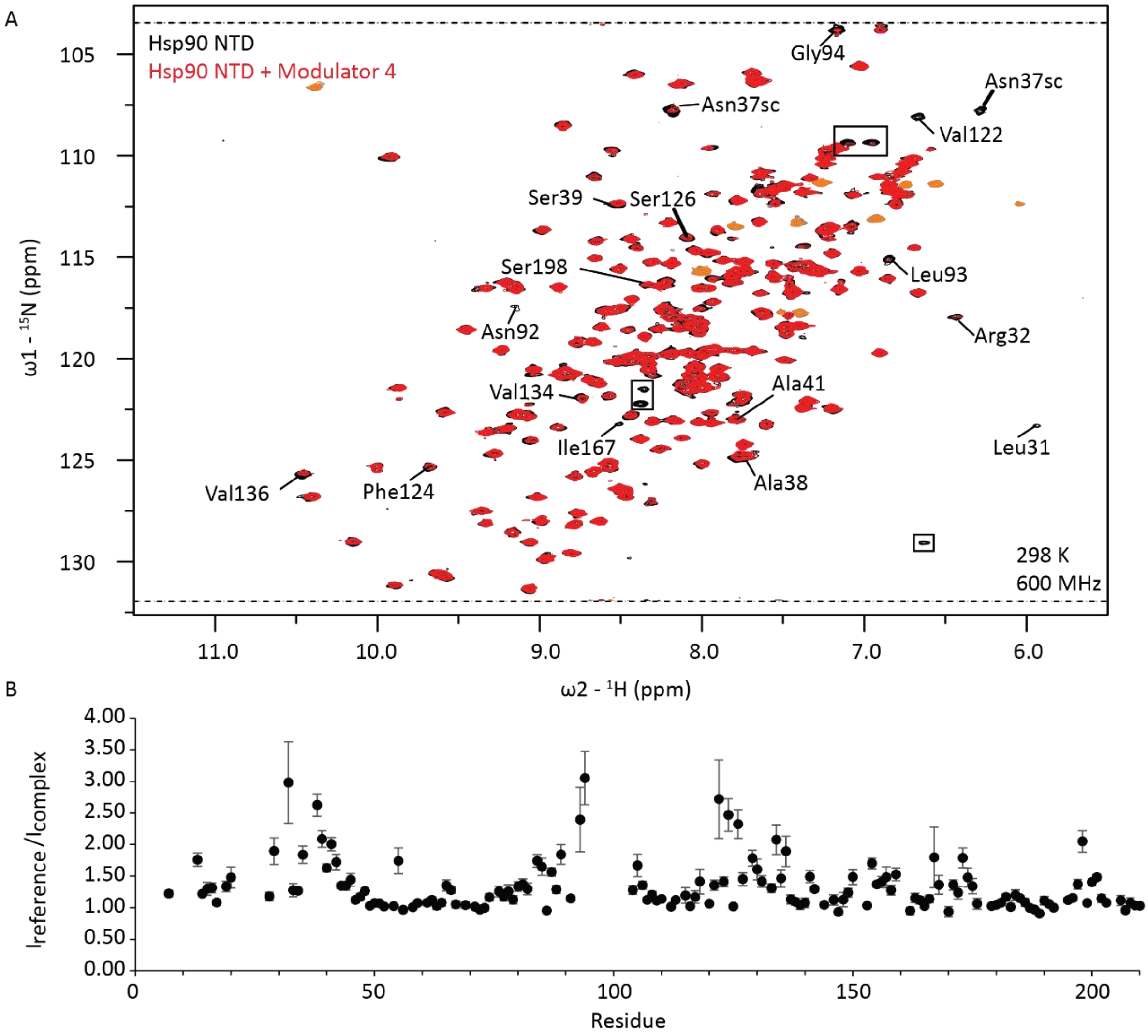


Figure 59 NMR analysis of the interaction of Hsp90 NTD with modulator 4: *A) Shown is a superimposition of 1H,15N-correlation spectra of the NTD of Hsp90 in absence (black) and presence (red) of modulator 4. Assigned peaks with reduced intensity in the presence of modulator 4 are labelled. Boxed peak have reduced peak intensity as well, but are not assigned. B) Ratio of the peak intensity of the spectra from A plotted against the residue number.*

For the NTD upon addition of the modulator 4 line broadening for several peaks can be observed. Peaks experiencing line broadening cluster in three regions on the primary sequence (residues 32-41, 89-94 and 121-132-134 (**Figure 59B**). These regions correspond to the nucleotide binding site and the interface between the NTD and the MD. Especially α-helix 2 is strongly affected (**Figure 60**). This data suggest that modulator 4 binds near α-helix 2 and changes the ATP binding site allosterically. The co-chaperone Aha1 also has the ability to increase the ATPase activity through allosteric binding. While the Hsp90 in the presence of Aha1 has a k_{cat} of 16 min^{-1} it can be further increased through modulator 4 to 20 min^{-1}. It seems therefore that Aha1 and modulator 4 can bind Hsp90 simultaneously. Indeed, the mapped binding site of modulator 4 does not overlap with the known binding site of Aha1 (Retzlaff et al. 2010).

The MD of Hsp90 shows no changes in chemical shift or peak intensity in presence of the modulator 4. Therefore the modulator seems to specifically interact with the NTD (**Figure 61**).

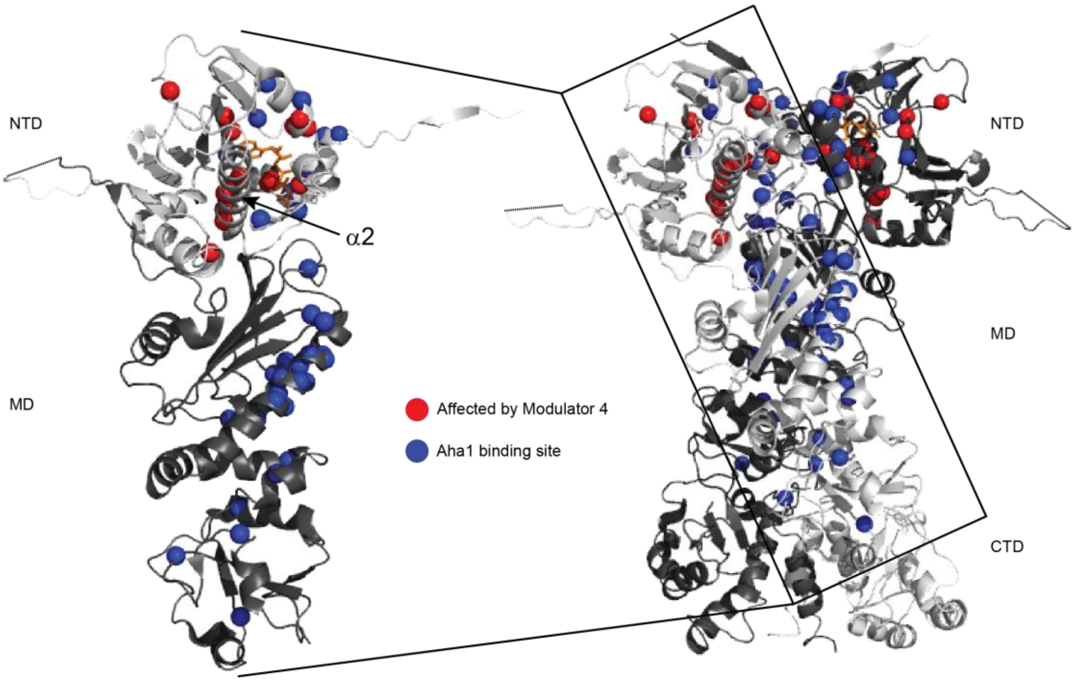


Figure 60 Mapping the binding site of modulator 4: *Residues with significant intensity changes upon binding of modulator 4 are indicated by red spheres on the structure of the closed conformation of full length Hsp90 (PDB: 2CG9). The left panel shows only the NTD (light grey) and MD of one monomer, while the right panel shows the full homodimer, with one monomer colored light and one in dark grey. In addition the known binding site of Aha1 is indicated by blue spheres (Retzlaff et al. 2010).*

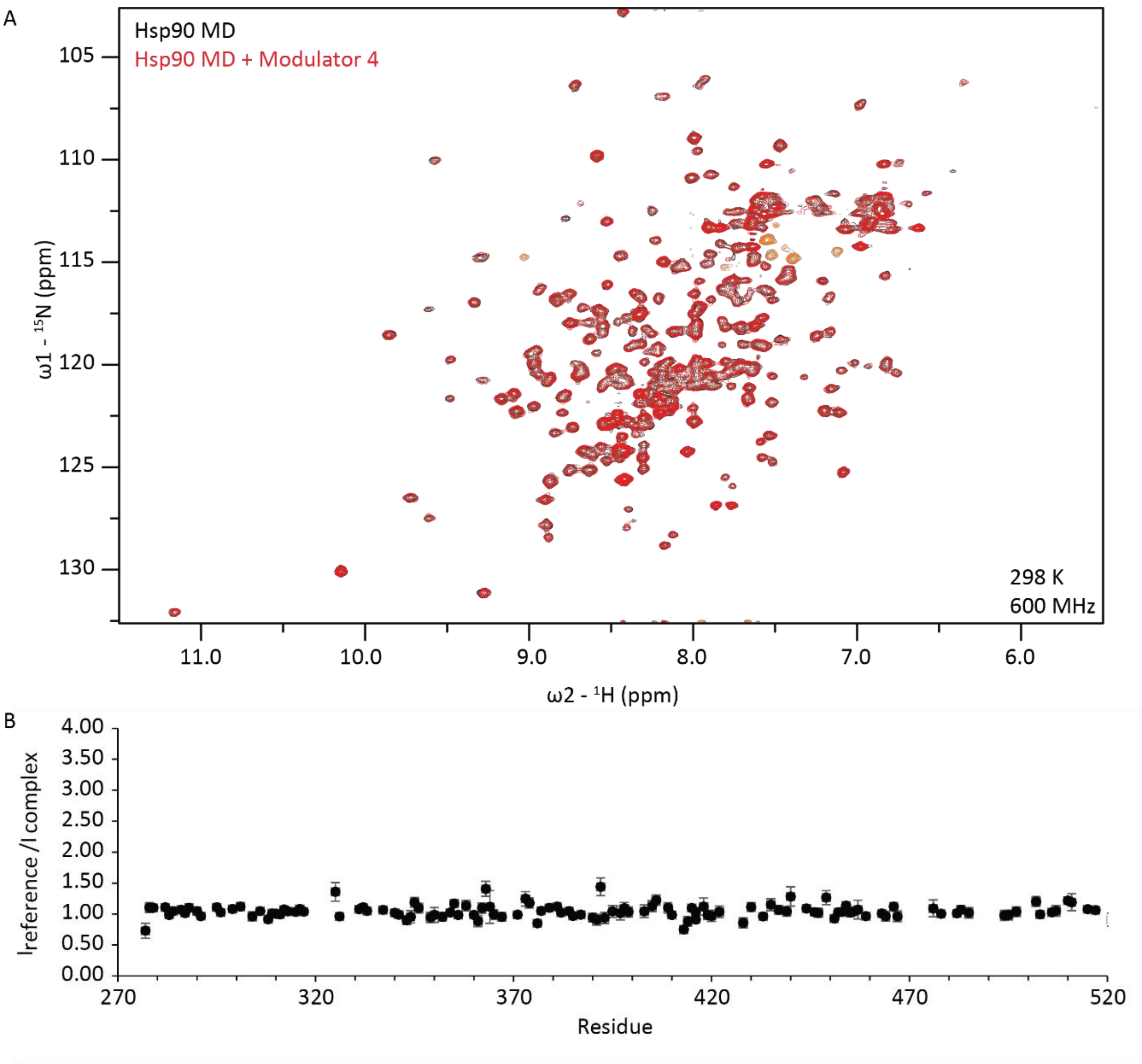


Figure 61 NMR analysis of the interaction of Hsp90 MD with modulator 4: *A) Shown is a superimposition of* ^{1}H,^{15}N*-correlation spectra of the MD of Hsp90 in absence (black) and presence (red) of modulator 4. B) Ratio of the peak intensity of the spectra from A plotted against the residue number.*

10.2 Hsp90-Aha1 Modulators

The data in this chapter have been submitted to the Journal of Biological Chemistry. The manuscript is currently revised.

10.2.1 Identification of Hsp90-Aha1 modulators by a FRET based screen

A similar approach as for the Hsp90 modulators was employed by Sandrine Stiegler from the Buchner group to identify modulators which are specific for Hsp90 in complex with Aha1. In this case the fluorophores for the FRET based assay were placed in the MD of Hsp90 and in Aha1. A FRET signal can only be observed from the complex of Aha1 and Hsp90 and therefore the FRET signal can be used to monitor complex formation. Since Aha1 binds Hsp90 only in the nucleotide bound closed form, closing was induced using the non-hydrolysable ATP analog AMP-PNP (**Figure 62A**). This assay was used to screen the same library that was used for the identification of the Hsp90 modulators. Interestingly none of the hits in the two screens overlapped. The most promising molecules were further characterized by an ATPase assay on the Hsp90-Aha1 complex to measure the stimulation of Hsp90 by Aha1 in presence of the compound. Some compounds stimulate the activation of Hsp90 by Aha1 further, while others inhibit the stimulation of Hsp90 by Aha1 (**Figure 62B**). With 93% inhibition modulator 1 (**Figure 62C**) shows the strongest inhibition of the activation of Hsp90 by Aha1, while it shows only a moderate effect on the ATPase activity of Hsp90 alone (22% inhibition). Modulator 1 has an apparent K_D for to the Hsp90:Aha1 complex of 23.5 µM.

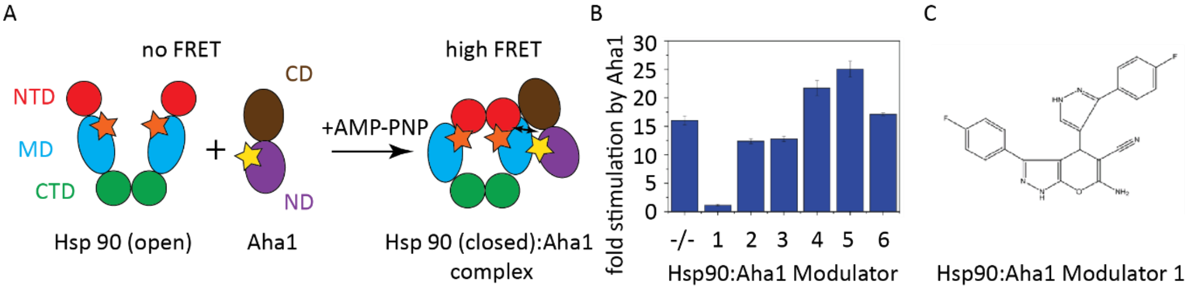


Figure 62 Identification of Hsp90-Aha1 modulator 1: *A) Hsp90 and Aha1 are labelled with FRET donor and acceptor fluorophores respectively. The positions of the fluorophores are designed in a way that in the complex a FRET signal is observed, thereby the binding of Aha1 can be monitored. Since Aha1 binds only the nucleotide bound closed form of Hsp90 the closure is induced by the non-hydrolysable ATP analog AMP-PNP. B) Stimulation of the ATPase activity of Hsp90 by Aha1 in presence of the modulators identified in the FRET based screen. C) Chemical structure of modulator1.*

10.2.2 NMR Analysis reveals that Modulator 1 can release ATP from the NTD of Hsp90

To identify the binding site of modulator 1 $^1H,^{15}N$-correlation spectra of the Hsp90 NTD and MD as well as the ND and CD of Aha1 in absence and in presence of modulator 1 were recorded. The MD of Hsp90 and the CD of Aha1 show small changes in chemical shift upon addition of modulator 1, which are probably caused by unspecific interactions (**Figure 64A and C**). The ND of Aha1 shows no changes in chemical shift or peak intensity and therefore no sign for binding of modulator 1 (**Figure 64B**).

The NTD of Hsp90 in contrast shows clear changes in chemical shift upon binding of modulator 1 (**Figure 63A**). The changes cluster around the ATP binding site, with α-helix 2 being especially strongly affected (**Figure 63B and C**). Interestingly the affected residues leigh partly in the known Aha1 binding site (Retzlaff et al. 2010). Since the changes in chemical shift cluster around the ATP binding site in Hsp90, binding of modulator 1 to the ATP bound form of the NTD of Hsp90 was tested. The $^1H,^{15}N$-correlation spectra of the NTD of Hsp90 in presence of ATP shows clear differences in chemical shifts to the spectrum in absence of ATP. Addition of modulator 4 to the Hsp90 NTD in presence of ATP causes peaks to appear in the positions of the spectrum of the Hsp90 NTD bound to modulator 4 in absence of ATP. Only a small portion of the peak intensity remains in the position of the Hsp90 NTD

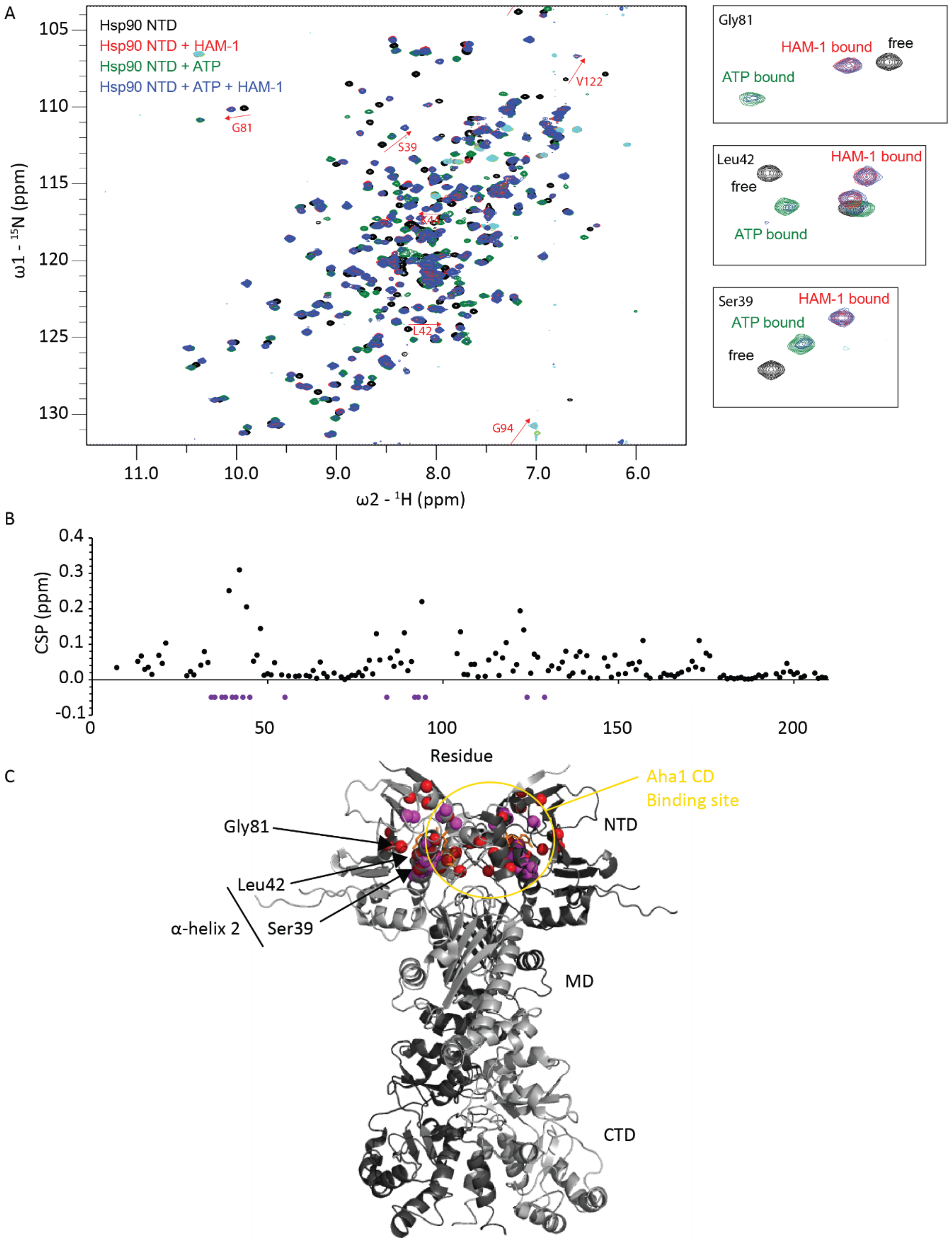


Figure 63 ***NMR analysis of modulator 1 binding by the NTD of Hsp90:*** *A) Superimposition of ^{1}H,^{15}N-correaltion spectra of the NTD of Hsp90 alone (black), in presence of modulator 1 (red), in presence of ATP (green) and in presence of modulator 1 and ATP (blue). On the right side zoom-ins are shown. B) Chemical shift perturbation between the NTD of Hsp90 in the absence and presence of modulator 1. C) Chemical shift perturbations larger 0.08 ppm from B indicated by red spheres on the structure of the closed full length Hsp90 heterodimer (PDB: 2CG9). Purple spheres indicate residues which experience resonance frequency changes that could not be traced.*

bound to ATP (**Figure 63A**). This clearly shows that modulator 4 can release ATP from the NTD of Hsp90. If this happens through direct competition or allosterically cannot be answered from the NMR data.

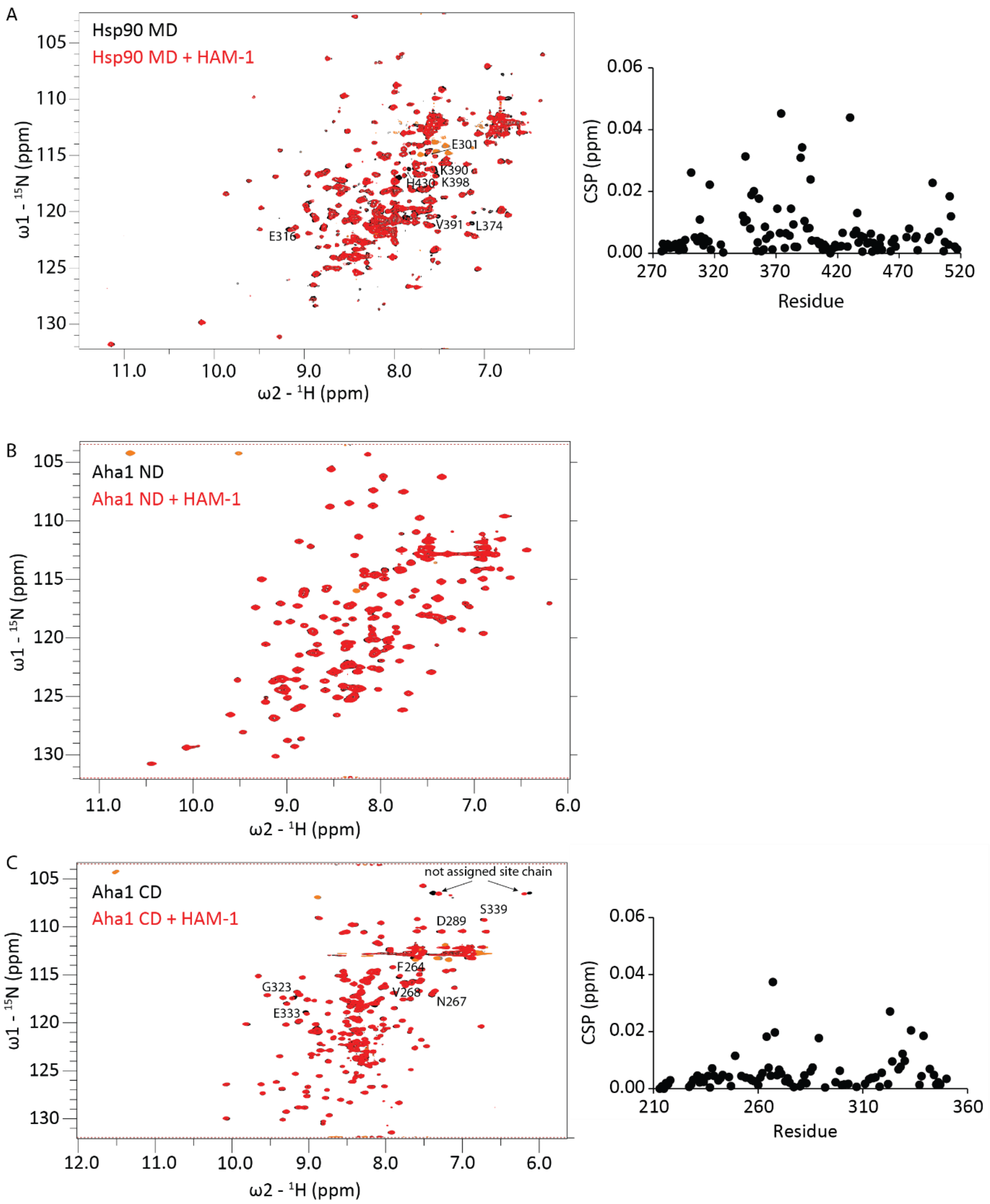


Figure 64 Testing Hsp90 MD and Aha1 domains for modulator 1 binding: *Shown are superimpositions of $^{1}H,^{15}N$-correlation spectra of the Hsp90 MD (A), the Aha1 ND (B), the Aha1 CD (C) in absence (black) and presence (red) of modulator 1. While for the Hsp90 MD and the Aha1 CD the changes in CSP are plotted against the residues number on the right side of the spectrum, this was analysis was not possible for the Aha1 ND since assignments are not available.*

10.3 Cycling timing of Hsp90 Chaperone Cycle

The data in this chapter has been published in:

“Zierer, Bettina K., Martin Rübbelke, Franziska Tippel, Tobias Madl, Florian H. Schopf, Daniel A. Rutz, Klaus Richter, Michael Sattler, and Johannes Buchner. “Importance of Cycle Timing for the Function of the Molecular Chaperone Hsp90.” Nature Structural & Molecular Biology 23, no. 11 (November 2016): 1020–28. “

10.3.1 ATPase activity of Hsp90 mutants does not correlate with viability in yeast

A variety of point mutations and a truncation mutation from the literature with known effects on the ATPase activity of Hsp90 were studied by Bettina Zierer and Franziska Tippel from the group of Johannes Buchner (**Figure 65A**). There are three groups of mutations. The A107N and to less extend also the Δ8 and T22I mutations show increased ATPase activity, the R346S and R380A mutants show reduced ATPase activity and the E33A and D79N mutations show no ATPase activity (**Figure 65B**). Interestingly in all of this three groups variants that can sustain viability in yeast and others that cannot can be found. The viability has been shown with a 5-fluoroorotic acid (5-FOA) shuffling assay (**Figure 65C**).

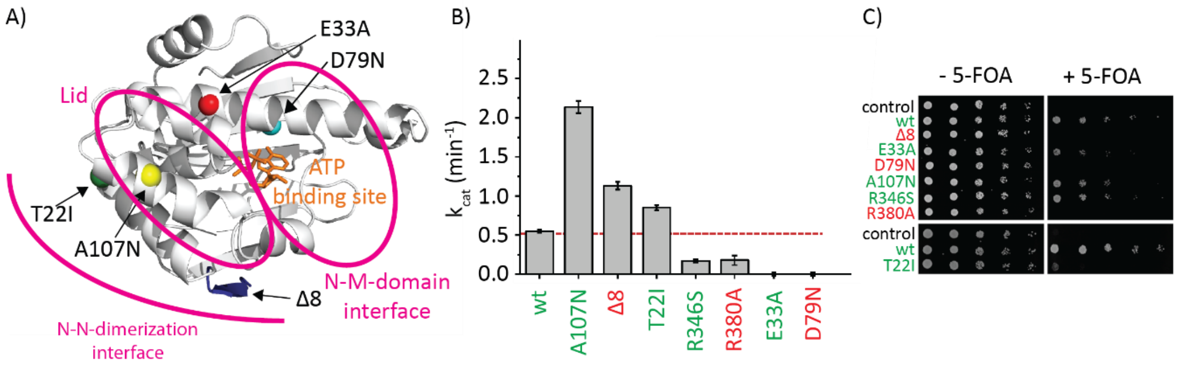


__Figure 65 ATPase activity and of Hsp90 mutants does not correlate with the survival in yeast:__ A) Overview of the mutants in the NTD of Hsp90 used in this study shown on the structure of the NTD of Hsp90 (PDB: 1AM1). B) ATPase activity of the mutants compared to the wt. The mutants labelled in green can support viability, while the mutants labelled in red cannot. C) 5-fluoroorotic acid (5-FOA) shuffling assay with the Hsp90 variants. Viability is indicated as in B.

10.3.2 NMR analysis of the mutants in the N-terminal domain of Hsp90

For the mutants in the N-terminal domain of Hsp90 ^{1}H,^{15}N-correlation spectra were recorded and compared to the spectrum of the wildtype protein to characterize structural changes by the mutation (**Figure 66**). The two ATPase deficient mutants E33A and D79N show very different changes in the ^{1}H,^{15}N-correlation spectra. While the E33A mutant shows differences in chemical shift spread over all α-helices, the D79N mutant has more local effect on the β-sheet. The wide spread effects of the E33A mutation show E33’s importance for the overall conformation and dynamics of the N-terminal domain.

The more local effects of D79N are in agreement with D79's role of making contacts to the bound ATP (Prodromou et al. 1997).

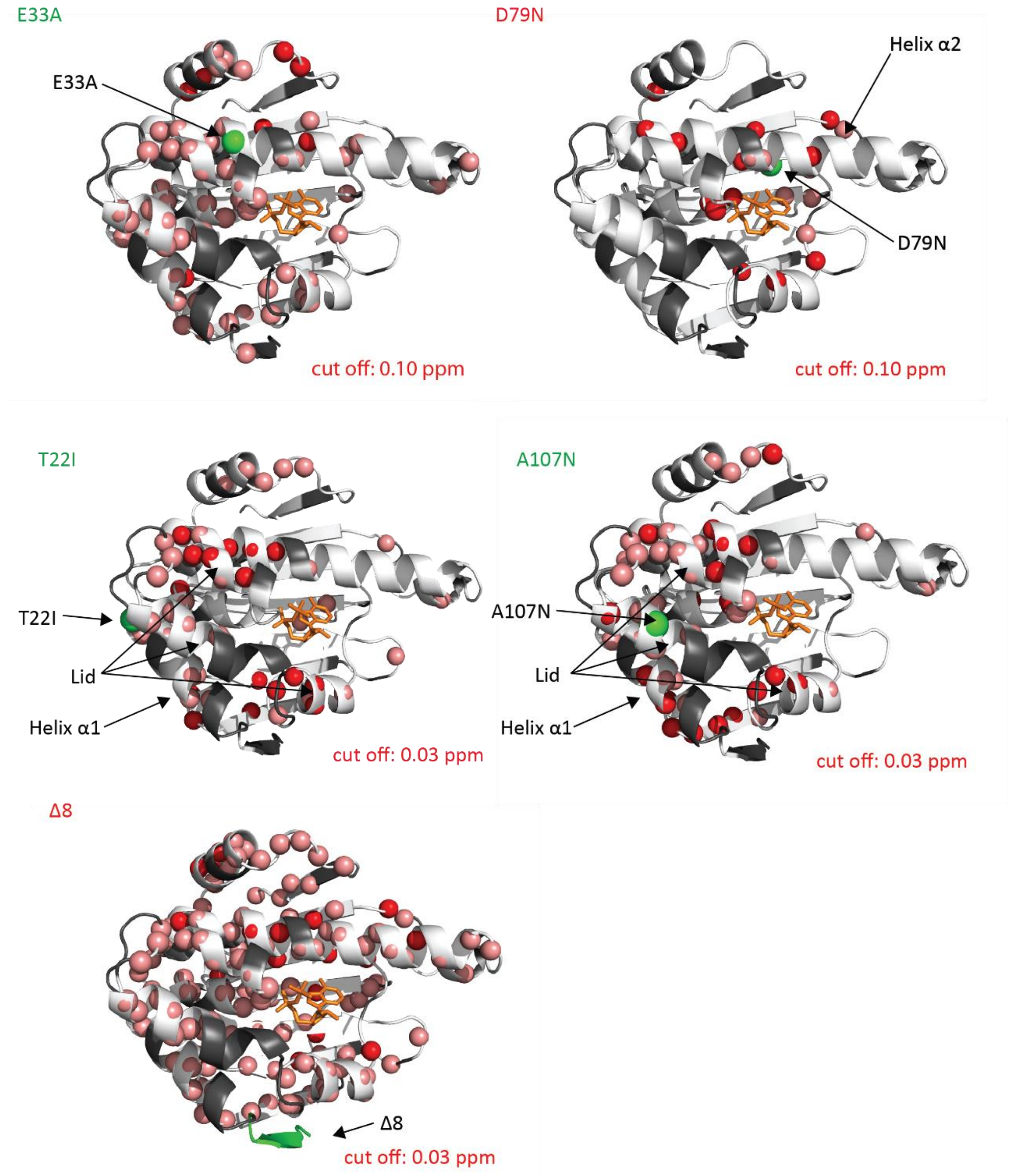

E33A
E33A
cut off: 0.10 ppm
D79N
Helix α2
D79N
cut off: 0.10 ppm
T22I
T22I
Lid
Helix α1
cut off: 0.03 ppm
A107N
A107N
Lid
Helix α1
cut off: 0.03 ppm
Δ8
Δ8
cut off: 0.03 ppm

Figure 66 NMR analysis of mutations in the NTD of Hsp90: *Chemical shift changes larger than the indicated cut-off extracted from ^1H,^{15}N-correlation spectra of the indicated variant compared to the wildtype are highlighted by red spheres. Purple spheres mark residues which could not be assigned in the mutant. Residues which are not assigned in the wildtype protein are colored grey. The labelling of the mutants is color coded. Green indicates mutants that can support viability in yeast, while red indicates those that cannot. The position of the mutation is indicated in green and bound ADP is shown in orange to mark the nucleotide binding site.*

The mutation in α-helix 1, which is involved in N-terminal dimerization, T22I and in the lid A107N interestingly show very similar differences in ^1H,^{15}N-correlation spectra to the wt. The differences in chemical shift, also far from the mutation site, indicate a complex network of allosteric interactions within in N-terminal domain.

The deletion of the first eight amino acids in the Δ8 variant of the Hsp90 NTD leads to substantial changes in chemical shift spread over the entire domain. This indicates a domain wide change in structure or domain integrity. The widespread effects fit to SAXS data that show a compaction of full length Hsp90 lacking the first eight residues compared to the wildtype (Zierer et al. 2014).

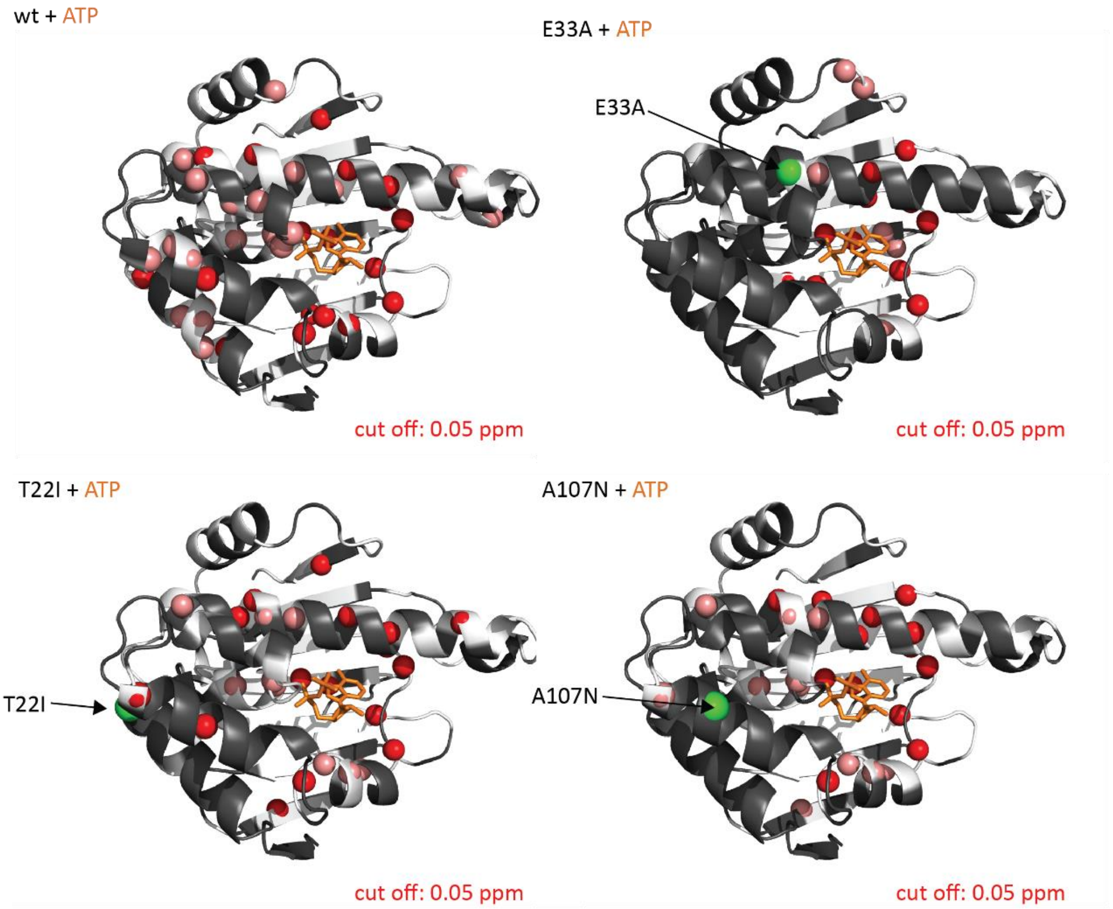


Figure 67 ATP binding of Hsp90 NTD variants: *Chemical shift changes extracted from ^1H,^{15}N-correaltion spectra of the indicate variant upon ATP binding lager than 0.05 ppm are highlighted on the structure of the NTD of Hsp90 (PDB: 1AM1) by red spheres. Salmon colored spheres indicate residues for that the peaks could not be tracked. Not assigned residues are shown in grey. Green spheres mark the mutation sites. The bound nucleotide is shown in orange to mark the nucleotide binding site.*

The addition of ATP to the NTD of Hsp90 leads to changes in chemical shift distributed through the entire domain, consistent with allosteric changes (**Figure 67**). Interestingly the ATPase deficient E33A mutant also shows changes in chemical shift, indicating that it can bind and respond to the binding of ATP. The ability to respond to the binding of ATP, might explain why E33A can support viability in yeast. The D79N mutant in contrast cannot bind nucleotide, as becomes clear from the addition of the non-hydrolysable ATP analog AMP-PNP to the Hsp90 NTD D79N variant (**Figure 68**), and therefore also cannot change its conformation in an ATP dependent manner. The two variants in helix α1 (T22I) and in the Lid (A107N) respond to ATP binding similar to the wt, which means the mutations do not hinder the normal conformational changes, which fits well to the increased ATPase activity of these variants.

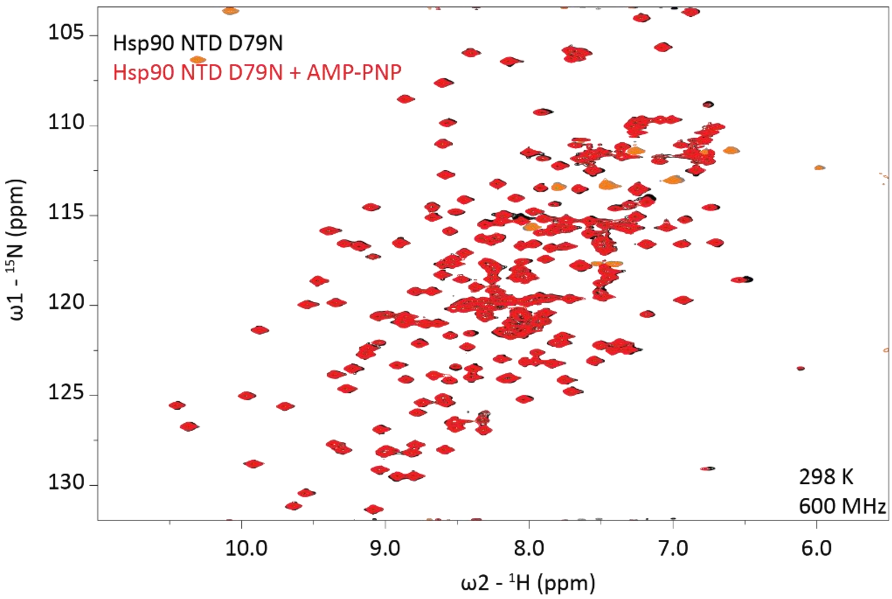


Figure 68 Hsp90 NTD D79N does not bin AMP-PNP: *Superimpositions of ^{1}H,^{15}N-correlation spectra of the Hsp90 NTD D79N variant in absence (black) and presence (red) of the non-hydrolysable ATP analog AMP-PNP.*

11 Hsp90 Discussion

11.1 Hsp90 and Hsp90-Aha1 modulators

Bettina Zierer from the group of Johannes Buchner identified a compound acting as an activator of the Hsp90 ATPase activity and its chaperone cycle by a FRET based assay. It could be shown by NMR spectroscopy that the activator binds the NTD of Hsp90 close to the ATP binding site and that the residues affected by the modulator binding do not overlap with the known Aha1 binding site. Since the binding of the co-chaperone p23 which binds the N-terminal closed state of Hsp90 is not facilitated by modulator 4, it seems to effect steps prior to the fully N-terminal dimerization (Zierer et al. 2014). The initial rearrangement of the dimerization helix α1 is a critical early step in the Hsp90 ATPase cycle (Richter et al. 2006). The modulator 4 is not able to stimulate the variant lacking the first eight amino acids, which facilitates already the initial rearrangement (Zierer et al. 2014)(Richter, Reinstein, and Buchner 2002). Therefore it is likely that modulator 4 facilitates this early rearrangements and thereby accelerates the ATPase activity. The relevance of this effect is demonstrated by a negative effect of modulator 4 on the maturation and activation of the glucocorticoid receptor, which is a known client protein of Hsp90 (Zierer et al. 2014).

Since Hsp90 has a lot of different client proteins, which are involved in various pathways and not all co-chaperones are important for all pathways, it might be possible to address a narrower range of targets by inhibiting not Hsp90 alone, but a specific Hsp90:co-chaperone complex. Sandrine Stiegler from the group of Johannes Buchner has identified an inhibitor (modulator 1) of the Hsp90:Aha1 complex using a FRET based assay. By NMR-spectroscopy using single domains of Hsp90 and Aha1 it could be shown that the inhibitor binds the NTD close or in the ATP binding site and is able to release ATP from the protein. If this release is done by competitive binding or allosterically could not be decided from the NMR spectroscopy data. Since it is known that the Hsp90-Aha1 complex is responsible for the degradation of the CFTR mutant Δ508, which is a known cause of cystic fibrosis, (X. Wang et al. 2006), the effect of the Hsp90-Aha1 modulator 1 on the stability of the Δ508 CFTR mutant was studied by Sandrine Stiegler. Indeed the presence of modulator 1 leads to a prolonged half-life of the Δ508 CFTR mutant, indicating a specific effect on the Hsp90-Aha1 activity by modulator 1 (manuscript in revision). This study is a proof of concept that it is possible to specifically inhibit Hsp90:co-chaperone complexes, with the potential benefit of more specific effects.

11.2 Importance of cycling time of the molecular chaperone Hsp90

Hsp90 undergoes conformational changes in an ATP dependent manner. Here several mutants that effect the ATPase activity are analyzed. Interestingly the ATP turnover rate is not a good indicator if these mutated variants can support viability in yeast. The strongly accelerated mutant Δ8 cannot support viability in yeast. According to NMR data the NTD of Hsp90 in the Δ8 variant has a different conformation in the apostate compared to the wildtype protein. SAXS data reveal that the full length Δ8 variant is more compact in the apostate than the wildtype and compacts further upon ATP binding. The further compaction probably corresponds to the formation of the closed-2 state, which could be shown through increased binding of the co-chaperone p23, which does not bind the closed-1 state (Zierer et al. 2016). The other striking example is the E33A variant, which is not able to hydrolyze ATP, but still supports viability. The NMR spectra in presence of ATP show that the E33A mutant can still react to ATP binding with conformational changes. The SAXS data show compaction and the p23 co-chaperone indicates that ultimately the closed-2 state is accumulated (Zierer et al. 2016). The D79N variant in contrast cannot bind nucleotides and is not able to change its conformation in an ATP dependent manner. In consequence the D79N variant cannot support viability in yeast. Finally there is the lid mutant A107N and the mutant in the N-terminal α-helix T22I. Both mutants effect the NTD in a similar way, both have slightly higher ATPase activity and both mutants can support viability. The similar changes indicate the cooperativity between the lid and the N-terminal α-helix, which is involved in dimerization. This study shows that the overall speed of the ATPase cycle is not crucial for the function of Hsp90. Instead the capability to undergo structural transitions is crucial for Hsp90 function.

12 Concluding Remarks

Studying highly dynamic systems like the RNA recognition of the splicing factor RBM10 or the molecular chaperone Hsp90 requires the use of solution based methods, since the domain rearrangement and the conformational changes upon RNA or nucleotide binding, which are crucial for the function of the two systems, are highly dynamic. For Hsp90 the importance of the structural rearrangement in the N-terminal ATP hydrolyzing domain could be shown by the analysis of mutated variants of the proteins biochemically, by SAXS and with ^{1}H,^{15}N-correlation NMR spectra in absence and presence of ATP. In case of the splicing factor RBM10 the rearrangement and compaction of the three RNA binding domains could be shown by a combination of NMR ^{15}N-Relaxation, PRE and SAXS data. NMR Spectroscopy is also a suitable method to study transient interactions or interactions with moderate affinity, like the domain-domain contacts in the apostate of RBM10 and the mapping of the binding site of the Hsp90 modulators, respectively.

13 Literature

Alberts, Bruce, Alexander Johnson, and Julian Lewis. 2007. *Molecular Biology of the Cell, W. CD-ROM*. 5 Har/Cdr edition. New York: Taylor & Francis.

Barash, Yoseph, John A. Calarco, Weijun Gao, Qun Pan, Xinchen Wang, Ofer Shai, Benjamin J. Blencowe, and Brendan J. Frey. 2010. "Deciphering the Splicing Code." *Nature* 465 (7294): 53–59. doi:10.1038/nature09000.

Battiste, John L., and Gerhard Wagner. 2000. "Utilization of Site-Directed Spin Labeling and High-Resolution Heteronuclear Nuclear Magnetic Resonance for Global Fold Determination of Large Proteins with Limited Nuclear Overhauser Effect Data." *Biochemistry* 39 (18): 5355–65. doi:10.1021/bi000060h.

Bechara, Elias G., Endre Sebestyén, Isabella Bernardis, Eduardo Eyras, and Juan Valcárcel. 2013. "RBM5, 6, and 10 Differentially Regulate NUMB Alternative Splicing to Control Cancer Cell Proliferation." *Molecular Cell* 52 (5): 720–33. doi:10.1016/j.molcel.2013.11.010.

Ben-Dov, Claudia, Britta Hartmann, Josefin Lundgren, and Juan Valcárcel. 2008. "Genome-Wide Analysis of Alternative Pre-mRNA Splicing." *Journal of Biological Chemistry* 283 (3): 1229–33. doi:10.1074/jbc.R700033200.

Berg, Jeremy M., John L. Tymoczko, and Lubert Stryer. 2010. *Biochemistry*. 7 edition. New York: W.H. Freeman & Company.

Boldon, Lauren, Fallon Laliberte, and Li Liu. 2015. "Review of the Fundamental Theories behind Small Angle X-Ray Scattering, Molecular Dynamics Simulations, and Relevant Integrated Application." *Nano Reviews* 6 (February). doi:10.3402/nano.v6.25661.

Case, David A. 1995. "Calibration of Ring-Current Effects in Proteins and Nucleic Acids." *Journal of Biomolecular NMR* 6 (4): 341–46. doi:10.1007/BF00197633.

Cavanagh, John, Wayne J. Fairbrother, Arthur G. Palmer, Nicholas J. Skelton, and Mark Rance. 2006. *Protein NMR Spectroscopy. Principles and Practice*. 2Rev ed. edition. Amsterdam ; Boston: Elsevier Ltd, Oxford.

Chen, Vincent B., W. Bryan Arendall, Jeffrey J. Headd, Daniel A. Keedy, Robert M. Immormino, Gary J. Kapral, Laura W. Murray, Jane S. Richardson, and David C. Richardson. 2010. "MolProbity: All-Atom Structure Validation for Macromolecular Crystallography." *Acta Crystallographica Section D: Biological Crystallography* 66 (Pt 1): 12–21. doi:10.1107/S0907444909042073.

Clamp, Michele, Ben Fry, Mike Kamal, Xiaohui Xie, James Cuff, Michael F. Lin, Manolis Kellis, Kerstin Lindblad-Toh, and Eric S. Lander. 2007. "Distinguishing Protein-Coding and Noncoding Genes in the Human Genome." *Proceedings of the National Academy of Sciences of the United States of America* 104 (49): 19428–33. doi:10.1073/pnas.0709013104.

Cléry, Antoine, Markus Blatter, and Frédéric H-T Allain. 2008. "RNA Recognition Motifs: Boring? Not Quite." *Current Opinion in Structural Biology*, Nucleic acids / Sequences and topology, 18 (3): 290–98. doi:10.1016/j.sbi.2008.04.002.

Collins, Katherine M., Yaroslav A. Kainov, Evangelos Christodolou, Debashish Ray, Quaid Morris, Timothy Hughes, Ian A. Taylor, Eugene V. Makeyev, and Andres Ramos. 2017. "An RRM–ZnF RNA Recognition Module Targets RBM10 to Exonic Sequences to Promote Exon Exclusion." *Nucleic Acids Research*. doi:10.1093/nar/gkx225.

Cordier, Florence, Andrew J. Dingley, and Stephan Grzesiek. 1999. "A Doublet-Separated Sensitivity-Enhanced HSQC for the Determination of Scalar and Dipolar One-Bond J-Couplings." *Journal of Biomolecular NMR* 13 (2): 175–80. doi:10.1023/A:1008301415843.

Delaglio, Frank, Stephan Grzesiek, Geerten W. Vuister, Guang Zhu, John Pfeifer, and Ad Bax. 1995. "NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes." *Journal of Biomolecular NMR* 6 (3): 277–93. doi:10.1007/BF00197809.

Güntert, P., C. Mumenthaler, and K. Wüthrich. 1997. "Torsion Angle Dynamics for NMR Structure Calculation with the New Program Dyana11Edited by P. E. Wright." *Journal of Molecular Biology* 273 (1): 283–98. doi:10.1006/jmbi.1997.1284.

Güntert, Peter, and Lena Buchner. 2015. "Combined Automated NOE Assignment and Structure Calculation with CYANA." *Journal of Biomolecular NMR* 62 (4): 453–71. doi:10.1007/s10858-015-9924-9.

Hansen, Mark R., Luciano Mueller, and Arthur Pardi. 1998. "Tunable Alignment of Macromolecules by Filamentous Phage Yields Dipolar Coupling Interactions." *Nature Structural & Molecular Biology* 5 (12): 1065–74. doi:10.1038/4176.

Hernández, Jordi, Elias Bechara, Doerte Schlesinger, Javier Delgado, Luis Serrano, and Juan Valcárcel. 2016. "Tumor Suppressor Properties of the Splicing Regulatory Factor RBM10." *RNA Biology* 13 (4): 466–72. doi:10.1080/15476286.2016.1144004.

Hertel, Klemens J. 2008. "Combinatorial Control of Exon Recognition." *Journal of Biological Chemistry* 283 (3): 1211–15. doi:10.1074/jbc.R700035200.

Huang, Jie-rong, Lisa R. Warner, Carolina Sanchez, Frank Gabel, Tobias Madl, Cameron D. Mackereth, Michael Sattler, and Martin Blackledge. 2014. "Transient Electrostatic Interactions Dominate the Conformational Equilibrium Sampled by Multidomain Splicing Factor U2AF65: A Combined NMR and SAXS Study." *Journal of the American Chemical Society* 136 (19): 7068–76. doi:10.1021/ja502030n.

Imielinski, Marcin, Alice H. Berger, Peter S. Hammerman, Bryan Hernandez, Trevor J. Pugh, Eran Hodis, Jeonghee Cho, et al. 2012. "Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing." *Cell* 150 (6): 1107–20. doi:10.1016/j.cell.2012.08.029.

Johnston, Jennifer J., Jamie K. Teer, Praveen F. Cherukuri, Nancy F. Hansen, Stacie K. Loftus, Karen Chong, James C. Mullikin, and Leslie G. Biesecker. 2010. "Massively Parallel Sequencing of Exons on the X Chromosome Identifies RBM10 as the Gene That Causes a Syndromic Form of Cleft Palate." *The American Journal of Human Genetics* 86 (5): 743–48. doi:10.1016/j.ajhg.2010.04.007.

Keeler, James. 2010. *Understanding NMR Spectroscopy*. 2 edition. Chichester, U.K: John Wiley & Sons.

Koradi, Reto, Martin Billeter, and Kurt Wüthrich. 1996. "MOLMOL: A Program for Display and Analysis of Macromolecular Structures." *Journal of Molecular Graphics* 14 (1): 51–55. doi:10.1016/0263-7855(96)00009-4.

Kornblihtt, Alberto R., Ignacio E. Schor, Mariano Alló, Gwendal Dujardin, Ezequiel Petrillo, and Manuel J. Muñoz. 2013. "Alternative Splicing: A Pivotal Step between Eukaryotic Transcription and Translation." *Nature Reviews Molecular Cell Biology* 14 (3): 153–65. doi:10.1038/nrm3525.

Maaskola, Jonas, and Nikolaus Rajewsky. 2014. "Binding Site Discovery from Nucleic Acid Sequences by Discriminative Learning of Hidden Markov Models." *Nucleic Acids Research* 42 (21): 12995–11. doi:10.1093/nar/gku1083.

Mackereth, Cameron D., Tobias Madl, Sophie Bonnal, Bernd Simon, Katia Zanier, Alexander Gasch, Vladimir Rybin, Juan Valcárcel, and Michael Sattler. 2011. "Multi-Domain Conformational Selection Underlies Pre-mRNA Splicing Regulation by U2AF." *Nature* 475 (7356): 408–11. doi:10.1038/nature10171.

Mackereth, Cameron D, and Michael Sattler. 2012. "Dynamics in Multi-Domain Protein Recognition of RNA." *Current Opinion in Structural Biology*, Nucleic acids/Sequences and topology, 22 (3): 287–96. doi:10.1016/j.sbi.2012.03.013.

Marcu, M. G., A. Chadli, I. Bouhouche, M. Catelli, and L. M. Neckers. 2000. "The Heat Shock Protein 90 Antagonist Novobiocin Interacts with a Previously Unrecognized ATP-Binding Domain in the Carboxyl Terminus of the Chaperone." *The Journal of Biological Chemistry* 275 (47): 37181–86. doi:10.1074/jbc.M003701200.

Meyer, Philippe, Chrisostomos Prodromou, Chunyan Liao, Bin Hu, S. Mark Roe, Cara K. Vaughan, Ignacija Vlasic, Barry Panaretou, Peter W. Piper, and Laurence H. Pearl. 2004. "Structural Basis for Recruitment of the ATPase Activator Aha1 to the Hsp90 Chaperone Machinery." *The EMBO Journal* 23 (3): 511–19. doi:10.1038/sj.emboj.7600060.

Misquitta-Ali, Christine M., Edith Cheng, Dave O'Hanlon, Ni Liu, C. Jane McGlade, Ming Sound Tsao, and Benjamin J. Blencowe. 2011. "Global Profiling and Molecular Characterization of

Alternative Splicing Events Misregulated in Lung Cancer." *Molecular and Cellular Biology* 31 (1): 138–50. doi:10.1128/MCB.00709-10.

Nguyen, Cuong D., Robyn E. Mansfield, Wilfred Leung, Paula M. Vaz, Fionna E. Loughlin, Richard P. Grant, and Joel P. Mackay. 2011. "Characterization of a Family of RanBP2-Type Zinc Fingers That Can Recognize Single-Stranded RNA." *Journal of Molecular Biology* 407 (2): 273–83. doi:10.1016/j.jmb.2010.12.041.

Panaretou, B, C Prodromou, S M Roe, R O'Brien, J E Ladbury, P W Piper, and L H Pearl. 1998. "ATP Binding and Hydrolysis Are Essential to the Function of the Hsp90 Molecular Chaperone in Vivo." *The EMBO Journal* 17 (16): 4829–36. doi:10.1093/emboj/17.16.4829.

Panaretou, Barry, Giuliano Siligardi, Philippe Meyer, Alison Maloney, Janis K. Sullivan, Shradha Singh, Stefan H. Millson, et al. 2002. "Activation of the ATPase Activity of Hsp90 by the Stress-Regulated Cochaperone Aha1." *Molecular Cell* 10 (6): 1307–18. doi:10.1016/S1097-2765(02)00785-2.

Pennisi, Elizabeth. 2003. "A Low Number Wins the GeneSweep Pool." *Science* 300 (5625): 1484–1484. doi:10.1126/science.300.5625.1484b.

Prestegard, J. H., C. M. Bougault, and A. I. Kishore. 2004. "Residual Dipolar Couplings in Structure Determination of Biomolecules." *Chemical Reviews* 104 (8): 3519–40. doi:10.1021/cr030419i.

Prodromou, Chrisostomos, Barry Panaretou, Shahzad Chohan, Giuliano Siligardi, Ronan O'Brien, John E. Ladbury, S.Mark Roe, Peter W. Piper, and Laurence H. Pearl. 2000. "The ATPase Cycle of Hsp90 Drives a Molecular 'clamp' via Transient Dimerization of the N-Terminal Domains." *The EMBO Journal* 19 (16): 4383–92. doi:10.1093/emboj/19.16.4383.

Prodromou, Chrisostomos, S. Mark Roe, Ronan O'Brien, John E. Ladbury, Peter W. Piper, and Laurence H. Pearl. 1997. "Identification and Structural Characterization of the ATP/ADP-Binding Site in the Hsp90 Molecular Chaperone." *Cell* 90 (1): 65–75. doi:10.1016/S0092-8674(00)80314-1.

Retzlaff, Marco, Franz Hagn, Lars Mitschke, Martin Hessling, Frederik Gugel, Horst Kessler, Klaus Richter, and Johannes Buchner. 2010. "Asymmetric Activation of the Hsp90 Dimer by Its Cochaperone Aha1." *Molecular Cell* 37 (3): 344–54. doi:10.1016/j.molcel.2010.01.006.

Richter, Klaus, Sandra Moser, Franz Hagn, Rainer Friedrich, Otmar Hainzl, Markus Heller, Sandra Schlee, Horst Kessler, Jochen Reinstein, and Johannes Buchner. 2006. "Intrinsic Inhibition of the Hsp90 ATPase Activity." *Journal of Biological Chemistry* 281 (16): 11301–11. doi:10.1074/jbc.M510142200.

Richter, Klaus, Jochen Reinstein, and Johannes Buchner. 2002. "N-Terminal Residues Regulate the Catalytic Efficiency of the Hsp90 ATPase Cycle." *Journal of Biological Chemistry* 277 (47): 44905–10. doi:10.1074/jbc.M208457200.

Rieping, Wolfgang, Michael Habeck, Benjamin Bardiaux, Aymeric Bernard, Thérèse E. Malliavin, and Michael Nilges. 2007. "ARIA2: Automated NOE Assignment and Data Integration in NMR Structure Calculation." *Bioinformatics* 23 (3): 381–82. doi:10.1093/bioinformatics/btl589.

Roe, S. Mark, Chrisostomos Prodromou, Ronan O'Brien, John E. Ladbury, Peter W. Piper, and Laurence H. Pearl. 1999. "Structural Basis for Inhibition of the Hsp90 Molecular Chaperone by the Antitumor Antibiotics Radicicol and Geldanamycin." *Journal of Medicinal Chemistry* 42 (2): 260–66. doi:10.1021/jm980403y.

Rückert, Markus, and Gottfried Otting. 2000. "Alignment of Biological Macromolecules in Novel Nonionic Liquid Crystalline Media for NMR Experiments." *Journal of the American Chemical Society* 122 (32): 7793–97. doi:10.1021/ja001068h.

Sattler, Michael, Jürgen Schleucher, and Christian Griesinger. 1999. "Heteronuclear Multidimensional NMR Experiments for the Structure Determination of Proteins in Solution Employing Pulsed Field Gradients." *Progress in Nuclear Magnetic Resonance Spectroscopy* 34 (2): 93–158. doi:10.1016/S0079-6565(98)00025-9.

Schnablegger, Heimo, and Yashveer Singh. 2011. "The SAXS Guide: Getting Acquainted with the Principles." *Anton Paar GmbH, Austria*.

Schopf, Florian H., Maximilian M. Biebl, and Johannes Buchner. 2017. “The HSP90 Chaperone Machinery.” *Nature Reviews Molecular Cell Biology* 18 (6): 345–60. doi:10.1038/nrm.2017.20.

Shen, Yang, Frank Delaglio, Gabriel Cornilescu, and Ad Bax. 2009. “TALOS+: A Hybrid Method for Predicting Protein Backbone Torsion Angles from NMR Chemical Shifts.” *Journal of Biomolecular NMR* 44 (4): 213–23. doi:10.1007/s10858-009-9333-z.

Simon, Bernd, Tobias Madl, Cameron D. Mackereth, Michael Nilges, and Michael Sattler. 2010. “An Efficient Protocol for NMR-Spectroscopy-Based Structure Determination of Protein Complexes in Solution.” *Angewandte Chemie International Edition* 49 (11): 1967–70. doi:10.1002/anie.200906147.

Stolc, Viktor, Zareen Gauhar, Christopher Mason, Gabor Halasz, Marinus F. van Batenburg, Scott A. Rifkin, Sujun Hua, et al. 2004. “A Gene Expression Map for the Euchromatic Genome of Drosophila Melanogaster.” *Science* 306 (5696): 655–60. doi:10.1126/science.1101312.

Sutherland, Leslie C., Nina D. Rintala-Maki, Ryan D. White, and Cory D. Morin. 2005. “RNA Binding Motif (RBM) Proteins: A Novel Family of Apoptosis Modulators?” *Journal of Cellular Biochemistry* 94 (1): 5–24. doi:10.1002/jcb.20204.

Tessier, Sarah J., Julie J. Loiselle, Anne McBain, Celine Pullen, Benjamin W. Koenderink, Justin G. Roy, and Leslie C. Sutherland. 2015. “Insight into the Role of Alternative Splicing within the RBM10v1 Exon 10 Tandem Donor Site.” *BMC Research Notes* 8: 46. doi:10.1186/s13104-015-0983-5.

Tolman, J. R., J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard. 1995. “Nuclear Magnetic Dipole Interactions in Field-Oriented Proteins: Information for Structure Determination in Solution.” *Proceedings of the National Academy of Sciences* 92 (20): 9279–83.

Vranken, Wim F., Wayne Boucher, Tim J. Stevens, Rasmus H. Fogh, Anne Pajon, Miguel Llinas, Eldon L. Ulrich, John L. Markley, John Ionides, and Ernest D. Laue. 2005. “The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline.” *Proteins: Structure, Function, and Bioinformatics* 59 (4): 687–96. doi:10.1002/prot.20449.

Wang, Xiaodong, John Venable, Paul LaPointe, Darren M. Hutt, Atanas V. Koulov, Judith Coppinger, Cemal Gurkan, et al. 2006. “Hsp90 Cochaperone Aha1 Downregulation Rescues Misfolding of CFTR in Cystic Fibrosis.” *Cell* 127 (4): 803–15. doi:10.1016/j.cell.2006.09.043.

Wang, Yongbo, Andreas Gogol-Döring, Hao Hu, Sebastian Fröhler, Yunxia Ma, Marvin Jens, Jonas Maaskola, et al. 2013. “Integrative Analysis Revealed the Molecular Mechanism Underlying RBM10-mediated Splicing Regulation.” *EMBO Molecular Medicine* 5 (9): 1431–42. doi:10.1002/emmm.201302663.

Westhoff, Britta, Ivan N. Colaluca, Giovanni D’Ario, Maddalena Donzelli, Daniela Tosoni, Sara Volorio, Giuseppe Pelosi, et al. 2009. “Alterations of the Notch Pathway in Lung Cancer.” *Proceedings of the National Academy of Sciences* 106 (52): 22293–98. doi:10.1073/pnas.0907781106.

Will, Cindy L., and Reinhard Lührmann. 2011. “Spliceosome Structure and Function.” *Cold Spring Harbor Perspectives in Biology* 3 (7): a003707. doi:10.1101/cshperspect.a003707.

Yamazaki, Toshio, Julie D. Forman-Kay, and Lewis E. Kay. 1993. “Two-Dimensional NMR Experiments for Correlating Carbon-13.beta. and Proton.delta./.epsilon. Chemical Shifts of Aromatic Residues in 13C-Labeled Proteins via Scalar Couplings.” *Journal of the American Chemical Society* 115 (23): 11054–55. doi:10.1021/ja00076a099.

Zhao, Jiawei, Yue Sun, Yin Huang, Fan Song, Zengshu Huang, Yufang Bao, Ji Zuo, et al. 2017. “Functional Analysis Reveals That *RBM10* Mutations Contribute to Lung Adenocarcinoma Pathogenesis by Deregulating Splicing.” *Scientific Reports* 7 (January): srep40488. doi:10.1038/srep40488.

Zhao, Rongmin, Elisa Leung, Stefan Grüner, Matthieu Schapira, and Walid A. Houry. 2010. “Tamoxifen Enhances the Hsp90 Molecular Chaperone ATPase Activity.” *PLOS ONE* 5 (4): e9934. doi:10.1371/journal.pone.0009934.

Zierer, Bettina K., Martin Rübbelke, Franziska Tippel, Tobias Madl, Florian H. Schopf, Daniel A. Rutz, Klaus Richter, Michael Sattler, and Johannes Buchner. 2016. “Importance of Cycle Timing for the Function of the Molecular Chaperone Hsp90." *Nature Structural & Molecular Biology* 23 (11): 1020–28. doi:10.1038/nsmb.3305.

Zierer, Bettina K., Matthias Weiwad, Martin Rübbelke, Lee Freiburger, Gunter Fischer, Oliver R. Lorenz, Michael Sattler, Klaus Richter, and Johannes Buchner. 2014. "Artificial Accelerators of the Molecular Chaperone Hsp90 Facilitate Rate-Limiting Conformational Transitions." *Angewandte Chemie International Edition* 53 (45): 12257–62. doi:10.1002/anie.201406578.

Zwahlen, Catherine, Pascale Legault, Sébastien J. F. Vincent, Jack Greenblatt, Robert Konrat, and Lewis E. Kay. 1997. "Methods for Measurement of Intermolecular NOEs by Multinuclear NMR Spectroscopy: Application to a Bacteriophage λ N-Peptide/boxB RNA Complex." *Journal of the American Chemical Society* 119 (29): 6711–21. doi:10.1021/ja970224q.

14 Acknowledgements

First of all I want to thank Michael Sattler for giving me the opportunity to work in his group and for first introducing me to the world of NMR during his lectures in the Biochemistry Bachelor and Master programs.

Special thanks goes to my teachers from who I learned the wet lab work and NMR spectroscopy. I want to thank Hamed Kooshapur, who supervised during my Bachelor thesis. From him I learned all the important practical things about protein expression and purification and he guided me in my first practical application of NMR spectroscopy. Gabriel Zoldak further helped me to gain experience in the field of biochemistry during my internship in the group of Matthias Rief. Than during my master studies I had the opportunity to deepen my knowledge about NMR in an internship in the group of Bernd Reif. I thank Riddhiman Sarkar who taught me how to read and write pulse sequences for NMR spectroscopy. My special gratitude goes to Lisa Warner, who supervised me during my Master thesis and supported me in the beginning of my doctorate. I want to thank her for all the NMR expertise she transferred to me and for all the great discussions we had in the office. It was very important for me to have her, who always showed interest in my progress and helped me to solve problems.

I am thankful for the help of Arie Geerlof. He not only performed the SLS measurements presented in this work, but also helped me with his advice with cloning, protein expression and protein purification. My thanks go also to Astrid Lauxen, who together with Arie kept the lab in a good shape and thereby made the experimental work of this thesis possible.

I want to thank Ralf Stehle, who performed the SAXS measurements presented in this thesis and helped me with IT problems. I also want to thank Rainer Häßner, for maintaining the NMR-spectrometers and the IT-system. For support with the NMR-spectrometers I thank Sam Asami and Gerd Gemmecker.

I am grateful for my collaborators from the group of Johannes Buchner, with whom I worked on the three Hsp90 projects presented in this thesis: Bettina Zierer for the work on the Hsp90 activator and the study of Hsp90 mutants, Franziska Tippel for the work on the Hsp90 mutants, Sandrine Stiegler for the work on the Hsp90:Aha1 modulators and of course Michael Sattler and Johannes Buchner who made this collaborations possible.

My thanks for the discussions goes to my collaborators in the RBM10 project from the CRG in Barcelona: Juan Valcárcel Juárez, Jordi Hernández Ribera, Sophie Bonnal.

I thank Laurin Waldmann, who helped me with the acquisition and evaluation of the relaxation data of the RRM1-ZF1 construct, while I supervised him during his bachelor thesis.

I want to thank all my colleagues from the Sattler lab. Especially Komal Soni, who works on the RBM10 homolog RBM5, for the great discussions about our related projects. I thank Grzegorz Popowicz and Ashish Kawale for the good time we had in our office.

I want to thank my parents for their support and their love throughout my entire live. And finally I express my biggest gratitude to my partner in life Carolina, who I got to know and love during the time of our doctorate. I am grateful to Carolina for complementing me, enriching my life and making me a better person.

16 Appendix

16.1 Buffers and Media

Table 12: M9 Minimal Medium

100 mL	M9 medium (10X)
10 mL	Trace elements solution
2g	Glucose
1g	NH_4Cl
1 mL	1 M $MgSO_4$
1 mL	0.1 M $CaCl_2$
1 mL	Biotin (1 mg/mL)
1 mL	Thiamin (1 mg/mL)
1 mL	Kanamycin (30 mg/mL)

Table 13: 10x M9 Medium - Ingredients per 1 L

67.8 g	Na_2HPO_4 (anhydrous)
30 g	KH_2PO_4
5 g	NaCl

Table 14: 100x Trace Elements Solution - Ingredients per 1 L

5 g	EDTA
0.83 g	$FeCl_3 \cdot 6\ H_2O$
84 mg	$ZnCl_2$
13 mg	$CuCl_2 \cdot 2\ H_2O$
10 mg	$CoCl_2 \cdot 6\ H_2O$
10 mg	H_3BO_3
1.6 mg	$MnCl_2 \cdot 6\ H_2O$

Table 15: Lysis Buffer

50 mM	Tris
1 M	NaCl
5 mM	β-mercaptoethanol
pH	8.0

Table 16: Washing Buffer

50 mM	Tris
50 mM	NaCl
5 mM	β-mercaptoethanol
pH	8.0

Table 17: Elution Buffer

50 mM	Tris
50 mM	NaCl
250 mM	Imidazole
5 mM	β-mercaptoethanol
pH	8.0

Table 18: TEV cleavage buffer

20 mM	Tris
300 mM	NaCl
5 mM	β-mercaptoethanol
pH	8.0

Table 19: NMR Buffer

50 mM	Na_2HPO_4 / NaH_2PO_4
50 mM	NaCl
1 mM	DTT
pH	6.5

Table 20: IPSL-dialysis buffer

20 mM	Tris
50 mM	NaCl
pH	8.0

Table 21: IPSL-reaction buffer

1 M	Tris
50 mM	NaCl
pH	8.0

Table 22: ITC buffer

50 mM	Na_2HPO_4 / NaH_2PO_4
50 mM	NaCl
5 mM	β-mercaptoethanol
pH	6.5

16.2 RBM 10 Constructs

Table 23: RBM10 constructs

Construct	**From residue**	**To residue**	**Artificial residues**	**Number of residues**	**Vector**
RRM1-ZF1	125	242	GAMA	122	pET-M11
RRM2 -V354	300	383	GAM	87	pET-M11
RRM2 +V354	300	384	GAM	88	pET-M11
RRM1-ZF1-RRM2 -V354	125	383	GAM	262	pET-M11
RRM1-ZF1-RRM2 +V354	125	384	GAM	263	pET-M11

16.3 Abbreviations:

AMP-PNP	Adenylyl-imidodiphosphate
ATP	Adenosine triphosphate
CD	C-terminal domain of Aha1
CFTR	Cystic fibrosis transmembrane conductance regulator
CLIP	Cross-linking immunoprecipitation
CLIP-seq	High-throughput sequencing of RNA isolated by CLIP
CPMG	Carr-Purcell-Meiboom-Gill sequence
CTD	C-terminal domain of Hsp90
DNAse	Deoxyribonuclease
DTT	Dithiothreitol
FRET	Förster resonance energy transfer
hnRNP	Heterogeneous nuclear RNP
Hsp	Heat shock protein
HSQC	Heteronuclear single quantum correlation
IMAC	Immobilized metal affinity chromatography

INEPT	Insensitive nuclei enhanced by polarization transfer
IPSL	3-(2-Iodoacetamido)-PROXYL
ITC	Isothermal titration calorimetry
MD	Middle domain of Hsp90
mRNA	Messenger RNA
ND	N-terminal domain of Aha1
NEB	New England Biolabs
NMR	Nuclear magnetic resonance
nOe	Nuclear Overhauser effect
NOESY	Nuclear Overhauser effect spectroscopy
NTD	N-terminal domain of Hsp90
PAR-CLIP	photoactivatable ribonucleoside-enhanced CLIP
PCR	Polymerase chain reaction
PDB	Protein data base
PFU	Pyrococcus furiosus
PRE	Paramagnetic relaxation enhancement
RBM	RNA binding motif
RDC	Residual dipolar coupling
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RRM	RNA recognition motif
SAXS	Small angle X-ray scattering
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SLS	Static light scattering
snRNA	Small nuclear RNA
SR-proteins	Serine/arginine rich proteins
TARP	Talipes equinovarus, Atrial septal defect, Robin sequence, and Persistent left superior vena cava
TOCSY	Total correlated spectroscopy
Tris	Tris(hydroxymethyl)aminomethane
wt	Wild type
ZF	Zinc finger

16.4 List of Figures

Figure 1 Splicing reaction 19
Figure 2 Spliceosome assembly 20
Figure 3 Schematic of regulation of alternative splicing. 22
Figure 4 Domain architecture of the homolog proteins RBM 5, 6 and 10. 23
Figure 5 RMB10 isoforms are generated through splicing 23
Figure 6 Effects of RBM10 on the alternative splicing of the NUMB gene and Notch signaling. 25
Figure 7: Restraints for structure calculation for RBM10 RRM1-ZF-RRM2 26
Figure 8 Backbone assignment experiments 34
Figure 9 Experiments for aliphatic side chain assignment 36
Figure 10 Experiments for aromatic side chain assignment 36
Figure 11 Pulse sequence for 2D-NOESY experiment 39
Figure 12 Schematic for 2D-NOESY spectrum 40
Figure 13 Schematic of an isotope filtered NOESY experiment 40
Figure 14 Assignment of RRM1-ZF1 57
Figure 15 Unfolding of ZF1 with EDTA 58
Figure 16: RBM10 RRM1-ZF1 NMR structure 59
Figure 17: Relaxation data RBM10 RRM1-ZF-1 60
Figure 18 NMR titration of RBM10 RRM1-ZF1 with NUMB derived 12-mer-RNA oligo 62
Figure 19 ITC-data for RRM1-ZF1 and RRM1 with NUMB derived 12-mer RNA oligo 64
Figure 20 Small angle light scattering data for RBM10 RRM-ZF1 RNA complex 64
Figure 21: Relaxation data forRBM10 RRM1-ZF1 NUMB RNA complex 66
Figure 22 Correlation times of RRM1-ZF1 domains free and bound to RNA 67
Figure 23 NUMB-RNA-TOCSY spectra with and without RBM10 RRM1-ZF1 68
Figure 24 ω1-filtered NOESY spectrum of the NUMB derived RNA 12-mer oligo saturated with RBM10 RRM1-ZF1 69
Figure 25 Assignment of RBM10 RRM2 70
Figure 26 RBM10 RRM2 NMR structure 71
Figure 27: NMR titration of RBM10 RRM2 with NUMB derived 12-mer-RNA oligo 72
Figure 28 ITC-data for RRM2 with NUMB derived 12-mer RNA oligo 74
Figure 29 Three-domain construct vs. RRM1-ZF1/RRM2. 75
Figure 30 RBM10 RRM-Z1 and RRM2 linker constructs 76
Figure 31 Relaxation data for RBM10 RRM1-ZF1-RRM2: 77
Figure 32 Average domain correlation times per domain for RRM1-ZF1-RRM2: 78

Figure 33 Small angle light scattering of RBM10 RRM-ZF1-RRM2 RNA complex................................ 79
Figure 34: NMR titration of RBM10 RRM1-ZF1-RRM2 with NUMB derived 12-mer-RNA oligo 80
Figure 35 ITC data RRM1-ZF1-RRM2 81
Figure 36 Relaxation data RBM10 RRM1-ZF1-RRM2 NUMB RNA complex 82
Figure 37 Correlation times of RRM1-ZF1-RRM2 domains free and bound to RNA 83
Figure 38 NUMB-RNA-TOCSY spectra with RBM10 RRM1-ZF1-RRM2 83
Figure 39 ω1-filtered NOESY spectrum of NUMB derived RNA 12-mer oligo saturated with RBM10 RRM1-ZF1-RRM2. 84
Figure 40 SAXS data for RBM10 RRM1-ZF1-RRM2. 85
Figure 41 Positions of cysteine mutants for PRE measurements... 87
Figure 42 PRE data for RBM10 RRM1-ZF1-RRM2 with spin labels in RRM1 88
Figure 43 PRE ratios of RBM10 RRM1-ZF1-RRM2 with spin label in position 143 89
Figure 44 PRE data for RBM10 RRM1-ZF1-RRM2 with the spin label in ZF1....................................... 90
Figure 45 PRE data for RBM10 RRM1-ZF1-RRM2 with spin labels in RRM2 91
Figure 46 ^{1}H,^{15}N-correlation spectra of the RRM2 + V354 isoform.. 92
Figure 47 RBM10 RRM2 +V354 NMR structure 93
Figure 48 RBM10 RRM2 +/- V354 comparison.. 94
Figure 49 ^{1}H,^{15}N-correlation spectra of the RRM2 + V354 isoform 96
Figure 50 RNA binding of RBM10 RRM2 +V354 isoform and V354E mutant....................................... 97
Figure 51: ITC data RRM1-ZF1-RRM2 +V354 isoform and V354E mutant... 98
Figure 52 -V354 isoform and V354E mutant in context of RRM1-ZF1-RRM2 99
Figure 53 RNA binding of RBM10 RRM1-ZF1 101
Figure 54 Comparison of RRM1-ZF1-RRM2 RNA binding... 102
Figure 55 RNA binding of RBM10 RRM1-ZF1-RRM2 103
Figure 56 Architecture of the Hsp90 homodimer 107
Figure 57 Architecture of the Hsp90:Aha1 complex 108
Figure 58: Identification of Hsp90 activator modulator 4.. 113
Figure 59 NMR analysis of the interaction of Hsp90 NTD with modulator 4 114
Figure 60 Mapping the binding site of modulator 4 115
Figure 61 NMR analysis of the interaction of Hsp90 MD with modulator 4 116
Figure 62 Identification of Hsp90-Aha1 modulator 1... 117
Figure 63 NMR analysis of modulator 1 binding by the NTD of Hsp90... 118
Figure 64 Testing Hsp90 MD and Aha1 domains for modulator 1 binding.. 119
Figure 65 ATPase activity and of Hsp90 mutants does not correlate with the survival in yeast 120

Figure 66 NMR analysis of mutations in the NTD of Hsp90 ... 121
Figure 67 ATP binding of Hsp90 NTD variants ... 122
Figure 68 Hsp90 NTD D79N does not bin AMP-PNP ... 123

16.5 List of Tables

Table 1: Ingredients used for PCR ... 47
Table 2: Temperature program used for PCR ... 47
Table 3: Temperature program used for quick change PCR ... 48
Table 4: Protein concentrations and Protein:RNA ratios used for NMR Titrations ... 51
Table 5: Experimental parameters of RNA TOCSY experiments ... 52
Table 6: Experimental parameters of ω1-filtered 2D-NOESY experiments ... 53
Table 7: Delays used for ^{15}N relaxation experiments ... 53
Table 8: Constructs and Concentrations used in ITC experiments ... 55
Table 9: Structure calculation statistics for the water refined RBM10 RRM1-ZF1 structure: ... 59
Table 10 Structure calculation statistics for water refined RBM10 RRM2 structure: ... 71
Table 11: Structure calculation statistics for RBM10 RRM2 +V354: ... 94
Table 12: M9 Minimal Medium ... 137
Table 13: 10x M9 Medium - Ingredients per 1 L ... 137
Table 14: 100x Trace Elements Solution - Ingredients per 1 L ... 137
Table 15: Lysis Buffer ... 137
Table 16: Washing Buffer ... 138
Table 17: Elution Buffer ... 138
Table 18: TEV cleavage buffer ... 138
Table 19: NMR Buffer ... 138
Table 20: IPSL-dialysis buffer ... 138
Table 21: IPSL-reaction buffer ... 138
Table 22: ITC buffer ... 139
Table 23: RBM10 constructs ... 139

17 Lebenslauf

Name: Martin Rübbelke

Adresse: Leutkircher Str. 9
80939 München

Geburtsdatum: 25.08.1987

Geburtsort: Gütersloh

Ausbildung:

seit 2013	Doktorand am Lehrstuhl für biomolekulare NMR-Spektroskopie, Technische Universität München.
2011-2013	Technische Universität München, Master of Science, Biochemie
2008-2011	Technische Universität München, Bachelor of Science, Biochemie
1998-2007	Marienschule Lippstadt, Gymnasium, Abitur

Preise:

2014	Jürgen Manchot Studienpreis
2016	Young Investigator Award, ICMRBS 2016, Kyoto, Japan

Ausgewählte Konferenzen und Workshops:

2013	35th FGMR Discussion Meeting and Joint Conference of the German, Italian and Slovenian Magnetic Resonance Societies., Fraueninsel Chiemsee
2014	8th NMR Winter Retreat, Parpan
2014	EMBO Practical Course: Multidimensional NMR in Structure Biology, Joachimsthal
2015	EMBO Practical Course: Structure, dynamics and function of biomacromolecules by NMR, Garching bei München
2016	International Conference on Magnetic Resonance in Biological Systems, Kyoto, Japan

Veröffentlichungen:

2014 Amo, Juan Miguel Lopez del, Vipin Agarwal, Riddhiman Sarkar, Justin Porter, Sam Asami, **Martin Rübbelke**, Uwe Fink, Yi Xue, Oliver F. Lange, and Bernd Reif. "Site-Specific Analysis of Heteronuclear Overhauser Effects in Microcrystalline Proteins." *Journal of Biomolecular NMR* 59, no. 4 (August 1, 2014): 241–49. doi:10.1007/s10858-014-9843-1.

2014 Zierer, Bettina K., Matthias Weiwad, **Martin Rübbelke**, Lee Freiburger, Gunter Fischer, Oliver R. Lorenz, Michael Sattler, Klaus Richter, and Johannes Buchner. "Artificial Accelerators of the Molecular Chaperone Hsp90 Facilitate Rate-Limiting Conformational Transitions." *Angewandte Chemie International Edition* 53, no. 45 (November 3, 2014): 12257–62. doi:10.1002/anie.201406578.

2016 Dallmann, Andre, Alexander V. Beribisky, Felix Gnerlich, **Martin Rübbelke**, Stefan Schiesser, Thomas Carell, and Michael Sattler. "Site-Specific Isotope-Labeling of Inosine Phosphoramidites and NMR Analysis of an Inosine-Containing RNA Duplex." Chemistry – A European Journal 22, no. 43 (October 17, 2016): 15350–59. doi:10.1002/chem.201602784.

2016 Zierer, Bettina K.*, **Martin Rübbelke***, Franziska Tippel*, Tobias Madl, Florian H. Schopf, Daniel A. Rutz, Klaus Richter, Michael Sattler, and Johannes Buchner. "Importance of Cycle Timing for the Function of the Molecular Chaperone Hsp90." Nature Structural & Molecular Biology 23, no. 11 (November 2016): 1020–28. doi:10.1038/nsmb.3305.

*contributed equally

Sprachen: Deutsch (Muttersprache), Englisch (fließend)