

**Biochemical and structural characterization of the Human Staufen1
protein in complex with the physiological mRNA target *ARF1***

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Lina María Bandholz-Cajamarca
aus Bogotá, Kolumbien

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 09.12.2019

Dekan

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Ass. Prof. Dr. Fulvia Bono

2. Berichterstatter:

Prof. Dr. Thilo Stehle

Preface

Parts of this thesis have been published:

Lazzaretti, D., **Bandholz-Cajamarca, L.**, Emmerich, C., Schaaf, K., Basquin, C., Irion, U. and Bono, F. (2018). The crystal structure of staufen1 in complex with a physiological RNA sheds light on substrate selectivity, *Life Science Alliance* 1(5): e201800187.

Contents

Abstract	1
Zusammenfassung	2
1 Introduction	5
1.1 Post-transcriptional gene regulation	5
1.2 RNA binding proteins	6
1.3 Structural classification of RNA binding proteins	7
1.3.1 RNA-Recognition Motif (RRM)	7
1.3.2 K-Homology domain (KH)	8
1.3.3 CCCH, CCHH and CCHC zinc finger (zf)	10
1.3.4 double-stranded RNA Binding domain (dsRBD)	11
1.4 Staufen is a conserved dsRNA binding protein	18
1.4.1 <i>Dm</i> Staufen: a central player for transport and localization	20
1.4.2 Transport and localization by mammalian Stau (maStau) (Stau1 and Stau2)	21
1.4.3 mRNA stability: Staufen mediated decay	23
1.5 General properties of Staufen associated RNA targets	25
2 Aim of this thesis	28
3 Results	29
3.1 hStau1 FL solubility and stability <i>in vitro</i> are influenced by ionic strength conditions	29
3.1.1 Purification of hStau1 FL protein	29
3.1.2 Purification of hStau1 truncations	30
3.2 hStau1 dimerization is concentration dependent <i>in vitro</i>	33
3.2.1 SEC-MALLS analysis of hStau1 FL	33
3.3 hStau1 FL binds the <i>ARF1</i> SBS with high affinity <i>in vitro</i>	36
3.3.1 Reconstitution of the hStau1 FL - <i>ARF1</i> SBS complex	36
3.3.2 Electrophoretic mobility shift assays of hStau1 FL with 5'FAM <i>ARF1</i> SBS _{20bp}	37
3.3.3 Fluorescence anisotropy of hStau1 FL and 5'FAM <i>ARF1</i> SBS _{20bp}	38
3.4 hStau1 RNA-binding activity induces dimerization independent of dsRBD2 and the dimerization region	40
3.4.1 Reconstitution of the hStau1 ^{dsRBD2-4} - <i>ARF1</i> SBS complex	40

3.4.2	Biochemical characterization of the hStau1 ^{dsRBD2-4} - <i>ARF1</i> SBS _{20bp} complex	41
3.4.3	Reconstitution of the hStau1 ^{ΔNΔC} - <i>ARF1</i> SBS complex	42
3.4.4	Reconstitution of the hStau1 ^{dsRBD3-5} - <i>ARF1</i> SBS complex	43
3.4.5	Biochemical characterization of the hStau1 ^{dsRBD3-5} - <i>ARF1</i> SBS _{20bp} complex	45
3.4.6	hStau1 ^{dsRBD3-5} - <i>ARF1</i> SBS complex crystallization	46
3.5	hStau1 dsRBD3 and 4 are sufficient for <i>ARF1</i> SBS binding <i>in vitro</i>	47
3.5.1	Reconstitution of the hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS complex	47
3.5.2	Biochemical characterization of the hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS _{20bp} complex	49
3.6	hStau1 ^{dsRBD3-4} recognizes and binds <i>ARF1</i> SBS by conformational readout of the minor and major groove.	50
3.6.1	hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS complex crystallization	50
3.6.2	hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS structure	52
3.7	The <i>ARF1</i> SBS structure is A-form in conformation	56
3.7.1	<i>ARF1</i> SBS crystallization	56
3.7.2	<i>ARF1</i> SBS conformational changes by binding to hStau1 ^{dsRBD3-4}	57
3.8	hStau1 ^{dsRBD3-4} recognition mode of <i>ARF1</i> SBS	58
3.8.1	Structure-based mutational analyses of hStau1 ^{dsRBD3-4} <i>ARF1</i> SBS	62
4	Discussion	65
5	Materials and Methods	73
5.1	Materials	73
5.1.1	Chemicals	73
5.1.2	Buffer solution	74
5.1.3	Consumables	75
5.1.4	Culture media	75
5.1.5	Resins and columns	76
5.1.6	Enzymes, Kits and Markers	76
5.1.7	Crystallization screens	77
5.1.8	Technical equipment	77
5.1.9	Bacterial strains	78
5.1.10	Oligonucleotides	78
5.1.11	RNA oligonucleotides	79
5.1.12	Software	79
5.2	Methods	79
5.2.1	Nucleic acid methods	79
5.2.2	Recombinant protein expression in <i>Escherichia coli</i>	81
5.2.3	Protein methods	82
5.2.4	Crystallization and data collection	87

Bibliography	88
List of figures	101
List of tables	103
Abbreviation	104
Acknowledgements	109

Abstract

mRNA localization is an essential process in the spatiotemporal regulation of gene expression. In this process, the cell restricts and controls different steps of the "mRNA life-cycle" by the dynamic assembly of ribonucleoprotein particles. These particles are composed of networks of mRNAs containing *cis*-acting elements that are recognized by *trans*-acting RNA-binding proteins. The RNA binding proteins can thus influence the mRNA fate by acting directly or indirectly on the RNA at every step of the process.

One of the best described *trans*-acting factors in mRNA localization is Staufen, a conserved double-stranded RNA binding protein (dsRBP). Staufen was first described as a determinant for the localized translation of maternal mRNAs in the *Drosophila melanogaster* oocyte. In mammals, two Staufen homologs have been identified Staufen1 and Staufen2. Both proteins regulate the transport and translational localization of specific mRNAs in different cell types. Additional to its role in localization, Staufen1 and Staufen2 affect target mRNA stability. Several mRNAs targets have been reported as specific targets of Staufen, some of the best-characterized are the *bicoid* mRNA (*bcd*) in *Drosophila melanogaster* and the ADP-ribosylation factor 1 (*ARF1*) mRNA in humans.

Staufen is a multidomain protein composed of four to five dsRNA binding domains (dsRBD). The canonical dsRBDs dsRBD3 and dsRBD4 have been shown to bind dsRNA *in vitro*. dsRBD2 and dsRBD5 are shorter domains that differ from the canonical dsRBD. Interestingly, dsRBD2 has RNA-binding activity whereas dsRBD5 is a non-RNA binding domain that mediates protein-protein interactions, and in the case of Staufen1 and Staufen2 is involved in protein dimerization. The fact that the canonical dsRBDs recognize specifically RNA secondary structures but lack sequence-specific discrimination opposes with Staufen specificity. Structural data of the *Drosophila melanogaster* dsRBD3 in complex with an artificial stem-loop, give some insight into the molecular basis of this interaction. However, the information is limited to a single dsRBD and a non-physiological relevant RNA.

In this thesis, I report the solution of the structure of dsRBD3 and dsRBD4 of the hStaufen1 in complex with a physiological dsRNA sequence encoding the *ARF1* Staufen binding site. The structure shows that the two domains wrap around the RNA duplex, that remains unchanged upon binding. Several interactions between the dsRBDs and the RNA sugar-phosphate backbone

were observed. Furthermore, both dsRBDs exhibit direct base interactions in both dsRNA minor grooves mediated by conserved Staufen residues. Evaluation of the affinity *in vitro*, by mutating the residues involved in the interaction suggests a stronger contribution by residues mediating nonspecific interactions compared to a negligible contribution of residues involved in base-directed interactions. However, *in vivo* experiments in *Drosophila melanogaster* showed that residues involved in base-directed interactions are significant contributors to accomplish binding accuracy. The biochemical and structural data presented here, indicate that affinity and specificity of RNA recognition by Staufen are based on the interplay of its canonical dsRBDs.

Zusammenfassung

mRNA-Lokalisierung ist ein essenzieller Mechanismus der räumlich und zeitlich regulierten Genexpression. Dieser Mechanismus ermöglicht der Zelle die Kontrolle und Überwachung der unterschiedlichen Phasen des mRNA-Lebenszyklus durch die dynamische Zusammenstellung von Ribonucleoprotein-Partikeln. Diese Partikel beinhalten Netzwerke von Komplexen aus mRNA-Molekülen und mRNA-bindenden Proteine. Die mRNA-Moleküle sind gekennzeichnet durch *cis*-wirkende Elemente, welche von *trans*-agierenden Faktoren, sogenannten mRNA-bindenden Proteinen, erkannt werden. Die mRNA-bindenden Proteine beeinflussen somit jede Phase des mRNA-Lebenszyklus, indem sie direkt oder indirekt mit der RNA interagieren.

Eines der am besten charakterisierten *trans*-agierenden Faktoren bei der mRNA Lokalisierung ist Staufen, ein konserviertes doppelsträngige-RNA-bindendes Protein (dsRBP). Staufen wurde als ein wichtiger Faktor bei der lokalen Translation von maternalen mRNAs in der Taufliege *Drosophila melanogaster* identifiziert. In Säugetieren, existieren zwei Homologe, Staufen1 und Staufen2. Beide Homologe werden benötigt, um den Transport und die lokale Translation von spezifischen Transkripten in verschiedenen Zelltypen zu regulieren. Neben ihrer Rolle in der Lokalisierung wirken sich Staufen1 und Staufen2 auf die mRNA-Stabilität aus. Mehrere spezifische Transkripte von Staufen wurden bereits identifiziert. Die am besten charakterisierten Transkripte sind die *bicoid* mRNA (*bcd*) in *Drosophila melanogaster* und die ADP-Ribosylierungsfaktor 1 mRNA (*ARF1*) in Menschen.

Staufen ist ein Multidomänenprotein und besteht aus vier bis fünf doppelsträngige-RNA bindenden Domänen (dsRBD). Davon gehören dsRBD1, dsRBD3 und dsRBD4 zu den kanonischen dsRBDs, die dsRNA *in vitro* binden können, im Unterschied dazu stehen dsRBD2 und dsRBD5, welche verkürzt sind. Interessanterweise konnte für dsRBD2, im Gegensatz zu dsRBD5, eine RNA-bindende Aktivität nachgewiesen werden. dsRBD5 beteiligt sich an Protein-Protein-Interaktionen und im Fall von Staufen1 und Staufen2 an der Dimerisierung der Proteine. Die Tatsache, dass die kanonischen dsRBD bestimmte Sekundärstrukturen erkennen, aber keine spezifische Sequenzen diskriminieren können, steht im Kontrast zur nachgewiesenen Spezifität von Staufen. Strukturelle Informationen aus dem Komplex von *Drosophila melanogaster* dsRBD3 mit einer künstlichen RNA-Haarnadelstruktur ermöglichen einen Einblick in die zugrunde liegenden molekularen Mechanismen der Interaktion. Allerdings beschränkt sich diese Information auf eine einzige dsRBD und

eine unphysiologische RNA.

In der vorliegenden Doktorarbeit, wurde die Struktur des Komplexes aus hStaufen1 dsRBD3 und dsRBD4 mit der physiologischen dsRNA Sequenz der *ARF1* Staufen Bindungsstelle bestimmt. In der Struktur binden die zwei Domäne das RNA-Duplex ein. Durch die Interaktion konnte keine signifikante Veränderung in der Struktur des RNA-Duplex festgestellt werden. In der Struktur ist deutlich zu erkennen, dass Interaktionen der dsRBDs mit dem Zucker-Phosphat-Rückgrat der RNA überwiegen. Zudem weisen beide Domänen, vor allem über konservierte Proteinreste im Staufen, direkte Interaktionen mit den Basen in beiden kleine Furchen der dsRNA auf. Analysen der Affinität *in vitro* haben gezeigt, dass Mutationen der Proteinreste, die an indirekten Interaktionen beteiligt sind, erhebliche negative Effekte auf die Affinität hatten, während Mutationen der Proteinreste, die an direkten Interaktionen beteiligt sind, nur marginale Effekte aufwiesen. *In vivo* Experimente an *Drosophila melanogaster* haben jedoch gezeigt, dass die an direkten Interaktionen beteiligte Proteinreste einen wesentlichen Beitrag leisten, um eine vollkommene Bindung mit der RNA zu erreichen. Die hier präsentierten biochemischen und strukturellen Daten legen nahe, dass die Affinität und Spezifität der dsRNA-Erkennung von Staufen auf das Zusammenspiel seiner kanonischen dsRBDs basieren.

1 Introduction

1.1 Post-transcriptional gene regulation

The transfer of genetic information via Ribonucleic acids (RNAs) to produce proteins is a conserved process across species. In eukaryotes, this process increases in complexity by the presence of new compartments in the cell and the requirement of cell differentiation. These changes may have led to the adaptation and specialization of regulatory mechanisms (Anantharaman, 2002).

At the post-transcriptional level, these regulatory mechanisms are present at all stages of the messenger RNA (mRNA) life cycle including transcription, maturation, transport, and degradation (Wilusz and Wilusz, 2004, Fasken and Corbett, 2005, 2009). This multi-step process starts in the nucleus when the new mRNA (pre-mRNA) molecule is transcribed by RNA polymerase II (Pol II). During transcription, the nascent pre-mRNA molecule is modified. These modifications include 5'-end capping, splicing (intron removal), 3'-end cleavage, polyadenylation and other regulatory modifications such as RNA editing (Colgan and Manley, 1997, Maniatis and Reed, 2002, Proudfoot et al., 2002, Proudfoot, 2011). The matured and polyadenylated mRNA is then exported to the cytoplasm (Fasken and Corbett, 2005, Carmody and Wentz, 2009, Rodríguez-Navarro and Hurt, 2011). In the cytoplasm, the mRNA is then translated by the ribosome. Before translation, the mRNA can be localized or stored (Martin and Ephrussi, 2009). Finally, the mRNA is targeted for degradation (Dreyfuss et al., 2002, Moore, 2005, Braunschweig et al., 2013, Bentley, 2014) (Fig. 1.1).

From birth to decay, the mRNA molecule is bound by a multitude of proteins. *Cis*-elements on the mRNAs are recognized by *trans*-acting factors (mainly mRNA binding proteins) to mediate the regulatory dynamics. A large variety of mRNA binding proteins (mRBPs) bind mRNA to form mRNA ribonucleoprotein particle (mRNP). The *trans*-acting factors contained in this particle can bind stably or transiently to other protein factors at different steps and so modulate the function in a combinatorial way.

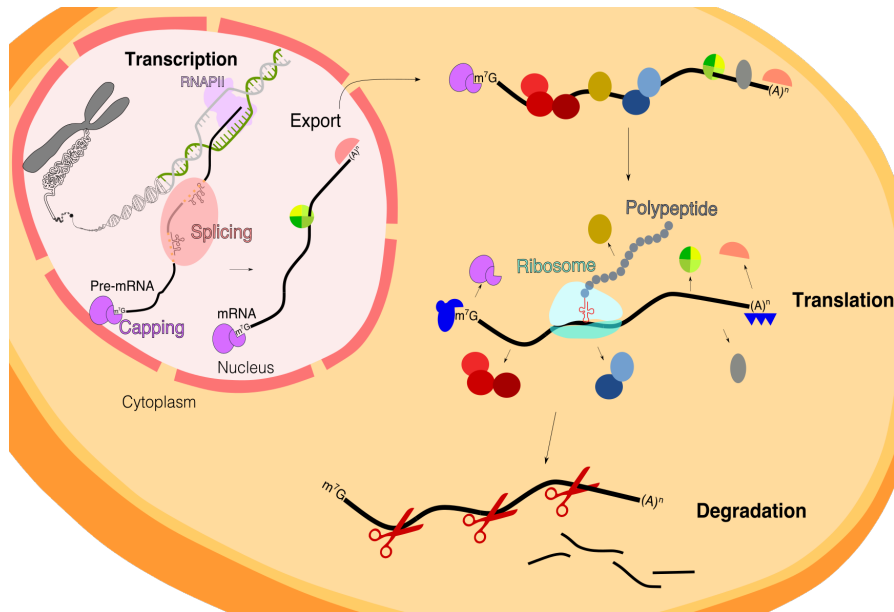


Figure 1.1. Schematic of post-transcriptional gene regulation PTGR in eukaryotes. In the nucleus, RBPs already bind the nascent pre-mRNAs (capping) and together with *cis*-elements facilitate their maturation and modification (splicing). Other mRBPs escort the mature mRNAs and enable their export to the cytoplasm. Once there, the mRNA can be translated. After translation is completed, the mRNA is then marked for degradation.

1.2 RNA binding proteins

In order to understand the mechanisms behind the events that regulate gene expression, it is necessary to elucidate not only the composition of mRNPs but also the molecular basis of recognition and interactions of *trans*-acting factors. Most of these *trans*-acting factors are mRBPs. The knowledge of the exact composition of mRNPs at different steps is still limited. However, in recent years various biochemical, genomics and bioinformatics studies have identified mRNP components in different cell types (Burd and Dreyfuss, 1994, Johnston et al., 1992, Lunde et al., 2007).

Different classes of mRBPs have been identified, some of them are universally conserved, others are only found in eukaryotes or in metazoans (Anantharaman, 2002). mRBPs are often modular proteins containing two classes of domains : (1) RNA binding domains (RBDs) that can recognize and bind mRNA; this binding can be sequence-specific and restricted to the recognition of structural elements within the mRNA. (2) Enzymatic and non-enzymatic domains related to functions other than mRNA binding (Anantharaman, 2002). Although most of these proteins contain only one RBD, there are groups that are characterized by the presence of multiple RBD repeats or a combination of class one and two. The combination of domains achieves an extended level of gene regulation, by allowing the protein to bind to the target with different affinities and specificity (Lunde et al., 2007).

1.3 Structural classification of RNA binding proteins

Structural and biochemical characterization gives insight into the recognition and binding mode of different mRBPs. This characterization allows the classification of the individual domains present in mRBPs. These domains have a particular fold and RNA binding activity, that reflects the adaptation of mRBPs to different biological functions during evolution (Anantharaman, 2002).

Different biochemical and structural methods together with bioinformatic approaches identified several RBD classes. The most common RNA binding domains are the RNA-Recognition Motif (RRM), the K-Homology domain (KH), the zinc finger (zf) and the double-stranded RNA Binding domain (dsRBD) (Fig. 1.2).

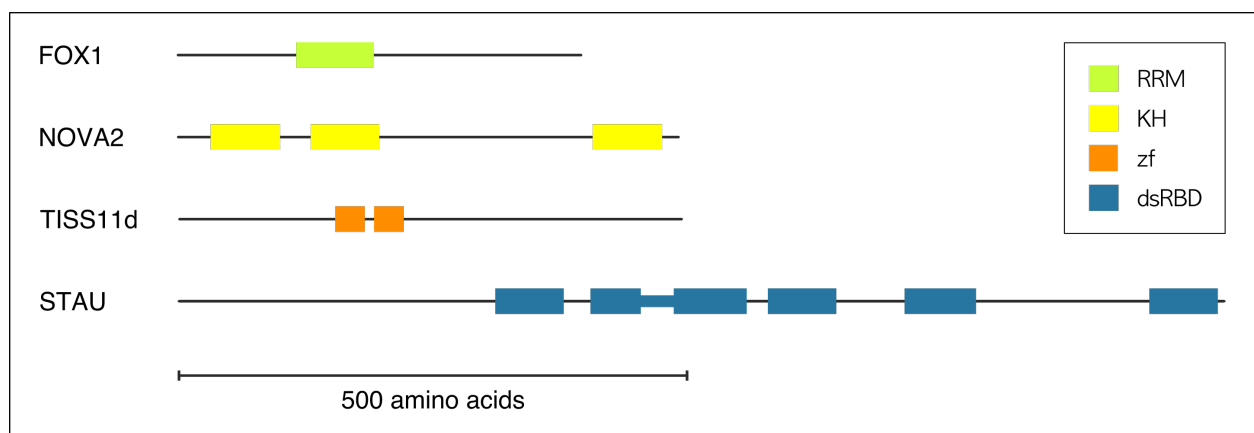


Figure 1.2. Architecture and domain composition of representative RNA binding domains of RBPs. RBPs can have one RBD or multiple copies connected by flexible loops. This modular architecture contributes to the target recognition. However, the topology differs for each class and allows RBPs to function in diverse aspects of PTGR.

1.3.1 RRM

The RRM (also known as the RNA-binding domain, RBD) is the most abundant eukaryotic RBD and the best characterized. RRM containing RBPs are involved in many aspects of RNA regulation and PTGR (Dreyfuss et al., 2002, Maris et al., 2005, Auweter et al., 2006). The RRM is about 90 amino acids long and adopts a $\beta\alpha\beta\beta\alpha\beta$ topology (Fig. 1.3). Two conserved regions characterize the RRM. Region one is present in β -strand 1 (known as RNP-2) and region two in β -strand 3 (known as RNP-1) (Adam et al., 1986, Swanson et al., 1987). The domain can bind single-stranded RNA (ssRNA) and single-stranded DNA (ssDNA) as well as mediate protein-protein interactions (Bandziulis et al., 1989, Kenan et al., 1991, Birney et al., 1993). Commonly, more than one copy is present in an RBP, and the interplay between the domains influences the function of the protein. Multiple RRMs can bind to ssRNA cooperatively, or have different specificities (Lunde et al., 2007, Anantharaman, 2002).

Conserved aromatic residues present in RNP-1 and RNP-2 are essential for RNA binding and recognition. Typically, these residues allow the domain to recognize specific sequences and to make stacking interactions with the RNA bases. Additionally, RNA binding specificity is obtained by the fourth β -sheet, where at least four nucleotides (nt) of ssRNA can be specifically recognized through hydrogen bonds and van der Waals interactions (Auweter et al., 2005, Lunde et al., 2007, Cléry et al., 2008). An example of this mode of recognition is shown by the structure of the Fox1 protein bound to the RNA heptamer UGCAUGU (Fig. 1.3)(Auweter et al., 2005). Three conserved residues (His120, Phe158, and Phe160) located in RNP-1 and RNP-2 bind specifically to two nucleotides (U5 and G6) (Fig. 1.3 B upper panel). Additionally, Fox1 RRM exhibits an extended RNA binding surface, where the loop between β -strand 1 and α -helix 1 recognizes four further nucleotides (U1, G2, C3, and A4) (Fig. 1.3 B bottom panel) (Auweter et al., 2005). As for the Fox1 protein, the length of the loops between secondary structure elements within the RRM, and extension of the N- C- termini can provide an extended binding region in different RRM-containing proteins. These extensions allow the domain to recognize variable length of ssRNA (up to eight nucleotides) and to modulate the binding affinity (Maris et al., 2005, Auweter et al., 2005, Lunde et al., 2007, Cléry et al., 2008, Muto and Yokoyama, 2012).

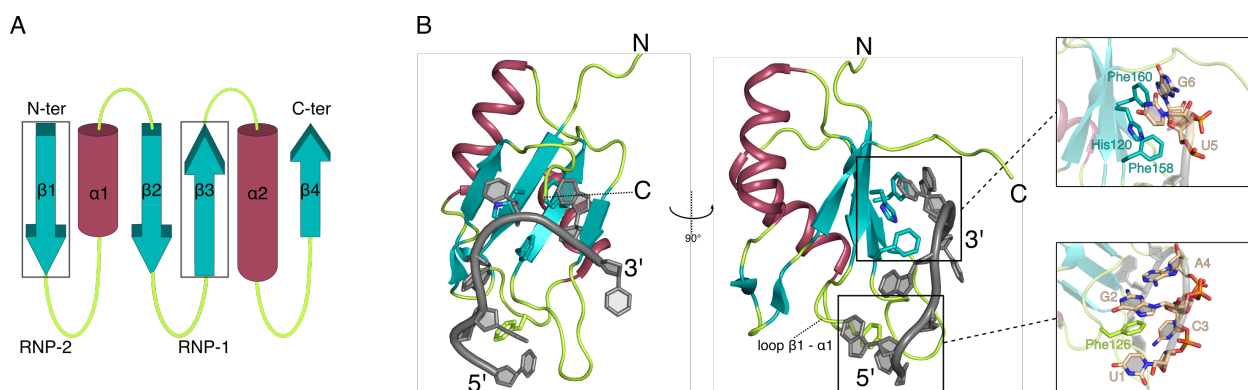


Figure 1.3. Representative RRM from the human Fox1 protein. **(A)** Topology of the RRM fold, colored by element: α -helices are colored in deep-magenta, β -sheets in teal and loops in lemon green. **(B)** nuclear magnetic resonance (NMR) structure of the RRM of human Fox1 in complex with a ssRNA heptamer. The domain shows the typical RRM fold (Protein Data Bank (PDB) : 2ERR) (Auweter et al., 2005). On the right equivalent view rotated by 90° , the upper panel shows an enlarged view of the canonical binding mode and the bottom panel shows extensions of the RNA binding mode. The structure is colored as in A, RNA is colored in gray, residues and bases involved in the interaction are represented as sticks, bases are colored in light orange. Illustrations of the protein structure used in all figures were generated with the PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC, 2015)

1.3.2 KH

The KH domain was first described in the human heterogeneous nuclear RNP K Protein (hNRNP K). This domain can bind to ssRNA and ssDNA and is involved in a variety of cellular functions such as splicing, transcription, and translation. The minimal KH motif ($\beta\alpha\alpha\beta$) has a topology that include a

consensus sequence (I/L/V)IGXXGXX(I/L/V) (GXXG motif) loop between α -helices 1 and 2, and a variable loop (3 to 60 aa) that connects β -strand 2 with an $\alpha\beta$ extension module ($\alpha'\beta'$) (Fig. 1.4 A) (Grishin, 2001, Valverde et al., 2008). This conserved domain evolved in two folds, type I (eukaryote-specific) and type II (prokaryote-specific), that differ in the position of the minimal KH motif (Fig. 1.4 A) (Grishin, 2001, Lunde et al., 2007, Valverde et al., 2008). Type I is characterized by a $\beta\alpha\alpha\beta\alpha$ topology; the minimal KH motif is located at the N-terminus and followed by the variable loop with the $\beta\alpha$ module. This type is usually present as a single domain in RBPs (Grishin, 2001, Valverde et al., 2008). In contrast, type II adopts a $\alpha\beta\beta\alpha\alpha\beta$ topology where the variable loop connects an $\alpha\beta$ module with the minimal KH motif at the C-terminus. This type is predominantly found as multiple repeats (Grishin, 2001, Valverde et al., 2008).

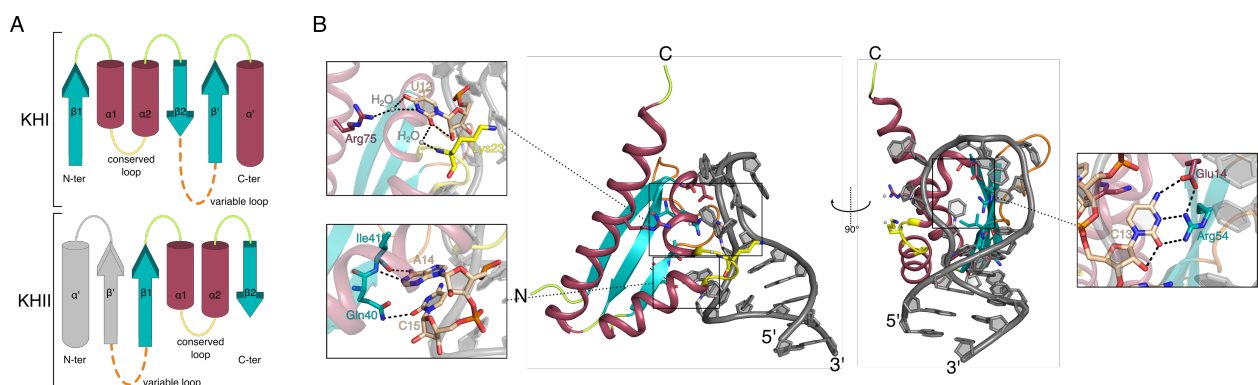


Figure 1.4. Representative KH domain from the Nova 2 protein. **(A)** Topology of type I and II KH folds colored by element: α -helices are colored in deep-magenta, β -sheets are colored in teal, loops are lemon green, GXXG motif is colored in yellow, and the variable loop is represented as an orange dotted line. **(B)** X-ray structure of the KH3 domain of the protein Nova 2 in complex with an RNA stem-loop (PDB: 1EC6) (Lewis et al., 2000). The panels show an expanded view of the interactions with the 5'-UCAC-3' sequence present in the RNA loop. On the right equivalent view rotated by 90° that shows the binding cleft formed by the minimal KH motif and the variable loop. The structure is colored as in A, RNA is colored in gray, residues and bases involved in the interaction are represented as sticks, bases are colored in light orange. Dotted lines indicate hydrogen-bonds, water molecules are represented as light grey spheres.

The RNA binding surface extends to the C-terminal region of the minimal KH motif. The two α -helices with the conserved GXXG motif are positioned face to face with the β -strand and the variable loop (Fig. 1.4 B). The formed cleft allows the recognition of a maximum of four nucleotides by van der Waals contacts, hydrogen bonds and hydrophobic interactions. This mode of binding is well illustrated by the structure of the KH3 of the Neuro-oncological ventral antigen 2 (Nova 2) protein bound to a ssRNA loop (Fig.1.4) (Lewis et al., 2000). This domain belongs to the KH type I, and recognizes the 5'-UCAC-3' sequence present in the RNA loop. For this, Lys23, in the GXXG motif, and Arg75, in α' -helix make a water-mediated interaction with U12. Residues in α -helix 1 and β' -strand, as well as residues in the β -strand 2, interact directly with C13 and C15 respectively. Finally, A14 is in hydrogen-bonding with the backbone of Ile41. Additional electrostatic interactions

with the RNA backbone stabilize the complex (Lewis et al., 2000). The RNA affinity of a single KH domain is low (micromolar range), compared to other RBDs. Nevertheless, the presence of multiple repeats in a protein increases the affinity of binding to the nanomolar range (Beuth et al., 2005, García-Mayoral et al., 2007).

1.3.3 CCCH, CCHH and CCHC zf

The zf is a conserved domain that can bind RNA as well as DNA (Miller et al., 1985, Clemens et al., 1993). The domain is about 20 to 30 amino acids long and contains four metal-binding residues; these can be cysteines and histidines in different combinations. Their number and order determine the classification of the zf (CCCH, CCHH or CCHC). The four metal-binding residues coordinate a zinc ion (Zn^{2+}) that forms and stabilizes the structure of the domain (Lu et al., 2003, Hudson et al., 2004, Lunde et al., 2007).

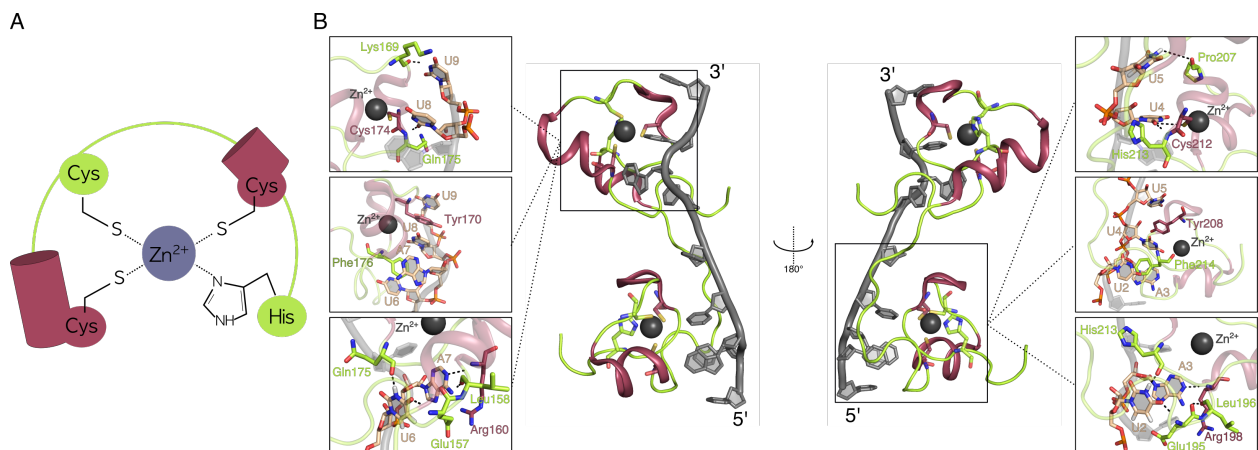


Figure 1.5. Representative zf from the human TIS11d protein. **(A)** The CCCH zinc finger motif (three Cys and one His residues bonded to a Zn^{2+} ion) colored by element: α -helices are colored in deep-magenta, and loops in lemon green. zinc ions are represented as grey spheres. **(B)** NMR structure of the two zf of TIS11d protein in complex with an AU-rich element (PDB: 1RGO) (Hudson et al., 2004). The panels show an expanded view of the interactions of the zf 1 (right) and zf 2 (left) with the RNA. The structure is colored as in A, RNA is colored in gray, residues and bases involved in the interaction are represented as sticks, bases are colored in light orange. Dotted lines indicate hydrogen-bonds.

In eukaryotes, Zn fingers have been shown to be involved in the regulation of different aspects of mRNA metabolism. The mode of recognition depends on the Zn fingers type. The CCHH type can interact with the backbone of double-stranded RNA (dsRNA), and also with bulged bases of the RNA (Lu et al., 2003). In contrast, the CCCH and CCHC type can bind ssRNA in a sequence-specific manner, through hydrogen bonds (Hudson et al., 2004). This sequence-specific mode of recognition is illustrated in the structure of the two zf domains (CCCH-type) of the human TIS11d protein in complex with an AU-rich element (5'-UUAUUUAUU-3') in Fig. 1.5. Both Zn fingers recognize the UAU motif through hydrogen bonds between the protein backbone and the Watson-Crick edges

of the bases (Fig. 1.5 top and bottom panel). Furthermore, stacking interactions (Fig. 1.5 middle panels) stabilize the interactions (Hudson et al., 2004). Typically, zf domains are present in multiple repeats or in combination with others RBD.

1.3.4 dsRBD

The dsRBD was first identified in a sequence-based database screen that showed that a common 65-70 amino acid region was present in proteins that specifically recognize dsRNA, such as the *Xenopus laevis* RNA-binding protein A (Xlrpba), the human TAR-binding protein (TRBP) and the *Drosophila* Staufen (*DmStau*) protein (Johnston et al., 1991). A subsequent search in the protein database identified further dsRBD-containing proteins with a high sequence similarity to the dsRBD motif, including *Escherichia coli* RNase III (March et al., 1985), responsible for dsRNA cleavage, dsRNA-dependent protein kinase (PKR), which control protein translation (Meurs et al., 1990, Green and Mathews, 1992) and dsRNA-dependent adenosine deaminase (ADAR), that catalyze the modification of adenosine to inositol in dsRNA molecules (Kim et al., 1994, O'Connell et al., 1995) among other proteins (Fig. 1.6).

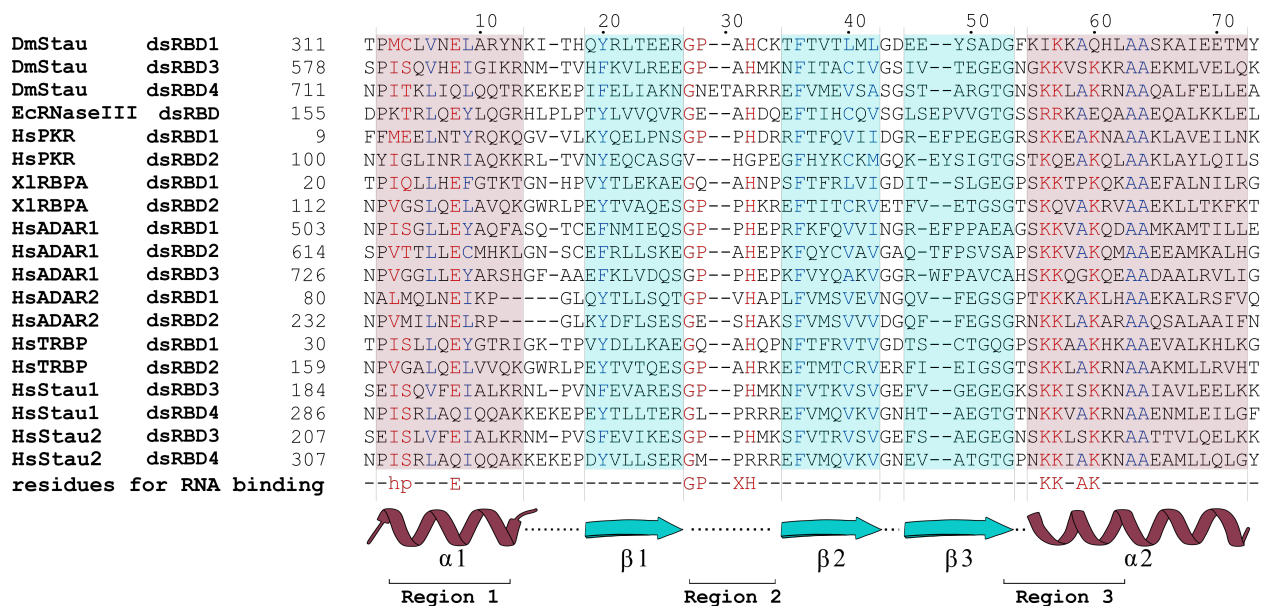


Figure 1.6. Multiple sequence alignment of dsRBDs in dsRBPs of different species: *Drosophila melanogaster* (*Dm*), *Escherichia coli* (*Ec*), *Homo sapiens* (*Hs*), *Xenopus laevis* (*Xl*). Conserved residues with structural function are highlighted in blue, residues involved in dsRNA binding are highlighted in red. Secondary structure elements are represented as cartoon below the alignment. The three major regions that make contact with dsRNA are indicated.

Binding experiments confirmed that dsRBDs in Xlrpba and *DmStau* bind with high affinity to dsRNA secondary structures but poorly if at all to ssRNA and double-stranded DNA (dsDNA) (Johnston et al., 1992). The identification and biochemical characterization of these dsRBPs confirmed the existence of a functional domain capable of interacting specifically with dsRNA (Johnston et al.,

1992, Green and Mathews, 1992). The list of dsRBPs is still growing today, and the vast variety of these proteins are active in almost all aspects of mRNA metabolism.

Table 1.1. Representatives members of dsRBPs

dsRBP	species	function	domains
RNAse III	<i>Escherichia coli</i>	dsRNA-specific endoribunoclease.	RNAse III domain and dsRBD
Staufen	<i>Drosophila melanogaster</i>	mRNA localization, translational regulation,	dsRBD A and B
	<i>H.sapiens</i>	mRNA decay.	dsRBD A and B, Tubulin binding domain (TBD)
TRBP	<i>H.sapiens</i>	Induce human immunodeficiency virus (HIV) expression, regulation of innate immune pathways,	dsRBD A and B
RBPA	<i>Xenopus laevis</i>	associated with ribosomes and hnRNPs.	
PKR	<i>H.sapiens</i>	transcriptional and translational regulation, regulation of innate immune pathways, control of cell growth, differentiation and apoptosis.	dsRBD A and Protein kinase
ADAR1	<i>H.sapiens</i>	adenosine to inosine RNA editing	Zalpha, dsRBD, deaminase
ADAR2			dsRBD A, deaminase
NF90	<i>H.sapiens</i>	Transcriptional, post-transcriptional and translational regulation; participates in innate antiviral response.	Domain associated with zinc fingers (DZF) and dsRBD A

The dsRNA binding domain consists of 65-70 amino acids and is present in a vast number of proteins in eukaryotes, prokaryotes, and viruses. The N-terminal part of the dsRBD varies in different species while the C-terminal part shows high conservation (Fig. 1.6). Mutation of conserved residues, present in three regions of the domain (Fig. 1.6), can decrease or abolish dsRNA-binding activity (Johnston et al., 1992, Bycroft et al., 1995, McMillan et al., 1995, Krovat and Jantsch, 1996, Ramos et al., 2000). The number of dsRBDs found in dsRBPs varies from one (more common in

viruses) to multiple copies. A large variety of dsRBPs contain a combination of dsRBDs with other domains (Table 1.1).

NMR studies of *E.coli* RNase III and of the third dsRBD of *DmStau* reveal a $\alpha\beta\beta\alpha$ topology (Bycroft et al., 1995, Kharrat et al., 1995) Fig. 1.7 B), in which the two α -helices at the N- and C- termini fold against the three stranded antiparallel β -sheet (Fig. 1.7) A). Mutagenesis experiments indicate that conserved basic residues present on the surface of the domain in dsRBD1 of the human PKR protein (McMillan et al., 1995) and the dsRBD3 of the *DmStau* protein (Bycroft et al., 1995) interact with the negatively charged backbone of the RNA. Successive studies in other dsRBPs confirmed that mutations in this conserved region of the domain mostly abolish or reduce RNA binding (Green and Mathews, 1992), (Bycroft et al., 1995, Kharrat et al., 1995, Krovat and Jantsch, 1996).

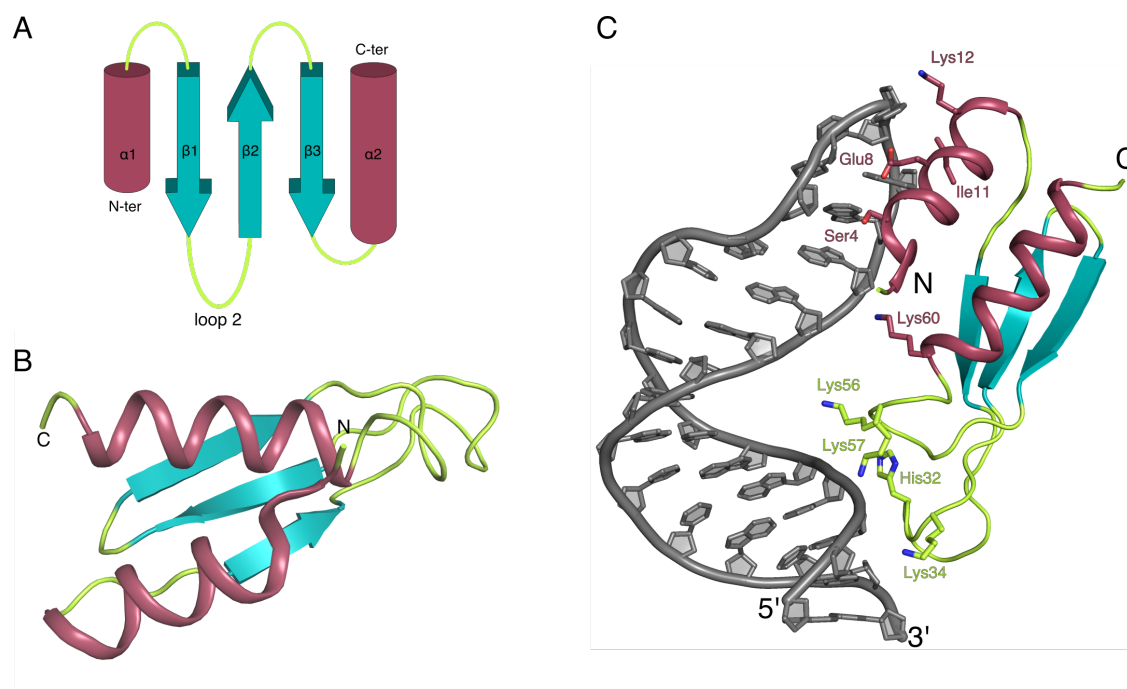


Figure 1.7. Structure, topology and substrate binding by a representative dsRBD domain: the dsRBD3 of *DmStaufen*. **(A)** Topology of the dsRBD colored by element: α -helices are colored in deep-magenta, β -sheets are colored in teal, loops are lemon green. **(B)** NMR structure of the dsRBD3 of *DmStaufen*. The hydrophobic core of the domain contributes to the stability and correct fold of the domain (PDB: 1STU) (Bycroft et al., 1995). **(C)** NMR structure of the dsRBD3 of *DmStaufen* in complex with an artificial RNA stem-loop (PDB: 1EKZ) (Ramos et al., 2000). Residues along helix α_1 , the loop between strands β_1 and β_2 (loop 2) and the conserved N-terminal part of helix α_2 recognize the ribose 2'OH groups and the RNA phosphate backbone. The structure is colored as in A. The RNA is colored in grey, residues involved in RNA binding are shown by stick representation.

In proteins with multiple dsRBDs, primary sequence derivation of some of the dsRBDs was observed by amino acid insertion, deletion or replacement. These domain variations were mostly unable to bind dsRNA *in vitro* (Johnston et al., 1992). Therefore, two types of dsRBD were de-

scribed: type A, the canonical dsRBD that binds to dsRNA *in vitro*, and type B, that lacks conserved motifs for dsRNA binding but is important for protein-protein interactions (Krovat and Jantsch, 1996, Johnston et al., 1991, 1992). In type A dsRBDs, conserved residues extend through the whole domain sequence, while in type B conservation is restricted to the C-terminal end (Johnston et al., 1992).

The hydrophobic core of dsRBDs is formed by conserved residues present in the α -helices 1 and 2 and in the β -strands 1 and 2. These are essential for the correct folding of the domain and stabilization of residues involved in RNA binding (Kharrat et al., 1995, Bycroft et al., 1995, Nanduri et al., 1998). In these regions, aliphatic residues allow a tight packing of the domain, while aromatic residues are important for the stabilization of residues implicated in RNA binding (Krovat and Jantsch, 1996, Bycroft et al., 1995, Ramos et al., 2000).

1.3.4.1 RNA recognition by dsRBDs

NMR and X-ray structures of different dsRBDs in complex with short dsRNAs (10 to 16 base pairs (bp)) provide insights into the RNA recognition and interaction of dsRBD (Bevilacqua and Cech, 1996, Ryter, 1998, Masliah et al., 2012). Structural information of these complexes together with mutagenesis experiments show a preference for the A-conformation of the RNA double helix. In this conformation, the RNA minor groove is broad and shallow (11Å wide) allowing interaction with the bases (Delarue and Moras, 1989, Steitz, 1990). In contrast, the major groove is deep and narrow (4Å wide); here the interaction can be only mediated by 2'OH groups and the phosphodiester backbone. The presence of the 2'OH groups in ribose sugars represents a potential interaction surface that could discriminate dsRNA from dsDNA (Delarue and Moras, 1989, Steitz, 1990, Green and Mathews, 1992, Bycroft et al., 1995, Ryter, 1998, Ramos et al., 2000).

Usually, the RNA interaction interface includes three regions within the dsRBD: region 1 is present in the helix α 1, region 2 corresponds to the motif GPxHx in loop 2 between strand β 1 and β 2, and region 3 corresponds to the motif KKxAK present in the N-terminal part of helix α 2. Region 1 and 2 interact with two successive minor grooves and region 3 contacts the major groove (Ryter, 1998) (Fig. 1.8 and 1.9)

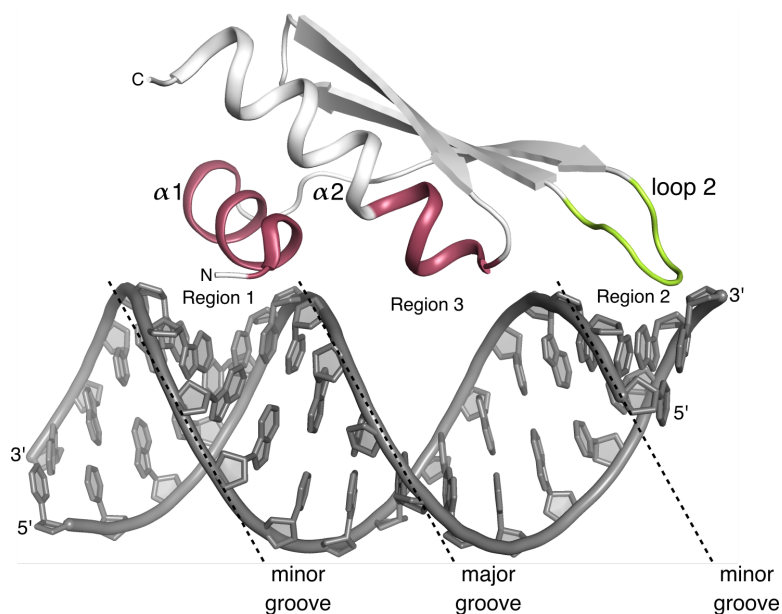


Figure 1.8. dsRBD2 of the Nuclear factors 90 (NF90) protein bound to a dsRNA molecule (PDB: 5DV7) (Jayachandran et al., 2015). The RNA A-conformation is recognized by three regions. Region 1, where helix $\alpha 1$ (deep-magenta) interacts with the minor groove, region 2 includes residues in loop 2 (green) that "inserts" into the subsequent minor groove and region 3 corresponds to the conserved residues in the N- terminal part of helix $\alpha 2$ (deep-magenta) that contacts phosphates of the RNA backbone of the major groove. Cartoon representation, color coding of structural elements is the same as in Fig. 1.3

Recognition by region 1: α helix 1 Region 1 was identified as a common dsRNA binding surface in different dsRBDs (Ryter, 1998). However, helix $\alpha 1$ is the least conserved region among dsRBDs of different proteins (Fig. 1.6) (Johnston et al., 1992, Krovat and Jantsch, 1996). This amphipathic helix interacts with the minor groove of the RNA. For this, the 2'OH of the RNA riboses are recognized mainly by water-mediated interactions and some direct interactions. The residues involved in RNA recognition and interaction are not well conserved (Fig. 1.6). However, the position where these residues are present is similar in different structures and also the nature of the side chains at this position is relevant for the interaction. So is the case for position 3, 4 and 8 from the sequence alignment on Fig. 1.6. Hydrophobic residues mainly occupy position 3. Residues in this position can place helix $\alpha 1$ in the minor groove by van der Waals interactions with the ribose and help to correctly position the motif KKxAK in helix $\alpha 2$ for dsRNA binding (Masliah et al., 2012). Position 4 is represented mainly by polar residues predicted to form hydrogen bonds with the ribose sugars with some exceptions. One of the exceptions is a methionine, which makes a sequence-specific hydrophobic interaction with an adenine in the structure of human ADAR2 with an RNA stem-loop of a natural substrate (Stefl et al., 2010) (Fig. 1.9). In mutation experiments, the replacement of A to G shows a five-fold binding reduction, confirming that this interaction can differentiate between AU from GC (Stefl et al., 2010). The last residue in position 8 is a conserved glutamate; this residue is essential for RNA binding and at the same time is part of the consensus

sequence (Johnston et al., 1992). The carboxylic group of the glutamate mediates the hydrogen bond interaction.

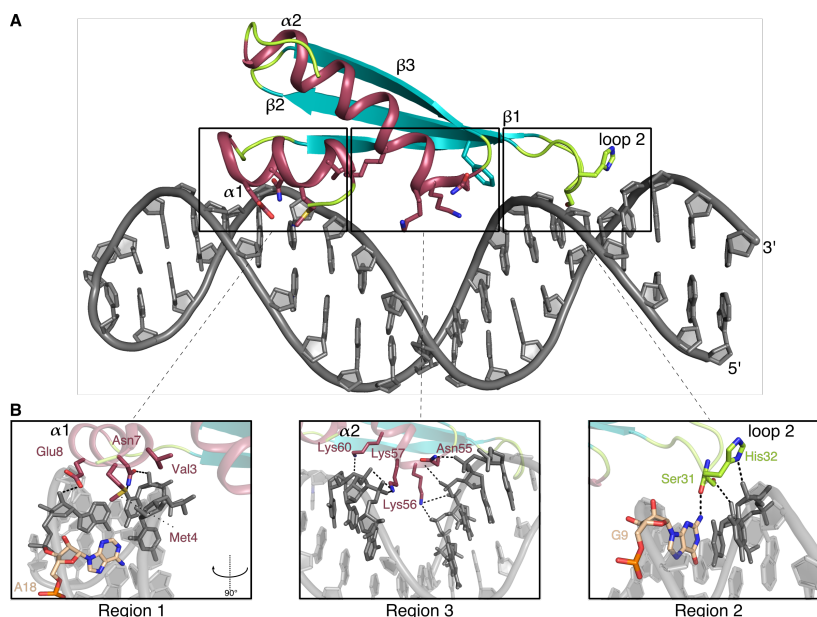


Figure 1.9. Global and detailed views of the NMR structure of the dsRBD2 of ADAR2 bound to the glutamate receptor 2 R/G lower stem-loop mRNA (GluR-2 R/G LSL) (PDB: 2L2K) (Stefl et al., 2010). **(A)** View of the dsRBD2 bound to the GluR-2 R/G LSL as cartoon representation. Color coding of structural elements is the same as in Fig. 1.3. The RNA is colored in gray. Residues that mediate recognition are shown as sticks and are numbered according to the sequence alignment of Fig. 1.6. Hydrogen bonds are indicated by black dotted lines. **(B)** Close-up view of the three regions of contact. In region 1, Met4 makes a hydrophobic and specific contact with A18 (light orange). Val3, Asn7, and Glu8 contact the bases surrounding this adenine and stabilize the interaction of helix $\alpha 1$; in region 3, the recognition of the major groove is achieved by Asn55 and the conserved Lys56, Lys57 and Lys60. These residues are hydrogen bonded with the phosphate backbone in both strands. In region 2, the main chain of Ser31 makes a hydrogen bond with the amino group of the G9 (light orange). This specific interactions together with other nonspecific interactions stabilize the binding to the minor groove. (Stefl et al., 2010).

Recognition by region 2: loop between β -strand 1 and 2 This region is well conserved and consists of a six amino acid long loop with the motif GP XHx (Fig. 1.6). In order to recognize the subsequent minor groove, the conserved His of the motif bridges the two RNA strands by binding to the 2'OH of the riboses through hydrogen bonds (Masliah et al., 2012). Structural data on different dsRBD-dsRNA complexes showed that this region can also contribute to substrate selectivity (Ryter, 1998, Stefl et al., 2010, Jayachandran et al., 2015). For this, the carbonyl group of the peptide backbone of the third residue in the motif (X) recognizes the amino group of a guanine base in the minor groove (Ryter, 1998, Stefl et al., 2010), (Masliah et al., 2012, Jayachandran et al., 2015). This recognition has been observed in the structure of dsRBD2 of Xlrbpa bound to a synthetic 10 bp dsRNA, the dsRBD2 of ADAR2 bound to the glutamate receptor 2 R/G and the dsRBD2 of NF90 with a synthetic 18 bp dsRNA (Ryter, 1998, Stefl et al., 2010), (Jayachandran et al., 2015) (Fig. 1.9).

Recognition by region 3: N-terminus of helix $\alpha 2$ In region 3, the consensus KKxAK motif present in the N-terminal part of helix $\alpha 2$, is essential for RNA binding (Ryter, 1998). This region recognizes the dsRNA major groove by making contact with the phosphodiester backbone of both strands. For this, the first and last Lys of the motif contact the phosphodiester backbone of one strand while the second Lys contacts the phosphodiester backbone of the other strand (Ryter, 1998). In this way, the domain can scan and recognize the width of the major groove. These interactions allow the domain to discriminate between A-form dsRNA and B-form dsDNA (Ryter, 1998, Bevilacqua and Cech, 1996). The required orientation of the Lys for RNA binding is predetermined by a set of van der Waals interactions and by the position of this residue in helix $\alpha 2$. Some variation is present in the last Lys of the motif: in some cases negative residues can replace this Lys. However, RNA binding is not affected given that in these cases a Lys from helix $\alpha 1$ compensates the binding (Masliah et al., 2012).

1.4 Staufen is a conserved dsRNA binding protein

One of the best-characterized dsRBP is Staufen (Stau). Stau was first isolated from *Drosophila melanogaster* (*Dm*) and identified as a multidomain dsRBP (Johnston et al., 1991, 1992). Further homologous of Stau were identified in *Drosophila virilis*, *Musca domestica*, *Caenorhabditis elegans*, *Mus musculus* and *Homo sapiens* (Marión et al., 1999, Wickham et al., 1999, Micklem, 2000, Ramos et al., 2000). Sequence alignment analysis showed that the conservation among the homologs is restricted to the dsRBDs, and the number of dsRBD copies varied from species to species (Micklem, 2000). Invertebrate Stau contains up to 5 dsRBD copies, while mammalian Stau has 4 dsRBDs copies (Micklem, 2000).

*Dm*Stau contains 5 copies of dsRBD (Fig. 1.10). From which dsRBD1, 3 and 4 are type A dsRBDs and dsRBD2 and 5 type B. Northwestern assays showed that dsRBD1, 3 and 4 bind to dsRNA structures *in vitro* (Johnston et al., 1992). However, dsRBD3 showed stronger dsRNA binding than dsRBD1 and 4 (Johnston et al., 1992, Ramos et al., 2000, Micklem, 2000). As for dsRBD2 and 5, both domains do not interact with dsRNA (Micklem, 2000). Structural and mutagenesis studies of dsRBD3 of *Dm* Stau bound to an artificial 12 bp stem-loop identified three RNA binding regions in the dsRBD, namely helix $\alpha 1$, loop 2 and the N-terminus of helix $\alpha 2$ (Ramos et al., 2000). Conserved residues in these three regions recognize and interact with two consecutive minor grooves (helix $\alpha 1$ and loop 2) and with the major groove (N-terminus of helix $\alpha 2$) of the A-form dsRNA. These residues are conserved through all dsRBD3 Stau homologs (Ramos et al., 2000). Some variation is present in dsRBD1 and 4, where these residues are either conserved or of the same type as dsRBD3 (Micklem, 2000). dsRBD2 of all Stau homologs are characterized by the insertion of a proline-rich region in loop 2 (Micklem, 2000). This insertion inhibits the recognition of the dsRNA major and minor groove by the domain (Micklem, 2000). As for dsRBD5, the residues involved in RNA binding are not conserved, while residues that maintain the hydrophobic core are conserved allowing the domain to fold as a classical dsRBD (Micklem, 2000).

In Mammals, two Stau homologs, Stau1 and Stau2, have been described (Fig. 1.10). Both homologs are present in several splicing isoforms. Stau1 generates three isoforms that differ at the N-terminus, and are characterized by the absence of the first dsRBD present in the *Dm* protein (DesGroseillers and Lemieux, 1996, Marión et al., 1999, Wickham et al., 1999). Up to 8 isoforms have been described for Stau2 (Buchner et al., 1999, Kiebler et al., 1999, Duchaîne et al., 2002). They differ at the C-terminus and are characterized by the absence of the last dsRBD present in the *Dm*Stau. Stau1 is ubiquitously expressed whereas Stau2 is restricted to the nervous system (Marión et al., 1999, Wickham et al., 1999, Duchaîne et al., 2002, Monshausen et al., 2008, Mallardo et al., 2003).

Although mammalian Stau (maStau) lacks one dsRBD, it has been shown by filter binding assays that it can bind dsRNAs with high-affinity *in vitro* (K_D of 10nM) (Marión et al., 1999, Wickham et al., 1999). As in *Dm* Stau, the mammalian Stau dsRBD3 is the main determinant of RNA binding whereas dsRBD4 has a weak binding activity (Marión et al., 1999, Wickham et al., 1999, Luo et al., 2002). In addition, mammalian Stau homologs contain a functional TBD between dsRBD4 and 5, for the interaction with the microtubule network (Wickham et al., 1999). As mentioned previously, dsRBD2 and 5 do not bind dsRNA, but could be important for protein-protein interactions (Marión et al., 1999, Kiebler et al., 1999, Wickham et al., 1999, Martel et al., 2010). Recently, it was reported that mStau2 dsRBDs1 and 2 could also mediate RNA binding. Nonetheless, the interplay with dsRBD3 and 4 is still necessary to accomplish high-affinity RNA binding (Heber et al., 2019). Finally, mammalian Stau includes a conserved motif formed by two α helices, the Staufen-swapping motif (SSM) (Gleghorn et al., 2013). A flexible linker connects the SSM to the N-terminus of dsRBD5. The SSM interacts with helix α 1 of the dsRBD5 of a second Stau protein to mediate dimerization (Gleghorn et al., 2013, Martel et al., 2010).

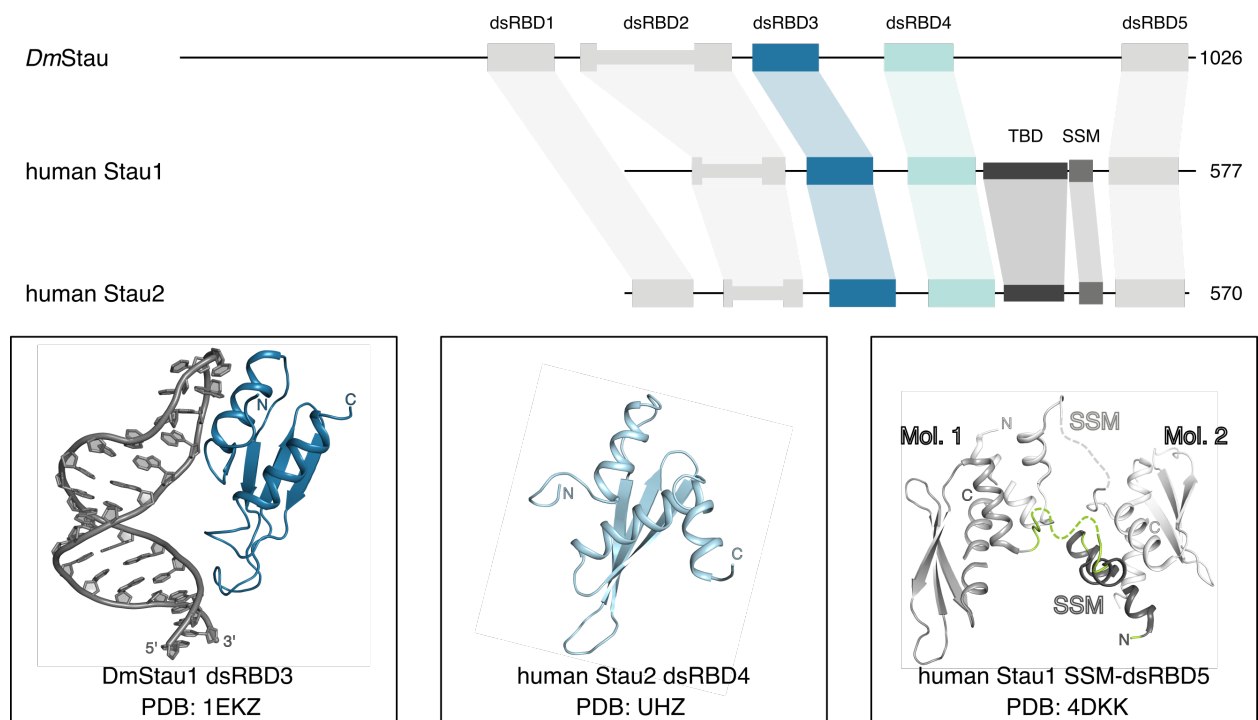


Figure 1.10. Domain organization of *Dm*Stau and *Hs*Stau. dsRBD implicated in dsRNA binding are colored in blue and turquoise. The TBD and the SSM are only present in mammals. Bottom left, NMR structure of the dsRBD3 (blue) of *Dm* Stau bound to an artificial RNA 12 bp stem-loop (gray) (Ramos et al., 2000); in the middle is the crystal structure of dsRBD4 (turquoise) of *Hs* Stau2 (He, F) and on the right is the crystal structure of the dimerization region of the *Hs* Stau1 protein, SSM and dsRBD5 (Molecule 1 in dark grey and molecule 2 in light grey) (Gleghorn et al., 2013).

1.4.1 *DmStaufen*: a central player for transport and localization

Stau was discovered in *Dm*, as a determinant for the correct anterior-posterior pattern development of the embryo (Schupbach and Wieschaus, 1986, St. Johnston et al., 1989, Johnston et al., 1991, Ephrussi et al., 1991, Kim et al., 2005, Kim-Ha et al., 1995, Broadus et al., 1998, Li et al., 1997). During oogenesis, *DmStau* is required for the correct transport and localization of *oskar* mRNA (*osk*) at the posterior pole of the oocyte (Ephrussi et al., 1991, Kim et al., 2005, Kim-Ha et al., 1995, Johnston et al., 1991). At the posterior pole *DmStau* is involved in the translational activation of *osk* (Micklem, 2000). The *Osk* protein is required for the formation of the pole plasm and the translational regulation of *nanos* mRNA, a determinant for abdomen formation (Ephrussi et al., 1991, Kim et al., 2005, Kim-Ha et al., 1995, Lehmann and Nüsslein-Volhard, 1991). The loop insertion on dsRBD2 and dsRBD5 of *DmStau* are required for the localization as well as the translational regulation of *osk* (Micklem, 2000). The function of these two domains was tested in *Drosophila* mutants by monitoring *osk* localization (Micklem, 2000). The deletion of the loop insertion (loop 2) in dsRBD2 (*Stau*^{Δloop2}) shows that the mutated protein is unable to transport and localize *osk* at the posterior pole (Micklem, 2000). A second mutant lacking dsRBD5 (*Stau*^{ΔdsRBD5}) rescues the wild-type localization of *osk*, and the colocalization with *DmStau*, but fails to activate translation of *osk* mRNA at the posterior of the oocyte (Micklem, 2000). This data suggests that loop 2 of dsRBD2 is required for microtubule-dependent localization of *osk* mRNA at the posterior pole. Once the mRNA has been localized, dsRBD5 is required for the translational activation of *osk* mRNA (Micklem, 2000).

At a later stage of development, *DmStau* is required for the localization of the maternal *bcd* at the anterior pole of the oocyte. *bcd* is required to produce a gradient of bicoid protein, that determines the pattern of the head and thorax (Driever and Nüsslein-Volhard, 1988a,b). *DmStau* binds to the 3' untranslated region (UTR) of *bcd* (stem-loops regions III, IV, and V) and is recruited into particles that are transported in a microtubule-dependent manner (Johnston et al., 1991, Ferrandon et al., 1994). Similar to *osk*, *bcd* localization is dependent on the loop insertion on dsRBD2 and dsRBD5 but independent of translation (Micklem, 2000).

During *Drosophila* neurogenesis, *DmStau* is required for the localization of *prospero* mRNA (*pros*). By the neuroblast asymmetric division, a larger neuroblast and a small ganglion mother cells (GMC) are generated. *DmStau* binds to *pros* to co-localize basally to the GMC. *DmStau* mediates the actin-dependent localization of *pros* by binding to Miranda through dsRBD5 (Fuerstenberg et al., 1998). Miranda is an adapter protein, that is also required in the asymmetric localization of Prospero (Shen et al., 1997). Miranda was the first described binding partner of *DmStau* (Shen et al., 1998, Fuerstenberg et al., 1998). In the GMC, the Prospero protein acts as a determinant for differentiation and stem cell fate suppression (Broadus et al., 1998, Li et al., 1997).

Other functions of *DmStau* arise in the adult fly, where *DmStau* is required during long-term memory formation. The exact mechanisms are not clear yet, but experiments with temperature-sensitive mutants show that at restrictive temperatures (*Stau* loss-of-function) short-term memory was abolished in adult flies, while long-term learning was maintained (Dubnau et al., 2003).

1.4.2 Transport and localization by maStau (*Stau1* and *Stau2*)

The role of maStau in transport and localization is well elucidated in the central nervous system (CNS). In neurons, both *Stau1* and *Stau2* are present in the somatodendritic domain of hippocampal neurons in mRNP (Kiebler et al., 1999, Köhrmann et al., 1999, Tang et al., 2001, Brendel et al., 2004). However, the localization of both *Stau* orthologues in hippocampal neurons is not restricted to the cytoplasm: *Stau1* was shown to be present in the nucleus where it interacts with nuclear RNAs (Le et al., 2000) and *Stau2* was shown to shuttle between the nucleus and the cytoplasm (Macchi et al., 2004, Miki and Yoneda, 2004). *Stau2* contains a nuclear localization signal (NLS) that is present in the unstructured region connecting dsRBD3 and 4 (Macchi et al., 2004, Miki and Yoneda, 2004). In the nucleus, *Stau2* binds dsRNA in order to be exported by exportin-5 (Exp-5) in a RanGTP-dependent manner (Miki and Yoneda, 2004).

Stau1 and *Stau2* were found to be distributed in distinct mRNPs, indicating non-redundant functions. Indeed each of them has different mRNA targets (Duchaîne et al., 2002, Thomas et al., 2005, Furic et al., 2007). Biochemical and localization studies of *Stau2* in hippocampal neurons identified two different mRNPs (Kiebler et al., 1999, Mallardo et al., 2003). The first class corresponds to large particles that are associated either with the nuclear membrane or the rough endoplasmic reticulum (RER) and are restricted to the soma of hippocampal neurons (Köhrmann et al., 1999, Mallardo et al., 2003). The second class corresponds to small particles that can be transported via microtubules and are distributed in the dendrites (Köhrmann et al., 1999, Mallardo et al., 2003).

Different components present in the maStau-containing RNP have been identified: mRNAs, large protein complexes that bind to *cis*-acting elements present in the mRNA 3'UTR, translation regulatory factors, ribosomes as well as cytoskeleton and motor proteins (Knowles et al., 1996, Brendel et al., 2004, Kanai et al., 2004, Villace P, 2004, Fritzsche et al., 2013). The presence of ribosomes and components required for transport is consistent with previous studies that showed that mammalian *Stau* co-localizes with the RER and associates with tubulin *in vitro* (Marión et al., 1999, Wickham et al., 1999). The RNA binding activity of dsRBD3 together with the protein-protein binding activity of dsRBD4-TBD mediate the *Stau*-RER and -tubulin interaction (Marión et al., 1999, Wickham et al., 1999, Duchaine et al., 2002).

Among the complex set of proteins found in maStau-containing mRNPs, different translational regulatory proteins have been identified. These include fragile X mental retardation protein (FMRP), nucleolin, purine-rich element binding protein A (Pur- α), cap binding protein 80 (CBP80), PABPN1, components of the RNA-induced silencing complex (RISC) and Pum2 (Brendel et al., 2004, Fritzsche et al., 2013). The mechanism by which Stau1 and 2 interact with these elements are not clear, but these elements are part of different mRNA regulatory processes in neurons. It is therefore likely that Stau works as an adapter protein in different regulatory processes.

The presence of nuclear proteins in Stau-containing particles indicates that the mRNP-complexes are previously assembled in the nucleus and exported to the cytoplasm for transport and localization (Le et al., 2000). These processes correlate with the nucleo-cytoplasmic distribution of Stau and with its capability of actin- and microtubules-dependent localization (Broadus and Doe, 1997, Micklem, 2000). Likewise, the high enrichment of translational regulators in Stau-containing mRNPs suggests that Stau might play a role in translational control, as described in *Dm* for *osk* mRNA. Indeed, it was shown in human cells that Stau1 up-regulates the translation of specific mRNAs that harbor structured elements in their 5'UTRs. For this, Stau1 binds to the structured 5'UTRs and destabilizes the mRNA to facilitate the recruitment of initiation factors and the ribosome (Dugre-Brisson et al., 2005).

As for *Dm*, mammalian Stau is also implicated in neurogenesis. During early neuronal development, neural stem cells also known as radial glial cells (RGCs) divide asymmetrically to form a new RGC and an intermediate progenitor cell (IPC) (Paridaen and Huttner, 2014). This asymmetric cell division is mediated by mRNP that contain Stau2 and other dsRBP (translational repressor Pum2 and RNA helicase DDX1) (Vessey et al., 2012). During this process, Stau2 regulates the transport and localization of proneurogenic mRNAs such as *Prox1*, *Trim32* and β -*actin* in the RGC to asymmetrical segregate into the differentiated cell IPC and suppress the RGC state (Vessey et al., 2012, Kusek et al., 2012).

In the mature nervous system, Stau1 and Stau2-containing RNP particles play a role in dendritic spine morphogenesis and synaptic plasticity. Both proteins are implicated in dendritic spine growth (Vessey et al., 2008). Experiments in rodents showed that knockdown of Stau1 or Stau2 results in morphologic changes in spines (Vessey et al., 2008, Lebeau et al., 2011). In addition knockdown of Stau1 lead to a significant reduction of the number of dendrites and a reduction in synapsis development (Vessey et al., 2008, Lebeau et al., 2011). Although these morphological alterations were present, learning or memory activity was not affected (Vessey et al., 2008). However, locomotor activity was decreased (Vessey et al., 2008).

As for synaptic plasticity, Stau1 and Stau2 exhibit different functions. Stau1 is required for the late form of long-term potentiation (L-LTP), where it acts as a regulator for transport and localization of specific mRNAs (Vessey et al., 2008, Lebeau et al., 2011). These specific mRNAs are essential for spine maturation and synapse function (Yang et al., 2008, Lebeau et al., 2011). Stau2 is essential for the derepression of the Microtubule-associated protein 1b mRNA (Map1b) to activate the metabotropic glutamate receptor (mGluR) long-term depression (LTD) (mGluR-LTD) (Lebeau et al., 2011). These data reinforce the idea that both proteins exhibit non-redundant functions.

The nucleocytoplasmic localization and the RNA binding activity of maStau play a central role in RNP formation (Kiebler and Bassell, 2006). Stau mediates RNP formation and organization through its RNA binding activity, protein-protein interactions as well as dimerization and oligomerization (Marión et al., 1999, Wickham et al., 1999, Luo et al., 2002, Martel et al., 2010, Gleghorn et al., 2013). Different regions of the protein are involved in RNP formation like dsRBD3 and 4 (RNA binding activity) and the SSM together with dsRBD5 (protein dimerization). However, it is still unclear how maStau recognizes its specific targets and how the oligomerization of the protein affects the RNA recognition.

Besides its functions in the nervous system, Stau was shown to be required for the assembly of viral mRNPs (Chatel-Chaix et al., 2004, 2007, de Lucas et al., 2010). Stau also influences the replication, transport, and translation of viral RNA (vRNA) (Chatel-Chaix et al., 2004, 2007, de Lucas et al., 2010, Banerjee et al., 2014, Dixit et al., 2016). During HIV-1 replication, Stau1 binds to the Gag precursor pr55 to regulate the production of new viral particles (Chatel-Chaix et al., 2004). Furthermore, Stau2 increases virus production by interacting with the viral regulatory protein, Rev (Banerjee et al., 2014). During influenza virus infection, Stau1 was shown to bind to the vRNA directly and to interact with the nonstructural protein 1 (NS1) protein of the virus (de Lucas et al., 2010). Stau is also required for the replication of Hepatitis C Virus (HCV) and for the transport of vRNA to facilitate the translation of viral proteins (Dixit et al., 2016).

1.4.3 mRNA stability: Staufen mediated decay

Stau1 has been implicated in the selective degradation of specific mRNAs in a process termed Staufen Mediated Decay (SMD) (Kim et al., 2005). The model is that Stau1 binds to the 3'UTR of specific mRNAs and recruits the ATPase/helicase Upf1, a Nonsense-Mediated mRNA Decay (NMD) component, to induce mRNA decay in a translation-dependent manner (Kim et al., 2007, Gong et al., 2008, Gong and Maquat, 2011, Gleghorn et al., 2013, Park et al., 2012) (Fig. 1.11). The targets of SMD are characterized by the presence of a Staufen Binding Site (SBS), located more than 25 nucleotides downstream of a termination codon in the mRNA 3'UTR. SBS are double-stranded motifs generated either by intramolecular base pairing or intermolecular base-pairing (Kim et al.,

2005, 2007) (Fig. 1.11). Intramolecular base pairing refers to the formation of secondary structures such as stems and loops within the mRNA 3'UTR. In contrast, intermolecular base-pairing is the interaction between a mRNA 3'UTR Alu element and a partially complementary long noncoding RNA (lncRNA) (Kim et al., 2007, Gong and Maquat, 2011).

The first identified physiological target of SMD encodes the human ADP-ribosylation factor 1 *ARF1* (Kim et al., 2007). *ARF1* is a Ras-related G protein that regulates vesicular transport and signaling molecules (Kim et al., 2007). The *ARF1* SBS is located in the first 300 nt of the mRNA 3'UTR. Within this region, a conserved 19 bp stem (nucleotides 75-93 and 212-194) is required for the binding of Stau1. The conservation of this duplex in mammals supports the function of this SBS (Kim et al., 2007). Other SMD target candidates that were found to be upregulated by Stau1 or Upf1 depletion are *c-JUN*, *SERPINE1*, *IL7R* and *GAP43* (Gong and Maquat, 2011).

In a yeast two-hybrid screen, the region of Stau1 that interacts with Upf1 was mapped within the dsRBD4 and the TBD (Kim et al., 2005). Although SMD was initially described for Stau1, it was shown that Stau2 could also bind Upf1 with higher affinity than Stau1 (Park et al., 2012). The binding region of Stau2 to Upf1 differed from the one in Stau1. For Stau2, dsRBD2 and dsRBD3 were mapped in pull-down experiments (Miki et al., 2011). Both protein Stau1 and Stau2 were shown to interact directly and efficiently with the CH domain of Upf1 in an RNA-independent manner to trigger mRNA decay (Park et al., 2012). However, these results are controversial since other studies in HEK293F cells, neuroblastoma cells, myoblast, and primary cortical neurons showed that Stau1 and Stau2 promote stabilization of its mRNA targets rather than their degradation (Miki et al., 2011, Ravel-Chapuis et al., 2012, Heraud-Farlow et al., 2013, Sugimoto et al., 2015, Parks et al., 2017).

The physiological function of SMD is based on cell differentiation. This function was tested in myogenesis by the differentiation of C2C12 Myoblasts (MBs) to Myotubes (MTs). Therefore, Stau1 was shown to down-regulate the SMD target *PAX3* mRNA that encodes a transcription factor responsible for the inhibition of MBs differentiation (Gong et al., 2008). Similarly, SMD regulates adipogenesis. During adipogenesis, Stau1 down-regulates the Krüppel-like factor 2 (*KLF2*) mRNA, which encodes the antiadipogenic factor KLF2 that inhibits the differentiation process of preadipocytes into adipocytes (Cho et al., 2012). Cho et al. (2012) additionally found that the proline-rich nuclear receptor coregulatory protein 2 (PNCR2) is involved in SMD. The authors then proposed the following SMD model, to induce SMD, Stau1 recruits Upf1. Once Upf1 is bound, it gets phosphorylated by the SMG1 protein and finally binds to PNCR2. Upf1 binds Stau1 and PNCR2 with its N-terminal region and with the C-terminal region respectively. The phosphorylated Upf1 acts as an adaptor protein between Stau1 and PNCR2. Finally, the bound PNCR2 induces decapping activity and 5' to 3' exonucleolytic activity (Cho et al., 2012).

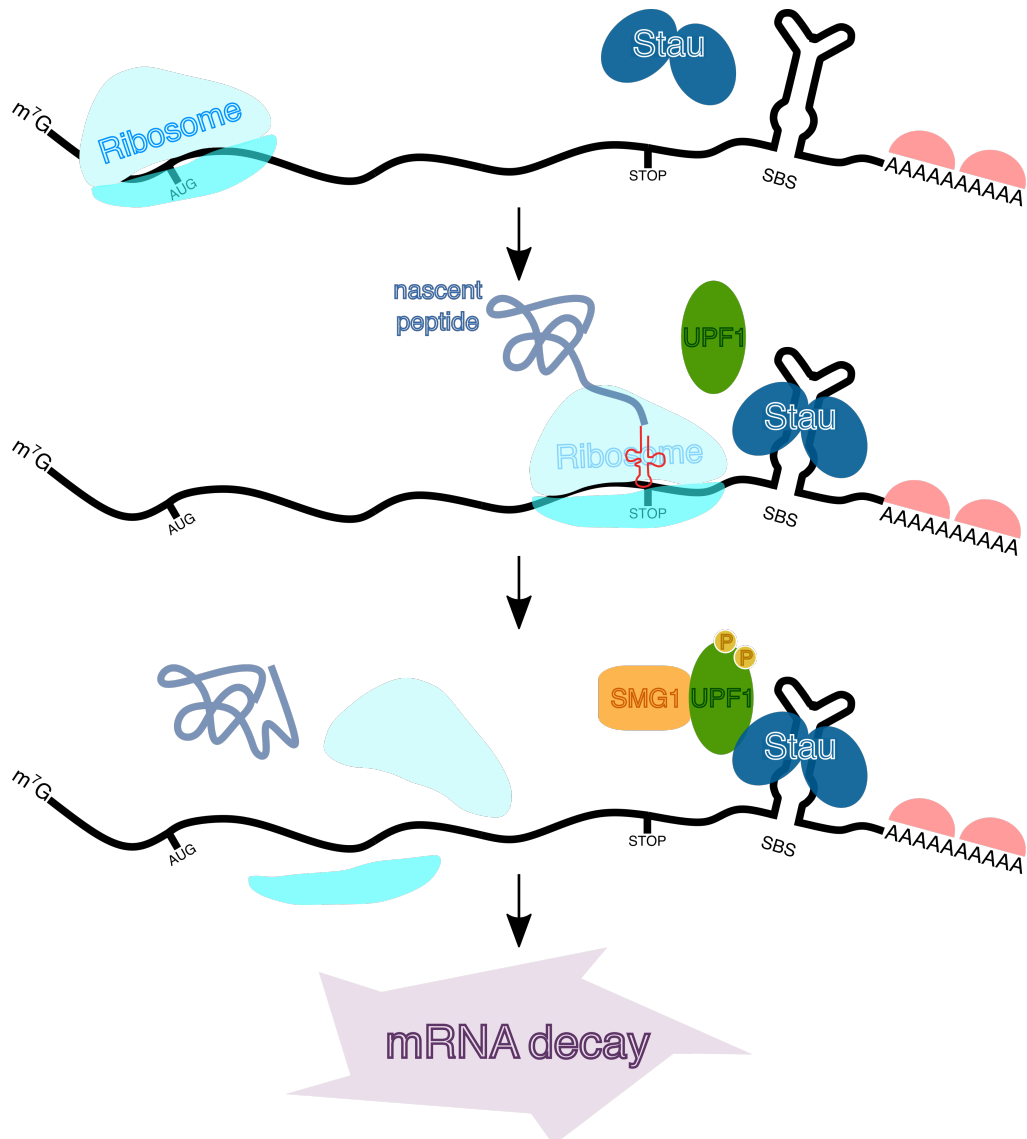


Figure 1.11. The Staufen mediated decay (SMD) The SMD is a translational dependent mRNA decay pathway. SMD activation is achieved by the binding of Stau1 or Stau2 (deep blue) to an SBS (25 nucleotides downstream of a termination codon) present in the 3'UTR of a specific mRNA. Stau1 then recruits Upf1. Upf1 is phosphorylated by SMG1 to activate the exonucleolytic RNA degradation pathway.

1.5 General properties of Staufen associated RNA targets

The function of Stau in PTGR has been described in different organisms and cell lines. During this process, several mRNAs have been identified to associate with Stau proteins (Johnston et al., 1992, Ferrandon et al., 1994, Mallardo et al., 2003, Kim et al., 2005, Furic et al., 2007, Kim et al., 2007, LeGendre et al., 2012, Heraud-Farlow et al., 2013, Laver et al., 2013, Ricci et al., 2013, de Lucas et al., 2014, Sugimoto et al., 2015). The identification of these mRNAs had been crucial in order to understand better the regulation mediated by Stau in different cell types and also helped to elucidate

novel Stau functions (Ferrandon et al., 1994, Mallardo et al., 2003, Kim et al., 2005, Furic et al., 2007, Kim et al., 2007, LeGendre et al., 2012, Heraud-Farlow et al., 2013, Laver et al., 2013, Ricci et al., 2013, de Lucas et al., 2014). Despite the identification of several mRNA targets in different organisms and cell lines, there are only a few well-characterized physiological targets: the three stem-loop regions III, IV, and V in the *bcd* 3'UTR mRNA in *Dm* and the 19 bp stem in the 3'UTR of the *ARF1* mRNA in humans (Johnston et al., 1992, Ferrandon et al., 1994, Kim et al., 2005, 2007).

In recent years, different studies have contributed to determine the sequence-binding preferences of Stau and to identify new sets of targets for Stau in different species and with different conditions. For this, different approaches have been implemented such as RNA coimmunoprecipitations (RIPs) followed by microarray analysis (RIP-Chip) or RIP in tandem with deep sequencing (RIPiT) (Ricci et al., 2013, Laver et al., 2013). Also, the novel technique RNA hybrid and individual-nucleotide resolution ultraviolet cross-linking and immunoprecipitation (hiCLIP) (Sugimoto et al., 2015) for the identification of RNA structural preferences of Stau. The results of these studies show good correlation with the already identifies sequences but also high variation, due to the different expression levels of Stau as well as different techniques and variation in the experimental conditions.

One of the first transcriptome-wide studies in *DmStau in vivo* identified three types of Staufen recognized structures (SRS) common in *DmStau* targets. All three stems were formed in *cis* and had a length between 19 and 12 with few mismatched bases, no unpaired bases or a small number of unpaired bases and short internal loops (Laver et al., 2013). Additionally, the authors used the available RIP microarray data from native mammalian Stau1 and Stau2 associated mRNAs to identify similar SRS (Furic et al., 2007). However, similar structural motifs could not be identified apart from the already described *ARF1* mRNA 19 bp stem (Laver et al., 2013).

Studies in human tissue-cultured cells identified two Stau1 common structure classes present in SBS; one class are stable secondary structures, such as inverted Alu sequences (Alu targets) or structures with high secondary structure formation propensity (non-Alu targets) (Ricci et al., 2013). These elements are found in extended 3'UTRs and can be present in the same transcript to induce simultaneous binding of multiple Stau1 molecules. The second class structure is more labile and present in GC-rich coding sequences (CDS), in this region Stau1 can transiently interact with the active ribosome (Ricci et al., 2013).

Recently, Sugimoto et al. (2015) conducted hiCLIP from cytoplasmic extracts of 293 cells in order to identify RNA duplexes bound by Stau1. The Stau1-hiCLIP analysis identified RNA duplexes from which half corresponded to mRNAs while a significant portion corresponded to ribosomal RNAs (rRNAs), in agreement with previous studies that indicated an interaction of Stau1 with translating ribosomes (Luo et al., 2002, Ricci et al., 2013). The identification of the SBS *ARF1*

mRNA 19 bp stem accomplished part of the hiCLIP data validation. Furthermore, the authors found other duplex regions within the *ARF1* 3'UTR. Interestingly, no significant enrichment of Alu elements was found. Sugimoto et al. (2015) concluded that this is an indication of predominant Stau-Alus interactions in the nucleus.

Different regions in the same RNA species were mapped to contain duplexes that bind to Stau1, in agreement with the preference of Stau1 for intramolecular interactions (Sugimoto et al., 2015). Nonetheless, the hiCLIP analysis identified some negligible intermolecular interactions with long non-coding RNAs. hiCLIP duplexes were found primarily in the 3'UTR region, follow of CDS region and a small portion in the 5'UTR region. The majority of the Stau1-bound duplexes identified by hiCLIP showed an average length between 5 to 14 nt, remarkably each strand incorporated stretches only of purines or pyrimidines. The duplex length was also decisive for the base content, GC pairs were more common in short duplexes whereas in long duplexes the common was AU pairs (Sugimoto et al., 2015).

Interestingly, the presence of Stau1-bound duplexes in a specific region affects the translational state of a transcript (Sugimoto et al., 2015). By ribosome profiling it was demonstrated that depletion of Stau1 increased the abundance of transcripts harboring duplexes at the 3'UTR region without affecting translational efficiency. In contrast, the abundance of transcripts with duplexes at the CDS region remained constant meanwhile translational efficiency increased (Sugimoto et al., 2015). Revealing the structural and sequence-binding preferences of Stau in different species and cell lines had been crucial to understand better the mechanism used by Stau in post-transcriptional regulation. However, the molecular mechanisms used by Stau to recognize structures of varying lengths and how Stau proteins achieve specific mRNA target recognition remain unclear.

2 Aim of this thesis

The Staufen protein was one of the first dsRBP identified. Since then, different studies have elucidated the vital role of the protein in various aspects of post-transcriptional gene regulation in different species. The functions of Staufen range from transport, localization to degradation of specific mRNAs. Staufen has a flexible multi-domain architecture that allows not only the binding of mRNA but also the interaction with other trans-acting factors. Until today biochemical and biophysical approaches have been used to identify physiological targets of the protein and to give some insights into the molecular determinants of the interactions with dsRNA. Although some physiological targets could be determined, the structural information still limited to only dsRBD3 in complex with an artificial RNA stem-loop. Therefore the aim of my thesis is to gain more structural information over human Staufen1 (hStau1) and by complexing the protein with the physiological target *ARF1* SBS try to extend the knowledge of the molecular mechanisms used by the protein to bind and recognize the targeted mRNA. To achieve this aim, I followed the next strategies:

1. Improve the purification protocol of the recombinantly produced hStau1 Full-Length (FL) protein and truncated hStau1 constructs, to avoid protein aggregation and degradation and to obtain a high protein yield for crystallization experiments.
2. Reconstitution and production of stable complexes of hStau1 FL and truncated hStau1 constructs with the physiological target *ARF1* SBS.
3. Biochemical characterization of the human Staufen1 protein, truncated hStau1 constructs as well as the complexes with the *ARF1* SBS.
4. Crystallization and structural characterization of the complexes that could be reconstituted.

3 Results

3.1 hStau1 FL solubility and stability *in vitro* are influenced by ionic strength conditions

3.1.1 Purification of hStau1 FL protein

To understand the mechanism of hStau1 FL regulation and to characterize the mechanisms of substrate recognition structurally, it is necessary to produce a considerable amount of pure protein. For this, different fusion tags and conditions were used in test expression experiments. Although hStau1 FL could be well expressed with all tags tested, the His₆-tagged protein was the only one that could be successfully purified.

Due to precipitation and degradation, I developed an ad hoc purification strategy to produce hStau1 FL in high amount and with high final purity. The protocol includes an extensive wash with high salt buffer and ion exchange chromatography (IEX) to remove further nucleic acids contamination. Nucleic acid contaminants contributed to aggregation and precipitation of hStau1 FL and could bias downstream experiments. The ionic strength of the buffer affected the stability of hStau1 FL significantly in the purification process. During dialysis, the protein tended to precipitate when the salt concentration of the buffer was lower than 300mM (Millimolar) Salt. As a final optimization step, hStau1 FL was subjected to size exclusion chromatography (SEC). With this protocol, the protein could be successfully expressed, and produced in a concentration of 10mg of pure protein per liter of cell culture.

The SEC profile of hStau1 FL showed a monodisperse peak, which eluted as a monomer (Fig. 3.1). The purified protein was sent for crystallization trials to the crystallization facility in the Max Planck Institute (MPI) in Martinsried. However, the inherent flexibility of the multidomain nature of hStau1 FL and the weak stability might have hindered crystallization. The recombinantly purified hStau1 FL was used for complex formation experiments and further biochemical analysis.

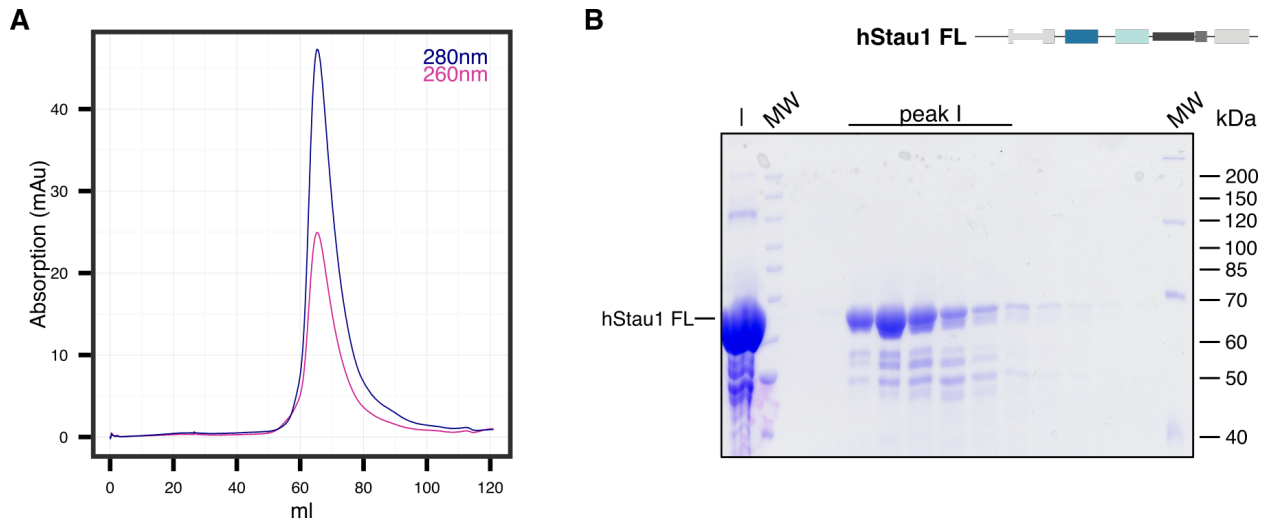


Figure 3.1. Representative final purification step and purity of hStau1 FL. **(A)** Size exclusion chromatography (SEC) profile of the hStau1 FL using a Superdex-200 16/60 GL column. Retention volume: 65 ml. **(B)** Coomassie stained Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel shows the fractions corresponding to the elution peak. Line I: Sample before the SEC run. The size of the Molecular Weight (MW) marker is in kDa.

3.1.2 Purification of hStau1 truncations

Different hStau1 deletions mutants were designed, to improve the stability of the protein and to reduce flexibility for structural characterization. These constructs were designed based on secondary structure predictions and conservation.

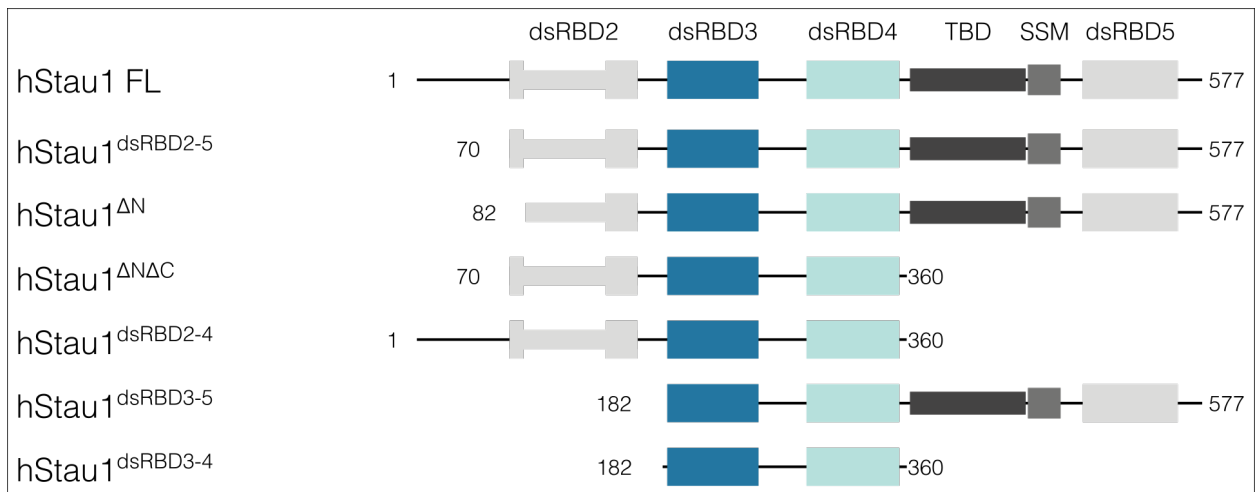


Figure 3.2. Schematic representation of the hStau1 FL and other truncations generated and analyzed in this thesis. Domains are shown as solid shapes and unstructured regions as lines.

The first construct lacks the N unstructured region and contains all dsRBD present in the FL protein (Fig. 3.2 hStau1^{dsRBD2-5}). A second construct was designed based on the short isoform (Uniprot identifier: O95793-2) described for the hStau1 protein. This construct lacks the first 81 amino acids (hStau1^{ΔN}). These two protein fragments were not only poorly expressed but also degradation issues at the second IEX step limited the purification and further biochemical characterization.

A set of truncations were also designed to analyze the oligomerization of the protein and to understand the requirements for mRNA-binding activity. Since SSM and dsRBD5 are the primary determinants involved in hStau1 dimerization (Martel et al., 2010, Gleghorn et al., 2013). A truncation that lacks these regions hStau1^{dsRBD2-4} is predicted to be monomeric (Fig. 3.2). A truncation that consists of dsRBD2, 3 and 4 (hStau1^{ΔNΔC} Fig. 3.2), was also designed with the purpose to remove the unstructured regions that may interfere with crystallization attempts.

Although the role of dsRBD2 in dimerization is not well understood, it has been shown that dsRBD2 alone can interact with the hStau1 FL (Martel et al., 2010). The proline-rich loop insertion present in dsRBD2 may affect mRNA binding activity, as in the case of *DmStaufen* protein (Micklem, 2000). The N terminus of hStau1 together with the dsRBD2 were removed to analyze the effect of the dsRBD2 in oligomerization and mRNA binding (hStau1^{dsRBD3-5} Fig. 3.2). Finally, a construct that included the dsRNA binding determinants dsRBD3 and 4 was obtained (hStau1^{dsRBD3-4} Fig. 3.2).

All fragments were successfully expressed and purified with similar conditions as the Full-Length protein with small variations. The deletions mutants constructs were stable during the whole purification. The ionic strength did not affect the stability of the majority of mutants. Minimal degradation issues were primarily observed in the mutants that contained the C-terminal region (SSM and dsRBD5).

The yield of the purified mutants ranges from 20-30mg of pure protein per liter of cell culture. After SEC, all recombinantly purified proteins showed a high homogeneity. The elutions profiles showed mostly a monodisperse peak (Fig. 3.3). However, in some cases, earlier eluting peaks probably corresponding to aggregation or nucleic acid contaminants were present in the sample. These impurities could be easily separated from the target protein. All recombinantly purified proteins were used for crystallization experiments, biochemical analysis, and interaction studies.

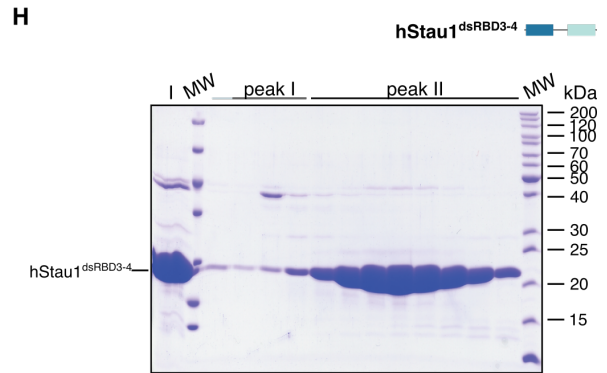
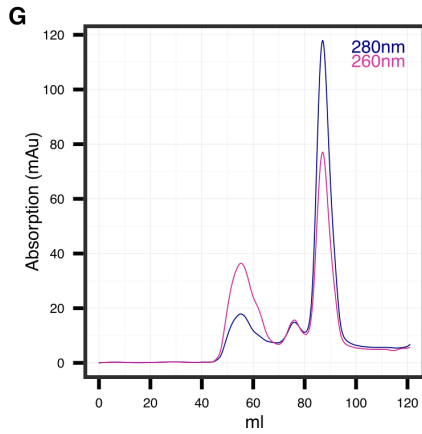
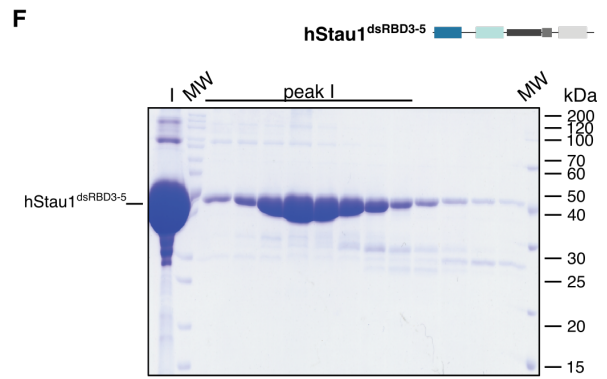
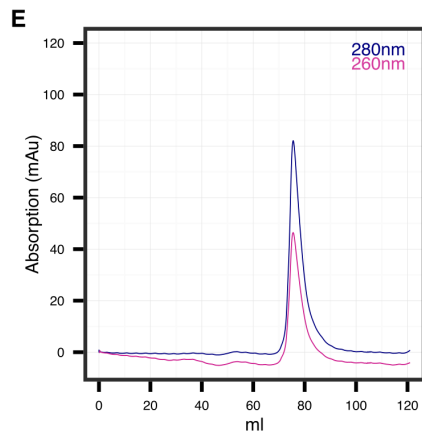
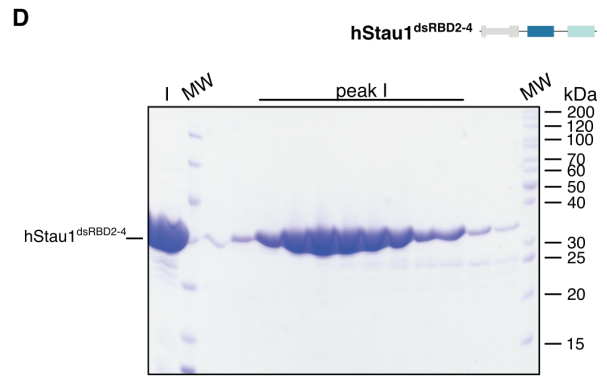
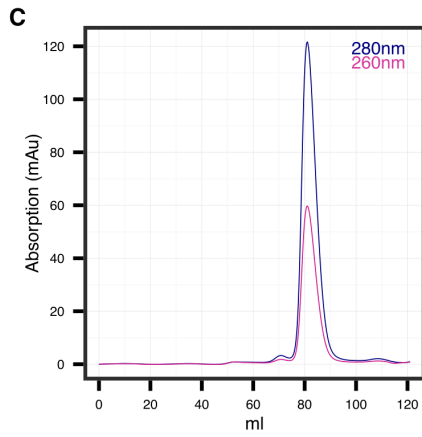
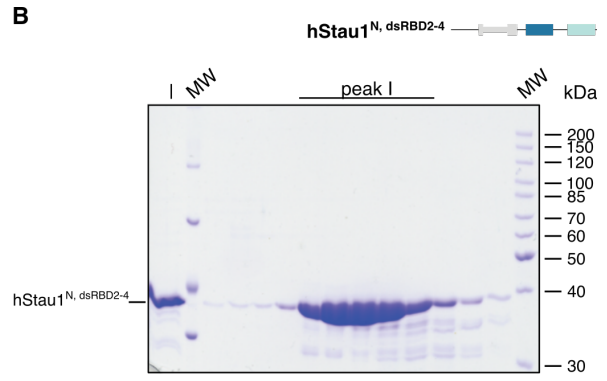
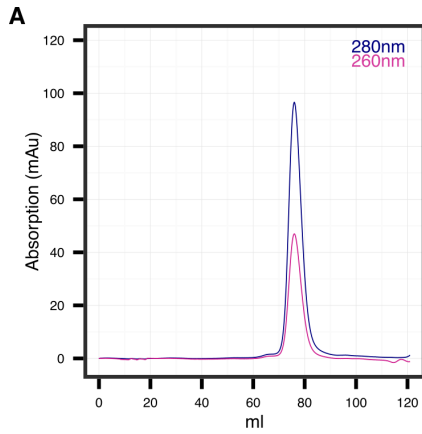



Figure 3.3b. Representative SEC results for the hStau1 truncations. (A, C, E, G) SEC elution profiles of the hStau1 mutants in physiological conditions. (B, D, F, H) Coomassie stained SDS-PAGE of the peak fractions (peak I) of the SEC experiments. Line I: Sample before SEC run. MW marker is in kDa.

3.2 hStau1 dimerization is concentration dependent *in vitro*

3.2.1 SEC-MALLS analysis of hStau1 FL

hStau1 was shown to dimerize via its SSM and dsRBD5 domain (Martel et al., 2010, Gleghorn et al., 2013). Both homophilic interactions between the SSM-dsRBD5 of two molecules but also an interaction of the dsRBD2 with this region through heterophilic interactions have been suggested (Martel et al., 2010, Gleghorn et al., 2013). This dimerization was not consistent with the initial SEC analysis of hStau1 that exhibited an elution profile compatible with a monomeric state. However, further SEC analysis showed altered elution profiles when conditions such as protein concentration or ionic strength varied.

In SEC experiments, when the concentration of the sample exceeded 10 mg/ml, two peaks were present in the chromatogram. Also, this was the case when the salt concentration of the buffer was higher than 150 mM. By analyzing the SEC fractions of the two elution peaks in a SDS-PAGE, two bands were present in the first elution peak (black arrow Fig. 3.5 D). The lower band had a MW that corresponded to the protein monomer (MW: 63 kDa), the second band had a MW between 120 and 150 kDa. The upper band was analyzed by Mass Spectrometry (TOPLAB GmbH). The Mass spectrometry results identified 75% of the human hStau1 protein indicating that the second peak corresponded to a multimer fraction of the protein (Fig. 3.4).



>NP_001309859.1 double-stranded RNA-binding protein Staufen homolog 1 isoform b [Homo sapiens]

1 MSQVQVQVQNPSAALSGSQILNKNQSLLSQPLMSIPSTTSSLPSENAGRPIQNSALPSASITSTSAAAESITPTVELNALCMKLGKMPMY
91 KPVPDYSRMQSTYNYNMRGGAYPPRYFYPFPVPLLYQVELSVGGQQFNGKGKTRQAAKHDAAKALRILQNEPLPERLEVNGRESEEEEN
181 LNKSEISQVFEIALKRNLPNFEVARESGPPHMKNFVTKVSVGEFVGELEGKSKKISKKNAAIAVLEELKKLPLPAVERVKPRIKKKTK
271 PIVKPQTSPEYGGQGINPISRLAQIQAKKEKEPEYTLTERGLPRRREFVMQVKVGNHTAEGTGTNKKVAKRNAENMLEILGFKVPQAQ
361 PTKPALKSEEKTIKPKGDGRKVTFFEPGSGDENGTSNKEDEFMPYLSHQQLPAGILPMVPEVAQAVGVSQGHHTKDFTRAAPNPAKAT
451 VTAMIARELLYGGTSPTAETILKNNISSGHVPHGLTRPSEQLDYLSRVQGFQVEYKDFPKNNKNEFVSLINCSSQPPLISHGIGKDVES
541 CHDMAALNILKLLSELDQQSTEMPRTGNGPMSVCGRC

Figure 3.4. Amino acid sequence of hStau1 FL. The Mass spectrometry analysis identified 75% of the human hStau1 FL. Identified peptides are highlighted in red.

This protein-concentration dependent multimerization effect was also observed in the hStau1^{dsRBD3-5} mutants that contained the dimerization region. For the hStau1^{dsRBD3-5}, the SEC profile showed a broad shoulder at a high molecular weight. Similar to the hStau1 FL, the fractions corresponding to the peak shoulder had two bands. This effect was increased by variation in protein concentration

and the ionic strength. However, SEC data is based on the shape of the molecule and is strongly affected by external experimental conditions (Lathe and Ruthven, 1956, Burgess, 2018). Because of this, SEC coupled with Multi-Angle Laser Light Scattering (SEC-MALLS) was used, to accurately characterize the oligomerization state of the Full-Length protein and the mutants.

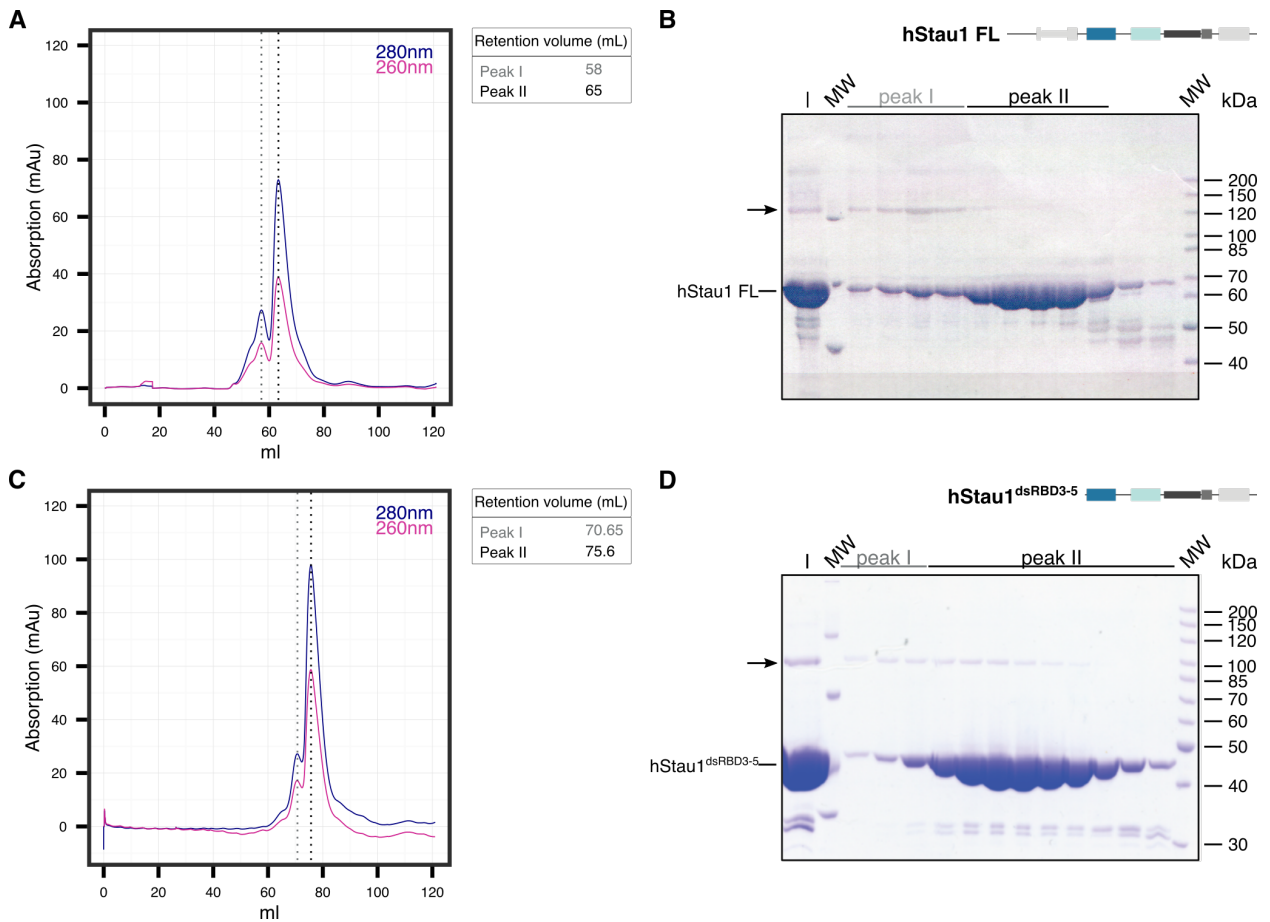


Figure 3.5. Representative SEC analysis of hStau1 FL and hStau1^{dsRBD3-5} in high ionic strength. **(A, C)** SEC profile of the hStau1 FL and hStau1^{dsRBD3-5} in high ionic strength (300mM NaCl buffer) Retention volume peak 1: 58 ml and peak 2: 65 ml. **(B, D)** Coomassie stained SDS-PAGE of the fractions corresponding to the elution peak. Two peaks are present, the first peak is characterized by a second band at 120 kDa for the hStau1 FL and hStau1^{dsRBD3-5} 90 kDa (black arrows). Line I: Sample before SEC run. The size of the MW is in kDa.

Dr. Claire Basquin conducted SEC-MALLS analysis at the MPI in Martinsried. The resulting experimental absolute molecular mass corresponded to a hStau1 FL monomer and a dimer with a MW of 68 ± 6 kDa (theoretical MW 63 kDa) and of 162 ± 10 kDa (theoretical MW 120 kDa), respectively (Fig. 3.6 A). In a second attempt, I performed SEC-MALLS experiments in the Dept. of protein evolution in the MPI in Tübingen. For this, different protein concentrations were tested. By increasing the concentration of the sample, the equilibrium shifted to the dimer population as indicated by the SEC experiments (Fig. 3.6 B).

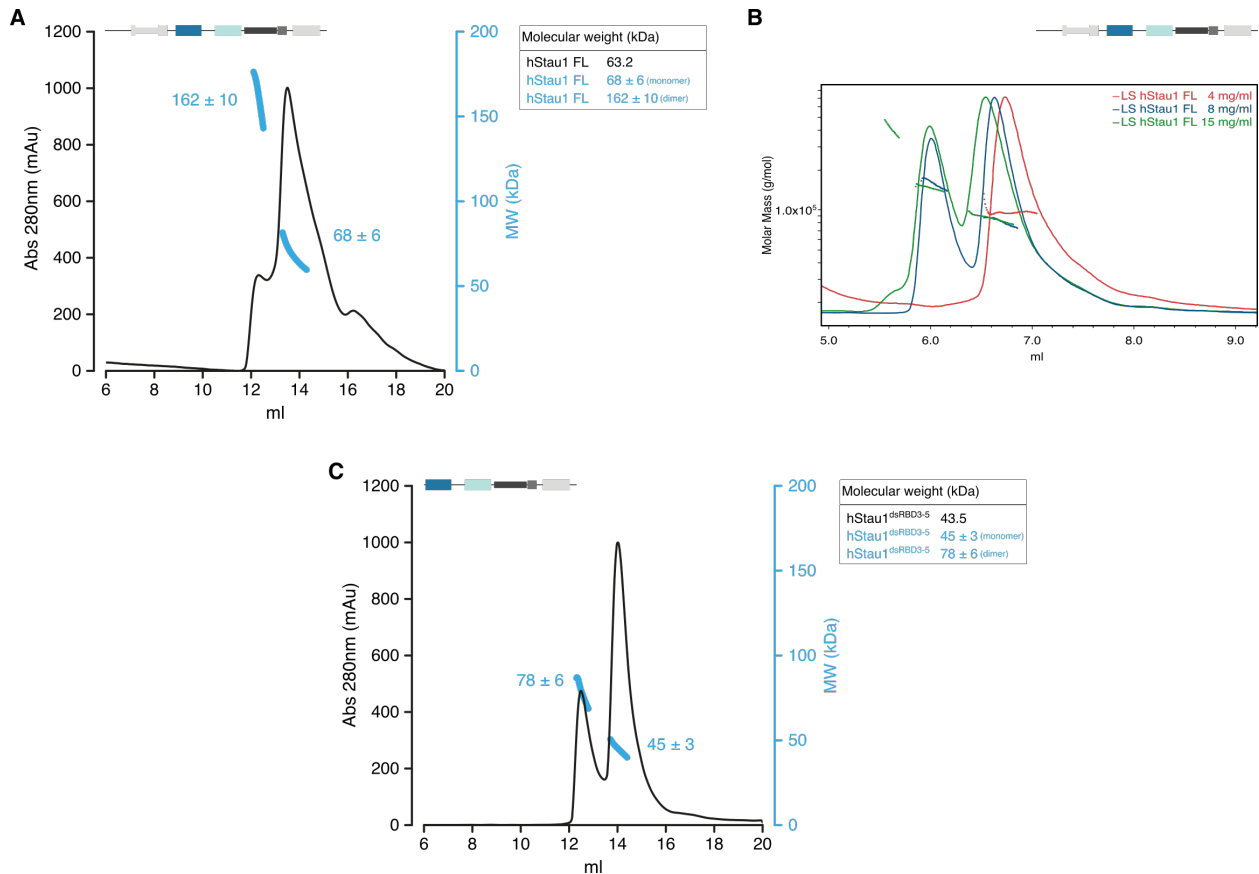


Figure 3.6. Representative SEC-MALLS results of hStau1 FL in two different set ups. **(A)** SEC-profile of hStau1 FL, the experimental calculated molecular mass of the eluted hStau1 FL is indicated in blue. **(B)** Overlay of light scattering data with measured molecular mass of hStau1 FL in three different concentrations: 4 mg/ml (red), 8 mg/ml (blue) and 15 mg/ml (green). Shifting of the elution peak at increasing concentrations of hStau1 FL suggest a concentration-dependent multimerization of the protein. **(C)** SEC-profile of the hStau1^{dsRBD3-5} protein, the calculated molecular mass of the eluted hStau1^{dsRBD3-5} is indicated in blue. Right box: theoretical MW in black and calculated molecular mass in blue.

As the hStau1 FL, the SEC profile of the hStau1^{dsRBD3-5} mutant showed two peaks. The SEC-MALLS elution profile showed a more clear separation of the two populations. The first peak was higher as it had on a regular SEC; the calculated molar masses of the two population were 78 ± 6 kDa for the first peak and 45 ± 3 kDa for the second peak, which corresponded to the mass of monomeric and dimeric hStau1^{dsRBD3-5} (Fig. 3.6 C). In contrast, the SEC-MALLS profile of hStau1^{dsRBD2-4} and hStau1^{dsRBD3-4} exhibited a single symmetric peak as in SEC. The experimental molecular mass of 39 kDa (theoretical mass 39.6 kDa) and 21 ± 1 (theoretical mass 20.2 kDa) for hStau1^{dsRBD2-4} and the hStau1^{dsRBD3-4}, respectively, support the monomeric state of both constructs in solution (Fig. 3.7 A-B). Altogether, these results support the role of the C-terminal region of hStau1, consisting of SSM and dsRBD5, as the main determinant of hStau1 dimerization. Moreover, this data shows the structural heterogeneity of the protein in solution and is consistent with previous studies in which hStau1 can form dimeric or multimeric states in live cells (Martel

et al., 2010).

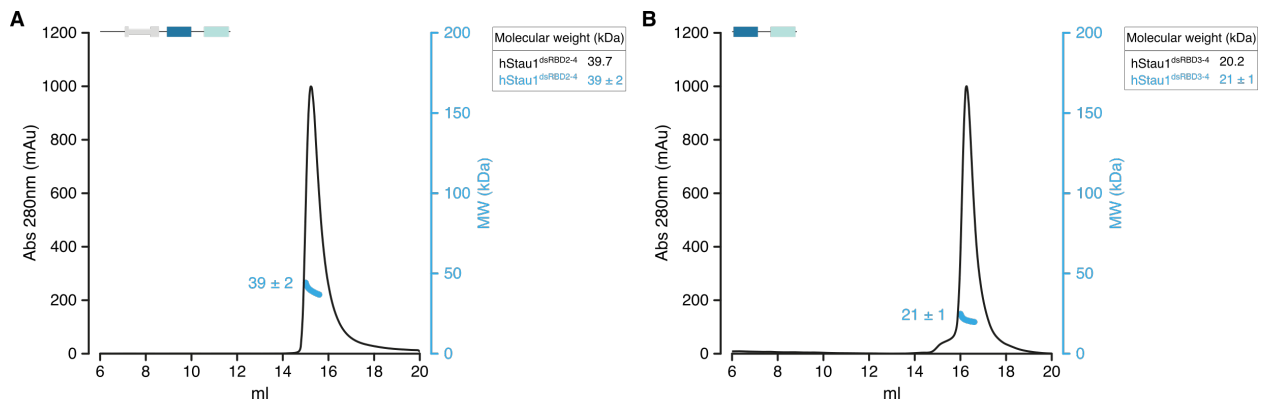


Figure 3.7. Representative SEC-MALLS results of the hStau1 mutants. The experimental calculated molecular mass is indicated in blue. **(A)** SEC-profile of the hStau1^{dsRBD2-4} protein, the calculated molecular mass of the eluted hStau1^{dsRBD2-4} is indicated in blue. **(B)** SEC-profile of the hStau1^{dsRBD3-4} protein, the calculated molecular mass of the eluted hStau1^{dsRBD3-4} is indicated in blue. Right box: theoretical MW in black and calculated molecular mass in blue.

3.3 hStau1 FL binds the *ARF1* SBS with high affinity *in vitro*

3.3.1 Reconstitution of the hStau1 FL - *ARF1* SBS complex

Studies have identified the different mRNA features that are likely to be bound by the protein (Johnston et al., 1992, Ferrandon et al., 1994, Kim et al., 2005, 2007, Laver et al., 2013, Ricci et al., 2013, Sugimoto et al., 2015). However, few candidates have been determined to bind specifically to the Staufen protein (Johnston et al., 1992, Ferrandon et al., 1994, Kim et al., 2005, 2007). At present, the best-described mRNA targets are the stem-loop regions III, IV, and V in the 3'UTR of the *bcd* mRNA for the *Dm* protein (Ferrandon et al., 1994) and the 19 bp stem in the 3'UTR of the *ARF1* mRNA for hStau1 (Kim et al., 2005, 2007).

Secondary structure prediction of the entire 3'UTR of the *ARF1* mRNA shows that the region mapped to bind hStau1 specifically, corresponds to a stem that forms between nucleotides 75-93 and 194-212 (Kim et al., 2005, 2007). Some relevant characteristics of the *ARF1* SBS are the presence of two wobble pairs and the high content of G-C pairs (G-C content: 53%, A-U content: 37% and the wobble G-U 2 10%). To test the mRNA binding capability of hStau1 FL *in vitro* I used the *ARF1* SBS. The two RNA strands were synthesised (Integrated DNA Technologies, Inc. (IDT)), resuspended in water and subsequently annealed. SEC experiments were performed to verify the homogeneity of the annealed *ARF1* SBS. The elution profile exhibited a monodisperse peak, analysis of the peak fractions showed the presence of a single band with no degradation products (Data not shown). Therefore, I could confirm the homogeneity of the annealed *ARF1* SBS.

The hStau1 FL - *ARF1* SBS complex was reconstituted in high ionic strength conditions, because of high precipitation, and subjected to SEC. The hStau1 FL- *ARF1* SBS complex eluted as a monodisperse peak. The peak shifted to higher molecular weight as compared with the elution profile of hStau1 unbound, and the 260/280 ratio was increased. The SEC profile was an indication of complex formation. SDS-PAGE analysis of the peak fractions showed a single band with some degradation products (Fig. 3.8). It was not possible to analyze the dimerization behavior of hStau1 FL on RNA-binding activity with low ionic strength conditions. These results indicated that hStau1 FL stably interacts with the *ARF1* SBS *in vitro* even in high salt condition. The resulting complex was set up for crystallization, but no hits could be observed.

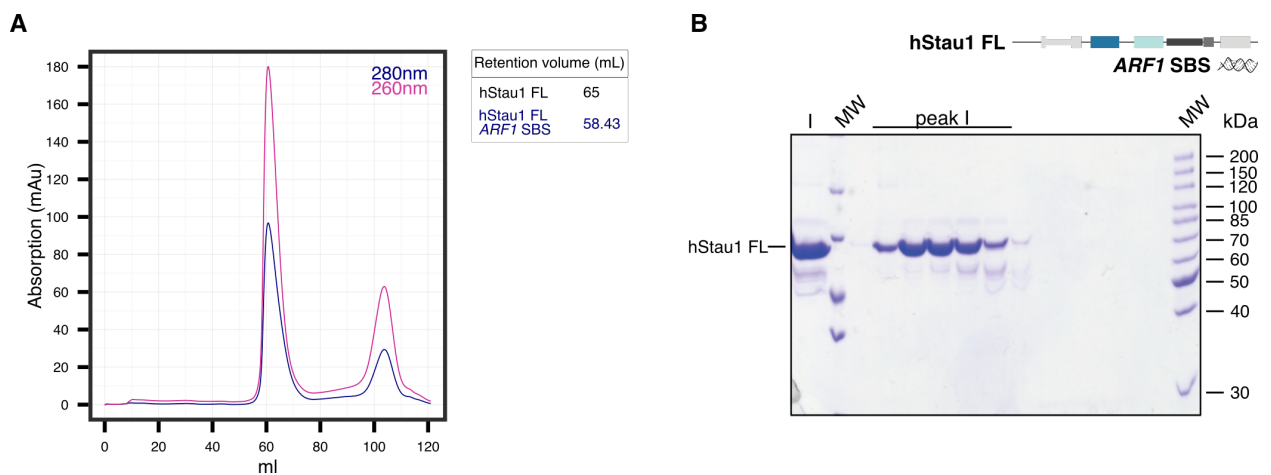


Figure 3.8. Representative SEC result of the complex of hStau1 FL with the *ARF1* SBS. **(A)** The chromatogram shows the elution profile of the complex. Retention volume of hStau1 FL apo (black) and the complex with the *ARF1* SBS (blue) are indicated in the table. The complex elutes as a single peak, the earlier elution volume together with, the higher 260nm absorbance confirms the presence of hStau1-*ARF1* SBS complex. The second peak corresponds to the RNA excess. **(B)** The Coomassie-stained SDS-PAGE shows the fractions that correspond to peak I and the second peak. Line I: Complex sample before SEC run. The size of the MW is in kDa.

3.3.2 Electrophoretic mobility shift assays of hStau1 FL with 5'FAM *ARF1* SBS_{20bp}

Electrophoretic mobility shift assays (EMSA) experiments were performed to verify the reconstitution of the complex and to determine the binding stoichiometry of the complex. For this, *ARF1* SBS labelled with a 5'-6-FAM fluorescein was used. One nucleotide had to be added to the sequence because of the fluorescence quenching by the adjacent guanosine of the *ARF1* SBS sequence (referred to as *ARF1* SBS_{20bp}). Different buffer systems were tested. The best running conditions were up to pH 9.8-10 because of the high pI of the protein (pI: 9.46).

The recombinantly purified hStau1 FL was able to shift the labeled *ARF1* SBS_{20bp}. One single band was present in the non-denaturing gel (Fig. 3.9 A). This band could correspond to one molecule hStau1 and one molecule *ARF1* SBS_{20bp}. Increasing the concentration of the protein did not affect band shifting of the complex. This result confirmed the hStau1 FL - *ARF1* SBS complex formation and was consistent with the SEC experiments. The stoichiometry of the hStau1 FL - *ARF1* SBS complex was analyzed by SEC-MALLS analysis. However, SEC-MALLS experiments failed because of the instability of the complex.

3.3.3 Fluorescence anisotropy of hStau1 FL and 5' FAM *ARF1* SBS_{20bp}

To assess the binding affinity of hStau1 FL to the *ARF1* SBS_{20bp} Dr. Daniela Lazzaretti conducted Fluorescence anisotropy (FA) measurements. For this, a range of concentrations of the recombinantly purified hStau1 FL was titrated against a constant concentration of labeled *ARF1* SBS_{20bp}. The binding of the protein changed the fluorescence anisotropy of the label *ARF1* SBS_{20bp}, and these changes were measured. By fitting the data, it was possible to calculate the dissociation constant (K_D). hStau1 FL bound the *ARF1* SBS_{20bp} with a K_D of 6 ± 1 nM (Fig. 3.9 B). Different RNAs substrates were tested to analyze the importance for binding of secondary structures in the RNA-binding activity of hStau1 FL.

Although hStau1 does not bind ssRNA *in vivo*, FA experiments showed, that hStau1 FL could bind single-stranded species though with weak affinity compared with dsRNA ($K_D > 100$ nM). The UG₁₀ oligonucleotide ($K_D 36 \pm 10$ nM) bound hStau1 with high affinity due to the formation of secondary structures like G-tetrad and quadruplex (Burge et al., 2006) and stabilized by the presence of monovalent cations.

The presence of ssRNA elements or wobble base pairs in the *ARF1* SBS did not affect the affinity (*ARF1* SBS_{20bp} K_D perfectly complementary (pc) 6 ± 2 nM, *ARF1* SBS_{20bp+hairpin} $K_D 7 \pm 0$ nM). Shortening the length of the *ARF1* SBS stem reduced the affinity slightly (14 bp *ARF1* mRNA $K_D 7 \pm 2$ nM). Other mRNA substrates were used, to test the mRNA binding activity of hStau1 FL. For this, we chose already characterized targets of the *Dm*Staufen protein. The 12 bp artificial stem-loop described to optimally bind to the dsRBD3 of the *Dm*Stau (referred as GC-rich hairpin in Fig. 3.9 B) (Ramos et al., 2000) showed a similar affinity to hStau1 FL as the *ARF1* SBS_{20bp} (K_D of 6 ± 1 nM). Next, we tested the stem-loop III and Vb of the *bcd* 3'UTR, these two regions were identified as the binding region of *Dm*Staufen. Both stems have complex secondary structures, and also a high GC content. hStau1 FL showed a higher affinity for the stem-loop III (*bcd* III K_D of 9 ± 2 nM) in comparison with the stem loop Vb (*bcd* Vb K_D of 12 ± 3 nM). A possible explanation could be the higher GC content in the stem-loop III in comparison with stem Vb.

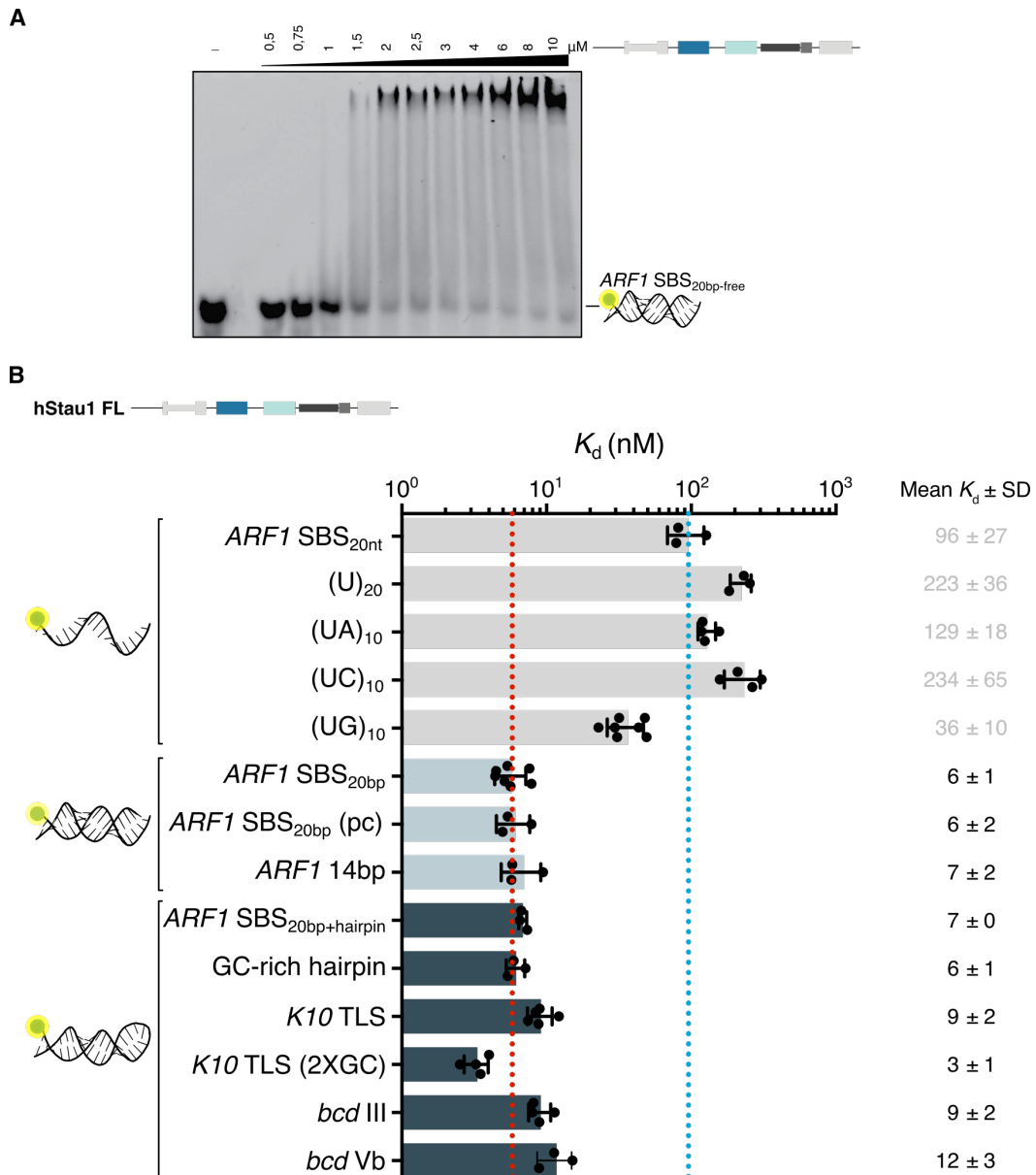


Figure 3.9. RNA-binding assays with hStau1 FL and the *ARF1* SBS_{20bp}. **(A)** EMSA analysis of the hStau1 FL interaction with the fluorescently labeled *ARF1* SBS_{20bp}. **(B)** The bar graph comparing the calculated binding affinity of the hStau1 FL for single (*ARF1* SBS 20nt sense strand) or double-stranded RNA. The calculated K_D and their corresponding error (error bars) are the mean and the standard deviations (SD) of a minimum of three experiments. (Mean K_D and SD in nM).

All substrates tested so far contained a high number of GC pairs (> 7GC). The RNA-binding affinity of hStau1 FL was further characterized by using the *K10* Transport and Localization Sequence (TLS) RNA that included secondary structures as well as lower GC content in the sequence. The *K10* TLS is a *Dm* mRNA required for the transport and localization of *K10* transcripts within the oocyte (Serano and Cohen, 1995). The *K10* TLS is a stem-loop composed of an 8nt loop and a 17 bp stem in which two nucleotide bulges are present; the AU content predominates

against the GC content (2 GC pairs) (Serano and Cohen, 1995). hStau1 FL bound the *K10* TLS with a K_D of 9 ± 2 nM. Even though *K10* TLS is not a physiological target of hStau1 FL, the relatively high affinity of hStau1 FL for *K10* TLS suggested that secondary structures within the mRNA play a considerable role for the hStau1 mRNA-binding activity. Finally, two GC pairs were added at the end of the *K10* TLS stem to determine the influence of GC pairs in RNA-binding affinity. As a result, the affinity improved remarkably with the K_D increasing by three-fold (Fig. 3.9 B).

Altogether, the SEC experiments with the biochemical characterization indicated that hStau1 binds the *ARF1* SBS with high affinity *in vitro*, and supported the idea that the GC content is a primary determinant for the binding rather than secondary structures elements in the mRNA substrate.

3.4 hStau1 RNA-binding activity induces dimerization independent of dsRBD2 and the dimerization region

The hStau1 truncations had been used in the previous section to analyze and characterize the biochemical and structural properties of the RNA-binding activity of the hStau1 protein.

3.4.1 Reconstitution of the hStau1^{dsRBD2-4}-*ARF1* SBS complex

The hStau1^{dsRBD2-4} was used for complex formation with the *ARF1* SBS. For this, hStau1^{dsRBD2-4} was mixed with a 1:1.2 *ARF1* SBS excess. After addition of the *ARF1* SBS, hStau1^{dsRBD2-4} tended to precipitate. The precipitation could be in part resolved by adding 300mM salt. Still, the majority of the complex precipitated. The resulting complex was then injected in a SEC column.

hStau1^{dsRBD2-4} could bind the *ARF1* SBS in SEC experiments. The elution profile showed a significant increased 260/280 ratio at the main peak that shifted to high molecular weight (Fig. 3.10). However, the main peak had a shoulder on the right side, the small peak at the end of the chromatogram represented the excess of mRNA. By analyzing the peak fractions on a SDS-PAGE, no contamination or degradation products were present, and the hStau1^{dsRBD2-4} band extended through the whole main peak. This result suggested that hStau1^{dsRBD2-4} could bind the *ARF1* SBS *in vitro*, despite the lack of the dimerization region. Nevertheless, the formed hStau1^{dsRBD2-4} - *ARF1* SBS complex was not homogeneous.

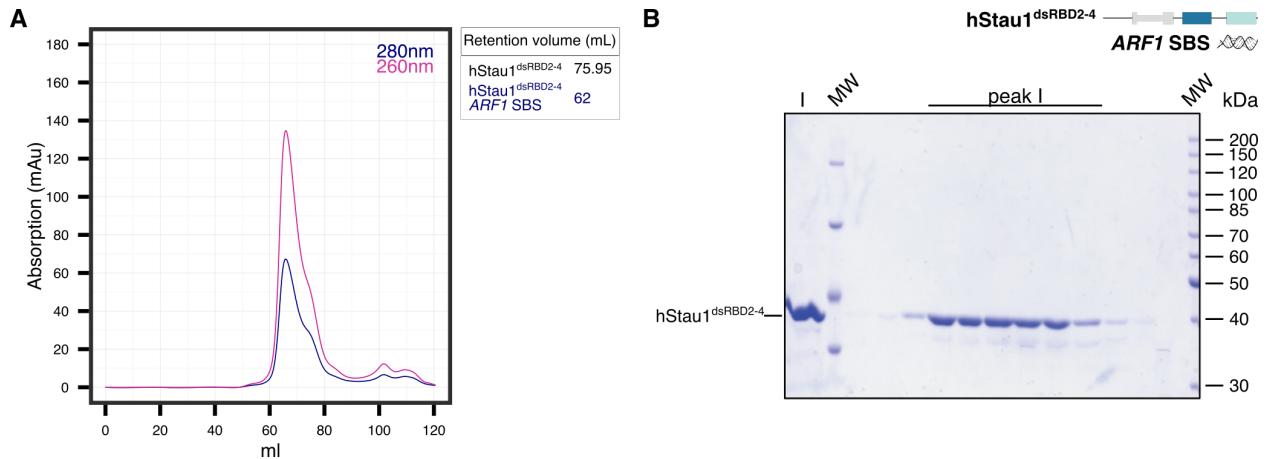


Figure 3.10. Representative SEC result for the hStau1^{dsRBD2-4} - ARF1 SBS complex. **(A)** SEC elution profile of the hStau1^{dsRBD2-4} - ARF1 SBS complex. Retention volume of the hStau1^{dsRBD2-4} apo (black) and the complex with the ARF1 SBS (blue) are indicated in the table **(B)** The Coomassie-stained SDS-PAGE shows the peak fractions (peak I) of the SEC. Line I: Complex sample before SEC run. The size of the MW is in kDa.

3.4.2 Biochemical characterization of the hStau1^{dsRBD2-4} - ARF1 SBS_{20bp} complex

EMSA assay was performed to confirm the formation of the hStau1^{dsRBD2-4} - ARF1 SBS_{20bp} complex and to determine the stoichiometry. hStau1^{dsRBD2-4} formed two distinct complexes at 1:1 and 1:2 (Fig. 3.11 B). Since hStau1^{dsRBD2-4} was shown to be monomeric in solution, the first shift might correspond to one molecule hStau1^{dsRBD2-4} bound to one molecule ARF1 SBS_{20bp}; the second shift might correspond either to one protein fragment binds to two ARF1 SBS_{20bp} molecules or possibly to one molecule ARF1 SBS_{20bp} binds to two protein molecules.

SEC-MALLS analysis was performed to clarify the molecular mass of the complexes. In contrast to the EMSA experiments, only one complex was present. The calculated molecular weight for the hStau1^{dsRBD2-4} - ARF1 SBS_{20bp} complex is 62 ± 4 kDa that corresponded to one protein molecule (MW: 39,7 kDa) and two ARF1 SBS_{20bp} molecules (MW: 12 kDa) (Fig. 3.11 A). This result was consistent with the second shift present in the EMSA experiment.

FA measurements showed that hStau1^{dsRBD2-4} bound to the ARF1 SBS_{20bp} with a K_D of 10 ± 2 nM, in comparison to the Full-Length protein the affinity was reduced only by two-fold (Fig. 3.11 C). This result indicated that the SSM, TBD, and dsRBD5 do not have a direct effect in mRNA binding activity, but rather might be required to stabilize the protein in order to fulfill the mRNA binding activity. Crystallization experiments were conducted to characterize the complex further, but no crystals could be obtained.

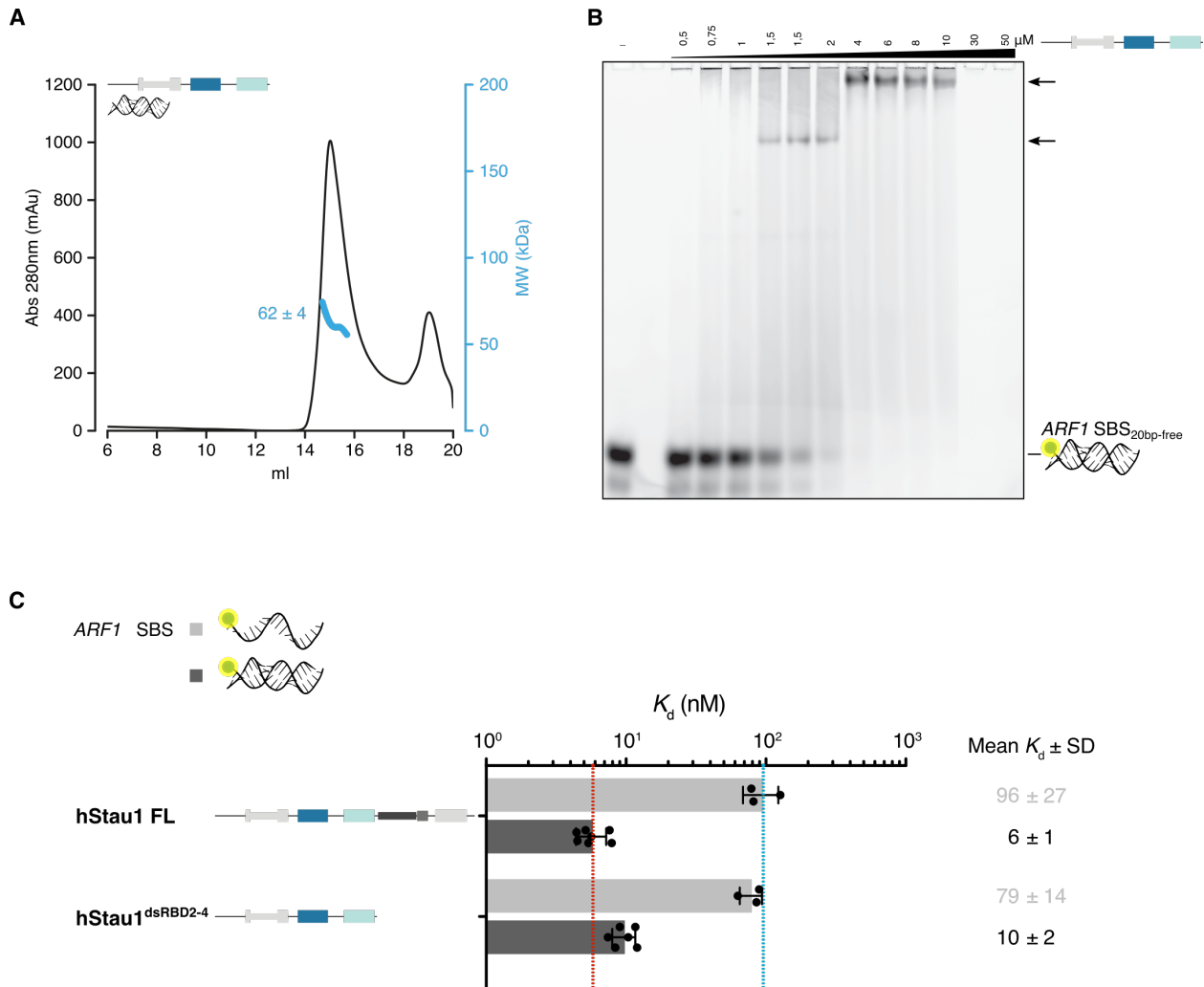


Figure 3.11. hStau1^{dsRBD2-4} formed stable complexes with ARF1 SBS *in vitro* (A) SEC-MALLS elution profile of hStau1^{dsRBD2-4} in complex with the ARF1 SBS, the calculated molecular mass is indicated in blue (B) EMSA analysis of the hStau1^{dsRBD2-4} interaction with the fluorescent label ARF1 SBS_{20bp}. The protein shifted the label ARF1 SBS_{20bp} and formed two RNA complexes (black arrows). (C) The bar graph shows the calculated binding affinity of the hStau1 FL and hStau1^{dsRBD2-4} for the single ARF1 SBS_{20bp} sense strand (light gray) and the double-stranded ARF1 SBS_{20bp} (dark gray). The calculated K_D and their corresponding error (error bars) are the mean and the SD of a minimum of three experiments. (Mean K_D and SD in nM).

3.4.3 Reconstitution of the hStau1^{ΔNΔC}- ARF1 SBS complex

The reconstitution of the hStau1^{ΔNΔC} - ARF1 SBS complex had been performed as described above using a 1:1.2 molar ratio (protein - RNA). Massive precipitation could be observed and persisted despite various optimization attempts that include: incubation at RT or 4°C, resuspension of both components in a high salt buffer, addition of glycerol or magnesium and dilution of the protein before the addition of the ARF1 SBS with subsequent complex formation and concentration. A final reconstitution attempt was performed in high salt buffer (300mM KCl). The reaction mixture

was centrifugated and applied to an SEC column.

The hStau1^{ΔNΔC} - *ARF1* SBS complex eluted as two separate peaks. The main peak shifted to the left, and the 260/280 ratio was slightly increased suggesting RNA incorporation. The SDS-PAGE showed a band corresponding to hStau1^{ΔNΔC} in the first eluted peak (Fig.3.12). Moreover, attempts to concentrate the fraction of the main peak failed because of the persistent heavy precipitation. Finally, due to precipitation no further characterization could be done with the hStau1^{ΔNΔC} - *ARF1* SBS complex.

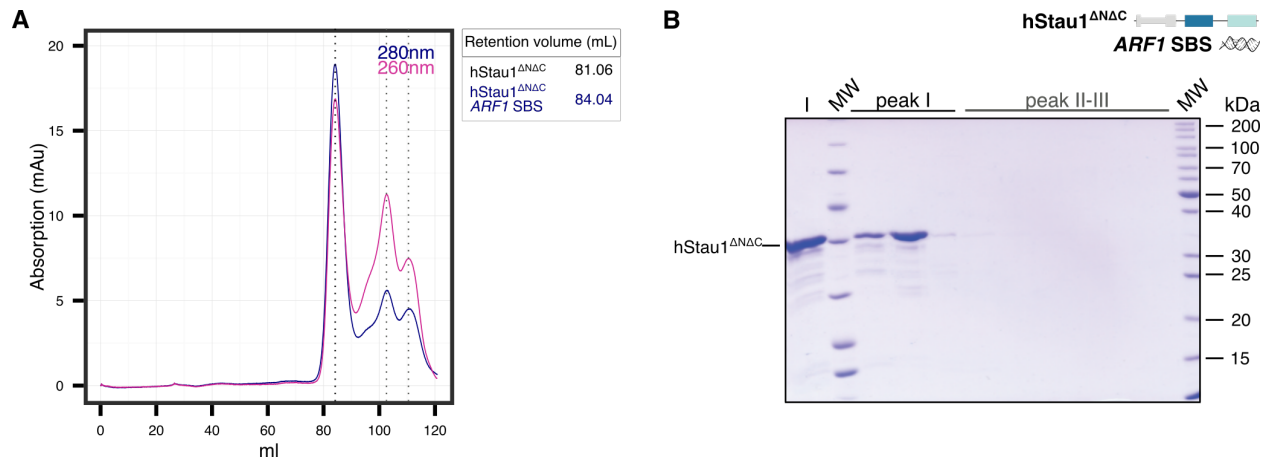


Figure 3.12. Representative SEC result of the hStau1^{ΔNΔC}-*ARF1* SBS complex. **(A)** SEC profile of the complex. Retention volume of the hStau1^{ΔNΔC} apo (black) and the complex with the *ARF1* SBS are indicated in the table (blue). There is no significant increase in the absorbance at 260 nm in the first peak. **(B)** The Coomassie-stained SDS-PAGE shows the fractions that correspond to the peak I containing the protein, peak II and III correspond to the unbound *ARF1* SBS. Line I: Sample before SEC run. The size of the MW is in kDa.

3.4.4 Reconstitution of the hStau1^{dsRBD3-5} - *ARF1* SBS complex

The recombinantly purified hStau1^{dsRBD3-5} was mixed with the *ARF1* SBS in a ratio of 1:1.2, heavy precipitation issues persisted after the components were mixed. The precipitation could be in part overcome by resuspending hStau1^{dsRBD3-5} in 300mM KCl before the *ARF1* SBS was added; in this way, the precipitation could be in part overcome. After 1h incubation on ice, the complex was centrifugated and subjected to SEC.

SEC experiments were performed with two different ionic strength conditions (100mM and 300mM NaCl), to test the effects of the ionic strength dependent multimerization of hStau1^{dsRBD3-5} (referred in section 3.2) in mRNA binding activity. In contrast to the FL protein, different ionic strength conditions could be used to test the capability of the hStau1^{dsRBD3-5} fragment to bind the *ARF1* SBS. The resulting profiles showed that hStau1^{dsRBD3-5} could bind *ARF1* SBS in the two ionic strength, but the stability of the complex is negatively affected by low ionic strength conditions. In

both cases, the elution profile showed a broad peak that shifted to higher molecular weight. Further, a significant increase of the 260/280 ratio confirmed complex formation in both cases (Fig. 3.13). SDS-PAGE analyses showed that hStau1^{dsRBD3-5} was present in all fractions of the peaks. The fractions corresponding to the main peak showed a second band that might correspond to the dimer. This result suggested that hStau1^{dsRBD3-5} could form a stable complex with the ARF1 SBS in *vitro* and the stability of the complex improve in a high ionic strength condition.

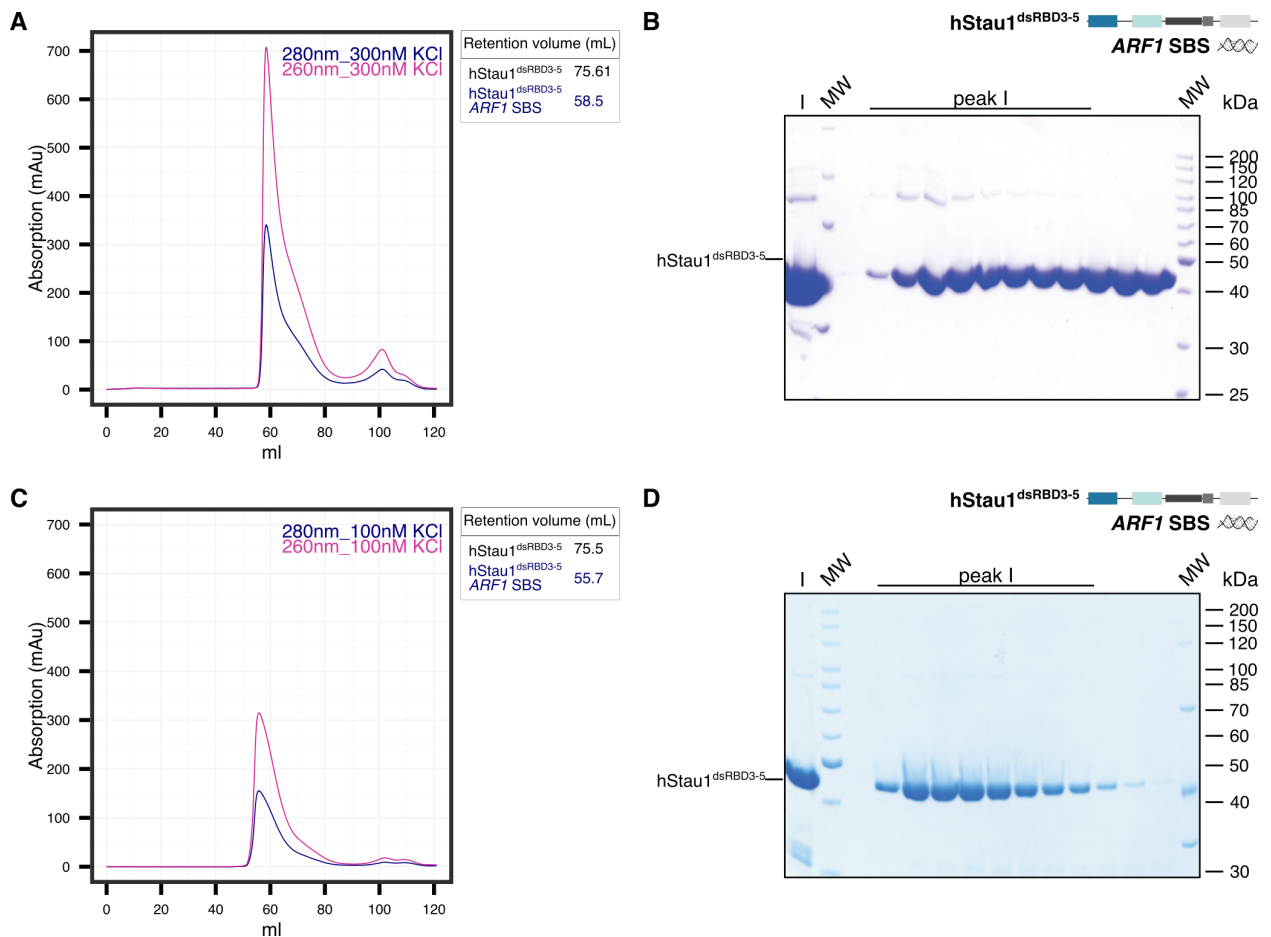


Figure 3.13. Representative SEC result of the hStau1^{dsRBD3-5}-ARF1 SBS complex in two different ionic strength conditions. Although for both experiments the same amount of protein was used, the peak absorption is significantly higher in the 300mM SEC run as well as the complex yield after SEC. **(A)** SEC profile of the hStau1^{dsRBD3-5} with the ARF1 SBS (300mM KCl), the complex elutes as a peak with a shoulder on the left side. Retention volume of the hStau1^{dsRBD3-5} protein alone and the complex with the 1ARF1 SBS are indicated in the table. **(B)** The Coomassie-stained SDS-PAGE exhibit a protein band that extend through the peak and the shoulder. **(C)** SEC profile of the hStau1^{dsRBD3-5} with the ARF1 SBS (100mM KCl). Retention volume of the hStau1^{dsRBD3-5} protein alone and the complex with the ARF1 SBS are indicated in the table. **(D)** The Coomassie-stained SDS-PAGE shows the fractions that correspond to the peak I. Line I shows the complex sample before SEC run. The size of the MW is in kDa.

3.4.5 Biochemical characterization of the hStau1^{dsRBD3-5} - ARF1 SBS_{20bp} complex

EMSA experiments were performed to confirm hStau1^{dsRBD3-5} - ARF1 SBS complex formation using the label ARF1 SBS_{20bp}. hStau1^{dsRBD3-5} shifted the ARF1 SBS_{20bp} and formed two protein complexes species (Fig. 3.14 B), since the SEC-MALLS result of the hStau1^{dsRBD3-5} alone showed that the protein was present as a monomer and dimer in solution; it could be possible that the two complexes observed corresponded to a 1:1 and a 2:1 (protein:RNA) ratio. To further determine the binding stoichiometries of the complexes, SEC-MALLS measurements were performed.

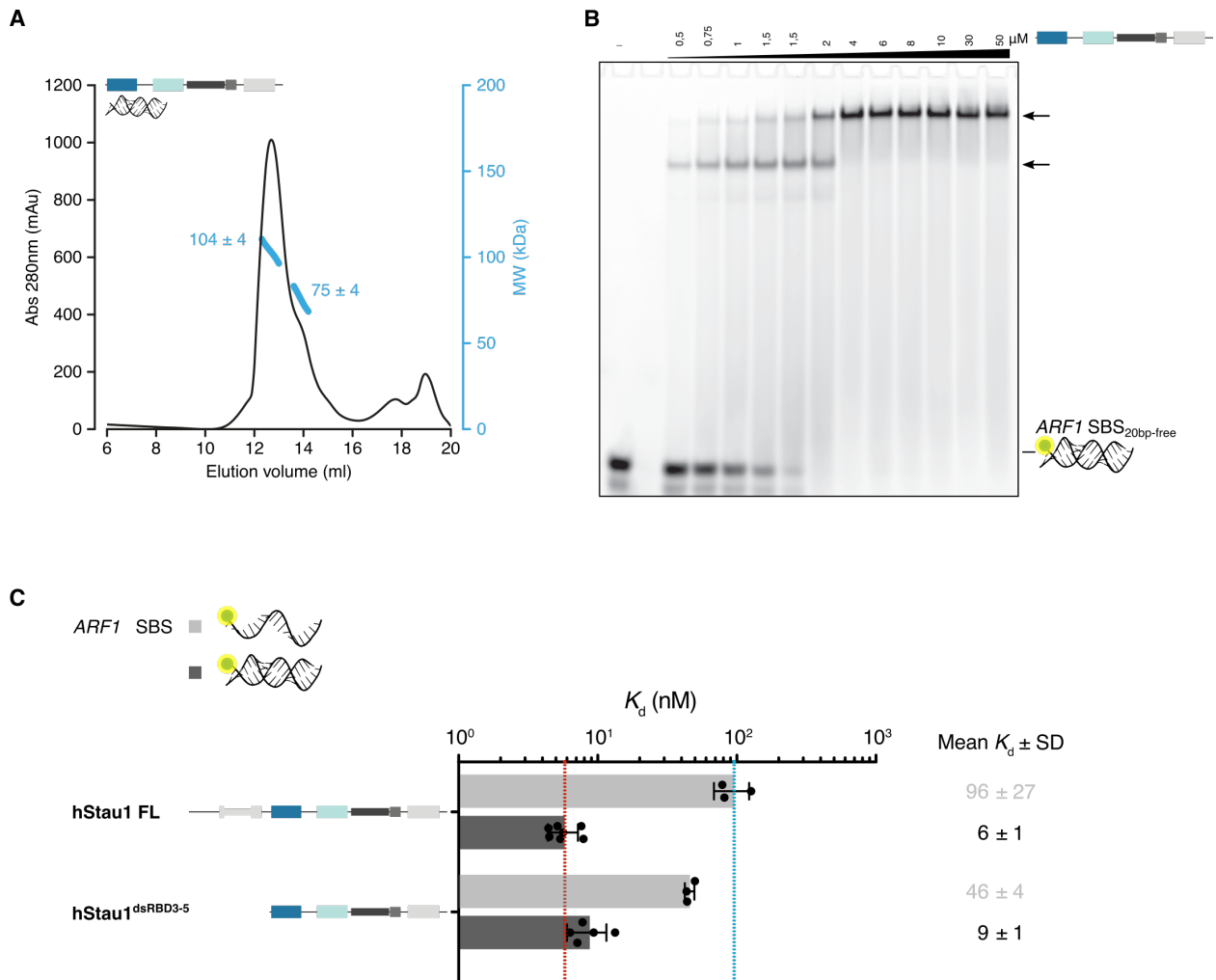


Figure 3.14. hStau1^{dsRBD3-5} formed stable complexes with ARF1 SBS *in vitro* (A) SEC-MALLS elution profile of the hStau1^{dsRBD3-5} - ARF1 SBS complex, the calculated molecular mass is indicated in blue (B) EMSA analysis of the hStau1^{dsRBD3-5} interaction with the fluorescent label ARF1 SBS_{20bp} on a native 6% Acrylamide gel. hStau1^{dsRBD3-5} shifted the label ARF1 SBS_{20bp} and formed two RNA complexes (black arrows). (C) The bar graph shows the calculated binding affinity of the hStau1 FL and the hStau1^{dsRBD3-5} for the single ARF1 SBS_{20bp} sense strand (light gray) and the ARF1 SBS_{20bp} (dark gray). The calculated K_D and their corresponding error (error bars) are the mean and the SD of a minimum of three experiments. (Mean K_D and SD in nM).

Examination of the SEC-MALLS data confirmed the existence of two complexes (Fig. 3.14 A). The experimentally measured molecular masses for peak I and II were 75 ± 4 kDa and 104 ± 4 kDa respectively. This result suggested that peak I might correspond to a complex with a binding ratio of 1:2 (one protein molecule and two *ARF1* SBS molecules) and peak II a second complex with a binding ratio of 2:2 (two protein molecules and two *ARF1* SBS molecules). To quantify the binding affinity of the hStau1^{dsRBD3-5} for *ARF1* SBS_{20bp}, we performed FA experiments with the purified hStau1^{dsRBD3-5} and the *ARF1* SBS_{20bp}. The calculated dissociation constant of K_D : 9 ± 1 nM was slightly weaker than the affinity of the FL protein. This result suggested that the N terminus region included dsRBD2 had a minor or no effect in the mRNA binding activity of the hStau1 protein (Fig. 3.14 C).

3.4.6 hStau1^{dsRBD3-5} - *ARF1* SBS complex crystallization

The purified hStau1^{dsRBD3-5} - *ARF1* SBS complex was concentrated and subjected to crystallization experiments. Preliminary crystals were observed in the Qiagen JSCG+, Qiagen PEGS, Qiagen Nucleix, Hampton Research Index, and Morpheus screens (Fig. 3.15 A). Hexagonal-shaped plates were obtained from an 8 mg/ml concentrated sample. Crystals grew in a period of 3 months. Most of the crystals could be successfully reproduced, but the majority were too thin and small for data collection, some crystals could be tested on the synchrotron beamline PXII of Swiss Light Source (SLS) (Paul Scherrer Institute, Villigen, Switzerland). The best crystals diffracted about to 14 Å resolution. The grew time could be shortened by different optimization attempts including micro-seeding. However, the crystal size still was not optimal for data collection.

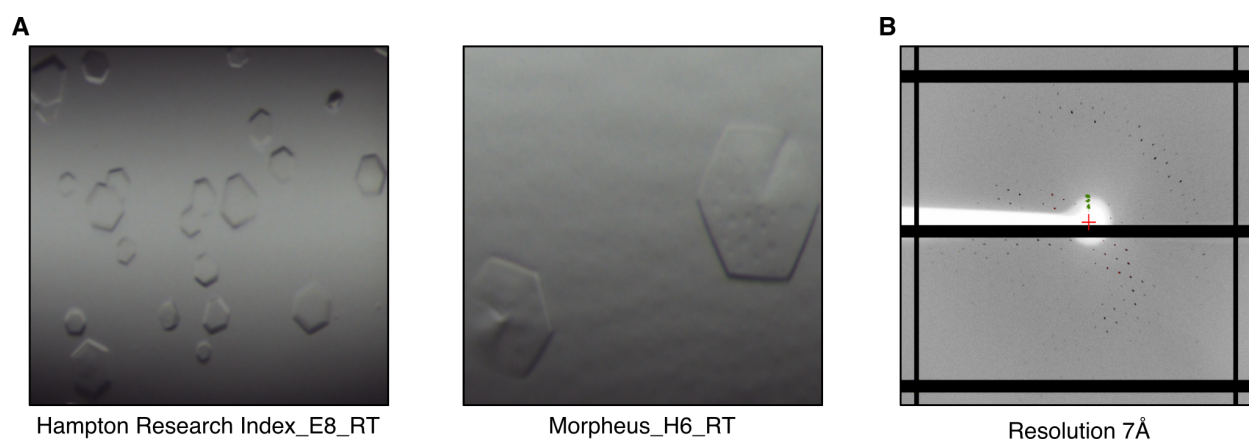


Figure 3.15. Crystals of hStau1^{dsRBD3-5} - *ARF1* SBS complex **(A)** Crystals obtained with the hStau1^{dsRBD3-5} - *ARF1* SBS complex. **(B)** Image of the best diffracting crystal (7 Å).

Attempts to remove flexible regions at the C terminus of hStau1^{dsRBD3-5} (ca. 20 aa) yielded crystals with the same morphology than the crystals obtained with the original hStau1^{dsRBD3-5} construct and resulted in poor diffraction (>20 Å). Further optimization was done by Christiane Emmerich

(laboratory technician). Although crystals from different conditions were tested, the resolution of the best diffracting crystals ranged from 9 to max. 7 Å (Fig. 3.15 B).

3.5 hStau1 dsRBD3 and 4 are sufficient for *ARF1* SBS binding *in vitro*

3.5.1 Reconstitution of the hStau1^{dsRBD3-4} - *ARF1* SBS complex

In the crystals optimization process of the hStau1^{dsRBD3-5} - *ARF1* SBS complex, we analyzed some of the crystals on a SDS-PAGE and denaturing urea/acrylamide gel. The denaturing urea/acrylamide gel verified that the *ARF1* SBS still bound to hStau1^{dsRBD3-5}. However, the SDS-PAGE showed a single protein band between the 20 and 25 kDa marker suggesting that the hStau1^{dsRBD3-5} construct undergo natural proteolytic degradation (Fig. 3.16 A) during crystallization.

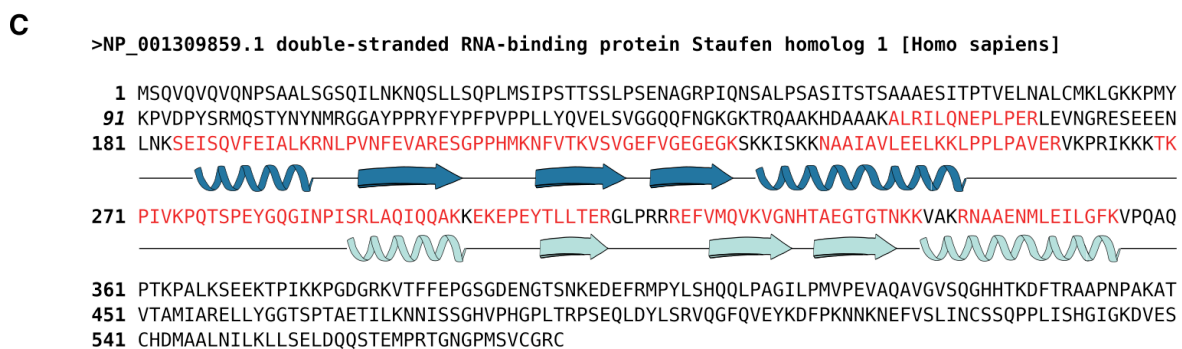
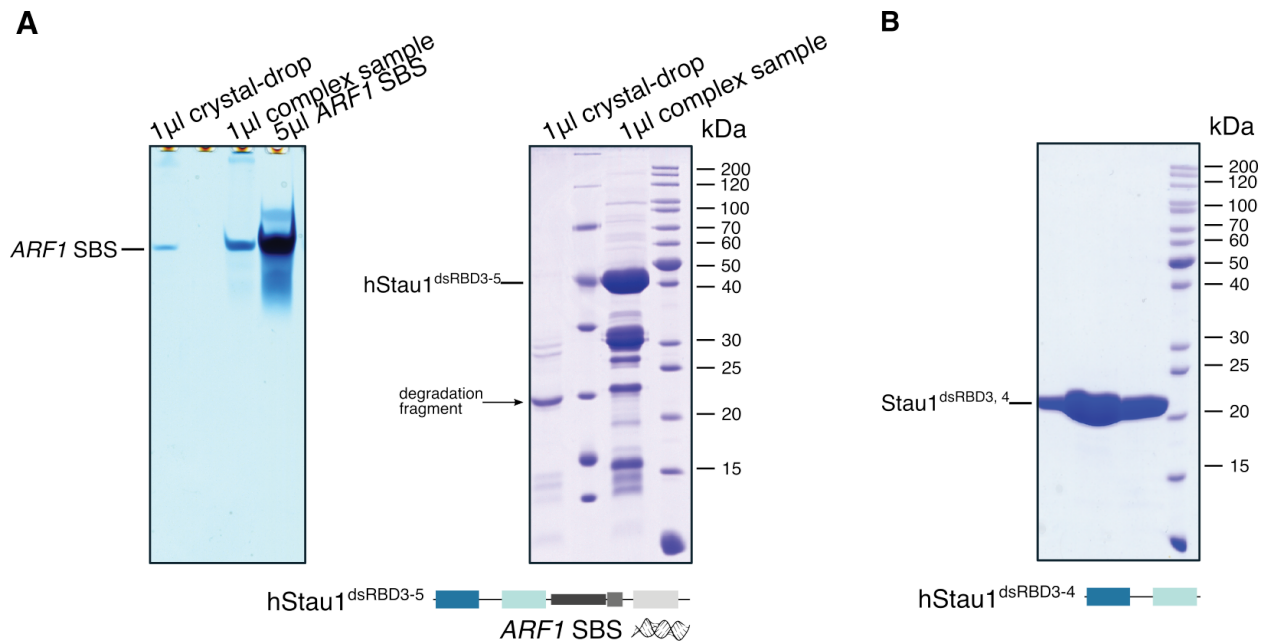


Figure 3.16b. hStau1^{dsRBD3-5} undergo natural proteolytic degradation during crystallization **(A)** 20% denaturing urea/acrylamide gel (left) and Coomassie-stained SDS-PAGE (right) of (I) dissolved crystals of the hStau1^{dsRBD3-5} - ARF1 SBS complex, (II) sample of hStau1^{dsRBD3-5} - ARF1 SBS complex and (III) sample of ARF1 SBS. The black arrow shows the degradation product present in the dissolved crystal. **(B)** Coomassie-stained SDS-PAGE of the recombinantly purified hStau1^{dsRBD3-4}. **(C)** MS analysis of the degradation fragment corresponding to the dissolved crystal of the hStau1^{dsRBD3-5} - ARF1 SBS complex (band indicated by the black arrow). The matched peptides correspond to 89% of the sequence of the hStau1^{dsRBD3-4} construct.

The band that corresponded to the degradation product was further analyzed and characterized by Mass spectrometry (TOPLAB GmbH). The mass spectrometry analysis identified the same region that comprised the hStau1^{dsRBD3-4} construct (Fig. 3.16 B and C). Since mRNA binding activity persisted after proteolysis of the hStau1^{dsRBD3-5} and the degradation product correspond to the hStau1^{dsRBD3-4} construct, we focused our work on this construct. For this Christiane Emmerich reconstituted the complex and conducted crystallization experiments.

Purified hStau1^{dsRBD3-4} was mixed with a 1.2 molar excess ARF1 SBS. The ARF1 SBS was resuspended in 3M KCl to avoid precipitation issues. After incubation, the complex was subjected to SEC. The elution profile was similar to the complex of hStau1^{dsRBD3-5} - ARF1 SBS. The complex eluted as a broad peak with a shoulder on the right side, indicating some heterogeneity of the sample. The retention volume shifted considerably to high molecular weight. Furthermore, the 260/280 ratio was significantly increased (Fig. 3.17 A). The SDS-PAGE of the peak fractions showed that the protein is present in all fractions (Fig. 3.17 B). This data demonstrated that dsRBD3 and 4 are sufficient to bind ARF1 SBS *in vitro*.

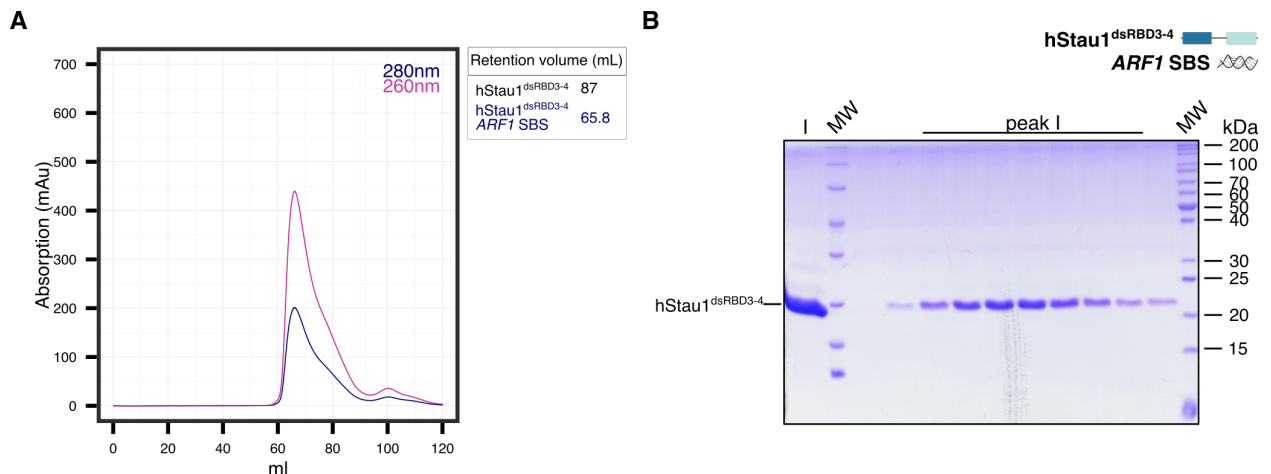


Figure 3.17. Size exclusion chromatography of hStau1^{dsRBD3-4} in complex with the ARF1 SBS using a Superdex 200 16/60 GL column. The size of the MW is in kDa. Line I shows the complex sample before the SEC run. **(A)** SEC elution profile of hStau1^{dsRBD3-4} in complex with the ARF1 SBS. The broader peak is significantly shifted, and the 260/280 ratio is increased considerably. Retention volume of hStau1^{dsRBD3-4} apo (black) and the complex (blue) are indicated in the table. **(B)** The Coomassie-stained SDS-PAGE loaded with the peak fractions (peak I) of the run show a single band of high purity.

3.5.2 Biochemical characterization of the hStau1^{dsRBD3-4} - *ARF1* SBS_{20bp} complex

To further characterize the mRNA binding activity of hStau1^{dsRBD3-4} EMSA analysis was performed. Similarly to hStau1^{dsRBD3-5}, hStau1^{dsRBD3-4} bound to *ARF1* SBS_{20bp} as observed by a shift and a supershift of the RNA band (Fig. 3.18 B). The binding stoichiometry of the complex was assessed by SEC-MALLS experiments. The experimentally determined molar mass was 49 ± 3 kDa, which corresponded to two molecules hStau1^{dsRBD3-4} and one *ARF1* SBS_{20bp} molecule in agreement with the second shift of the EMSA experiments (Fig. 3.18 A).

The binding affinity was quantified by FA using the *ARF1* SBS_{20bp}. The hStau1^{dsRBD3-4} construct bound the *ARF1* SBS_{20bp} with a K_D of 14 ± 4 nM. The affinity was reduced more than two-fold compared to the FL protein (Fig. 3.18 C). These data showed the possible role in mRNA binding regulation mediated by dsRBD2 together with the TBD and dimerization region (SSM and dsRBD5). Both regions could be required to fulfill the recognition and binding of targets mRNA more efficiently.

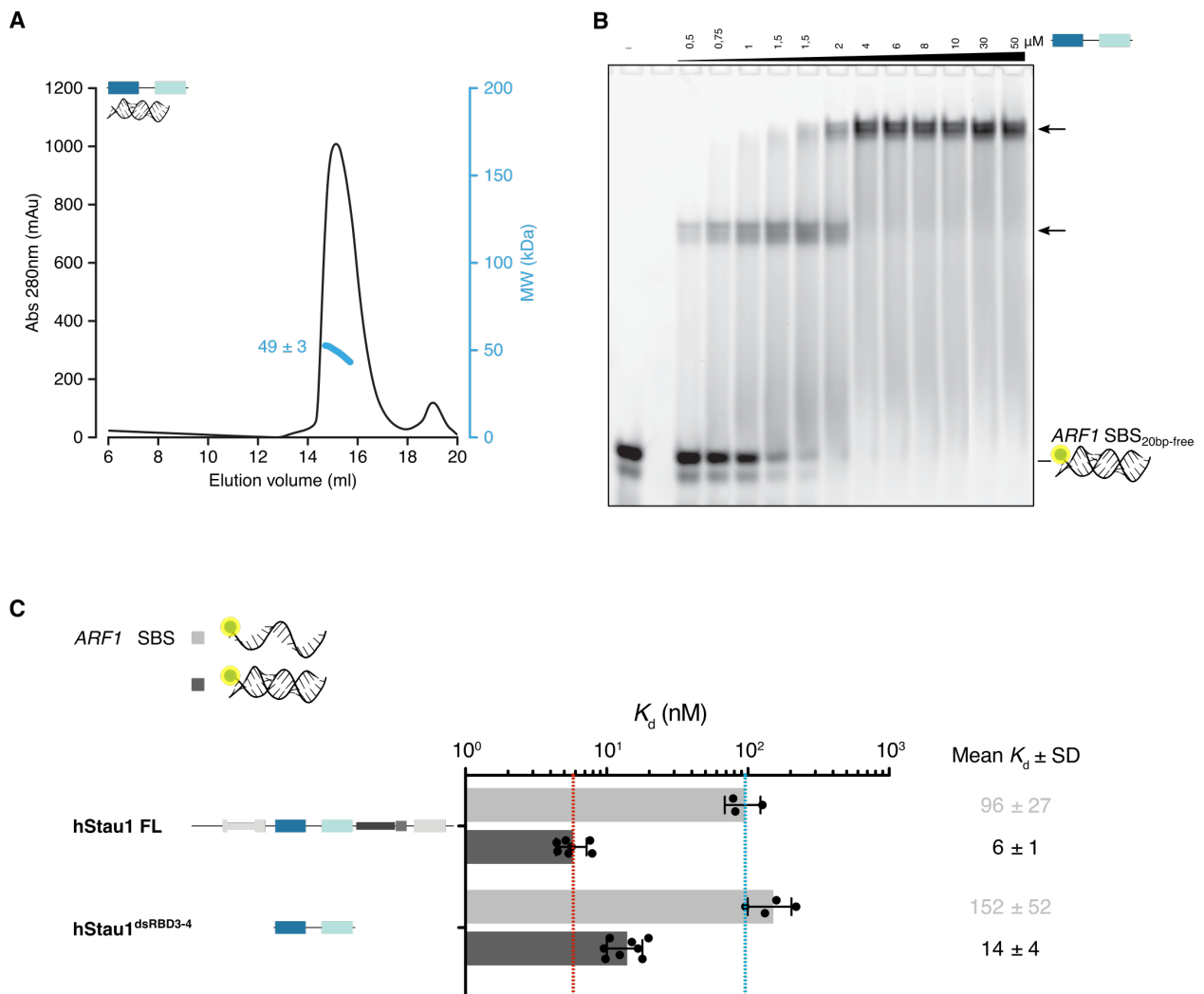


Figure 3.18b. hStau1^{dsRBD3-4} formed stable complexes with *ARF1* SBS *in vitro* (A) SEC-MALLS elution profile of hStau1^{dsRBD3-4} in complex with the *ARF1* SBS, the calculated molecular mass is indicated in blue (B) EMSA analysis of the hStau1^{dsRBD3-4} interaction with the fluorescent label *ARF1* SBS_{20bp} on a native 6% Acrylamide gel. The hStau1^{dsRBD3-4} shifted the label *ARF1* SBS_{20bp} and formed two RNA complexes (black arrows). (C) The bar graph shows the calculated binding affinity of the hStau1 FL and the hStau1^{dsRBD3-4} for the single *ARF1* SBS_{20bp} sense strand (light gray) and the *ARF1* SBS_{20bp} (dark gray). The calculated K_D and their corresponding error (error bars) are the mean and SD of a minimum of three experiments. (Mean K_D and SD in nM).

3.6 hStau1^{dsRBD3-4} recognizes and binds *ARF1* SBS by conformational readout of the minor and major groove.

3.6.1 hStau1^{dsRBD3-4} - *ARF1* SBS complex crystallization

The purified complex was used for crystallization experiments. The complex crystallized in different conditions. All crystals could be reproduced and promising conditions were optimized. The crystals appeared after one day and showed a significant increase in size. Different crystals were tested at the synchrotron beamline PXII of SLS. The best crystals diffracted to 3-2.9 Å resolution.

The collected data set was processed with the X-ray Detector Software (XDS) software (Kabsch, 2010). The images were indexed in the tetragonal space group $P4_1 2_1 2$ with unit cell dimensions of $a = b = 105.9 \text{ \AA}$ and $c = 169.2 \text{ \AA}$ and $\alpha = \beta = \gamma = 90^\circ$. The volume of the asymmetric unit (ASU) contained one complex subunit. The structure of hStau1^{dsRBD3-4} - *ARF1* SBS complex was determined at 2.9 Å resolution and solved by Molecular replacement (MR) using PHASER (McCoy et al., 2007). The model was generated by the superposition of two known structures of Staufen dsRBD (*Dm*Stau dsRBD5 PDB ID: 5CFF (Jia et al., 2015) and human hStau1 dsRBD5 PDB ID: 4DKK), with this model two dsRBD3 molecules could be found. The density map that was obtained after the first refinement cycles revealed an additional electron density on the back side of the structure. This density was used to place a homology model of dsRBD4 manually, that was built using the SWISS-MODEL server (swissmodel.expasy.org) (Arnold et al., 2005). The second dsRBD4 could not be fitted in the electron density.

A seleno-methionine (SeMet) data set was collected to verify the exact arrangement of the three dsRBDs found. The SeMet data set and the refined phases of the native crystal were used to calculate difference Fourier map. Based on the difference Fourier map, it was possible to trace the methionine sites of each dsRBD. From the 4 Met (Met213 for the two dsRBD3, Met321, and Met348 in dsRBD4) present in the sequence four were identified (Fig. 3.19).

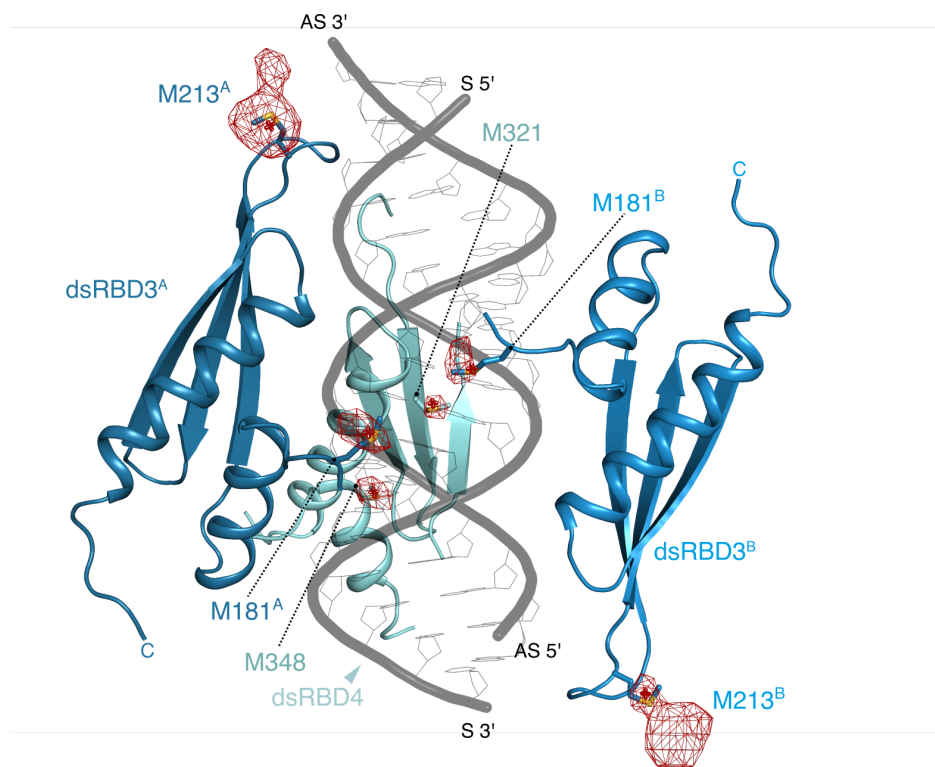


Figure 3.19. Difference Fourier map of the hStau1^{dsRBD3-4}-ARF1 SBS complex. Six methionine were observed after SeMet substitution (M181 in dsRBD3 molecules A and B do not belong to the wild-type sequence). The anomalous difference Fourier map of the complex contoured at 4σ (shown in red) is superimposed on the structure of the hStau1^{dsRBD3-4}-ARF1 SBS complex. Met residues in dsRBD3 (molecules A and B) and in dsRBD4 are indicated.

After rebuilding and refinement in COOT and PHENIX, the R_{work} was refined to 21.7% and the R_{free} to 24.0% with good stereochemistry. Five residues at the C terminus were missing from the electron density map. Furthermore, the loop between dsRBD3 and 4, together with two loops in the dsRBD4 were disordered in the structure. The final statistics for the structure were summarized in the table 3.1.

Table 3.1. Data collection and refinement statistics.

	Native <i>ARF1</i> SBS-hStau1 ^{dsRBD3-4}	SeMet <i>ARF1</i> SBS-hStau1 ^{dsRBD3-4}	<i>ARF1</i> SBS
Data collection			
Space group	P4 ₁ 2 ₁ 2	P6 ₁ 2 ₁ 2	H32
Cell dimensions			
a, b, c (Å)	105.9, 105.9, 169.2	79.9, 79.9, 237.3	43.8, 43.8, 452.1
α , β , γ (°)	90, 90, 90	90, 90, 120	90, 90, 120
Wavelength	1.045	0.980	1.000
Resolution (Å)	50-2.89 (3.06-2.89)	50-4.23 (4.48-4.23)	50-1.9 (2.02-1.90)
R _{sym} or R _{merge}	0.133 (2.864)	0.274 (1.685)	0.089 (1.341)
I / σ I	20.52 (1.32)	10.81(2.03)	13.52 (1.70)
Completeness (%)	99.4 (96.8)	99.7 (99)	99.9 (99.6)
Redundancy	27.5 (25.5)	27.7 (25.5)	10.3 (10.3)
CC(1/2)	100 (81.4)	100 (75.1)	100 (69.5)
Refinement			
Resolution (Å)	47.35-2.89		38-1.9
No. reflections	40874		13828
R _{work} / R _{free}	21.7/24.0		24.4/26.2
No. atoms			
Protein	4278		1883
RNA	3058		1828
Water	1217		40
B-factors (Å ²)			
Protein	3		54.4
RNA	109.7		53.5
Water	126.7		46.7
R.m.s deviations			
Bond lengths (Å)	0.004		0.019
Bond angles (°)	0.58		2.38

One native crystal for each construct and one SeMet crystal were used for data collection. Values in parentheses are for the highest-resolution shell.

3.6.2 hStau1^{dsRBD3-4} - *ARF1* SBS structure

A view of the structure is given in Fig. 3.20. In the structure, the *ARF1* SBS was clamped in the middle of the three dsRBDs. All three dsRBDs adopted the canonical $\alpha\beta\beta\beta\alpha$ topology, where a short α -helix 1 and a longer α -helix 2 were packaged against the hydrophobic face of the anti-

parallel β -sheet. Hydrophobic interactions between the two α -helices with the β -sheet allowed a tight packaging of the dsRBD structure, while at the same time this compact conformation allowed charged residues to form a positively charged surface to recognize and bind the A-conformation of the RNA.

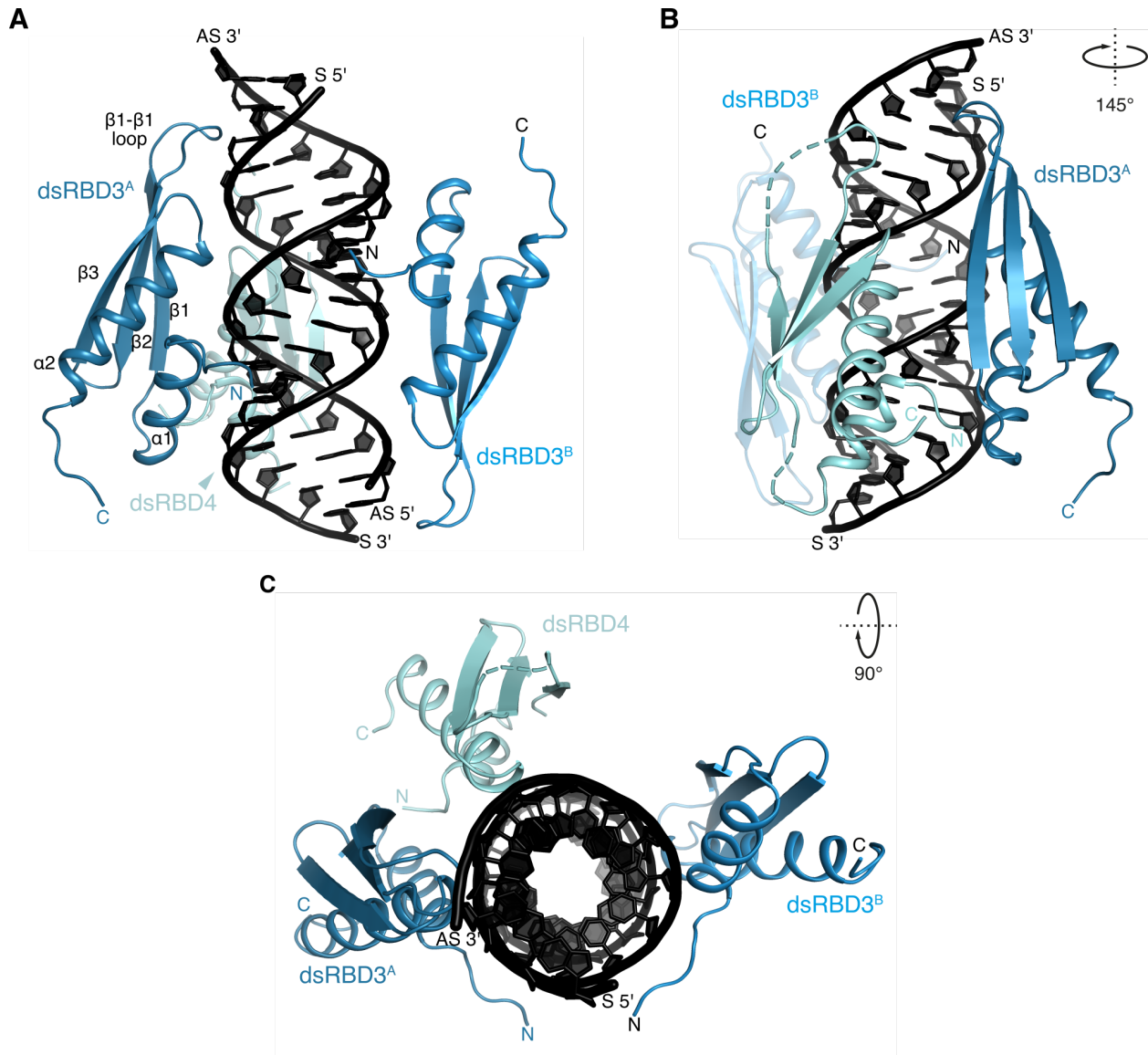


Figure 3.20. Structure of the hStau1^{dsRBD3-4}- ARF1 SBS complex **(A)** Cartoon representation of the structure of the hStau1^{dsRBD3-4}- ARF1 SBS complex. dsRBD3 molecules A and B are in blue, RNA is in black. N and C terminus are indicated in both domains as well as secondary structure elements in dsRBD3^A. **(B)** Side view of the structure (145° rotation over the vertical axis). dsRBD4 is in cyan, N and C terminus are indicated. The regions that are not visible in the electron density are represented as dotted lines. **(C)** Top view of the structure (90° rotation over the horizontal axis).

The RNA recognition mode of the three dsRBDs was similar to that described for other dsRBD (Bycroft et al., 1995, Nanduri et al., 1998, Ryter, 1998, Ramos et al., 2000, Blaszczyk et al., 2004, Stefl

et al., 2010, Jayachandran et al., 2015). The interaction surface on the RNA extended to the major groove and the two successive minor grooves of the RNA. I used the Protein Interfaces Surfaces and Assemblies (PISA) server, to analyze the interaction surfaces of the hStau1^{dsRBD3-4}-*ARF1* SBS complex.

Two individual dsRBD3 molecules (dsRBD3 molecule A and dsRBD3 molecule B referred to as dsRBD3^A and dsRBD3^B respectively) were positioned in anti-parallel fashion along the *ARF1* SBS. The two dsRBD3 molecules dsRBD3^A and dsRBD3^B are very similar and therefore with a Root Mean Square Deviation (RMSD) of 0.8 Å over 61 C α atoms (residues 184- 253). dsRBD3^A bound the *ARF1* SBS with a buried surface area of 803² and had a complex-formation significance (CSS) score of 1 (CSS ranges from 0 to 1 as the significance of the interface in complex formation) indicating that the interaction surface is biological relevant and not a crystal-packing artifact. In contrast, dsRBD3^B had a smaller binding surface (586,2 Å²) nevertheless the CSS scored 0.6, this indicates that the interaction surface of dsRBD3^B is also biologically significant.

Only one dsRBD4 was visible in the electron density map. In the structure, dsRBD4 interacted with dsRBD3^A and with the *ARF1* SBS. dsRBD4 bound the *ARF1* SBS with a buried surface area of 444 Å² (PISA CSS score 1). The electron density for the second dsRBD4 was not visible. Binding of the missing dsRBD4 with the RNA can be excluded because steric clashes with the already bound dsRBD4 limited the interaction. The missing electron density also suggested that the second dsRBD4 molecule could be unbound and disordered. PISA analysis of the interaction between dsRBD4 and dsRBD3^A indicated that the interface region involved a 273 Å² (5,1 %) area. The CSS score for this interface was 0.01, suggesting that the interaction might not be of biological relevance. Four hydrogen bonds contributed to the interactions between the domains (Fig.3.21 A) and engaged mainly the β -strand 1 of dsRBD3^A and the α - helix 2 of dsRBD4. To determine the effect of this interaction in mRNA-binding activity, we mutated two residues involved in the interactions (hStau1^{dsRBD3-4} N197A, R342A) and performed FA experiments with the hStau1^{dsRBD3-4} N197A, R342A mutant. The apparent affinity of the mutants for *ARF1* SBS was similar to the wild-type hStau1^{dsRBD3-4}, indicating that the mRNA-binding activity of the domains was not influenced by the inter-domain interaction (Fig. 3.21 A and B).

The linker segment connecting dsRBD3 and 4 was not visible in the electron density map, suggesting that this 29 amino acid long linker might be flexible. Sequence-based calculations predicted that the length of the inter-domain linker varied from 100 to 120 Å. This distance could be plausible to connect the dsRBD4 with any of the dsRBD3. The residues present in the linker were mainly positively charged (27% of the residues in the linker were positively charged) of these six are lysines and two arginines and play a role in mRNA-binding activity. To determine the impact of the inter-domain linker in binding affinity, we designed two mutants and measured the binding affinity by

FA experiments. One mutant was designed to alter the length of the linker and the second mutant was designed to reduce positive charge of the linker.

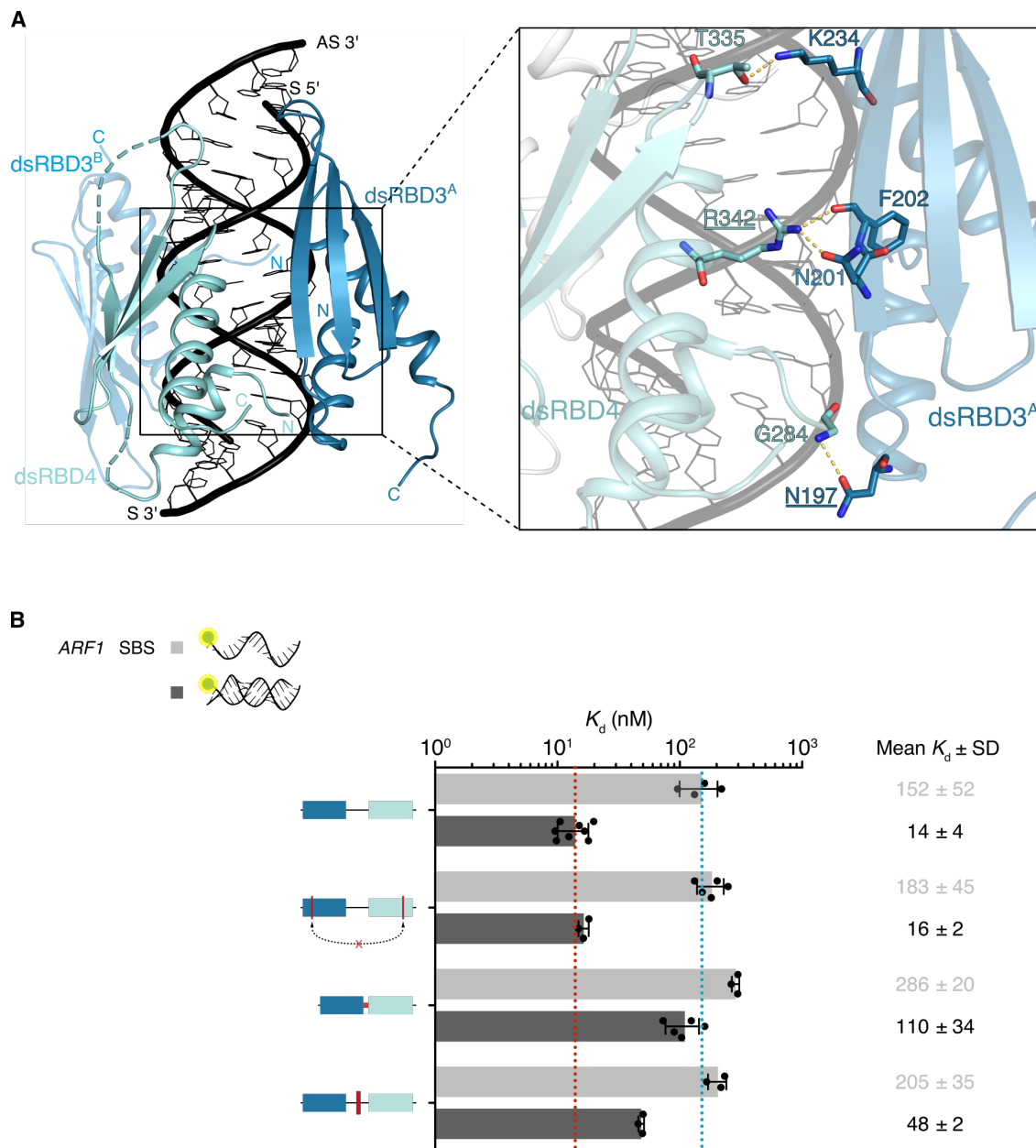


Figure 3.21. dsRBD3^A and dsRBD4 interaction interface **(A)** Side view of the structure with a close-up view of the interaction interface between dsRBD3^A and dsRBD4. Interacting residues are shown as sticks. Hydrogen bonds are indicated as dashed lines. Residues mutated for FA experiments are underlined. **(B)** The bar graph shows the calculated binding affinity of the hStau1^{dsRBD3-4} construct and the designed mutants (interface mutant, length linker mutant, and linker composition mutant) for the single 20 nt *ARF 1* mRNA sense strand (light gray) and the double-stranded 20 bp *ARF 1* mRNA (dark gray). The calculated K_D and their corresponding error (error bars) are the mean and SD of a minimum of three experiments. (Mean K_D and SD in nM).

When the linker region (amino acid 254 to 283) was replaced with a GSSGSGSS sequence, the affinity for *ARF1* SBS was reduced more than eightfold compared to the wild-type protein (Fig.3.21 B). However, when four lysins of the linker were mutated to alanine, the affinity for *ARF1* SBS was decreased only by three-fold. These results indicated that the inter-domain linker contributed to RNA binding. Both the length of the linker and also the amino acid composition might stabilize the binding between the two dsRBD with the *ARF1* SBS.

3.7 The *ARF1* SBS structure is A-form in conformation

3.7.1 *ARF1* SBS crystallization

To address the conformational changes of the *ARF1* SBS upon binding, we crystallized and determined the apo-state *ARF1* SBS structure at a resolution of 1.9Å. The images were processed with the XDS software (Kabsch, 2010). The structure of the *ARF1* SBS was solved by MR using a 19 bp A-form RNA as a model (PDB: 1QC0 Klosterman P.S. (1999)). The crystal lattice corresponded to the trigonal space group (R32) H32, with unit cell parameters $a = b = 43.8 \text{ \AA}$ and $c = 452.1 \text{ \AA}$ and $\alpha = \beta = 90^\circ \gamma = 120^\circ$. In this space group, the ASU comprised one and a half RNA molecules (fig. 3.22 A). The model exhibited a pseudo-infinite helix along the c-axis (fig. 3.22 B). The 19bp length of the *ARF1* SBS induced six statically disordered conformations (Lazzaretti et al., 2018). The refinement of the model yielded a R_{free} of 26.2% and a R_{work} of 24.4% with good stereochemistry. The final statistics for the structure were summarized in the table 3.1.

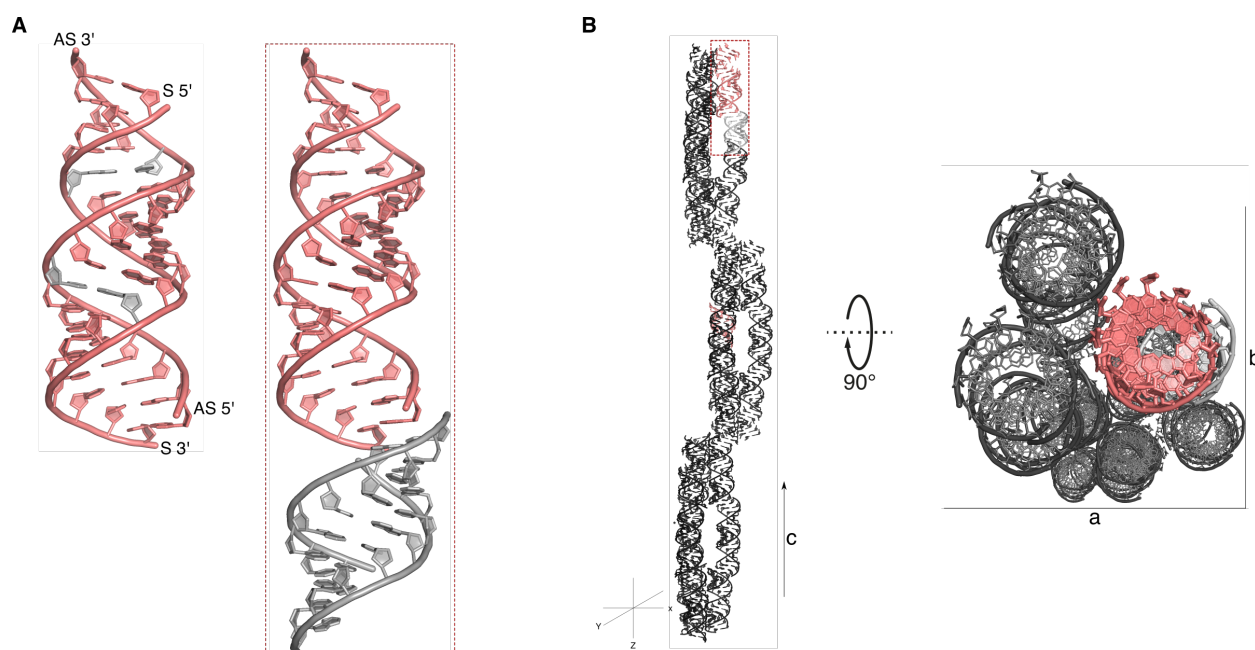


Figure 3.22b. Structure of the apo-state *ARF1* SBS (**A**) On the left structure of the *ARF1* SBS. GU wobble base pairs are in gray. On the right, an asymmetric unit consisting of one complete *ARF1* SBS molecule plus half *ARF1* SBS molecule. (**B**) View of the crystal packing where the pseudo-infinite helix extends through the *c*-axis. All RNA and protein structure figures were generated with PyMOL.

The *ARF1* SBS structure has a canonical A-form double-stranded helices with blunt ends. The RNA contains 17 Watson-Crick base-pairs and two G-U wobble base pairs. The C1'-C1' distance of the two wobble pairs (10.2Å and 10.3Å) did not differ of the average distance of Watson-Crick base-pairs (10.5Å (Seeman et al., 1976)). The glycosidic bond angle between the C1' sugar atom and the base was different from the ~54° angle of Watson-Crick base-pairs (Fig. 3.23). Still, the two wobble base pairs did not significantly distort the RNA-backbone.

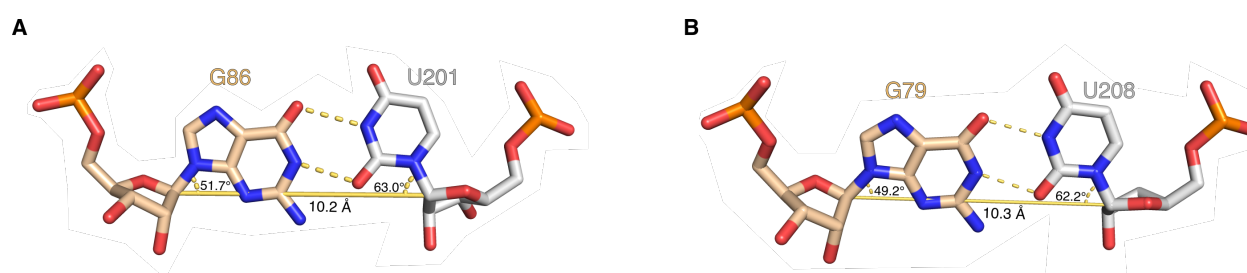


Figure 3.23. Glycosidic angle and distance of the C1'-C1' of the two wobble pairs present in the *ARF1* SBS structure. Bases are shown as sticks. Hydrogen bonds are indicated as dashed lines. Distance between the ribose C1' atoms and glycosidic bond angles are indicated in black. (**A**) Interaction of the wobble pair G₈₆-U₂₀₁. (**B**) Interaction of the wobble pair G₇₉-U₂₀₈.

3.7.2 *ARF1* SBS conformational changes by binding to hStau1^{dsRBD3-4}

The superposition of the apo and bound *ARF1* SBS structures showed an RMSD of 1.32 Å. This low RMSD value indicated that the RNA underwent barely significant conformational changes upon protein binding. Although bound *ARF1* SBS retained the characteristic A-conformation, numerous sites in the RNA revealed slight differences. The binding of hStau1^{dsRBD3-4} led to elongation of the *ARF1* SBS. Consequently, the major groove widened to facilitate interactions between the α -Helix 2 of both dsRBD3 and the RNA backbone (Fig. 3.24 A). Additionally, the interaction between the β 1- β 2 loop of dsRBD3^A and the *ARF1* SBS induced a deviation of 3.9 Å at the 5' end and 3' end of the sense strand (Fig. 3.24 A). Further changes in the bound structure were the bending of the helical axis of the symmetry-related RNA molecule bounded by stacking end to end. This favorable stacking arrangement formed a semi-continuous major/minor groove conformation that allows the β 1- β 2 loop of dsRBD3^B to interact with the subsequent *ARF1* SBS molecule (Fig. 3.24 B).

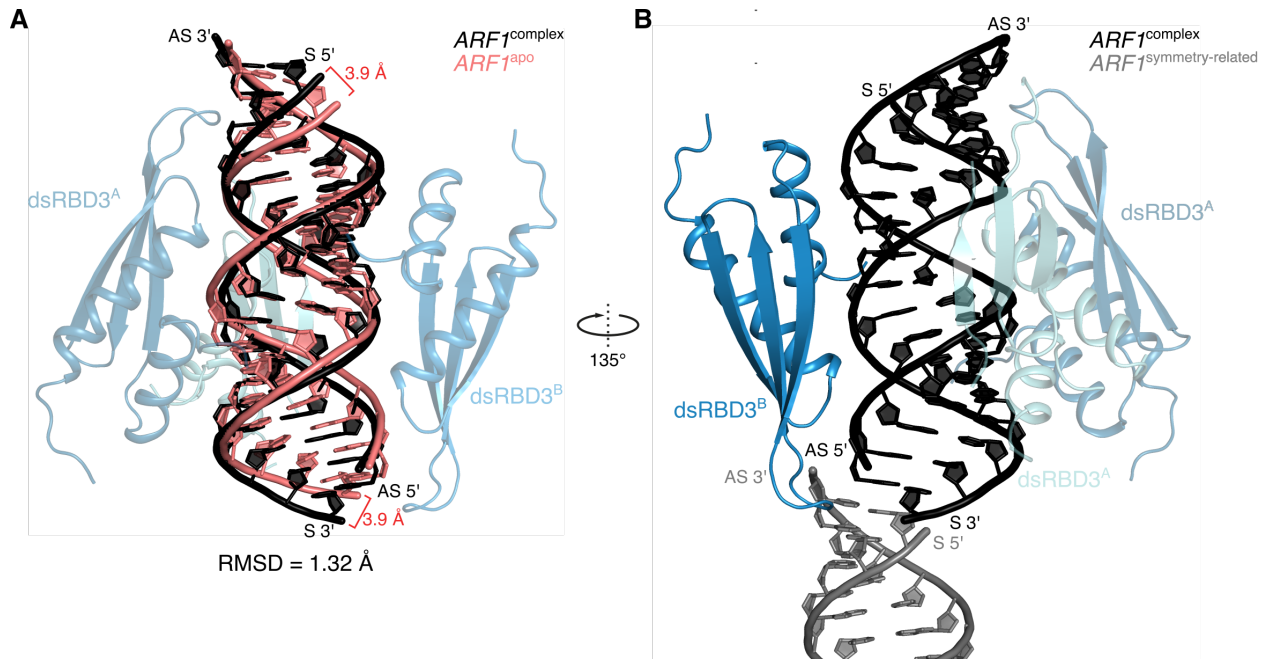


Figure 3.24. *ARF1* SBS conformational changes by binding to hStau1^{dsRBD3-4} (A) Superposition of the structure of the unbound *ARF1* SBS (deep salmon) and in complex with hStau1^{dsRBD3-4} (black) in the same orientation. The RMSD is indicated. (B) Structure of the hStau1^{dsRBD3-4} - *ARF1* SBS complex and the simetry-related *ARF1* SBS molecule.

3.8 hStau1^{dsRBD3-4} recognition mode of *ARF1* SBS

The structure of the hStau1^{dsRBD3-4} - *ARF1* SBS complex showed that hStau1^{dsRBD3-4} bound over nearly the entire length of the *ARF1* SBS. The three visible dsRBDs primarily recognized and bound the backbone through three different regions. Region 1 corresponded to α -helix 1, region 2 was present in the loop between β 1 and β 2 strand and region 3 located in the N-terminal end of α -helix 2. The recognition mode observed in the structure was similar to those described for other crystal structures of dsRBD-RNA complexes (Ryter, 1998, Ramos et al., 2000, Blaszczyk et al., 2004, Stefl et al., 2010). Interestingly, all three domains exhibited direct interactions with G:C base pairs. These interactions included direct and water-mediated hydrogen bonds (Fig. 3.25 and 3.26).

Region 1

In Region 1, polar and charged residues of the α -helix 1 made hydrogen bonds with both strands in the minor groove of the *ARF1* SBS primary (Fig. 3.25 and 3.26 region 1). In dsRBD3^A residues K183, S184, S187, and E191 formed hydrogen bonds with the ribose groups of U88 (2'OH'), G200 (2'OH), and U201 (O4'). Another binding contribution was observed at the C-terminus of α -helix 1. Residue K195 formed a salt bridge with the phosphate group of C90 stabilizing the interaction of α -helix 1 (Fig. 3.26 A). In addition to the non-sequence specific contacts, a sequence directed

interaction was formed between S187 and G200. The side chain of S187 was hydrogen bonded to the N² of G200 through a water molecule (Fig. 3.25 D). dsRBD3^B showed similar interactions as dsRBD3^A. The side chains of S184, S187, and E191 made hydrophobic contacts with A84 (O4'), G83 (2'OH'), and U205 (2'OH') (Fig. 3.26 A). The sequence-specific interaction of S187 was also present in this domain. As in molecule A, the side chain of S187 made a water-mediated hydrogen bond with the N² of G83 (Fig. 3.25 C). dsRBD4 contacted the minor groove only through polar residues of α -helix 1 (Fig. 3.26 B). The side chains of N286, S289, Q293, and Q296 were hydrogen bonded to the ribose groups of A199 (O4'), C198 (2'OH'), C91 (O4'), and C91 (2'OH') respectively. The NE2 atom of Q293 docked into the minor groove of the *ARF1* SBS to come in a closed position to the O² of C90 (Fig. 3.25 B).

Region 2

In region 2 the minor groove of the *ARF1* SBS was recognized by the loop connecting the β 1 and β 2 strands (Fig. 3.25 and 3.26 region 2). The GPxHxx motif of dsRBD3^A contacted the minor groove at base G77 and U211. In dsRBD3^A, the backbone of P211 was in hydrogen bonding to the exocyclic amino group of G77 (Fig. 3.25 E). In addition, the ribose 2'OH' group of U211 was hydrogen-bonded to the side chain of H212. In dsRBD3^B residues, P211 and H212 were close to the minor groove formed by base stacking interaction with the symmetry-related RNA molecule, but specific hydrogen bonds could not be observed. In contrast, the loop between the β 1 and β 2 strand of dsRBD4 did not interact with the minor groove and appeared to be partially unstructured.

Region 3

The positive charged surface in region 3 of all dsRBDs interacted with the sugar-phosphate backbone of the major groove (Fig. 3.25 and 3.26 region 3). The conserved motif, KKxAK in region3 of dsRBD3^A and RBD3^B all three Lys (K234, K235, and K238) contacted the RNA (Fig. 3.26 A). The side chain of K234 formed a salt bridge with the backbone of G80 (molecule A) and G197 (molecule B); while it simultaneously formed a hydrogen bond with the O3' atom of G79 (molecule A) and G196 (molecule B). Besides, the amide proton of K234 was hydrogen-bonded to the phosphate of G79 (molecule A) and G196 (molecule B). The side chain of K235 protruded into the major groove and interacted with the RNA backbone by electrostatic interactions. In contrast, the electron density of this Lys was not well defined in molecule B. Finally, the last Lys K238 formed a salt bridge with the phosphate group of U201(molecule A) and A84 (molecule B) (Fig. 3.26 A).

In dsRBD4 four residues made contact with the RNA backbone. The side chain of N336 was hydrogen-bonded to the phosphate group of C81. The conserved K337 was packaged parallelly to the sense strand to form a hydrogen-bond with the ribose of C81 as well as a salt bridge with the phosphate group of A82. At the same time, the main chain formed a hydrogen bond with the phosphate group of C81. The second Lys (K338) of the conserved motif was interacting with the

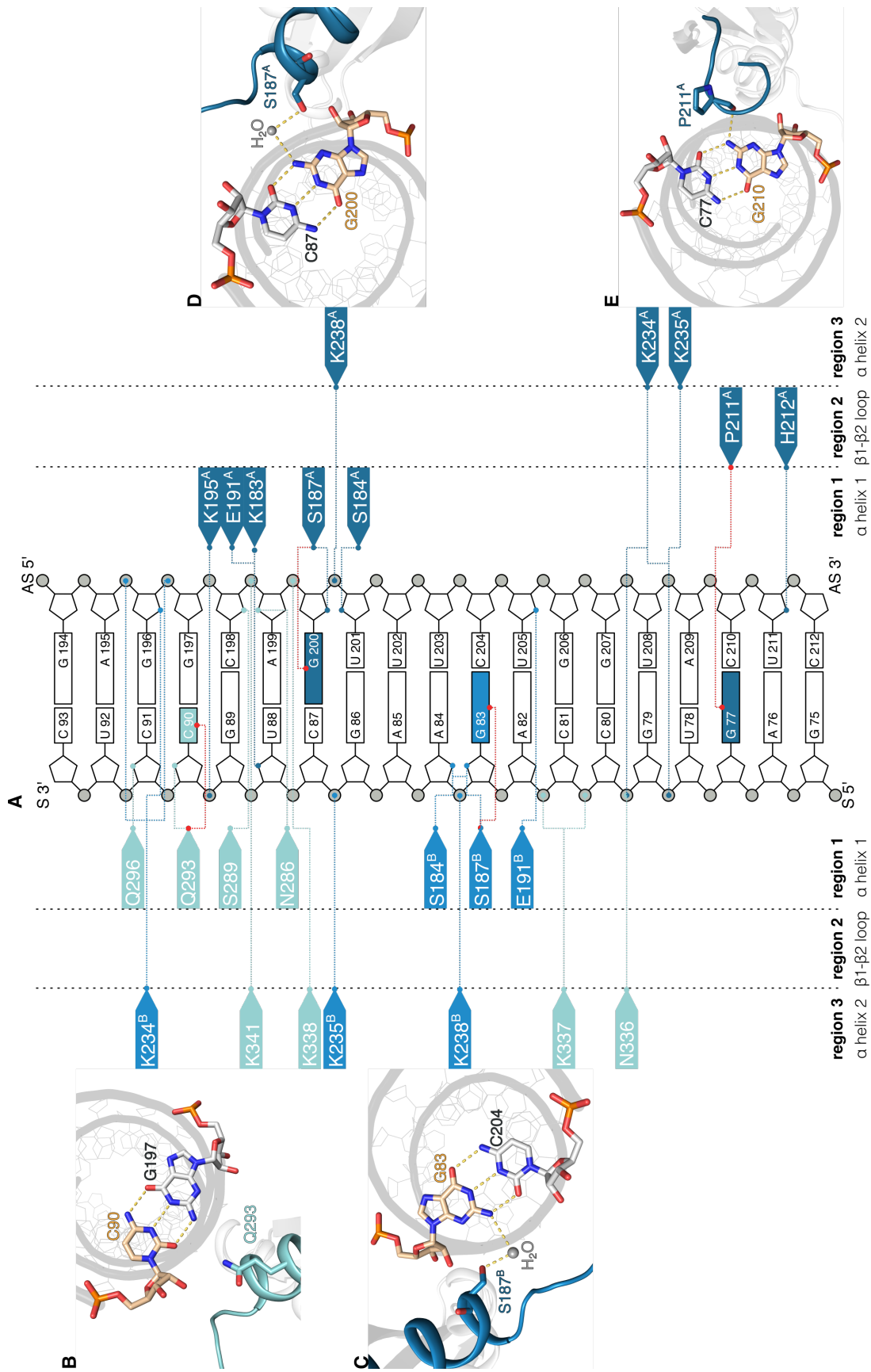


Figure 3.25. Schematic representation of the ARF1 SBS showing the interactions with the hStau1^{dsRBD3-4} (A) Schematic of the ARF1 SBS interactions with the hStau1^{dsRBD3-4}. Protein residues that interact with the RNA are colored as the domains in the structure. Residues that mediate direct contacts are indicated with a red dot. (B-E) An expanded view of the direct interactions between the hStau1^{dsRBD3-4} and the RNA minor groove. Protein residues and GC base pairs involved in the interaction are shown in sticks. Dashed lines indicate hydrogen-bonds.

phosphate group of G200 by a hydrogen-bond and salt bridges. Finally, the last Lys K341 formed a salt bridge with the phosphate group of A199, bridging the two RNA strands across the major groove (Fig. 3.26 B).

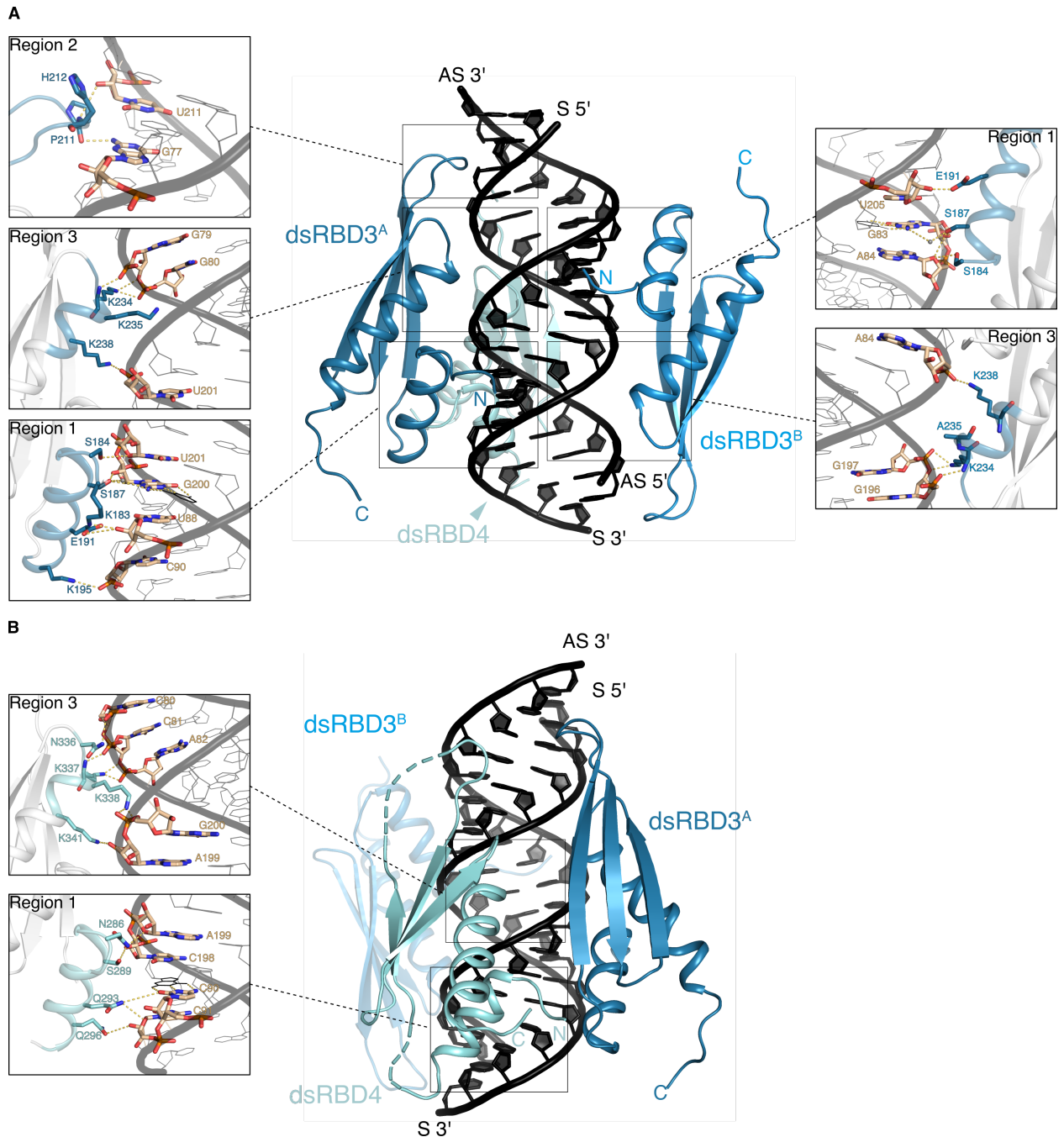


Figure 3.26. Recognition of the *ARF1* SBS by hStau1^{dsRBD3-4} **(A)** Front view and **(B)** side view of the structure. Interaction region in dsRBD3^A, dsRBD3^B and dsRBD4 are indicated in the structure (Region 1: α -helix 1, region 2: loop between β -strand 1 and 2 and region 3: α -helix 2). For each region, an extended view of the side chain residues interacting with the RNA is shown. Side chains and bases are shown as sticks. Dashed lines indicate hydrogen-bonds.

3.8.1 Structure-based mutational analyses of hStau1^{dsRBD3-4} ARF1 SBS

Three mRNA binding regions in dsRBD were identified through biochemical and structural studies of dsRBD-RNA complexes (Bycroft et al., 1995, Kharrat et al., 1995, Nanduri et al., 1998, Ryter, 1998, Ramos et al., 2000). Substitution of residues within these regions either severely reduced RNA binding or entirely abolished the dsRNA binding (Bycroft et al., 1995, Kharrat et al., 1995, Nanduri et al., 1998, Ryter, 1998, Ramos et al., 2000). For the Staufen protein, biochemical and structural data of the complexes of *Dm*Staufen dsRBD3 bound to an artificial 12 bp stem-loop identified five highly conserved residues present in region 2 and 3 essential for RNA-binding activity *in vitro* (Ramos et al., 2000). In the structure of the hStau1^{dsRBD3-4} - ARF1 SBS complex four of these residues in dsRBD3^A and ^B were implicated in dsRNA recognition (H212 in region 2 and K234, K235, K238 in region 3). In the case of dsRBD4, only residues in region 3 were present (K337, K338, and K341 in region 3). We wanted to analyze the significance of the residues present in the binding regions of the two domains for the ARF1 SBS binding and to determine the effect of a base directed read-out of the minor groove in the context of a physiological target. For this, we designed and generated a set of mutations in both domains.

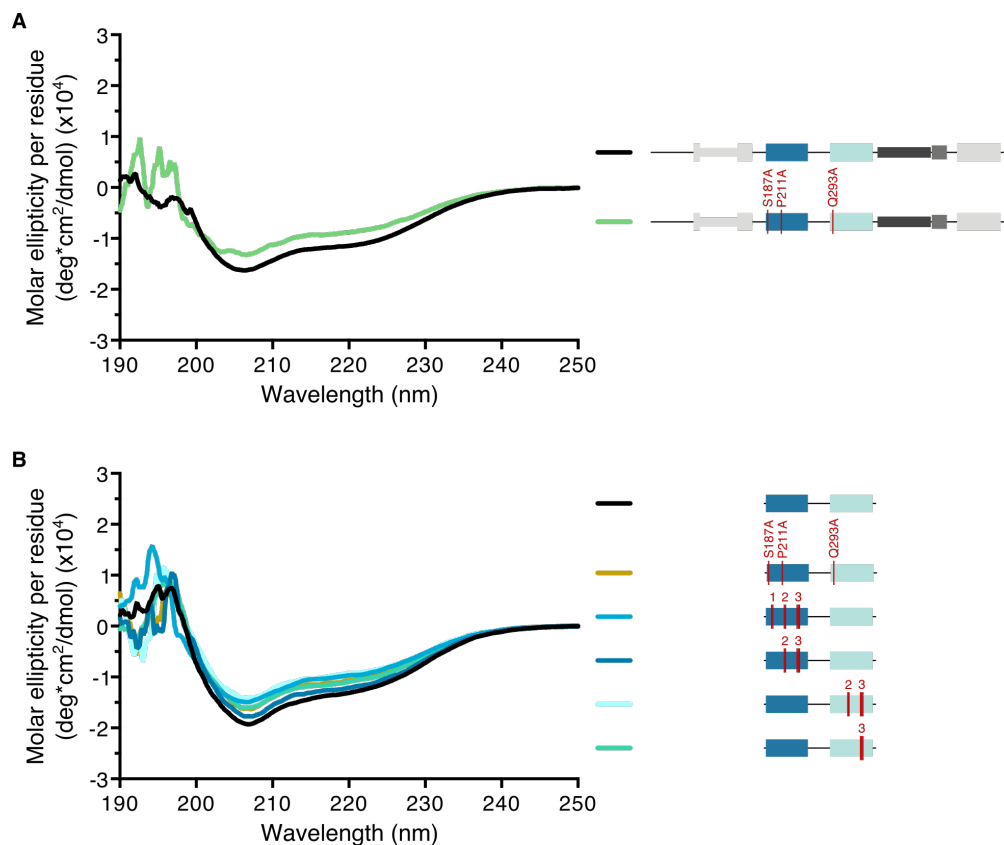


Figure 3.27. Circular Dichroism (CD) spectra of **(A)** the hStau1 FL, the sequence specific mutants, **(B)** hStau1^{dsRBD3-4} and mutants set used for the FA analysis.

One set of mutations encompassed all three regions of dsRBD3 (region 1: I186A, S187A, region 2: H212A, K214A and region 3: K234A, K235A, K238); a second set included region 2 and 3 of dsRBD3 and dsRBD4 (dsRBD4 region 2: R315A, R317A and region 3: K337A, K338A, K341A) and, the third set comprised region 3 in dsRBD4. Finally, for the last mutant construct, all three residues that contribute to the direct base recognition were mutated to alanine. The same amino acid substitution was tested in the context of the FL protein (Fig. 3.27 and 3.28). All mutants were expressed and purified as the wild-type protein and eluted as a mono-monodispersed peak in SEC (Data not shown). Dr. Daniela Lazzaretti performed Circular Dichroism (CD) analysis of the purified mutants. The CD spectra of the mutants did not show any significant conformational changes (Fig. 3.27). This data indicated that a possible reduction or loss of dsRNA binding activity was not a consequence of protein fold destabilization. With FA experiments we determined the dissociation constant of the mutants with the *ARF1* SBS_{20bp}.

Mutations in all three regions of dsRBD3 showed a stronger effect on mRNA binding affinity. In this mutant, *ARF1* SBS_{20bp} binding is reduced more than 14-fold (Fig. 3.28 Set 1). However, mutations in only two regions of the dsRBD3 exhibited a less severe effect on mRNA binding (8-fold) (Fig. 3.28 Set 2 top construct). Mutations in region 2 and 3 of dsRBD4 showed a considerable reduction of 4-fold in binding affinity, whereas a mutation only in region 3 had a even milder effect on dsRNA binding (3-fold) (Fig. 3.28 Set 2 bottom to construct and set 3). This data indicated that mutations in dsRBD3 showed a significant reduction in *ARF1* SBS binding, but did not entirely abolish mRNA binding activity. At the same time, mutations in dsRBD4 had a less drastic effect on mRNA binding. Altogether, these results stressed the significance of the interplay of the two domains in the complete recognition and interactions of the *ARF1* SBS.

The hStau1^{dsRBD3-4} direct base recognition mutant had a modest effect in binding affinity with a K_D of 22 ± 2 nM (1.6-fold), the same direct base recognition mutations in the context of the FL protein had a stronger effect in RNA binding affinity with a 3-fold reduction (K_D of 18 ± 2 nM) in comparison with the wild-type protein (Fig. 3.28). These results suggested that the hStau1 binding affinity for *ARF1* SBS was mainly mediated by sequence-independent interactions, whereas the observed base directed interactions had a minor contribution in binding. These residues might provide a mechanism of fine-tuned substrate identification.

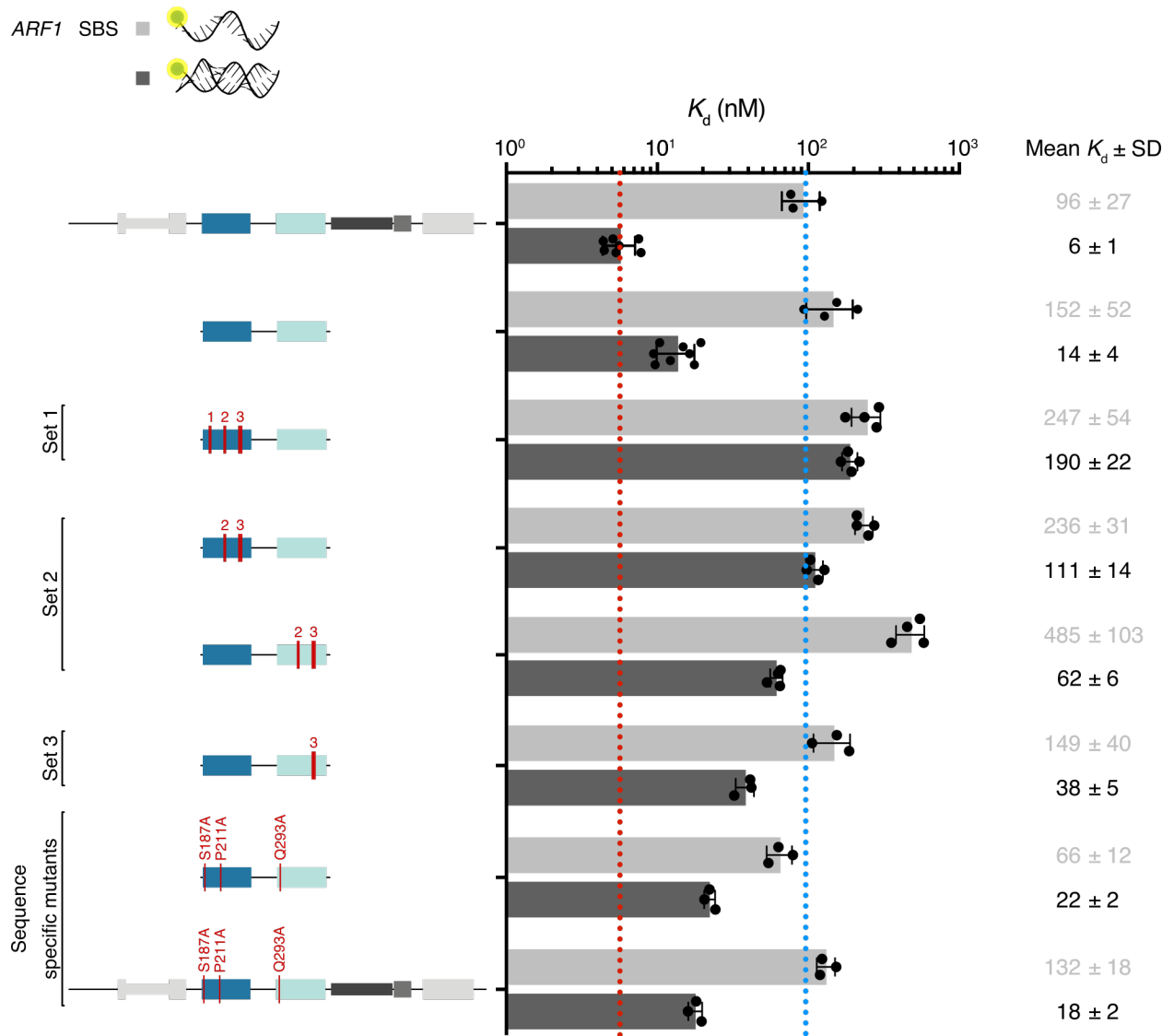


Figure 3.28. Fluorescence anisotropy assay with interactions mutants. The bar graph shows the calculated binding affinity of hStau1 FL, the hStau1^{dsRBD3-4} construct and the sets of designed mutants for the single 20 nt *ARF1* SBS sense strand (light gray) and the double-stranded 20 bp *ARF1* SBS (dark gray). The calculated K_D and their corresponding error (error bars) are the mean and SD of a minimum of three experiments. (Mean K_D and SD in nM).

4 Discussion

Based on the biochemical analysis we could elucidate some insights of the mRNA binding properties of hStau1, and by structural analysis as well as structure-based mutagenesis, the molecular principles of dsRNA recognition of dsRBD3 and 4 of hStau1 bound to the *ARF1* SBS.

Protein dimerization and oligomerization is a common mechanism in proteins to regulate their function (Marianayagam et al., 2004). Different dsRBPs had been shown to dimerize or multimerize to activate and execute diverse functions. This is the case of Xlrpba, PACT, ADAR2, PKR, TRBP, and hStau1 (Daher et al., 2001, Hitti et al., 2004, Lemaire et al., 2005, Valente and Nishikura, 2007, Martel et al., 2010, Gleghorn et al., 2013). hStau1 was shown to dimerize and multimerize *in vitro* and *in vivo* (Martel et al., 2010, Gleghorn et al., 2013). hStau1 dimerization depended on the SSM and dsRBD5 regions of the protein with a minor contribution assigned to dsRBD2 (Martel et al., 2010, Gleghorn et al., 2013). Consistently, our data confirmed the findings that the main multimerization region resided in the C-terminus of hStau1. In contrast, we could not detect dimerization activity induced by dsRBD2; even the lack of this domain did not affect dimerization. It is still possible that heterophilic interactions between dsRBD2 and 5 contribute to hStau1 self-association.

The presented SEC and SEC-MALLS assays indicate also that hStau1 can also co-exist in a monomer-dimer population in solution. This result is consistent with studies in living cells where hStau1 monomers have been shown to interact in order to form larger complexes *in vivo* (Martel et al., 2010). It is not clear, however, which factor or conditions determine and regulate dimerization and the biological function of hStau1 dimerization.

hStau1 binds *ARF1* SBS *in vitro* independently from the dimerization domain

Previous studies identified a conserved 19 bp stem within the *ARF1* 3'-UTR as a specific hStau1 binding site *in vivo*. In this thesis, I have demonstrated that hStau1 could directly interact with the *ARF1* SBS and form a stable complex *in vitro*. hStau1 showed high affinity not only for the physiological target *ARF1* SBS but also for DmStau targets (*bcd* III and *bcd* Vb) and randomly RNA substrates that contain secondary structures and predominant GC content. This is consistent with the hStau1 binding preference for RNA duplex substrates (Johnston et al., 1992, Ferrandon et al., 1994, Mallardo et al., 2003, Kim et al., 2005, Furic et al., 2007, Kim et al., 2007, LeGendre et al., 2012, Heraud-Farlow et al., 2013, Laver et al., 2013, Ricci et al., 2013, de Lucas et al., 2014,

Sugimoto et al., 2015).

Based on the biochemical analysis of the different hStau1 truncated constructs, we concluded that dsRBD3 and 4 are the minimum structure capable of binding the *ARF1* SBS. This is consistent with previous studies that identified dsRBD3 and 4 as the minimal region that binds dsRNA (Johnston et al., 1992, Marión et al., 1999, Wickham et al., 1999, Luo et al., 2002). Deletion of dsRBD2 and dsRBD5 did not affect RNA binding activity, suggesting that dimerization of the hStau1 is not necessary for RNA binding activity *in vitro*.

However, the semi-quantitative EMSA analysis showed that all hStau1 truncated constructs formed two distinct complexes in solution independently of the dimerization region. Although *ARF1* SBS was described as a specific target, hStau1 apparently binds simultaneously at different regions of the *ARF1* SBS. This suggests alternative binding modes of hStau1 and endorses the idea that hStau1 recognize primary structural elements rather than the *ARF1* SBS sequence. Furthermore, it has been reported that additional hStau1-bound duplexes are present in the *ARF1* 3'-UTR (Sugimoto et al., 2015).

It is tempting to speculate that the dsRNA targets of hStau1 could have a primary function in the formation of hStau1 contained mRNP by interacting simultaneously with multiples hStau1 that because of the proximity dimerize or oligomerize. hStau1 oligomerization could lead to structural rearrangements that position other dsRBDs such as dsRBD2 for hStau1 and hStau2, or dsRBD1 for *DmStau* and hStau2, in an orientation that allows the recognition of larger binding motifs or other secondary structures within the mRNA. Thus, Stau could regulate more accurately the affinity but also the specificity of its targets. This idea is consistent with the observations that Stau is involved in the formation of mRNP by regulating different targets or by binding to varying regions into the same RNA target (Ferrandon et al., 1994, Micklem, 2000, Duchaîne et al., 2002, Mallardo et al., 2003, Thomas et al., 2005, Kretz et al., 2012).

hStau1 dsRBD3 and 4 adopt the canonical dsRBD fold

Our structure revealed that the fold of hStau1 dsRBD3 and 4 resemble the canonical dsRBD with some differences at the helix $\alpha 1$. In comparison to other dsRBD, helix $\alpha 1$ is the region with the highest sequence variability (Masliah et al., 2012). Structural superposition of hStau1 dsRBD3 and 4 with other dsRBD presents in the protein data bank revealed two characteristic helix $\alpha 1$, an elongated and a shorter helix $\alpha 1$ (Fig. 4.1).

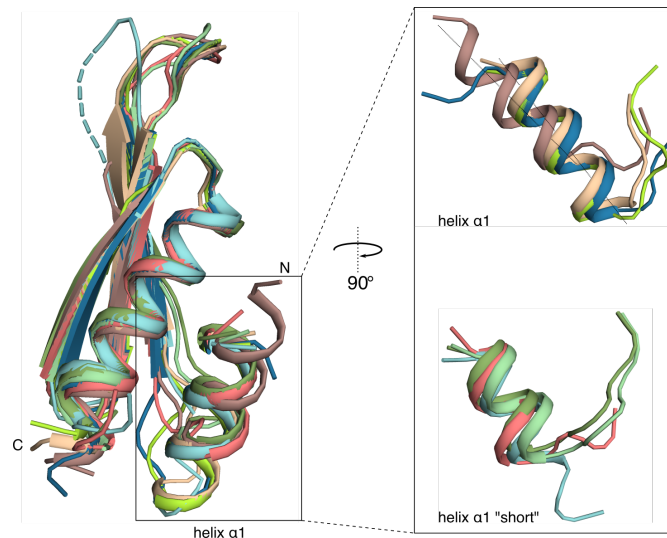


Figure 4.1. Comparison of the structures of RNA-bound dsRBDs. All structures were aligned using the secondary-structure elements $\beta 1\beta 2\beta 3\alpha 2$. Superposition of hStau1 dsRBD3^A (dark blue), dsRBD3^B (blue, rmsd 0.213 Å) and dsRBD4 (cyan, rmsd 0.723 Å), PDB: 6HTU (Lazzaretti et al., 2018)) with representative RNA bound dsRBD structures present in the protein data bank. Close up view of the helix $\alpha 1$ (rotated by 90°). The dsRBD of hStau1 (dsRBD3), ADAR2 (dsRBD1), Xlrpba, and RNase III exhibit a longer helix $\alpha 1$. The position of the extended helix $\alpha 1$ is similar in all dsRBD, except for ADAR2 dsRBD1 that adopts a different angle (dotted lines, top panel). In contrast, hStau1 dsRBD4, ADAR2 dsRBD2, and both dsRBD of NF90 have a shorter version of helix $\alpha 1$ (bottom panel). Structures used: *Xenopus laevis* Xlrpba dsRBD (lemon, rmsd 0.513 Å), PDB: 1DI2 (Ryter, 1998). *Aquifex aeolicus* RNase III dsRBD (beige, rmsd 0.621 Å) PDB: 2NUF (Gan et al., 2006). *Rattus norvegicus* ADAR2 dsRBD1 (dark rose, rmsd 0.848 Å) and dsRBD2 (rose, rmsd 0.792 Å), PDB: 2L3J (Stefl et al., 2010). *Mus musculus* NF90 dsRBD1 (dark green, rmsd 0.738 Å), dsRBD2 (green, rmsd 0.583 Å), PDB: 5DV7 (Jayachandran et al., 2015).

The helix $\alpha 1$ of hStau1 dsRBD3 as well as ADAR2 dsRBD1, Xlrpba dsRBD2, and *Aquifex aeolicus* RNase III dsRBD is extended by one turn (Ryter, 1998, Gan et al., 2006, Stefl et al., 2010). Meanwhile, helix $\alpha 1$ of hStau1 dsRBD4, ADAR2 dsRBD2, NF90 dsRBD1 and 2 is shorter (Stefl et al., 2010, Jayachandran et al., 2015). As a consequence of the shorter helix $\alpha 1$, the subsequent loop is longer and may be more flexible. This structural variation of helix $\alpha 1$ could be a determinant factor in the specific RNA recognition by dsRBD given that helix $\alpha 1$ participate in recognition of the dsRNA minor groove.

ARF1 SBS recognition by hStau1 dsRBD3 and 4

The recognition mode of hStau1 dsRBD3 and 4 resemble the canonical binding mode, observed in other dsRBD - dsRNA complexes (Ryter, 1998, Ramos et al., 2000, Stefl et al., 2010, Jayachandran et al., 2015, Masliah et al., 2018). In the structure, two dsRBD3 were bound to the RNA in an anti-parallel arrangement, and one dsRBD4 wrapped one side of the ARF1 SBS. The mode of binding of the molecules of dsRBD3 demonstrated that this domain used three canonical mRNA binding regions to recognize the A-conformation of the physiological ARF1 SBS target. Superposition of the two dsRBD3 domains presented in the structure revealed that only $\beta 1$ – $\beta 2$ loop in dsRBD3 (in

the canonical RNA binding region two) adopted different conformations to engage the minor groove of the *ARF1* SBS. This conformational change is an indication of how the flexibility of the $\beta 1$ – $\beta 2$ loop could be used by the domain to recognize different targets.

In our structure, the only visible dsRBD4 engaged just two regions to bind the *ARF1* SBS (canonical region one and three). In this domain, the $\beta 1$ – $\beta 2$ loop was not visible in the electron density supporting the idea that region two of dsRBD4 might adopt different conformations to screen between different secondary structures of the mRNA target. This is consistent with other dsRBD - dsRNA complexes in which this $\beta 1$ – $\beta 2$ loop was disordered (Ryter, 1998) or demonstrated by the relaxation properties of this region in *DmStau* in NMD data (Ramos et al., 2000, Castrignanò et al., 2002). Mutations of conserved residues in hStau1 involved in RNA binding in region two and three of dsRBD4 confirmed that dsRBD4 mutants affected RNA binding affinity less than the corresponding mutations in dsRBD3 in FA experiments. Consistently, previous studies showed weaker RNA-binding activity of dsRBD4 in comparison with dsRBD3 in filter binding assays (Wickham et al., 1999, Luo et al., 2002). It is plausible that the weak binding of dsRBD4 could facilitate a transient interaction with the RNA for screening the RNA for the correct SBS and subsequently induce anchoring of dsRBD3. This idea is supported further by the finding that hStau1 exhibited sliding activity on mRNA targets *in vivo* (Wang et al., 2015).

As for other dsRBD domains, our structural and mutagenesis data showed that the dsRNA recognition mode of hStau1 is primarily determined by non-sequences specific interactions with the RNA backbone. This reinforced the idea that the recognition mode of dsRBD is primarily determined by the dsRNA A-form conformation (Ryter, 1998, Chang and Ramos, 2005, Stefl et al., 2005). In addition, dsRBD3 exhibited a base-specific interaction mediated by helix $\alpha 1$ at the minor groove. This interaction was present in the two dsRBD3 molecules in the structure. Ser187 bound the amino group of G200 and G83 (for molecule A and B respectively). These interactions are water-mediated and possibly the result of an adaptation of the helix $\alpha 1$ to induce a specific binding. Two other dsRBD containing proteins ADAR2 and NF90 in complex with RNA exhibited sequence-specific interactions with a Met at the corresponding position (Stefl et al., 2010, Jayachandran et al., 2015). The $\beta 1$ – $\beta 2$ loop of dsRBD3^A mediated the next additional sequence-specific interaction at the subsequent RNA minor groove. Here, the backbone of Pro211 interacted with the exocyclic amino group of G77. Similar sequence-specific interactions by $\beta 1$ – $\beta 2$ loop have been described for other dsRNA-dsRBD complexes (Ryter, 1998, Chang and Ramos, 2005, Stefl et al., 2005, 2010, Jayachandran et al., 2015).

The base-directed contacts mediated by hStau1 helix $\alpha 1$ and $\beta 1$ - $\beta 2$ loop favor GC base pairs. The structure revealed that the base-directed interactions involved the guanine amino group of the GC pair. Thus, AU base pairs would be probably discriminated against GC. The principal

reason is the absence of hydrogen bond donor/acceptor of AU base pairs at the minor groove. This could constrain not only the interactions with the protein but also allowed discrimination over GC pairs. At the same time, it is likely that the domain would fail to discriminate between base pairs reversal GC or CG because of the relative same center position of the amino group in both cases (Seeman et al., 1976). Mutation of the residues involved in base-directed interactions showed no significant effect on RNA binding in FA. However, in the context of the FL protein, a small negative effect on mRNA binding activity could be observed, suggesting that two GC pairs in the continuous minor groove of the *ARF1* SBS A-conformation contribute to *ARF1* SBS readout by hStau1. Although the contribution to the overall binding affinity of the GC recognition by helix $\alpha 1$ and $\beta 1$ - $\beta 2$ loop for the binding was negligible, it could be critical for target discrimination.

Comparison of the hStau1^{dsRBD3-4} - *ARF1* SBS structure with the *DmStau* dsRBD3 - 12 bp stem-loop complex

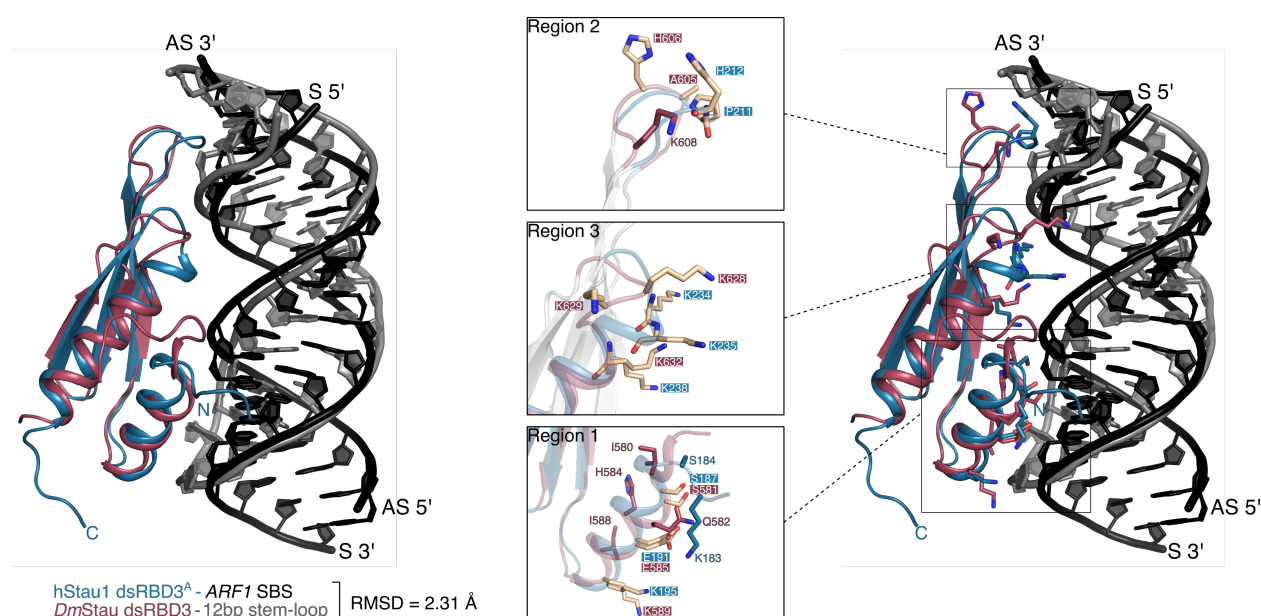


Figure 4.2. Superposition of hStau1 dsRBD3^A - *ARF1* SBS complex (dsRBD3^A dark blue, *ARF1* SBS black) with the structures of *DmStau* dsRBD3 - 12 bp stem-loop (dsRBD3 darkred, stem-loop light gray). For each region, an extended view of the side chain residues interacting with the RNA is shown. Side chains are shown as sticks. Common interacting residues are colored salmon. Structures used: *DmStau* - 12 bp stem-loop complex (NMR State 33), PDB: 1EKZ (Ramos et al., 2000). hStau1 dsRBD3^A - *ARF1* SBS complex, PDB: 6HTU (Lazzaretti et al., 2018).

Globally, the crystallographic structure of hStau1 dsRBD3^A complexed with the *ARF1* SBS resembles the NMR structure of the complex between *DmStau* dsRBD3 and a 12 bp stem-loop RNA (Fig. 4.2 RMSD 2.31 Å). Surprisingly, the optimized stem-loop RNA can adopt a similar fold as the *ARF1* SBS despite a 7 bp difference and the presence of a tetraloop. For this, the tetraloop mimics the fold and shape of a consecutive minor groove. As a consequence, most of the residues involved

in the binding pointed toward the same recognition region as in the dsRBD3^A - *ARF1* SBS structure.

DmStau dsRBD3 and dsRBD3^A adopt a similar conformation. Some differences arise from the position of $\beta 1$ - $\beta 2$ loop and the angle of helix $\alpha 1$. On the first minor groove (Fig. 4.2 Region 2), $\beta 1$ - $\beta 2$ loop of dsRBD3^A is pulling toward the bases to interact with the GC pair. In contrast, *DmStau* dsRBD3 $\beta 1$ - $\beta 2$ loop is less flexed to extends over the minor groove and interact with the RNA backbone. Both domains use the same conserved residues in $\beta 1$ - $\beta 2$ loop to interact with the minor groove. *DmStau* dsRBD3 exhibited an additional contact as a consequence of the extension of the loop over the groove. In region 3, the same conserved residues in *DmStau* and hStau1 dsRBD3 interact with the major groove similarly by binding across the major groove (Fig. 4.2 Region 3).

Significant differences are seen in region 1. The different angle of *DmStau* helix $\alpha 1$ is an induced fit caused by the different geometry of the tetraloop in comparison with the minor groove of the *ARF1* SBS duplex. The conformation of the loop extends the binding surface allowing a reorientation of *DmStau* helix $\alpha 1$. This reorientation of *DmStau* helix $\alpha 1$ facilitates more extensive contacts with the RNA as for the hStau1 helix $\alpha 1$. A conserved Ser in helix $\alpha 1$ (*DmS581*- C, hS187 - G) is the only common residue that mediates a base directed interaction in both complexes (Fig.4.2 Region 1). Two additional residues in the *DmStau* helix $\alpha 1$ Glu and Ile (E585 and I588) make base contacts through stacking and van de Waals interactions. The equivalent Glu in hStau1 helix $\alpha 1$ (E191) mediate a nonspecific interaction with the RNA. The other residues involved in the RNA interaction maintain and stabilize the position of helix $\alpha 1$ by interacting with the RNA backbone. Consistently, *DmStau* helix $\alpha 1$ exhibited more contacts as hStau1 helix $\alpha 1$ due to the tetraloop fold.

Comparison of the hStau1^{dsRBD3-4} - *ARF1* SBS structure with other dsRBD/RNA Complexes

Comparison of the structure of dsRBD3^A with other dsRNA-dsRBD complexes (where direct read out of the RNA minor groove is also achieved by helix $\alpha 1$ and $\beta 1$ - $\beta 2$ loop) revealed a similar mode of binding (Fig. 4.3). However, the flexible conformations of the $\beta 1$ - $\beta 2$ together with sequence variability and orientation of helix $\alpha 1$ confer each dsRBD with domain specific binding. The variations in helix $\alpha 1$ and $\beta 1$ - $\beta 2$ loop led to differences in the number of bases between these regions that in turn determine a different "register length" to discriminate between targets (Ryter, 1998, Ramos et al., 2000, Chang and Ramos, 2005, Stefl et al., 2005, 2010, Masliah et al., 2012, Jayachandran et al., 2015). In principle, dsRBD could discriminate longer RNA substrates by the interplay of two or more dsRBD. For example, *DmStau* and hStau2 have three and four dsRBD, respectively that bind RNA (Johnston et al., 1992, Ramos et al., 2000, Mickle, 2000, Heber et al., 2019). It is possible that the proteins could recognize not only more extended regions but also tertiary structures into the same RNA. Additionally, a certain grade of influence in target discrimination is also attributed to further secondary structures present in the dsRNA that could alter the binding register of the Stau protein through dimerization as in the case of *bcd* in *DmStau* (Ferrandon et al., 1994).

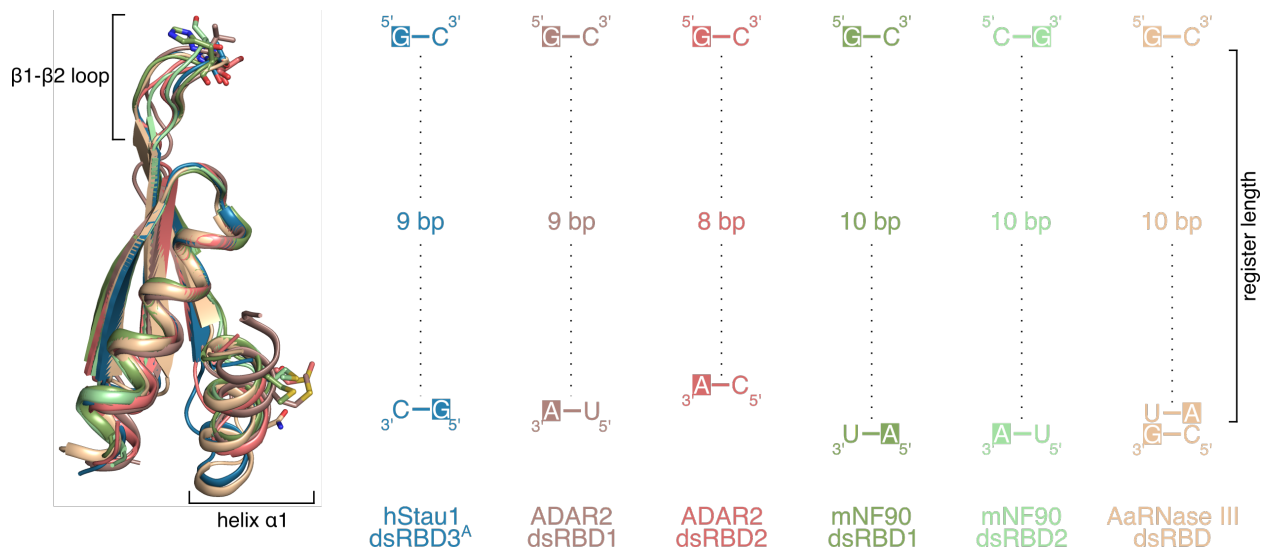


Figure 4.3. Superposition of hStau1 dsRBD3^A (dark blue) with structures of RNA bound dsRBDs that exhibit sequence-specific recognition mediated by helix $\alpha 1$ and $\beta 1$ - $\beta 2$ loop. Residues involve in the interactions are shown as sticks. Bases in direct contact with the dsRBD are highlighted in white. The register length is determined by the x number of bases in between of the base directed interactions. Structures used: *Aquifex aeolicus* RNase III dsRBD (beige, RMSD 0.621 Å) PDB: 2NUF (Gan et al., 2006). *Rattus norvegicus* ADAR2 dsRBD1 (dark rose, RMSD 0.848 Å) and dsRBD2 (rose, RMSD 0.792 Å), PDB: 2L3J (Stefl et al., 2010). *Mus musculus* NF90 dsRBD1 (dark green, RMSD 0.738 Å), dsRBD2 (green, RMSD 0.583 Å), PDB: 5DV7 (Jayachandran et al., 2015)

Dr. Daniela Lazzaretti introduced directed mutations in the *DmStau* protein to test the *in vitro* and *in vivo* significance of the interactions determined by the structural data. For the *DmStau*-FA analysis, the stem-loop III from *bcd* 3'UTR was used to test the affinity in the context of a physiological target. The FA data of the *DmStau*1 RNA-backbone-binding mutants support the apparent dominance of dsRBD3 over dsRBD4 in RNA-binding activity observed for hStau1 (Lazzaretti et al., 2018). Consistently *in vivo* dsRBD3 mutation of residues interacting with the RNA backbone (conserved residues H606A and K608A in region 2; K628E and K629A in region 3) could not rescue the Stau-null phenotype; quite contrary to the similar mutations in dsRBD4 that restore the Staufen phenotype. These result, raise the question if other domains of the *DmStau* compensated the dsRBD4 mutation effect or if the function of dsRBD4 is not essential for RNA binding, but it could be important in other regulatory mechanisms later in development as for dsRBD2 and 5 with *osk* (Ferrandon et al., 1994, Micklem, 2000).

Interestingly, base-directed mutations in either dsRBD3 or 4 (S581 and Q718 corresponding to S187A and Q293A in hStau1 respectively) had a mild effect *in vitro* but exhibited a clear additive negative effect in rescue experiments *in vivo*. These observations suggest that shape-specific, as well as sequence-specific recognition, are essential for the RNA binding activity of the Stau

protein. At the same time, these results showed that the *in vitro* determined binding affinity of Stau is insufficient to predict the biological affinity or specificity. It will be necessary to take account not only affinity distribution methods with different dsRNA duplexes (*in vitro*) but also hiCLIP methods (*in vivo*) to have a more global and accurate view of the biological affinity and specificity of Stau.

The here present data evidence the promiscuous binding nature of Stau a protein that was described to be mostly shape-specific than sequence-specific. However, it has been shown in different species that Stau regulates various specific targets in different stages of development (St. Johnston et al., 1989, Johnston et al., 1991, 1992, Broadus and Doe, 1997, Broadus et al., 1998, Ephrussi et al., 1991, Ferrandon et al., 1994, Fuerstenberg et al., 1998, Kim-Ha et al., 1995, Li et al., 1997, Schuldt et al., 1998, Kiebler et al., 1999, Duchaine et al., 2002, Macchi et al., 2004, Vessey et al., 2012, Kusek et al., 2012). It is still unclear how the promiscuous binding of Stau can fit the specific requirements in a biological context. Our *in vitro* and *in vivo* measurements, as well as the complex structure, are probably not sufficient to answer this question. Though our data indicate that although each hStau1 domain (dsRBD3 and 4) can mediate a shape-dependent and a base directed recognition, is the interplay of them both what accomplish full RNA binding and probably a more accurate sequence-specific recognition. This interdependency between the domains could be the regulatory factor that connects the promiscuous- and the specific-binding nature of Stau. Consequently, the affinity and specificity context is likely to be affected by the presence of multiple dsRBDs that can bind RNA. So is the case of *DmStau* or hStau2 where further dsRBD (dsRBD1 in *DmStau*, dsRBD1 and 2 in hStau2) could enhance binding affinity (Heber et al., 2019) but also contribute to the discrimination of larger substrates.

Based on previously published data and the presented thesis, a possible model of RNA recognition by Stau integrates the RNA binding ability and specificity of the active canonical dsRBD. In this model, a minimum of two canonical dsRBDs mediate dsRNA binding, where a "low-affinity" dsRBD is compensated by a "high-affinity" dsRBD that specifically recognized the RNA substrate. We speculate that the "low-affinity" dsRBD first screens the dsRNA substrate to find GC pairs at the minor groove; for this, the protein could slide over the RNA as previously reported (Wang et al., 2015). That way the "high-affinity" dsRBD could be accommodated at the binding sites, where the specific register length is expected, to accomplish full recognition and binding.

5 Materials and Methods

5.1 Materials

5.1.1 Chemicals

Standard chemicals that are not listed here were purchased from Sigma-Aldrich, Germany, Merck, Germany or Roth, Germany.

Table 5.1. List of chemicals used for this thesis

Chemical	Supplier
Acetic acid	Merck, Germany
Agarose Standard, for electrophoresis	Roth, Germany
Ampicillin sodium salt	Roth, Germany
Ammonium persulfate (APS)	Sigma-Aldrich, Germany
Ammonium sulfate ((NH ₄) ₂ SO ₄)	Merck, Germany
Boric acid	Merck, Germany
Bovine serum albumin (BSA)	Sigma-Aldrich, Germany
Bradford assay reagent	Bio-Rad, Germany
Bromphenol Blue sodium salt	Sigma-Aldrich, Germany
Calcium chloride dihydrate	Merck, Germany
Complete Protease Inhibitor, EDTA-free	Roche, Germany
Di-Potassium hydrogen phosphate (K ₂ HPO ₄)	Merck, Germany
1,4-Dithiothreitol (DTT)	Biomol
Ethanol, absolut	VWR
Ethylene glycol	SERVA, Germany
Glycerol, 86 %	Sigma-Aldrich, Germany
Glycerol, 100 %	Merck, Germany
Guanidine hydrochloride	Roth, Germany
Hydrogen chloride (HCl)	VWR
Imidazole, low UV-absorption	Merck, Germany
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	Roth, Germany
Isopropanol	VWR
Kanamycin sulfate	SERVA, Germany
Magnesium chloride hexahydrate	Fluka, Switzerland
Magnesium sulfate (MgSO ₄)	VWR
Methanol	VWR

Chemical	Supplier
2-Mercaptoethanol (β -ME)	Roth, Germany
PhastGel Blue R-350	GE Healthcare, Germany
Polyethylene glycol 200-35000	Sigma-Aldrich, Germany
Polyethylene glycol 3350	Sigma-Aldrich, Germany
Potassium chloride (KCl)	Merck, Germany
Potassium di-hydrogen phosphate (KH ₂ PO ₄)	Sigma-Aldrich, Germany
Rotiphorese [®] Gel 30 (29.2% acrylamide, 0.8% N,N'-methylene-bis-acrylamide)	Roth, Germany
SafeView TM Nucleic Acid Stain	Gentaur
Sodium dodecyl sulfate (SDS)	Roth, Germany
Sodium hydroxide (NaOH)	Merck, Germany
Sodium chloride (NaCl)	VWR
Spectinomycin-2HCl pentahydrate	SERVA, Germany
Tetramethylethylenediamin (TEMED) Triton X-100	Fluka, Switzerland
Triton X-100	Merck, Germany
Trizma [®] base	Roth, Germany
TRIZOL [®]	Sigma-Aldrich, Germany
Merck, Germany	Invitrogen, Germany
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Roth, Germany

5.1.2 Buffer solution

Milli-Q water was used, for the preparation of all buffers. The pH was adjusted by titration with 37% HCl or 10 M NaOH. The prepared buffers were then filter-sterilized (0.22 μ m) or autoclaved.

Table 5.2. List of buffers used for this thesis

Buffer	Composition
10×TBE	0.89 M Tris pH 8.0, 0.89 M Boric Acid, 25 mM EDTA
10×SDS running buffer	250 mM Tris pH 6.8, 2 mM Glycine, 1% (w/v) SDS
5×SDS loading buffer	250 mM Tris pH 6.8, 8% (w/v) SDS, 10% (v/v) β -ME, 30% (v/v) Glycerol, 0.02% (w/v) Bromophenol blue
Destaining solution	40% (v/v) ethanol, 20% (v/v) acetic acid

5.1.3 Consumables

Table 5.3. List of chemicals used for this thesis

Consumable	Supplier
Amersham Hybond-P PVDF membrane	GE Healthcare, Germany
Clear Sealing Tape	3M Hampton Research, USA
ClearVue™ Sealing Sheets	Molecular Dimensions, UK
CleneGlass™ Coverslips	Molecular Dimensions, UK
Concentrators (Amicon Ultra)	Millipore, USA
Crystallization plates MRC, 96 well, sitting drop	Molecular Dimensions, UK
Dialysis membranes	Spectra/Por, USA
Electroporation cuvettes	Bio-Rad, Germany
Eppendorf safe-lock micro test tubes 1.5 ml and 2 ml	Eppendorf, Germany
Mounted Cryo loops	3M Hampton Research, USA
Sterile filters 0.22 µm, 0.45 µm	Sarstedt, Germany
Syringes	Braun, Germany
Tubes (15 ml and 50 ml)	Greiner-Bio-One, Germany
24-well plates, "Linbro style"	Molecular Dimensions, UK
Weighting dishes	Roth, Germany

5.1.4 Culture media

All types of media and additional buffers were sterilized by autoclaving (121 °C, 20 min) before usage. Components were dissolved in deionized and filtered water. The media used in this work are listed in table 5.4.

Table 5.4. Culture media

Medium	
LB medium (for 1 l)	10 g Tryptone 5 g Yeast extract Sodium chloride
TB medium (for 0.9 l)	12 g Bacto™ Tryptone 24 g Bacto™ Yeast extract 4.7 mL Glycerol, 86 % After cooling down, 100mL Potassium phosphate buffer (0.17 M KH ₂ PO ₄ , 0.72 M KH ₂ PO ₄) is added.
ZY medium (for 0.93 l)	10 g Tryptone 5 g Yeast extract After cooling down, 1 mL MgSO ₄ (1 M), 20 mL 5052 (50×) and 50 mL NPS (20×) is added.

5.1.5 Resins and columns

Table 5.5 lists the resins and columns used routinely in this thesis.

Table 5.5. Beads and columns

Beads and column types	Name	Supplier
Cobalt beads	HIS-Select [®] Co Affinity Gel	Sigma-Aldrich, Germany
Nickel beads	Ni-NTA Agarose	Qiagen, Germany
Cation exchange beads	Macro-Prep [®] High S Support	Bio-Rad, Germany
Gelfiltration	HiLoad [™] 16/60 Superdex [™] 200	GE Healthcare, Germany

5.1.6 Enzymes, Kits and Markers

All enzymes, Kits and Markers used in this work are listed in table 5.6

Table 5.6. Enzymes, commercial kits and markers

Type	Name	Supplier
Bovine serum albumin (BSA)	Sigma-Aldrich, Germany	
DNA marker	1 kb DNA Ladder	NEB
	100 bp DNA Ladder	NEB
Ligase	T4 DNA Ligase	Thermo Scientific, Germany
Plasmid preparation	Qiagen, Germany Plasmid Midi Kit	Qiagen, Germany
Polymerase	<i>Pfu</i> DNA Polymerase	Thermo Scientific, Germany
Proteases	Trypsin	Sigma-Aldrich, Germany
	Subtilisin	Sigma-Aldrich, Germany
	GluC	Roche, Germany
	Elastase	Sigma-Aldrich, Germany
	TEV	Home-made, recombinant
Protein marker	PageRuler Prestained Protein Ladder	Thermo Scientific, Germany
	Pierce Unstained Protein MW Marker	Thermo Scientific, Germany
	PageRuler Unstained Protein Ladder	Thermo Scientific, Germany
Restriction enzymes	BamHI-HF [™]	NEB
	NdeI	NEB

5.1.7 Crystallization screens

Table 5.7. List of crystallization screens used for this thesis

Consumable	Supplier
Classics Suite, PEGs Suite, PEGs II Suite, PACT Suite, Nucleix Suite, JCSG+ Suite	Qiagen, Germany
Additives, Index, SaltRX, complex screen 1, complex screen 2	Hampton Research, USA
Crystal platform Magic 1 , AJ1, Morpheus	MPI-Martinsried Crystallization Facility

5.1.8 Technical equipment

All devices used in this work are listed in table 5.8

Table 5.8. Technical equipment

Machine	Name	Manufacturer
Chromatography system	ÄKTApri [®] plus	GE Healthcare, Germany
Dispensing robot	mosquito [®] crystal	TTP Labtech Ltd, UK
Electroporator	MicroPulser [™]	Bio-Rad, Germany
Gel system	Mini-PROTEAN [®] Tetra Cell	Bio-Rad, Germany
Homogenizer	M-110L Pneumatic	Microfluidics
Incubator	Heraeus [®]	Thermo Scientific, Germany
Phosphorimager	Phosphorimager Typhoon 8600	GE Healthcare, Germany
Pipettes	Pipetman	Eppendorf, Germany
Power supplies	Power supplies	Bio-Rad, Germany
Screen optimizer	dragonfly [®] crystal	TTP Labtech Ltd, UK
Shakers	Innova [®] Laboratory Shakers	New Brunswick
Sonicator	Sonopuls HD 3100	Bandelin
Spectrometer	NanoDrop 2000	Thermo Fisher Scientific, USA
	Microtip MS 73	Bandelin
Tabletop centrifuge	Centrifuge 5424	Eppendorf, Germany
Tabletop centrifuge, refrigerated	Centrifuge 5415 R	Eppendorf, Germany
	Allegra [®] X-15R	Beckman Coulter
Thermocycler	MJ Mini [™] Thermal Cycler	Bio-Rad, Germany
Thermomixer	Thermomixer comfort	Eppendorf, Germany
Ultracentrifuge, refrigerated	Avanti [®] J26XPI	Beckman Coulter
Ultracentrifuge rotors	JA-25.50	Beckman Coulter
	JLA-8.1000	Beckman Coulter
	JLA-16.250	Beckman Coulter

5.1.9 Bacterial strains

E. coli strains were used for plasmid production and protein expression. Transformation of plasmids was carried out using electroporation.

Table 5.9. *E. coli* strains used for cloning and protein expression

	Genotype and Characteristics	Supplier
<i>E. coli</i> XL1-Blue	<i>endA1 supE44 thi-1 hsdR17 recA1 gyrA96 relA1 lac [F' proAB lacIq ZΔM15 Tn10 (Tetr)]</i> Typical host strain for cloning applications and for DNA preparations; can be used in blue-white colony screenings.	Stratagene
<i>E. coli</i> BL21 (DE3) Gold	<i>F- ompT hsdS(rB- mB-) dcm+ Tetr gal λ (DE3) endA Hte</i> Features the phenotype which increases transformation efficiency over BL21; can be used for DNA preparations.	Stratagene
<i>E. coli</i> BL21 (DE3) Star	<i>F- ompT hsdSB (rB- mB-) gal dcm rne131 (DE3)</i> RNase deficient strain; reduction in mRNA degradation and increase the yield of recombinant protein	Invitrogen, Germany

5.1.10 Oligonucleotides

Primers used for cloning contain a NdeI site prior to an ATG start codon in the forward primer and a BamHI site after the TAA stop codon in the reverse primers. All primers used in this work are listed in table 5.10

Table 5.10. Primer sequences used for mutagenesis

Name	Sequence 5' -> 3'
STAU _co 182-end mut for	GGCGAAAACCTTTACTTCCAGGGCCATATGAACAAGAGCGAAATTTCTCAGGTC
STAU _co 182-end mut rev	GACCTGAGAAATTTTCGCTCTTGTTCATATGGCCCTGGAAGTAAAGGTTTTCGCC
STAU _co 360-end mut for	CTGGGCTTCAAAGTCCGCAGCGTCAATAAGGATCCTAATAGTCTAGAGCTAGC
STAU _co 360-end mut rev	GCTAGCTCTAGACTATTAGGATCCTTATTGACGCTGCGGAACCTTTGAAGCCAG
STAU _co 70-end mut for	GGCGAAAACCTTTACTTCCAGGGCCATATGAGCATTACCCCGACGGTCGAACTG
STAU _co 70-end mut rev	CAGTTCGACCGTCGGGGTAATGCTCATATGGCCCTGGAAGTAAAGGTTTTCGCC

5.1.11 RNA oligonucleotides

For the RNA interaction experiments we ordered the *ARF1* SBS, the 5'-fluorescein-labeled *ARF1* SBS_{20bp} and the tested RNA strands from Integrated DNA Technologies, Inc.(IDT).

Table 5.11. RNA oligonucleotides used for RNA interaction experiments

Name	Sequence 5' -> 3'
<i>ARF1</i> SBS sense	GAGUGCCAGAAGCUGCCUC
<i>ARF1</i> SBS antisense	GAGGCAGUUUCUGGUACUC
<i>ARF1</i> SBS antisense (pc)	GAGGCAGcUUCUGGcACUCA
<i>ARF1</i> SBS _{20bp} sense	UGAGUGCCAGAAGCUGCCUC
<i>ARF1</i> SBS _{20bp} antisense	GAGGCAGUUUCUGGUACUCA
PolyU ₂₀ (U ₂₀)	U ₂₀
PolyUA ₁₀ (UA ₁₀)	UA ₁₀
PolyUC ₁₀ (UC ₁₀)	UC ₁₀
PolyUG ₁₀ (UG ₁₀)	UG ₁₀
20 bp <i>ARF1</i> hairpin	UGAGUGCCAGAAGCUGCCUCCGUAU GAGGCAGUUUCUGGUACUCA
GC-rich hairpin	GAGGCAGUUUCUGGUACUCA
K10 TLS	CUUGAUUGUAUUUUUAAAUAUUUCUUA AAAACUACAAUUAAG
K10 TLS (2xGC)	CGCUUGAUUGUAUUUUUAAAUAUUUCUUA AAAACUACAAUUAAGCG
<i>bcd</i> III	GGGCGCUAUUCGCCUUAUGAUGUAUCUGGGUGGCUGCUCACUAAAGCCC GGGAAUAUGCAACCAGUUACAUUUGAGGCCAUUUGGGCUUAAGCG
<i>bcd</i> Vb	CCCAAAUUGAAAAUUGUUUCUCUUGGGCGUAAUCUCAUACAUAUGAUUACC CUUAAAGAUCGAACAUUUAACAUAUAUUUGGG

5.1.12 Software

Table 5.12. RNA oligonucleotides used for RNA interaction experiments

Software	Reference
Coot	Emsley and Cowtan, 2004; Emsley et al., 2010
GraphPad Prism	GraphPad Software, Inc., USA
PyMOL	Schrödinger, LLC (2015)
XDS	Kabsch, 2010

5.2 Methods

5.2.1 Nucleic acid methods

DNA amplification

The human Staufen 1 protein sequence was ordered as a codon-optimized construct in the pnC57 plasmid from GenScript. The hStau1 codon-optimized gene was amplified by Polymerase chain

reaction (PCR) using *Pfu* DNA Polymerase (Thermo Scientific, Germany) according to the instructions of the manufacturer. The PCR product was analyzed by agarose gel electrophoresis (1.5% w/v containing 1 x SYBR Safe DNA Stain (Invitrogen)) and purified from the gel using *QIAquick Gel Extraction Kit* (Qiagen, Germany). The DNA was resuspended in 30 μ l H₂O.

Cloning in pET-MCN vectors

The amplified hStau1 codon-optimized gene was cloned into a pET-MCN vector. For this purpose pET-MCN vectors with an N-terminal GST tag and with a hexahistidine (His) tag were linearized using restriction enzymes NdeI and BamHI (NEB, Germany) and incubated for 2-3 h at 37 °C (table 5.13). The digested vectors were loaded into an agarose gel and further purified using *QIAquick Gel Extraction Kit* (Qiagen, Germany). The extracted DNA vectors were resuspended in 30 μ l H₂O.

Table 5.13. Restriction digest of pET-MCN vectors

Component	Volume
pET-MCN vector	150-200 ng
10x NEBuffer IV	5 μ l
NdeI	1 μ l
BamHI	1 μ l
BSA	0.5 μ l
ddH ₂ O	to 50 μ l

Ligation

The ligation was performed with the T4 DNA Ligase (Thermo Scientific, Germany) and according to the manufacturer's instructions. For the ligation, 100ng of the digested pET-MCN vector was mixed with two to five fold molar excess of the DNA insert. The reaction was incubated for 2h at 20 °C or overnight at 16 °C.

Transformation in *E.coli*

For transformation purpose, circa (ca.) 150 ng DNA were mixed with 50 μ l electro-competent *E. coli* cells on ice. The cells were subjected to a 4.8 ms pulse of 2.5 kV in an ice-cold electroporation cuvette. Afterward, the cells were resuspended in 300 μ l of LB medium and incubated for 1 h at 37 °C while shaking. The transformed cells were streaked out on an agar plate containing the selective antibiotics and incubated overnight at 37 °C.

Plasmid preparation and verification

A single colony from the grown LB-agar plate was isolated and used to inoculate 5mL LB medium with the correct antibiotics. The cultures were grown overnight at 37 °C while shaking. Plasmid purification was performed using *QIAGEN Plasmid Midi Kit* (Qiagen), according to the manufacturer's

instructions. All plasmids were sequenced to verify the cloning success (Sequence Core Facility MPI in Tübingen).

Site-directed mutagenesis

To generate the shorter constructs and to introduce point mutations site-directed mutagenesis was performed. To this end, complementary oligonucleotide primers were designed that contain the desired deletion or base substitution and ordered from Integrated DNA Technologies, Inc.(IDT). The reaction mixture for a typical *PfuUltra* high-fidelity DNA polymerase-based PCR are shown in table 5.14.

Table 5.14. Mutagenesis reaction

Component	Volume
Plasmid DNA	100-150 ng
10x <i>PfuUltra</i> HF reaction buffer Buffer	5 μ l
dNTPs (10 mM each dNTP)	1 μ l
sense oligonucleotide primer (2 μ M)	1 μ l
antisense oligonucleotide primer (2 μ M)	1 μ l
<i>PfuUltra</i> HF DNA polymerase (2.5 U/ μ l)	1 μ l
ddH ₂ O	to 50 μ l

Parameters in the amplification program such as temperature and number of cycles were adjusted according to the oligonucleotide primers and plasmid template. The methylated parental plasmid was digested with 1 μ l DpnI enzyme for 3h at 37°C. Finally, 5 μ l of the mutagenesis reaction was transformed into 50 μ l of *E.coli* XL1-Blue strain.

5.2.2 Recombinant protein expression in *Escherichia coli*

Pull-down experiments

Pull-down experiments were performed to analyze the solubility of the protein and to determine the best conditions to produce high amounts of recombinant hStau1. For this, different cell strains (*E.coli* BL21 (DE3) Star, *E.coli* BL21 (DE3) Star and *E.coli* BL21 (DE3) Plyss) and media (LB-, TB- and ZY-medium) were tested. 20 mL cultures were inoculated and incubated at 37°C while shaking until the optical density (OD600) reached 0.6 for LB and TB media and 2 for the ZY medium. The cultures were cooled to 20°C while shaking. In the case of the LB and TB culture, 0.1 mM Isopropyl- β -D-thiogalactopyranoside (IPTG) was added to induce protein expression. The cultures were incubated for 16 to 18 h. For each culture two samples of 2ml were centrifugated. The pellet was resuspended in lysis buffer (Table 5.15) and briefly sonicated (Bandelin Sonoplus, tip VS70T, pulse ON/OFF 0.2/0.2, 25% amplitude, Bandelin electronic, Berlin, Germany).

After centrifugation, the supernatant was incubated with the preequilibrated gel (Glutathione Sepharose 4B for GST tagged proteins and HIS-Select Cobalt Affinity Gel for HIS tagged proteins) for 1h under constant agitation. The gel was centrifugated (500g for 1min) to remove the supernatant and washed three times with the lysis buffer. The washed gel was stained with SDS loading dye and loaded on SDS-PAGE.

Large-scale expression

The large-scale expression was performed using the conditions found to be suitable in the expression optimization screen to produce high amounts of recombinant protein. Typically *E.coli* BL21 (DE3) Star cells were grown in 6L of TB medium containing the corresponding antibiotic at 37°C while shaking. The large-scale expression was performed using the conditions found to be suitable in the expression optimization screen to produce high amounts of recombinant protein. Typically *E.coli* BL21 (DE3) Star cells were grown in 6L of TB medium containing the corresponding antibiotic at 37°C while shaking. For SeMet substituted protein expression, *E.coli* DL41 (DE3) cells were grown in 3L 1X M9 mineral medium (37.4 mM NH₄Cl, 44.1 mM KH₂PO₄, and 95.8 mM anhydrous Na₂HPO₄) supplemented with 0.4% w/v glucose, 40 µg/ml amino acids (with seleno-L- methionine instead of methionine), and 1 µg/ml vitamins (riboflavin, niacinamide, pyridoxine monohydrochloride, and thiamine) and antibiotics.

The pre-cooled cells (20°C) were induced with 1mM IPTG when the OD600 reached 0.6-0.7. The protein expression was carried out at 20°C for 16-18h. The cells were harvested by centrifugation (4000rpm 20min) and the supernatant discarded. The pellets were frozen in liquid Nitrogen and stored at -80°C until further use.

5.2.3 Protein methods

All protein purification procedures were conducted at 4°C and performed on an ÄKTA purification systems (GE Healthcare). All buffers were prepared with deionized H₂O and filtered (0.22 µm). Purification buffers used in this work are listed in table 5.15.

Determination of protein concentration

Protein concentration was determined by the Bradford assay (Bradford, 1976). A suitable volume of recombinant protein (1- 20µl) was diluted with Milli-Q H₂O to a final volume of 800µl and mixed with 200µl of Bradford solution. After 5 min incubation at RT, the absorbance of the recombinant protein was measured at the wavelength 595nm on a spectrophotometer. The concentration of each sample was determined using the calibration curve according to its absorbance by interpolation (Bradford, 1976).

Denaturing SDS-Gelelectrophoresis

According to the protein size, proteins samples were analyzed on 10% or 12% SDS-polyacrylamide gels. The samples were mixed with protein loading buffer and heated at 95°C for 5min. The samples were loaded into the gel with an unstained protein marker and run in 1x running buffer at 200V until the bromophenol blue border nearly run out the gel. The SDS-PAGE was rinsed with water and incubated in Coomassie staining solution for 10-15min and subsequently destained.

Recombinant protein purification

The purification protocol consists of three steps: first affinity enrichment AC (by Cobalt Affinity Gel or Glutathione Sepharose media), followed by IEX and SEC. An extended wash with lysis buffer (1-1.2M NaCl) after affinity gel binding was introduced to reduce the nucleic acids that are bound to the protein and could bias further experiments. All steps were monitored and analyzed by SDS-PAGE.

The cell pellets were re-suspended in lysis buffer supplemented with EDTA-free protease inhibitor and lysed by two passes through a microfluidizer (18,000 psi). The lysate was centrifugated to remove cell debris (18000rpm for 1h). The supernatant was carefully collected and filtered using a 5 µm filter. The filtered lysate was added to the pre-equilibrated affinity chromatography media (Glutathione Sepharose 4B for GST tagged proteins and HIS-Select Cobalt Affinity Gel for HIS tagged proteins) and incubated for 1h by constant agitation.

The slurry was packed into a column and connected to the ÄKTA purification systems. The column was washed with 10 column volume (CV) of lysis buffer or until the UV absorbance at 260 nm significantly sank and equilibrated with 5 to 10 CV AC buffer. Depend on the tag the recombinant protein was eluted either with a linear gradient to 300 mM imidazole for HIS-tagged constructs (5.15 HIS-AC elution buffer) or with a linear gradient to 40 mM reduced glutathione (5.15 GST-AC elution buffer). The recombinant protein was dialyzed against dialysis buffer (5.15), and Tobacco Etch Virus (TEV) protease was added to cleave the tag (HIS or GST).

The cleaved protein was loaded on a HiTrap CantoS and washed with 10 CV IEX buffer. The recombinant protein was then eluted from the cation-exchanger using a gradient to 100% NaCl (IEX elution buffer). As a final purification step, the protein was subjected to SEC on a pre-equilibrated HiLoad 16/600 Superdex 200pg column (GE Healthcare) with GF buffer.

Table 5.15. Purification buffer hStau1 and truncated constructs

Construct	Buffer	composition
hStau1 FL		
AC		
HIS	Lysis buffer	20mM Tris-HCl pH 7.5 at 4 °C, 1M NaCl, 1mM β -Mercaptoethanol
	AC buffer	20mM Tris-HCl pH 7.5 at 4 °C, 300mM NaCl, 1mM β -Mercaptoethanol
	AC elution buffer	Affinity buffer supplemented with 300mM Imidazol
GST	Lysis buffer	20mM Tris-HCl pH 7.5 at 4 °C, 1M NaCl, 1mM DTT
	AC buffer	20mM Tris-HCl pH 7.5 at 4 °C, 300mM NaCl, 1mM DTT
	AC elution buffer	Affinity buffer supplemented with 40mM reduced glutathione
IEX		
	Dialysis and IEX buffer	20mM Tris-HCl pH 7.5 at 4 °C, 300mM NaCl, 1mM DTT
	IEX elution buffer	20mM Tris-HCl pH 7.5 at 4 °C, 1.2M NaCl, 1mM DTT
SEC or GF		
	GF buffer	20mM HEPES pH 7.5 at 4 °C, 150mM NaCl or KCl, 1mM DTT
hStau1 truncations		
AC		
HIS	Lysis buffer	20mM Tris-HCl pH 7.5 at 4 °C, 1M NaCl, 1mM β -Mercaptoethanol, 5mM Imidazol
	AC buffer	20mM Tris-HCl pH 7.5 at 4 °C, 100mM NaCl, 1mM β -Mercaptoethanol, 5mM Imidazol
	AC elution buffer	Affinity buffer supplemented with 300mM Imidazol
GST	Lysis buffer	20mM Tris-HCl pH 7.5 at 4 °C, 1M NaCl, 1mM DTT
	AC buffer	20mM Tris-HCl pH 7.5 at 4 °C, 100mM NaCl, 1mM DTT
	AC elution buffer	Affinity buffer supplemented with 40mM reduced glutathione
IEX		
	Dialysis and IEX buffer	20mM Tris-HCl pH 7.5 at 4 °C, 100mM NaCl, 1mM DTT
	IEX elution buffer	20mM Tris-HCl pH 7.5 at 4 °C, 1.2M NaCl, 1mM DTT
SEC or GF		
	GF buffer	20mM HEPES pH 7.5 at 4 °C, 150mM KCl, 1mM DTT

Affinity chromatography (AC), Ion exchange chromatography (IEX), size exclusion chromatography (SEC) or gel filtration (GF).

Circular dichroism spectroscopy (CD)

CD measurements of hStau1 and mutants were recorded on a Jasco J-810 spectropolarimeter in the spectra range 190–240 nm, at 20 °C. All measurements were performed with a protein concentration of 0.2 mg/ml in the CD buffer (Table. 5.16). The average mean spectrum was calculated from data of five scans.

Table 5.16. Protein methods buffers

Method	Buffer	composition
CD	CD buffer	20 mM Tris–HCl, pH 7.5, at 4 °C, 150 mM KCl
SEC-MALLS	SM buffer	20 mM Tris–HCl, pH 7.5 AT RT, 150 mM KCl
Complex reconstitution	CR buffer	20 mM Tris, pH 7.5, at 4 °C, 100 mM KCl, 1 mM DTT
FA	FA buffer	20 mM HEPES, pH 7.5 at RT, 150 mM KCl
EMSA	TG buffer	25 mM Tris, 192 mM Glycine pH 9.78 at RT
	NG loading buffer	2.5X TG buffer, 50% Glycerol

Circular dichroism (CD), Fluorescence Anisotropy (FA), native gel (NG).

Complex reconstitution

For complex reconstitution, the recombinantly purified protein was resuspended in high salt buffer (>300mM to 3M KCl) to prevent precipitation. The *ARF1* SBS was purchased from IDT. The sense and antisense strands were resuspended in deionized and sterilized H₂O and mixed equimolarly. The *ARF1* SBS were heated at 95 °C for 5 min and cooled to RT. The annealed *ARF1* SBS was resuspended in CR buffer (Table. 5.16) and mixed in a ratio 1:1.2M (RNA excess) with the purified protein. The complex was incubated at 4 °C for 1h, centrifuged and injected onto a pre-equilibrated HiLoad 16/600 Superdex 200pg column (GE Healthcare) with CR buffer.

SEC-MALLS

SEC-MALLS was performed using a Superdex 200 Increase 10/300 GL column (GE Healthcare). light scattering (LS) intensity at three angles was measured on a miniDAWN TREOS detector (Wyatt Technologies, Santa Barbara, USA) operating at 659nm and refractive index (RI) was monitored with an Optilab T-rEX differential refractometer (Wyatt Technologies). The equilibration of the system and the runs were performed with the SM buffer (Table. 5.16). For each construct, the sample volume injected into the systems was: 200 µg for hStau1 FL, 430 µg for hStau1^{dsRBD3-4}, 400 µg for hStau1^{dsRBD2-4}, and 350 µg hStau1^{dsRBD3-5}. For the RNA–protein complexes, 140 µg of each construct was mixed with a 1.2 M excess of *ARF1* SBS. The collected data were recorded and analyzed using the protein conjugate procedure implemented in ASTRA software (Wyatt Technologies).

Fluorescence Anisotropy

The binding affinity of hStau1 FL and mutants were determined by fluorescence anisotropy using 5'-6-fluorescein amidite (FAM)-labeled RNA at RT (19–21 °C). The 5'-6-FAM-labeled sense strand and unlabeled antisense strand were purchased from IDT (The sequence are listed in Table. 5.12) and annealed as described before (Complex reconstitution). Fluorescence polarization and intensity were measured in an Infinite F200 plate reader (Tecan). The excitation and emission wavelength were set to 485 nm and 535 respectively.

For the anisotropy measurements, 10nM of label RNA were incubated with increasing concentrations of purified protein in FA buffer (Table. 5.16). The binding reactions were incubated for 10min at RT in a total volume of 50 μ l. Each titration point was consecutively scanned three times, with an integration time of 40 μ s. The anisotropy data was analyzed using Prism 6 software (GraphPad). The K_D was calculated with the anisotropy average by nonlinear regression fitting to the following equation:

$$Y = Y_f + (Y_b - Y_f) \cdot \frac{L + K_d + X - \sqrt{(L + K_d + X)^2 - 4 \cdot L \cdot X}}{2 \cdot L}$$

where L = concentration of labeled RNA, X = protein concentration, Y_f = anisotropy of free RNA = ($X = 0$), and Y_b = anisotropy of bound RNA (Fluorescence Polarization Technical Resource Guide, fourth edition, Invitrogen).

Electrophoretic Mobility Shift Assay

EMSA experiments were performed with the 5'-6-FAM-labeled dsRNA. Typically, 1 μ M of 5'-6-FAM-labeled dsRNA was mixed with increasing concentration of purified protein (hStau1 FL and mutants) in a 10 μ l reaction, with FA buffer (Table. 5.16). Reactions were incubated on ice for 45 min. Afterward, the reactions were supplemented with 2 μ l of NG buffer (Table. 5.16) and loaded on a pre-equilibrated 6% polyacrylamide gel (60min in 1X TG buffer at 100V at 4 °C).

Table 5.17. 6% native acrylamide gels

Component	Volume
TG buffer 10X	7.5 ml
30% Acrylamide	15 ml
TEMED	75 μ l
APS	750 μ l
Glycerol (5%)	4.31 ml
ddH ₂ O	to 75 μ l

Electrophoresis was performed at 4°C in 1X TG buffer and ran at 220V for 3-4h. The gel was analyzed and imaged at 488nm using the Amersham Typhoon gel imager (GE Healthcare).

5.2.4 Crystallization and data collection

Initial crystallization attempts of hStau1 and mutants were performed at 4°C and 20 °C in 96-well MRC sitting drop plates with different crystallisations screens (crystallisations screens are listed in Table. 5.7). The crystallization setups were based on 200nl droplets (100nl protein or complex plus 100nl reservoir) over 70nl reservoir solution. The drops were dispensed using a mosquito[®] crystal dispensing robot (TTP Labtech Ltd, UK) with 8 channels.

Initial hits were usually optimized in 96 well plates using the dragonfly[®] crystal screen optimizer (TTP Labtech Ltd, UK). The plates were automatically imaged using the imaging system of the crystallization facility in house.

hStau1^{dsRBD3-4} - ARF1 SBS complex

Crystals of hStau1^{dsRBD3-4} in complex with ARF1 SBS were obtained at 22°C by mixing 1 µl of complex solution at 7 mg/ml with 1 µl of reservoir solution (50 mM MgCl₂, 120 mM KCl, 50 mM sodium cacodylate, pH 6.33, 5% 1,6-hexanediol) using the sitting drop vapor diffusion. The crystals were cryo-protected with mother liquor supplemented with 30% 1,6-Hexanediol and flash frozen in liquid nitrogen.

ARF1 SBS

Crystals of ARF1 SBS were obtained at 22°C using the sitting drop vapor diffusion (1.7 M (NH₄)₂SO₄ and 100 mM Tri-sodium citrate, pH 6.2). The crystal was cryo-protected by transfer into mother liquor supplemented with 20% glycerol and then flash frozen in liquid nitrogen.

The diffraction data were collected at the PXIII (hStau1^{dsRBD3-4} - ARF1 SBS complex native and derivative) and PXII (ARF1 SBS unbound) beamlines of the Swiss Light Source.

Data processing and structure determination

The diffraction data were indexed, integrated, and scaled with the XDS program (Kabsch, 2010). For the hStau1^{dsRBD3-4} - ARF1 SBS complex, the determined space group P41212 correspond to the tetragonal crystal system with cell dimension $a = b = 105.9 \text{ \AA}$, $c = 169.2 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$. In the case of the ARF1 SBS the space group was H32 with cell dimensions of $a = b = 43.8 \text{ \AA}$, $c = 452.1 \text{ \AA}$, $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$. The structure of the hStau1¹⁸²⁻³⁶⁰ (dsRBD3-4 construct) - ARF1 SBS complex and the apo ARF1 SBS was solved by MR using PHASER (McCoy et al., 2007). The phases of the complex were obtained with an optimized dsRBD search model by the

superposition of *DmStau* dsRBD5 PDB ID: 5CFF Jia et al. (2015) and human hStau1 dsRBD5 PDB ID: 4DKK. For the *ARF1* SBS the phases were determined with a 19 bp A-form RNA as a search model (PDB: 1QC0 Klosterman P.S. (1999)). The asymmetric unit was assumed to contain one complex in the ASU for the complex and one and a half molecules for the *ARF1* SBS. Both models were built manually using COOT and subjected to iterative cycles of restrained refinement with PHENIX. Interactions within the structure were analyzed with the program PISA.

In attempting to verify the model and find anomalous scatterers, a single-wavelength anomalous diffraction (SAD) data were collected from the hStau1^{dsRBD3-4} - *ARF1* SBS complex with SeMet-substituted hStau1^{dsRBD3-4}. The calculated anomalous-difference Fourier map (using log-likelihood gradient maps in PHASER) allowed unambiguous determination of the positions of the antiparallels dsRBD3^A and dsRBD3^B as well as dsRBD4.

Bibliography

- Adam, S. A., Nakagawa, T., Swanson, M. S., Woodruff, T. K. and Dreyfuss, G. (1986). mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleo-protein consensus sequence., *Molecular and Cellular Biology* **6**(8): 2932–2943.
- Anantharaman, V. (2002). Comparative genomics and evolution of proteins involved in RNA metabolism, *Nucleic Acids Research* **30**(7): 1427–1464.
- Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2005). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics* **22**(2): 195–201.
- Auweter, S. D., Fasan, R., Reymond, L., Underwood, J. G., Black, D. L., Pitsch, S. and Allain, F. H.-T. (2005). Molecular basis of RNA recognition by the human alternative splicing factor fox-1, *The EMBO Journal* **25**(1): 163–173.
- Auweter, S. D., Oberstrass, F. C. and Allain, F. H.-T. (2006). Sequence-specific binding of single-stranded RNA: is there a code for recognition?, *Nucleic Acids Research* **34**(17): 4943–4959.
- Bandziulis, R. J., Swanson, M. S. and Dreyfuss, G. (1989). RNA-binding proteins as developmental regulators., *Genes & Development* **3**(4): 431–437.
- Banerjee, A., Benjamin, R., Balakrishnan, K., Ghosh, P. and Banerjee, S. (2014). Human protein staufin-2 promotes HIV-1 proliferation by positively regulating RNA export activity of viral protein rev, *Retrovirology* **11**(1): 18.
- Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space, *Nature Reviews Genetics* **15**(3): 163–175.
- Beuth, B., Pennell, S., Arnvig, K. B., Martin, S. R. and Taylor, I. A. (2005). Structure of a mycobacterium tuberculosis NusA–RNA complex, *The EMBO Journal* **24**(20): 3576–3587.
- Bevilacqua, P. C. and Cech, T. R. (1996). Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR†, *Biochemistry* **35**(31): 9983–9994.
- Birney, E., Kumar, S. and Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors, *Nucleic Acids Research* **21**(25): 5803–5816.

- Blaszczyk, J., Gan, J., Tropea, J. E., Court, D. L., Waugh, D. S. and Ji, X. (2004). Noncatalytic assembly of ribonuclease III with double-stranded RNA, *Structure* **12**(3): 457–466.
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry* **72**(1-2): 248–254.
- Braunschweig, U., Gueroussov, S., Plocik, A. M., Graveley, B. R. and Blencowe, B. J. (2013). Dynamic integration of splicing within gene regulatory pathways, *Cell* **152**(6): 1252–1269.
- Brendel, C., Rehbein, M., Kreienkamp, H.-J., Buck, F., Richter, D. and Kindler, S. (2004). Characterization of staufen 1 ribonucleoprotein complexes, *Biochemical Journal* **384**(2): 239–246.
URL: <http://www.biochemj.org/content/384/2/239>
- Broadus, J. and Doe, C. Q. (1997). Extrinsic cues, intrinsic cues and microfilaments regulate asymmetric protein localization in drosophila neuroblasts, *Current Biology* **7**(11): 827–835.
- Broadus, J., Fuerstenberg, S. and Doe, C. Q. (1998). Staufen-dependent localization of prospero mRNA contributes to neuroblast daughter-cell fate, *Nature* **391**(6669): 792–795.
- Buchner, G., Bassi, M. T., Andolfi, G., Ballabio, A. and Franco, B. (1999). Identification of a novel homolog of the drosophila staufen protein in the chromosome 8q13–q21.1 region, *Genomics* **62**(1): 113–118.
- Burd, C. and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins, *Science* **265**(5172): 615–621.
- Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. and Neidle, S. (2006). Quadruplex dna: sequence, topology and structure, *Nucleic Acids Research* **34**(19): 5402–5415.
URL: <http://dx.doi.org/10.1093/nar/gkl655>
- Burgess, R. R. (2018). A brief practical review of size exclusion chromatography: Rules of thumb, limitations, and troubleshooting, *Protein Expression and Purification* **150**: 81–85.
- Bycroft, M., Grünert, S., Murzin, A., Proctor, M. and Johnston, D. S. (1995). NMR solution structure of a dsRNA binding domain from drosophila staufen protein reveals homology to the n-terminal domain of ribosomal protein s5., *The EMBO Journal* **14**(14): 3563–3571.
- Carmody, S. R. and Wente, S. R. (2009). mRNA nuclear export at a glance, *Journal of Cell Science* **122**(12): 1933–1937.
- Castrignanò, T., Chillemi, G., Varani, G. and Desideri, A. (2002). Molecular dynamics simulation of the RNA complex of a double-stranded RNA-binding domain reveals dynamic features of the intermolecular interface and its hydration, *Biophysical Journal* **83**(6): 3542–3552.

- Chang, K.-Y. and Ramos, A. (2005). The double-stranded RNA-binding motif, a versatile macromolecular docking platform, *FEBS Journal* **272**(9): 2109–2117.
- Chatel-Chaix, L., Abrahamyan, L., Frechina, C., Mouland, A. J. and DesGroseillers, L. (2007). The host protein stau1 participates in human immunodeficiency virus type 1 assembly in live cells by influencing pr55gag multimerization, *Journal of Virology* **81**(12): 6216–6230.
- Chatel-Chaix, L., Clement, J.-F., Martel, C., Beriault, V., Gatignol, A., DesGroseillers, L. and Mouland, A. J. (2004). Identification of stau1 in the human immunodeficiency virus type 1 gag ribonucleoprotein complex and a role in generating infectious viral particles, *Molecular and Cellular Biology* **24**(7): 2637–2648.
- Cho, H., Kim, K. M., Han, S., Choe, J., Park, S. G., Choi, S. S. and Kim, Y. K. (2012). Stau1-mediated mRNA decay functions in adipogenesis, *Molecular Cell* **46**(4): 495–506.
- Clemens, K., Wolf, V., McBryant, S., Zhang, P., Liao, X., Wright, P. and Gottesfeld, J. (1993). Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein, *Science* **260**(5107): 530–533.
- Cléry, A., Blatter, M. and Allain, F. H.-T. (2008). RNA recognition motifs: boring? Not quite, *Current Opinion in Structural Biology* **18**(3): 290–298.
- Colgan, D. F. and Manley, J. L. (1997). Mechanism and regulation of mRNA polyadenylation, *Genes & Development* **11**(21): 2755–2766.
- Daher, A., Longuet, M., Dorin, D., Bois, F., Segéral, E., Bannwarth, S., Battisti, P.-L., Purcell, D. F., Benarous, R., Vaquero, C., Meurs, E. F. and Gatignol, A. (2001). Two dimerization domains in the trans-activation response rna-binding protein (trbp) individually reverse the protein kinase r inhibition of hiv-1 long terminal repeat expression, *Journal of Biological Chemistry* **276**(36): 33899–33905.
URL: <http://www.jbc.org/content/276/36/33899.abstract>
- de Lucas, S., Oliveros, J. C., Chagoyen, M. and Ortín, J. (2014). Functional signature for the recognition of specific target mRNAs by human stau1 protein, *Nucleic Acids Research* **42**(7): 4516–4526.
- de Lucas, S., Peredo, J., Marion, R. M., Sanchez, C. and Ortín, J. (2010). Human stau1 protein interacts with influenza virus ribonucleoproteins and is required for efficient virus multiplication, *Journal of Virology* **84**(15): 7603–7612.
- Delarue, M. and Moras, D. (1989). RNA structure, *Nucleic Acids and Molecular Biology*, Springer Berlin Heidelberg, pp. 182–196.

- DesGroseillers, L. and Lemieux, N. (1996). Localization of a human double-stranded RNA-binding protein gene (STAU) to band 20q13.1 by fluorescence in situ hybridization, *Genomics* **36**(3): 527–529.
- Dixit, U., Pandey, A. K., Mishra, P., Sengupta, A. and Pandey, V. N. (2016). Staufen1 promotes HCV replication by inhibiting protein kinase R and transporting viral RNA to the site of translation and replication in the cells, *Nucleic Acids Research* **44**(11): 5271–5287.
- Dreyfuss, G., Kim, V. N. and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry, *Nature Reviews Molecular Cell Biology* **3**(3): 195–205.
- Driever, W. and Nüsslein-Volhard, C. (1988a). The bicoid protein determines position in the drosophila embryo in a concentration-dependent manner, *Cell* **54**(1): 95–104.
- Driever, W. and Nüsslein-Volhard, C. (1988b). A gradient of bicoid protein in drosophila embryos, *Cell* **54**(1): 83–93.
- Dubnau, J., Chiang, A.-S., Grady, L., Barditch, J., Gossweiler, S., McNeil, J., Smith, P., Buldoc, F., Scott, R., Certa, U., Broger, C. and Tully, T. (2003). The staufen/pumilio pathway is involved in drosophila long-term memory, *Current Biology* **13**(4): 286–296.
- Duchaîne, T. F., Hemraj, I., Furic, L., Deitinghoff, A., Kiebler, M. A. and DesGroseillers, L. (2002). Staufen2 isoforms localize to the somatodendritic domain of neurons and interact with different organelles., *Journal of cell science* **115**: 3285–3295.
- Dugre-Brisson, S., Elvira, G., Boulay, K., Chatel-Chaix, L., Mouland, A. J. and DesGroseillers, L. (2005). Interaction of staufen1 with the 5' end of mRNA facilitates translation of these RNAs, *Nucleic Acids Research* **33**(15): 4797–4812.
- Ephrussi, A., Dickinson, L. K. and Lehmann, R. (1991). oskar organizes the germ plasm and directs localization of the posterior determinant nanos, *Cell* **66**(1): 37–50.
- Fasken, M. B. and Corbett, A. H. (2005). Process or perish: quality control in mRNA biogenesis, *Nature Structural & Molecular Biology* **12**(6): 482–488.
- Fasken, M. B. and Corbett, A. H. (2009). Mechanisms of nuclear mRNA quality control, *RNA Biology* **6**(3): 237–241.
- Ferrandon, D., Elphick, L., Nüsslein-Volhard, C. and Johnston, D. S. (1994). Staufen protein associates with the 3'UTR of bicoid mRNA to form particles that move in a microtubule-dependent manner, *Cell* **79**(7): 1221–1232.
- Fritzsche, R., Karra, D., Bennett, K. L., yee Ang, F., Heraud-Farlow, J. E., Tolino, M., Doyle, M., Bauer, K. E., Thomas, S., Planyavsky, M., Arn, E., Bakosova, A., Jungwirth, K., Hörmann, A.,

- Palfi, Z., Sandholzer, J., Schwarz, M., Macchi, P., Colinge, J., Superti-Furga, G. and Kiebler, M. A. (2013). Interactome of two diverse RNA granules links mRNA localization to translational repression in neurons, *Cell Reports* **5**(6): 1749–1762.
- Fuerstenberg, S., Peng, C.-Y., Alvarez-Ortiz, P., Hor, T. and Doe, C. Q. (1998). Identification of miranda protein domains regulating asymmetric cortical localization, cargo binding, and cortical release, *Molecular and Cellular Neuroscience* **12**(6): 325–339.
- Furic, L., Maher-Laporte, M. and DesGroseillers, L. (2007). A genome-wide approach identifies distinct but overlapping subsets of cellular mRNAs associated with stau1- and stau2-containing ribonucleoprotein complexes, *RNA* **14**(2): 324–335.
- Gan, J., Tropea, J. E., Austin, B. P., Court, D. L., Waugh, D. S. and Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III, *Cell* **124**(2): 355–366.
- García-Mayoral, M. F., Hollingworth, D., Masino, L., Díaz-Moreno, I., Kelly, G., Gherzi, R., Chou, C.-F., Chen, C.-Y. and Ramos, A. (2007). The structure of the c-terminal KH domains of KSRP reveals a noncanonical motif important for mRNA degradation, *Structure* **15**(4): 485–498.
- Gleghorn, M. L., Gong, C., Kielkopf, C. L. and Maquat, L. E. (2013). Stau1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay, *Nature Structural & Molecular Biology* **20**(4): 515–524.
- Gong, C., Kim, Y. K., Woeller, C. F., Tang, Y. and Maquat, L. E. (2008). SMD and NMD are competitive pathways that contribute to myogenesis: effects on PAX3 and myogenin mRNAs, *Genes & Development* **23**(1): 54–66.
- Gong, C. and Maquat, L. E. (2011). lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via alu elements, *Nature* **470**(7333): 284–288.
- Green, S. R. and Mathews, M. B. (1992). Two RNA-binding motifs in the double-stranded RNA-activated protein kinase, DAI., *Genes & Development* **6**(12b): 2478–2490.
- Grishin, N. V. (2001). KH domain: one motif, two folds, *Nucleic Acids Research* **29**(3): 638–643.
- Heber, S., Gáspár, I., Tants, J.-N., Günther, J., Moya, S. M. F., Janowski, R., Ephrussi, A., Sattler, M. and Niessing, D. (2019). Stau2-mediated RNA recognition and localization requires combinatorial action of multiple domains, *Nature Communications* **10**(1).
- Heraud-Farlow, J. E., Sharangdhar, T., Li, X., Pfeifer, P., Tauber, S., Orozco, D., Hörmann, A., Thomas, S., Bakosova, A., Farlow, A. R., Edbauer, D., Lipshitz, H. D., Morris, Q. D., Bilban, M., Doyle, M. and Kiebler, M. A. (2013). Stau2 regulates neuronal target RNAs, *Cell Reports* **5**(6): 1511–1518.

- Hitti, E. G., Sallacz, N. B., Schoft, V. K. and Jantsch, M. F. (2004). Oligomerization activity of a double-stranded rna-binding domain, *FEBS Letters* **574**(1): 25 – 30.
URL: <http://www.sciencedirect.com/science/article/pii/S0014579304009731>
- Hudson, B. P., Martinez-Yamout, M. A., Dyson, H. J. and Wright, P. E. (2004). Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d, *Nature Structural & Molecular Biology* **11**(3): 257–264.
- Jayachandran, U., Grey, H. and Cook, A. G. (2015). Nuclear factor 90 uses an ADAR2-like binding mode to recognize specific bases in dsRNA, *Nucleic Acids Research* **44**(4): 1924–1936.
- Jia, M., Shan, Z., Yang, Y., Liu, C., Li, J., Luo, Z.-G., Zhang, M., Cai, Y., Wen, W. and Wang, W. (2015). The structural basis of miranda-mediated staufen localization during drosophila neuroblast asymmetric division, *Nature Communications* **6**(1).
- Johnston, D. S., Beuchle, D. and Nüsslein-Volhard, C. (1991). staufen, a gene required to localize maternal RNAs in the drosophila egg, *Cell* **66**(1): 51–63.
- Johnston, D. S., Brown, N. H., Gall, J. G. and Jantsch, M. (1992). A conserved double-stranded RNA-binding domain., *Proceedings of the National Academy of Sciences* **89**(22): 10979–10983.
- Kabsch, W. (2010). XDS, *Acta Crystallographica Section D Biological Crystallography* **66**(2): 125–132.
- Kanai, Y., Dohmae, N. and Hirokawa, N. (2004). Kinesin transports RNA, *Neuron* **43**(4): 513–525.
- Kenan, D. J., Query, C. C. and Keene, J. D. (1991). RNA recognition: towards identifying determinants of specificity, *Trends in Biochemical Sciences* **16**: 214–220.
- Kharrat, A., Macias, M., Gibson, T., Nilges, M. and Pastore, A. (1995). Structure of the dsRNA binding domain of e. coli RNase III., *The EMBO Journal* **14**(14): 3572–3584.
- Kiebler, M. A. and Bassell, G. J. (2006). Neuronal RNA granules: Movers and makers, *Neuron* **51**(6): 685–690.
- Kiebler, M. A., Hemraj, I., Verkade, P., Köhrmann, M., Fortes, P., Marión, R. M., Ortín, J. and Dotti, C. G. (1999). The mammalian staufen protein localizes to the somatodendritic domain of cultured hippocampal neurons: Implications for its involvement in mRNA transport, *The Journal of Neuroscience* **19**(1): 288–297.
- Kim-Ha, J., Kerr, K. and Macdonald, P. M. (1995). Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential, *Cell* **81**(3): 403–412.

- Kim, U., Wang, Y., Sanford, T., Zeng, Y. and Nishikura, K. (1994). Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing., *Proceedings of the National Academy of Sciences* **91**(24): 11457–11461.
- Kim, Y. K., Furic, L., DesGroseillers, L. and Maquat, L. E. (2005). Mammalian stau1 recruits upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay, *Cell* **120**(2): 195–208.
- Kim, Y. K., Furic, L., Parisien, M., Major, F., DesGroseillers, L. and Maquat, L. E. (2007). Stau1 regulates diverse classes of mammalian transcripts, *The EMBO Journal* **26**(11): 2670–2681.
- Klosterman P.S., Shah S.A., S. T. (1999). Crystal structures of two plasmid copy control related rna duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution †, *Biochemistry* **38**: 14784–92.
- Knowles, R. B., Sabry, J. H., Martone, M. E., Deerinck, T. J., Ellisman, M. H., Bassell, G. J. and Kosik, K. S. (1996). Translocation of RNA granules in living neurons, *The Journal of Neuroscience* **16**(24): 7812–7820.
- Kretz, M., Siprashvili, Z., Chu, C., Webster, D. E., Zehnder, A., Qu, K., Lee, C. S., Flockhart, R. J., Groff, A. F., Chow, J., Johnston, D., Kim, G. E., Spitale, R. C., Flynn, R. A., Zheng, G. X. Y., Aiyer, S., Raj, A., Rinn, J. L., Chang, H. Y. and Khavari, P. A. (2012). Control of somatic tissue differentiation by the long non-coding RNA TINCR, *Nature* **493**(7431): 231–235.
- Krovat, B. C. and Jantsch, M. F. (1996). Comparative mutational analysis of the double-stranded RNA binding domains of *Xenopus laevis* RNA-binding protein a, *Journal of Biological Chemistry* **271**(45): 28112–28119.
- Kusek, G., Campbell, M., Doyle, F., Tenenbaum, S., Kiebler, M. and Temple, S. (2012). Asymmetric segregation of the double-stranded rna binding protein stau2 during mammalian neural stem cell divisions promotes lineage progression, *Cell Stem Cell* **11**(4): 505 – 516.
URL: <http://www.sciencedirect.com/science/article/pii/S1934590912003682>
- Köhrmann, M., Luo, M., Kaether, C., DesGroseillers, L., Dotti, C. G. and Kiebler, M. A. (1999). Microtubule-dependent recruitment of stau1-green fluorescent protein into large RNA-containing granules and subsequent dendritic transport in living hippocampal neurons, *Molecular Biology of the Cell* **10**(9): 2945–2953.
- Lathe, G. H. and Ruthven, C. R. J. (1956). The separation of substances and estimation of their relative molecular sizes by the use of columns of starch in water, *Biochemical Journal* **62**(4): 665–674.
- Laver, J. D., Li, X., Ancevicus, K., Westwood, J. T., Smibert, C. A., Morris, Q. D. and Lipshitz, H. D. (2013). Genome-wide analysis of stau1-associated mRNAs identifies secondary structures that confer target specificity, *Nucleic Acids Research* **41**(20): 9438–9460.

- Lazzaretti, D., Bandholz-Cajamarca, L., Emmerich, C., Schaaf, K., Basquin, C., Irion, U. and Bono, F. (2018). The crystal structure of stau1 in complex with a physiological RNA sheds light on substrate selectivity, *Life Science Alliance* **1**(5): e201800187.
- Le, S., Sternglanz, R. and Greider, C. W. (2000). Identification of two RNA-binding proteins associated with human telomerase RNA, *Molecular Biology of the Cell* **11**(3): 999–1010.
- Lebeau, G., Miller, L. C., Tartas, M., McAdam, R., Laplante, I., Badeaux, F., DesGroseillers, L., Sossin, W. S. and Lacaille, J.-C. (2011). Stau2 regulates mGluR long-term depression and map1b mRNA distribution in hippocampal neurons, *Learning & Memory* **18**(5): 314–326.
- LeGendre, J. B., Campbell, Z. T., Kroll-Conner, P., Anderson, P., Kimble, J. and Wickens, M. (2012). RNA targets and specificity of stau1, a double-stranded RNA-binding protein in *Caenorhabditis elegans*, *Journal of Biological Chemistry* **288**(4): 2532–2545.
- Lehmann, R. and Nusslein-Volhard, C. (1991). The maternal gene nanos has a central role in posterior pattern formation of the drosophila embryo, *Development* **112**(3): 679–691.
URL: <http://dev.biologists.org/content/112/3/679>
- Lemaire, P. A., Lary, J. and Cole, J. L. (2005). Mechanism of pkr activation: Dimerization and kinase activation in the absence of double-stranded rna, *Journal of Molecular Biology* **345**(1): 81 – 90.
URL: <http://www.sciencedirect.com/science/article/pii/S0022283604013233>
- Lewis, H. A., Musunuru, K., Jensen, K. B., Edo, C., Chen, H., Darnell, R. B. and Burley, S. K. (2000). Sequence-specific RNA binding by a nova KH domain, *Cell* **100**(3): 323–332.
- Li, P., Yang, X., Wasser, M., Cai, Y. and Chia, W. (1997). Inscuteable and stau1 mediate asymmetric localization and segregation of prospero RNA during drosophila neuroblast cell divisions, *Cell* **90**(3): 437–447.
- Lu, D., Searles, M. A. and Klug, A. (2003). Crystal structure of a zinc-finger–RNA complex reveals two modes of molecular recognition, *Nature* **426**(6962): 96–100.
- Lunde, B. M., Moore, C. and Varani, G. (2007). RNA-binding proteins: modular design for efficient function, *Nature Reviews Molecular Cell Biology* **8**(6): 479–490.
- Luo, M., Duchaine, T. F. and DesGroseillers, L. (2002). Molecular mapping of the determinants involved in human stau1–ribosome association, *Biochemical Journal* **365**(3): 817–824.
- Macchi, P., Brownawell, A. M., Grunewald, B., DesGroseillers, L., Macara, I. G. and Kiebler, M. A. (2004). The brain-specific double-stranded RNA-binding protein stau2, *Journal of Biological Chemistry* **279**(30): 31440–31444.

- Mallardo, M., Deitinghoff, A., Muller, J., Goetze, B., Macchi, P., Peters, C. and Kiebler, M. A. (2003). Isolation and characterization of staufen-containing ribonucleoprotein particles from rat brain, *Proceedings of the National Academy of Sciences* **100**(4): 2100–2105.
- Maniatis, T. and Reed, R. (2002). An extensive network of coupling among gene expression machines, *Nature* **416**(6880): 499–506.
- March, P. E., Ahnn, J. and Inouye, M. (1985). The DNA sequence of the gene (*rnc*) encoding ribonuclease III of *Escherichia coli*, *Nucleic Acids Research* **13**(13): 4677–4685.
- Marianayagam, N. J., Sunde, M. and Matthews, J. M. (2004). The power of two: protein dimerization in biology, *Trends in Biochemical Sciences* **29**(11): 618–625.
- Marión, R. M., Fortes, P., Beloso, A., Dotti, C. and Ortín, J. (1999). A human sequence homologue of staufen is an RNA-binding protein that is associated with polysomes and localizes to the rough endoplasmic reticulum, *Molecular and Cellular Biology* **19**(3): 2212–2219.
- Maris, C., Dominguez, C. and Allain, F. H.-T. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression, *FEBS Journal* **272**(9): 2118–2131.
- Martel, C., Dugre-Brisson, S., Boulay, K., Breton, B., Lapointe, G., Armando, S., Trepanier, V., Duchaine, T., Bouvier, M. and Desgroseillers, L. (2010). Multimerization of staufen1 in live cells, *RNA* **16**(3): 585–597.
- Martin, K. C. and Ephrussi, A. (2009). mRNA localization: Gene expression in the spatial dimension, *Cell* **136**(4): 719–730.
- Masliah, G., Barraud, P. and Allain, F. H. T. (2012). RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence, *Cellular and Molecular Life Sciences* .
- Masliah, G., Maris, C., König, S. L., Yulikov, M., Aeschmann, F., Malinowska, A. L., Mabile, J., Weiler, J., Holla, A., Hunziker, J., Meisner-Kober, N., Schuler, B., Jeschke, G. and Allain, F. H.-T. (2018). Structural basis of siRNA recognition by trbp double-stranded rna binding domains, *The EMBO Journal* **37**(6).
URL: <http://emboj.embopress.org/content/37/6/e97089>
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007). Phaser crystallographic software, *Journal of Applied Crystallography* **40**(4): 658–674.
- McMillan, N. A. J., Carpick, B. W., Hollis, B., Toone, W. M., Zamanian-Daryoush, M. and Williams, B. R. G. (1995). Mutational analysis of the double-stranded RNA (dsRNA) binding domain of the dsRNA-activated protein kinase, PKR, *Journal of Biological Chemistry* **270**(6): 2601–2606.

- Meurs, E., Chong, K., Galabru, J., Thomas, N. B., Kerr, I. M., Williams, B. R. and Hovanessian, A. G. (1990). Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon, *Cell* **62**(2): 379–390.
- Micklem, D. R. (2000). Distinct roles of two conserved staufer domains in oskar mRNA localization and translation, *The EMBO Journal* **19**(6): 1366–1377.
- Miki, T., Kamikawa, Y., Kurono, S., Kaneko, Y., Katahira, J. and Yoneda, Y. (2011). Cell type-dependent gene regulation by staufer2 in conjunction with upf1, *BMC Molecular Biology* **12**(1): 48.
- Miki, T. and Yoneda, Y. (2004). Alternative splicing of staufer2 creates the nuclear export signal for CRM1 (exportin 1), *Journal of Biological Chemistry* **279**(46): 47473–47479.
- Miller, J., McLachlan, A. and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from xenopus oocytes., *The EMBO Journal* **4**(6): 1609–1614.
- Monshausen, M., Putz, U., Rehbein, M., Schweizer, M., DesGroseillers, L., Kuhl, D., Richter, D. and Kindler, S. (2008). Two rat brain staufer isoforms differentially bind RNA, *Journal of Neurochemistry* **76**(1): 155–165.
- Moore, M. J. (2005). From birth to death: The complex lives of eukaryotic mRNAs, *Science* **309**(5740): 1514–1518.
- Muto, Y. and Yokoyama, S. (2012). Structural insight into RNA recognition motifs: versatile molecular lego building blocks for biological systems, *Wiley Interdisciplinary Reviews: RNA* **3**(2): 229–246.
- Nanduri, S., Carpick, B. W., Yang, Y., Williams, B. R. and Qin, J. (1998). Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation, *The EMBO Journal* **17**(18): 5458–5465.
- O'Connell, M. A., Krause, S., Higuchi, M., Hsuan, J. J., Totty, N. F., Jenny, A. and Keller, W. (1995). Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine deaminase., *Molecular and Cellular Biology* **15**(3): 1389–1397.
- Paridaen, J. T. and Huttner, W. B. (2014). Neurogenesis during development of the vertebrate central nervous system, *EMBO reports* **15**(4): 351–364.
- Park, E., Gleghorn, M. L. and Maquat, L. E. (2012). Staufer2 functions in staufer1-mediated mRNA decay by binding to itself and its paralog and promoting UPF1 helicase but not ATPase activity, *Proceedings of the National Academy of Sciences* **110**(2): 405–412.
- Parks, T. E. C., Ravel-Chapuis, A., Bondy-Chorney, E., Renaud, J.-M., Côté, J. and Jasmin, B. J. (2017). Muscle-specific expression of the RNA-binding protein staufer1 induces progressive skeletal muscle atrophy via regulation of phosphatase tensin homolog, *Human Molecular Genetics* **26**(10): 1821–1838.

- Proudfoot, N. J. (2011). Ending the message: poly(a) signals then and now, *Genes & Development* **25**(17): 1770–1782.
- Proudfoot, N. J., Furger, A. and Dye, M. J. (2002). Integrating mRNA processing with transcription, *Cell* **108**(4): 501–512.
- Ramos, A., Grünert, S., Adams, J., Micklem, D. R., Proctor, M. R., Freund, S., Bycroft, M., Johnston, D. S. and Varani, G. (2000). RNA recognition by a staufen double-stranded RNA-binding domain, *The EMBO Journal* **19**(5): 997–1009.
- Ravel-Chapuis, A., Bélanger, G., Yadava, R. S., Mahadevan, M. S., DesGroseillers, L., Côté, J. and Jasmin, B. J. (2012). The RNA-binding protein staufen1 is increased in DM1 skeletal muscle and promotes alternative pre-mRNA splicing, *The Journal of Cell Biology* **196**(6): 699–712.
- Ricci, E. P., Kucukural, A., Cenik, C., Mercier, B. C., Singh, G., Heyer, E. E., Ashar-Patel, A., Peng, L. and Moore, M. J. (2013). Staufen1 senses overall transcript secondary structure to regulate translation, *Nature Structural & Molecular Biology* **21**(1): 26–35.
- Rodríguez-Navarro, S. and Hurt, E. (2011). Linking gene regulation to mRNA production and export, *Current Opinion in Cell Biology* **23**(3): 302–309.
- Ryter, J. M. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA, *The EMBO Journal* **17**(24): 7505–7513.
- Schrödinger, LLC (2015). The PyMOL molecular graphics system, version 2.0.
- Schuldt, A. J., Adams, J. H., Davidson, C. M., Micklem, D. R., Haseloff, J., Johnston, D. S. and Brand, A. H. (1998). Miranda mediates asymmetric protein and RNA localization in the developing nervous system, *Genes & Development* **12**(12): 1847–1857.
- Schupbach, T. and Wieschaus, E. (1986). Germline autonomy of maternal-effect mutations altering the embryonic body pattern of drosophila, *Developmental Biology* **113**(2): 443–448.
- Seeman, N. C., Rosenberg, J. M., Suddath, F., Kim, J. J. P. and Rich, A. (1976). RNA double-helical fragments at atomic resolution: I. the crystal and molecular structure of sodium adenylyl-3',3'-uridine hexahydrate., *Journal of Molecular Biology* **104**(1): 109–144.
- Serano, T. and Cohen, R. (1995). A small predicted stem-loop structure mediates oocyte localization of drosophila k10 mrna, *Development* **121**(11): 3809–3818.
URL: <http://dev.biologists.org/content/121/11/3809>
- Shen, C.-P., Jan, L. Y. and Jan, Y. N. (1997). Miranda is required for the asymmetric localization of prospero during mitosis in drosophila, *Cell* **90**(3): 449–458.

- Shen, C.-P., Knoblich, J. A., Chan, Y.-M., Jiang, M.-M., Jan, L. Y. and Jan, Y. N. (1998). Miranda as a multidomain adapter linking apically localized inscuteable and basally localized staufer and prospero during asymmetric cell division in drosophila, *Genes & Development* **12**(12): 1837–1846.
- St. Johnston, D., Driever, W., Berleth, T., Riehlstein, S. and Nüsslein-Volhard, C. (1989). Multiple steps in the localization of bicoid rna to the anterior pole of the drosophila oocyte, *Development* **107**(Supplement): 13–19.
URL: <http://dev.biologists.org/content/107/Supplement/13>
- Steffl, R., Oberstrass, F. C., Hood, J. L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R. B. and Allain, F. H.-T. (2010). The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove, *Cell* **143**(2): 225–237.
- Steffl, R., Skrisovska, L. and Allain, F. H.-T. (2005). RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle, *EMBO reports* **6**(1): 33–38.
- Steitz, T. A. (1990). Structural studies of protein–nucleic acid interaction: the sources of sequence-specific binding, *Quarterly Reviews of Biophysics* **23**(3): 205–280.
- Sugimoto, Y., Vigilante, A., Darbo, E., Zirra, A., Militti, C., D’Ambrogio, A., Luscombe, N. M. and Ule, J. (2015). hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by staufer 1, *Nature* **519**(7544): 491–494.
- Swanson, M. S., Nakagawa, T. Y., LeVan, K. and Dreyfuss, G. (1987). Primary structure of human nuclear ribonucleoprotein particle c proteins: conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins., *Molecular and Cellular Biology* **7**(5): 1731–1739.
- Tang, S. J., Meulemans, D., Vazquez, L., Colaco, N. and Schuman, E. (2001). A role for a rat homolog of staufer in the transport of RNA to neuronal dendrites, *Neuron* **32**(3): 463–475.
- Thomas, M. G., Tosar, L. J. M., Loschi, M., Pasquini, J. M., Correale, J., Kindler, S. and Boccaccio, G. L. (2005). Staufer recruitment into stress granules does not affect early mRNA transport in oligodendrocytes, *Molecular Biology of the Cell* **16**(1): 405–420.
- Valente, L. and Nishikura, K. (2007). Rna binding-independent dimerization of adenosine deaminases acting on rna and dominant negative effects of nonfunctional subunits on dimer functions, *Journal of Biological Chemistry* **282**(22): 16054–16061.
URL: <http://www.jbc.org/content/282/22/16054.abstract>

- Valverde, R., Edwards, L. and Regan, L. (2008). Structure and function of KH domains, *FEBS Journal* **275**(11): 2712–2726.
- Vessey, J. P., Amadei, G., Burns, S. E., Kiebler, M. A., Kaplan, D. R. and Miller, F. D. (2012). An asymmetrically localized staufen2-dependent RNA complex regulates maintenance of mammalian neural stem cells, *Cell Stem Cell* **11**(4): 517–528.
- Vessey, J. P., Macchi, P., Stein, J. M., Mikl, M., Hawker, K. N., Vogelsang, P., Wieczorek, K., Vendra, G., Riefler, J., Tubing, F., Aparicio, S. A. J., Abel, T. and Kiebler, M. A. (2008). A loss of function allele for murine staufen1 leads to impairment of dendritic staufen1-RNP delivery and dendritic spine morphogenesis, *Proceedings of the National Academy of Sciences* **105**(42): 16374–16379.
- Villace P, Marion RM, O. J. (2004). The composition of staufen-containing rna granules from human cells indicates their role in the regulated transport and translation of messenger rnas., *Nucleic Acids Research* **32**(8): 2411–2420.
- Wang, X., Vukovic, L., Koh, H. R., Schulten, K. and Myong, S. (2015). Dynamic profiling of double-stranded rna binding proteins, *Nucleic Acids Research* **43**(15): 7566–7576.
URL: <http://dx.doi.org/10.1093/nar/gkv726>
- Wickham, L., Duhaîne, T., Luo, M., Nabi, I. R. and DesGroseillers, L. (1999). Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum, *Molecular and Cellular Biology* **19**(3): 2220–2230.
- Wilusz, C. J. and Wilusz, J. (2004). Bringing the role of mrna decay in the control of gene expression into focus., *Trends in genetics : TIG* **20**: 491–497.
- Yang, Y., b. Wang, X., Frerking, M. and Zhou, Q. (2008). Spine expansion and stabilization associated with long-term potentiation, *Journal of Neuroscience* **28**(22): 5740–5751.

List of Figures

1.1	PTGR	6
1.2	RNA binding domains of RBP	7
1.3	RNA-recognition motif of the Fox1 protein in complex with ssRNA	8
1.4	KH domain of the Nova 2 protein in complex with ssRNA	9
1.5	zinc fingers motif of the TIS11d protein in complex with ssRNA	10
1.6	Multiple sequence alignment of dsRBP	11
1.7	dsRBD3 of the <i>DmStau</i> protein	13
1.8	dsRNA recognition by dsRBD domain	15
1.9	dsRNA recognition by ADAR2	16
1.10	Schematic representation of Staufen proteins	19
1.11	Staufen mediated decay (SMD)	25
3.1	SEC analysis of hStau1 FL	30
3.2	hStau1 constructs used in this thesis	30
3.3	Representative SEC results of the hStau1 mutants	33
3.4	hStau1 FL sequence	33
3.5	Representative SEC analysis of hStau1 FL and hStau1 ^{dsRBD3-5} in high ionic strength	34
3.6	Representative SEC-MALLS results of hStau1 FL in two different set ups.	35
3.7	Representative SEC-MALLS results of the hStau1 mutants	36
3.8	Gel filtration of hStau1 FL	37
3.9	EMSA and FA of hStau1 FL in complex with <i>ARF1</i> SBS	39
3.10	SEC of hStau1 ^{dsRBD2-4} - <i>ARF1</i> SBS complex	41
3.11	Biochemical characterization of the hStau1 ^{dsRBD2-4} - <i>ARF1</i> SBS complex	42
3.12	SEC of hStau1 ^{ΔNΔC}	43
3.13	SEC of hStau1 ^{dsRBD3-5} - <i>ARF1</i> SBS complex	44
3.14	SEC-MALLS of hStau1 ^{dsRBD3-5} - <i>ARF1</i> SBS complex	45
3.15	Crystals of hStau1 ^{dsRBD3-5} - <i>ARF1</i> SBS complex	46
3.16	SDS-PAGE and Urea gel of hStau1 ^{dsRBD3-5}	48
3.17	SEC of hStau1 ^{dsRBD3-4}	48
3.18	SEC-MALLS of hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS complex	50
3.19	Difference Fourier map of the hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS complex	51
3.20	Structure of the hStau1 ^{dsRBD3-4} - <i>ARF1</i> SBS complex	53

3.21 dsRBD3 ^A and dsRBD4 interaction interface	55
3.22 Structure of the apo-state <i>ARF1</i> SBS	57
3.23 The G-U wobble pairs <i>ARF1</i> SBS	57
3.24 Superposition of the structure of the <i>ARF1</i> SBS unbound state and in complex with Stau1 ^{dsRBD3-4}	58
3.25 Schematic of the <i>ARF1</i> SBS showing the interactions with the hStau1 ^{dsRBD3-4}	60
3.26 Interactions regions of hStau1 ^{dsRBD3-4} with the <i>ARF1</i> SBS	61
3.27 Circular Dichroism of interactions mutants	62
3.28 Fluorescence anisotropy of interactions mutants	64
4.1 Comparison of the Structures of RNA-bound dsRBDs	67
4.2 Comparison of the Structures of bound hStau1 dsRBD3 and <i>DmStau</i> dsRBD3	69
4.3 Binding register of various dsRBD	71

List of Tables

1.1	Representatives members of dsRBPs	12
3.1	Data collection and refinement statistics.	52
5.1	List of chemicals used for this thesis	73
5.2	List of buffers used for this thesis	74
5.3	List of chemicals used for this thesis	75
5.4	Culture media	75
5.5	Beads and columns	76
5.6	Enzymes, commercial kits and markers	76
5.7	List of crystallization screens used for this thesis	77
5.8	Technical equipment	77
5.9	<i>E. coli</i> strains used for cloning and protein expression	78
5.10	Primer sequences used for mutagenesis	78
5.11	RNA oligonucleotides used for RNA interaction experiments	79
5.12	RNA oligonucleotides used for RNA interaction experiments	79
5.13	Restriction digest of pET-MCN vectors	80
5.14	Mutagenesis reaction	81
5.15	Purification buffer hStau1 and truncated constructs	84
5.16	Protein methods buffers	85
5.17	6% native acrylamide gels	86

Abbreviation

ADAR	dsRNA-dependent adenosine deaminase
ARF1	ADP-ribosylation factor 1
ASU	asymmetric unit
bcd	<i>bicoid</i> mRNA
bp	base pairs
ca.	circa
CBP80	cap binding protein 80
CD	Circular Dichroism
CDS	coding sequences
CNS	central nervous system
CSS	complex-formation significance
CV	column volume
Dm	<i>Drosophila melanogaster</i>
DmStau	<i>Drosophila</i> Staufen
dsDNA	double-stranded DNA
dsRBD	double-stranded RNA Binding domain
dsRNA	double-stranded RNA
Ec	<i>Escherichia coli</i>
EMSA	Electrophoretic mobility shift assays
Exp-5	exportin-5
FA	Fluorescence anisotropy

FAM	fluorescein amidite
FL	Full-Length
FMRP	fragile X mental retardation protein
GMC	ganglion mother cells
HCV	Hepatitis C Virus
hiCLIP	RNA hybrid and individual-nucleotide resolution ultraviolet cross-linking and immunoprecipitation
HIV	human immunodeficiency virus
hNRNP K	human heterogeneous nuclear RNP K Protein
Hs	<i>Homo sapiens</i>
hStau1	human Staufen1
IDT	Integrated DNA Technologies, Inc.
IEX	ion exchange chromatography
IPC	intermediate progenitor cell
IPTG	Isopropyl- β -D-thiogalactopyranoside
KD	dissociation constant
KH	K-Homology domain
KLF2	Krüppel-like factor 2
lncRNA	long noncoding RNA
L-LTP	long-term potentiation
LS	light scattering
LTD	long-term depression
Map1b	Microtubule-associated protein 1b mRNA
maStau	mammalian Stau
mGluR	metabotropic glutamate receptor
MBs	Myoblasts

mM	Millimolar
MPI	Max Planck Institute
MR	Molecular replacement
mRBP	mRNA binding protein
mRNA	messenger RNA
mRNP	mRNA ribonucleoprotein particle
MTs	Myotubes
MW	Molecular Weight
NF90	Nuclear factors 90
NLS	nuclear localization signal
NMD	Nonsense-Mediated mRNA Decay
NMR	nuclear magnetic resonance
Nova 2	Neuro-oncological ventral antigen 2
NS1	nonstructural protein 1
nt	nucleotide
osk	<i>oskar</i> mRNA
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PISA	Protein Interfaces Surfaces and Assemblies
PKR	dsRNA-dependent protein kinase
PNCR2	proline-rich nuclear receptor coregulatory protein 2
PoI II	RNA polymerase II
pros	<i>prospero</i> mRNA
PTGR	post-transcriptional gene regulation
Pur-alpha	purine-rich element binding protein A

RBD	RNA binding domain
RER	rough endoplasmatic reticulum
RGCs	radial glial cells
RI	refractive index
RISC	RNA-induced silencing complex
RIPs	RNA coimmunoprecipitations
RMSD	Root Mean Square Deviation
RNA	Ribonucleic acid
RRM	RNA-Recognition Motif
rRNAs	ribosomal RNAs
SAD	single-wavelength anomalous diffraction
SBS	Staufen Binding Site
SD	standard deviations
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SEC-MALLS	SEC coupled with Multi-Angle Laser Light Scattering
SeMet	seleno-methionine
SLS	Swiss Light Source
SMD	Staufen Mediated Decay
SRS	Staufen recognized structures
ssDNA	single-stranded DNA
SSM	Staufen-swapping motif
ssRNA	single-stranded RNA
Stau	Staufen
TBD	Tubulin binding domain

TEV	Tobacco Etch Virus
TLS	Transport and Localization Sequence
TRBP	human TAR-binding protein
UTR	untranslated region
vRNA	viral RNA
XDS	X-ray Detector Software
XI	<i>Xenopus laevis</i>
Xlrbpa	<i>Xenopus laevis</i> RNA-binding protein A
zf	zinc finger

Acknowledgements

I want to thank my supervisor Fulvia Bono for allowing me to work in your research group. Thank you a lot for your support, trust, and patience. I am also grateful to the members of my thesis committee, my second supervisor Prof. Dr. Thilo Stehle and Dr. Remco Sprangers, for their helpful discussions and suggestions. I want to thank the Ph.D. Program coordinator Dagmar Sigurdardottir, for help and support in different aspects during my Ph.D.

My sincere thanks to Dr. Claire Basquin at the Max Planck Institute of Biochemistry in Martinsried for assistance with biophysical experiments and scientific support.

I thank all the former and current members of the AG Bono for the friendly working atmosphere. Thanks to Dr. Zebin Hong and Christiane Emmerich for their support with crystallography experiments. Special thanks to Dr. Daniela Lazzaretti for your expertise, your support and collaboration during my Ph.D. I have appreciated your company and all the funny moments together.

Schließlich möchte ich mich bei meinem Ehemann Michael und meiner Tochter Alina bedanken. Danke für eure unendliche Geduld, eure liebevolle Unterstützung und bedingungslose Liebe.

Danke liebe Alina dass du da bist. Es gab sehr schwierige Momente auf diesem Weg aber du warst immer da, so fröhlich und positiv. Du hast immer an mich geglaubt und zu mir gesagt: „Ja Mami, du schaffst das!“. Dieser Satz stärkt mich noch bis heute. Ich liebe dich unendlich.