



universität
wien

DISSERTATION

Titel der Dissertation

“Nuclear import by an ADAR1 double stranded RNA
binding domain“

Verfasserin

Silpi Banerjee

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, 2014

Studienkennzahl lt. Studienblatt:

A 094 490

Dissertationsgebiet lt. Studienblatt:

Molekulare Biologie

Betreut von:

Ao. Univ. Prof. Dr. Michael F. Jantsch

Dedicated to you Mom,

Your battle with cancer has been the driving force behind my research interest.

Table of Contents

Summary.....	3
Zusammenfassung.....	4
1. Introduction	5
1.1. RNA binding proteins	5
1.2. The double stranded RNA binding domains (dsRBDs)	5
1.2.1 Canonical dsRBDs	7
1.2.2 dsRBDs and RNAs	10
1.2.3 Additional functions of dsRBDs	11
1.3. Nuclear targeting of proteins	12
1.3.1 NLS and Import factors.....	14
1.3.2 Transportin-1 as import factor.....	16
1.4. RBPs harboring NLSs	19
1.5. ADAR proteins: dsRBD containing protein harboring NLSs.	20
1.5.1 Substrates of ADAR	24
1.5.2 dsRBDs of ADARs.....	26
1.5.3 An RNA regulated shuttling motif.....	28
2. Specific Aims.....	30
3. Materials and Method.....	31
3.1 Cloning of cDNA for recombinant protein expression.....	31
3.2 Expression and Purification of recombinant proteins	31
3.3 In vitro mutagenesis	32
3.4 Transfection and immunofluorescence localization.....	32
3.5 Immunoprecipitation using recombinant proteins	32
3.6 RNA samples for Immunoprecipitation.....	33
3.7 Import assay.....	33
3.8 NMR Spectroscopy and protein structure determination.....	34

3.9	Protein structure determination	34
3.10	RNA samples and isothermal titration calorimetry	34
3.11	Modelling the interaction with Transportin 1.....	35
3.12	Accession Codes	35
4.	Results	36
4.1	ADAR1-dsRBD3 with extended borders served ideal for NMR	36
4.2	Nuclear localization of ADAR1 requires the region N-terminal to dsRBD3..	38
4.3	ADAR1-dsRBD3 is an extended domain.....	38
4.4	Extended features of ADAR1-dsRBD3 embody the nuclear localization elements	41
4.5	dsRBD3 acts as a scaffolding domain for the bimodular NLS	45
4.6	Designing a minimal putative NLS	49
4.7	ADAR1-dsRBD3 uses the same binding surface as that of known Transportin-1 –PY NLSs.....	53
4.8	Modulation of nuclear import by dsRNA binding	56
5.	Discussion	62
5.1	ADAR1-dsRBD3 is an extended dsRBD allowing a dual function	63
5.2	Comparison of the ADAR1-NLS with classical bipartite NLSs	64
5.3	Comparison of the ADAR1-NLS with PY-NLSs.....	67
5.4	Regulation of nuclear Import	68
5.5	Nuclear import regulation <i>via</i> an RNA-sensing NLS.....	69
6.	References	72
	Appendix	92
	Acknowledgements	94
	Curriculum Vitae.....	95

Summary

Adenosine deaminases that act on RNAs (ADAR) influence the stability, structure and localization of RNA molecules by RNA editing. Mammalian ADAR1 is a nucleo-cytoplasmic shuttling protein containing three double stranded RNA binding domains (dsRBDs). The third dsRBD is the first domain of this type which shows nuclear localization signal (NLS) activity. We could show previously that nuclear import is mediated by Transportin 1 (Trn1) and is inhibited by RNA-binding. However, the specific structural features allowing this dsRBD to act as a NLS were unknown.

In this PhD thesis in a collaborative effort the solution structure of this domain, ADAR1-dsRBD3 was determined. The structure revealed that ADAR1-dsRBD3 has an additional α -helix at its amino terminus and flanking regions on either side of the domain. *In vitro* mutagenesis analysis demonstrated that the two flanking regions constitute a bimodular nuclear localization signal. Importantly, the additional N-terminal α -helix radically changes the relative position of the N- and C-termini of the dsRBD thereby bringing the N-terminus in proximity to the C-terminus. Furthermore I could design new bimodular NLSs by replacing ADAR1-dsRBD3 by another unrelated dsRBD or even by a small peptide linker maintaining the positioning of the N- and C-terminal regions. This indicates that the dsRBD itself is not required for Trn1 interaction, but only helps to juxtapose the two NLS-modules that are distantly spaced in the primary structure. These observations were verified by both in-vitro import assay and pull down assays. Consequently, the two flexible fragments flanking ADAR1-dsRBD3 that constitute the active NLS are exposed on one surface and can be recognized by the same surface of Trn1 as is known for other Trn1 substrates. In the presence of double stranded RNA, Trn-1 interaction is disturbed. Modeling suggests that RNA-binding prevents the dsRBD-NLS to enter the Trn-1 pouch in its RNA-bound state.

Zusammenfassung

RNA-spezifische Adenosin-Deaminasen (ADARs) beeinflussen die Stabilität, Struktur und Lokalisation von RNAs durch RNA-Editierung. Säuger-ADAR1 ist ein nukleo-cytoplasmatisch shuttlendes Protein, welches drei doppelstrang-RNA Bindungsdomänen (dsRBDs) beinhaltet. Die dritte dsRBD in ADAR1 ist die erste Domäne dieses Typs, welche Kernlokalisierung mediert. In früheren Studien konnten wir zeigen, dass Kernimport durch Transportin 1 (Trn1) vermittelt wird und durch RNA-Bindung inhibiert wird. Die strukturellen Eigenschaften dieser Domäne waren jedoch unbekannt.

In dieser Doktorarbeit wurde im Rahmen einer Kollaboration die Struktur dieser Domäne mittels NMR ermittelt. Die Struktur zeigt eine dsRBD Faltung mit einer zusätzlichen alpha-Helix am Aminoterminal sowie weitere unstrukturierte Elemente an beiden Enden der Domäne. In vitro Mutagenese zeigt, dass die zwei flankierenden Regionen ein bimodulares Kernlokalisierungssignal bilden. Die aminoterminal alpha Helix verändert dabei die relative Lage des N- und C-Terminus der dsRBD und bringt beide Enden nahe zueinander. Ich konnte ein neues bimodulares NLS erzeugen, indem ich die ADAR1-dsRBD3 durch eine andere dsRBD oder ein kurzes Peptid ersetzte, wobei die Lage der N- und C-terminalen Segmente unverändert blieb.

Dies zeigt, dass die dsRBD nicht für die Trn1 Interaktion benötigt wird, sondern lediglich die Position der zwei NLS-Module zueinander gewährleistet. Diese Beobachtung wird durch in-vitro Import - und Kopräzipitationsversuche verifiziert. Es werden also die zwei flexiblen Fragmente welche die ADAR1-dsRBD bestimmen an einer Seite der Domäne exponiert. Diese werden von Trn1 über die gleichen Kontaktpunkte gebunden, wie andere Trn1 Substrate. In der Gegenwart von doppelstrang RNA ist die Interaktion mit Trn1 inhibiert. Modelle zeigen, dass die dsRBD in Gegenwart von RNA vermutlich nicht in die Bindungstasche von Trn1 eindringen kann.

1. Introduction

1.1. RNA binding proteins

The fate of a nascent RNA transcript is defined largely by the protein complex associated with it. A RNA species from its very birth is almost never freely available but is always surrounded by protein complexes. The protein complexes associated with the RNA then guide the function of the RNA by getting involved in various biological processes beginning from transcription to localization to either being translated to a protein or by being involved in gene expression or by getting degraded at the end. To successfully enable any one of these processes the participating proteins need to recognize and bind to a RNA transcript either specifically or in a non-specific way. The complex and dynamic association of a protein with RNA defines the lifespan, processing and localization of the RNA (Martínez-Salas et al., 2013) (Figure 1.1). All these proteins are RNA-binding proteins (RBPs) and harbor RNA binding modules (RBMs). The diversity of functions these RBPs accomplish suggests a considerably large heterogeneity in the structural determinants they possess (Lunde et al., 2007). However, most RBPs are made from a few RBMs to successfully achieve the various functions. This high variability is achieved by multiple copies of RBMs and their different arrangements (Figure 1.2). RBMs are very conserved and vary from their mode of action. Together with the regulatory catalytic domains and various RBMs of similar or different structural types the versatile specificity of RBPs is achieved.

1.2. The double stranded RNA binding domains (dsRBDs)

Many eukaryotic and viral proteins interact with double-stranded RNA to regulate numerous biological processes such as RNA-localization, RNA-processing, RNA-interference and RNA-editing. Many of these RBPs recognize their substrate RNA specifically by either using a sequence specific or a structure specific signature. These proteins sense the double-stranded RNA structure through their double stranded RNA binding domains (dsRBD). dsRBD is a conserved protein motif of approximately 65-75 amino acids in length that binds double-stranded or highly structured RNAs. It was first discovered in *Drosophila*

Staufen and in *Xenopus laevis* RNA binding protein-A (XlrpbA)(St Johnston et al., 1992) .

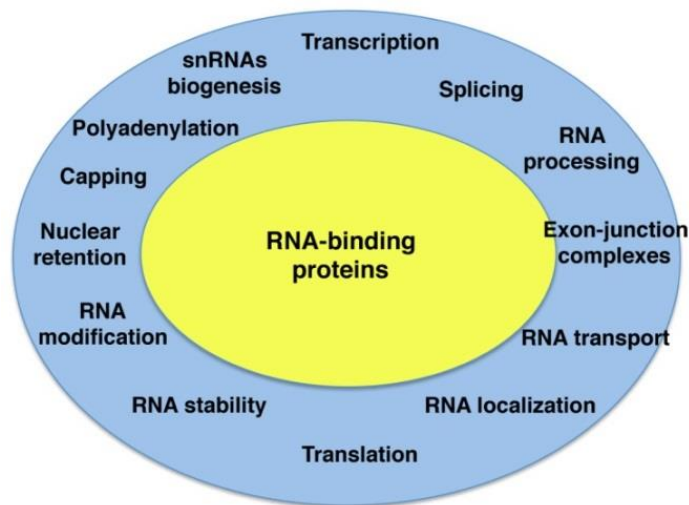


Figure 1.1 RNA binding proteins in action

Each step of gene expression is controlled by RNA-binding proteins (*yellow*), such that altering one step affects the rest (*blue ring*) impacting on translation efficiency of mRNAs.(Martínez-Salas et al., 2013)

Structural and biochemical analyses indicate that dsRBDs recognize the typical A-form helix assumed by double-stranded RNA (dsRNA) and require at least 1.5 helical turns for binding (Ramos et al., 2000; Ryter and Schultz, 1998b). dsRBDs can be found in all kingdoms of life and are the second most abundant RNA binding motif identified today (Fierro-Monti and Mathews, 2000). Consistently, proteins harboring this domain vary considerably in function, depending on catalytic domains and interaction partners. Well studied proteins containing this motif are RNase III, the RNA activated kinase PKR, the RNA localizing protein Staufen, the RNAi enzyme Dicer, or the RNA editing protein family ADAR (Chang and Ramos, 2005). Figure 1.2 shows the diversity in the number of copies (as many as 14 in vigilin) and the arrangement of dsRBDs that exist. For example, in Dicer and RNase III the dsRBD is located after the catalytic endonuclease domain. Although, both of these proteins recognize dsRNA, Dicer has evolved to interact specifically with RNA species that are produced through the RNA interference pathway through additional domains that recognize the unique structural features of these RNAs (Jantsch and Gall, 1992; Lunde et al., 2007).

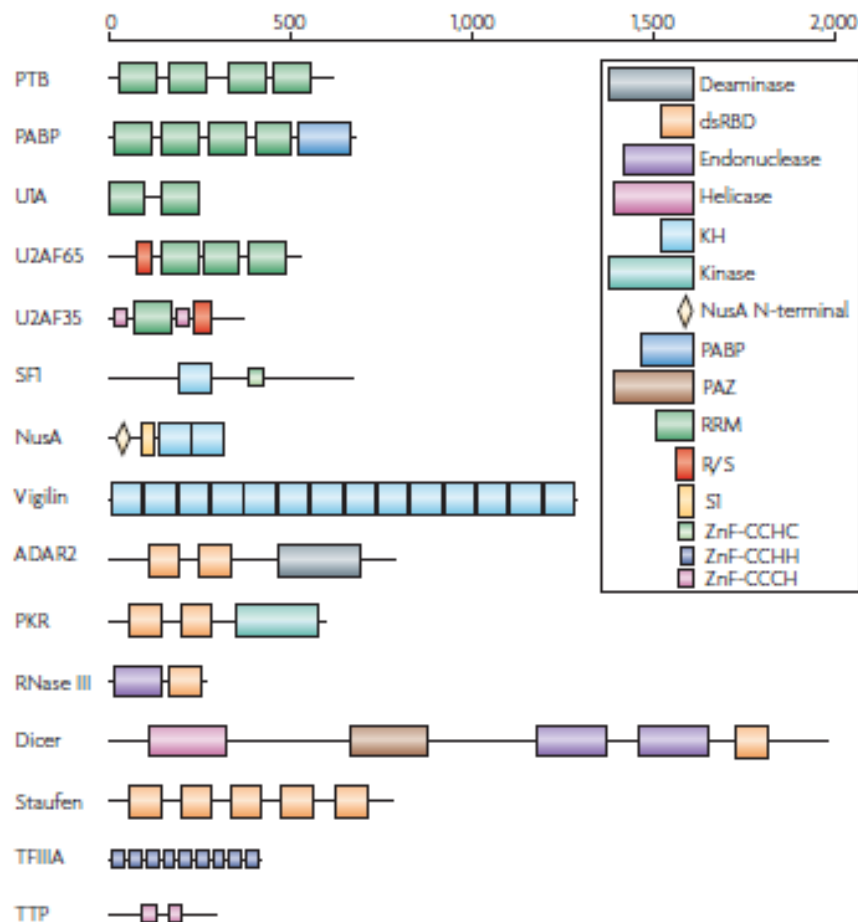


Figure 1.2 RNA binding protein with RNA binding modules

Schematic representation of examples from some of the most common RNA-binding protein families, are illustrated here. Different domains are represented as colored boxes. These include the RNA-recognition motif (RRM), the K-homology (KH) domain (which can bind both single-stranded RNA and DNA), the dsRBD (a sequence-independent dsRNA-binding module) and RNA binding zinc-finger (ZnF) domains. Enzymatic domains and less common functional modules are also shown. PABP (poly(A)-binding protein); PTB (polypyrimidine-tract binding); R/S (Arg/Ser-rich domain); SF1 (splicing factor-1); TTP (tristetraprolin); U2AF (U2 auxiliary factor). Figure taken from (Lunde et al., 2007).

1.2.1 Canonical dsRBDs

So far about 30 dsRBD structures have been reported with a mixed α/β fold having a conserved $-\alpha\beta\beta\beta\alpha-$ topology (Masliah et al., 2013). The second dsRBD of XlrbpA forms the prototype of a canonical dsRBD with three anti-parallel β sheets packed in between two α -helices (Ryter and Schultz, 1998a) (Figure 1.3).

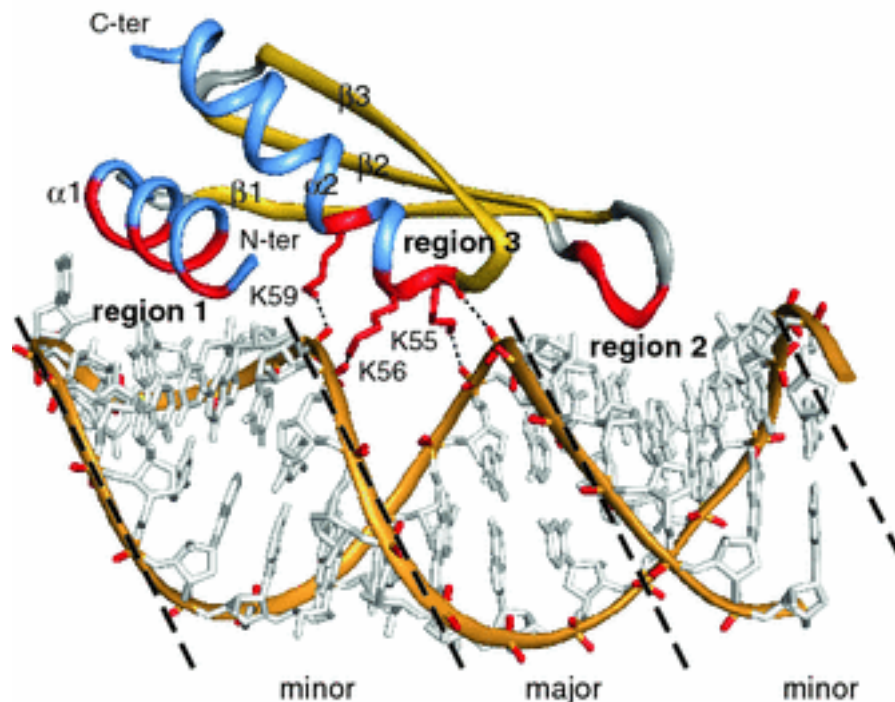


Figure 1.3 Structure of canonical dsRBD with dsRNA interacting surface

A typical dsRBD has three contact regions with a dsRNA substrate with two minor grooves and a major groove as depicted. Region 1 (helix $\alpha 1$) and region 2 (loop between $\beta 1$ - $\beta 2$ or loop 2) insert into successive RNA minor grooves. Region 3 (loop between $\beta 3$ - $\alpha 2$ or loop 4) contacts the phosphodiester backbone of the intermediate RNA major groove. The canonical KKxAK-dsRNA recognition motif of dsRBD is represented here by red sticks in the ADAR2-dsRBD1-RNA stem-loop complex (PDB code 2L3C). Figure taken from (Masliah et al., 2013).

Overall the dsRBD structures are all conserved but the majority of the amino acids that are highly conserved span around the last third of the domain at the C-terminal end of the domain. The first two third from the N-terminus of the domain showing maximum variations among dsRBDs and those dsRBDs which differ in structure considerably in these regions have been termed type B dsRBDs due to their low binding affinity to their respective dsRNA substrates (Krovat and Jantsch, 1996a). As depicted in Figure 1.3 the residues appear to be conserved for maintaining a stable hydrophobic core and for optimal dsRNA binding. The conserved hydrophobic residues comprise from all the secondary structured elements namely, $\alpha 1$, $\beta 1$, $\beta 2$ and $\alpha 2$ structure of dsRBD. Aliphatic side chains namely L6, L9, V39, V41, A58, A62, A63, A66, L67 and L70 (residue are numbered as per alignment shown in (Figure 1.4) located in $\alpha 1$, $\beta 2$ and $\alpha 2$ helices

respectively are highly conserved. The small side chains in A58 and A62 in helix $\alpha 2$ along with the conserved motif GxG at the end of $\beta 3$ forms a tight packing in this region making a remarkably close proximity between the $\alpha 2$ helix and $\beta 3$ sheet (Masliah et al., 2013). Apart from the aliphatic side chains the aromatic rings are predominantly in the β -strands. Y21 in the $\beta 1$ and F35 in the $\beta 2$ are the most conserved aromatic side chains among all dsRBDs known so far as in Figure 1.4. These aromatic rings are involved in interaction with dsRNA substrate by maintaining positively charged residues in optimum orientation (Krovat and Jantsch, 1996a; Ramos et al., 2000; Ryter and Schultz, 1998a). The positioning of the aromatic rings at the edge of the hydrophobic core possibly imparts additional stability to the entire domain.

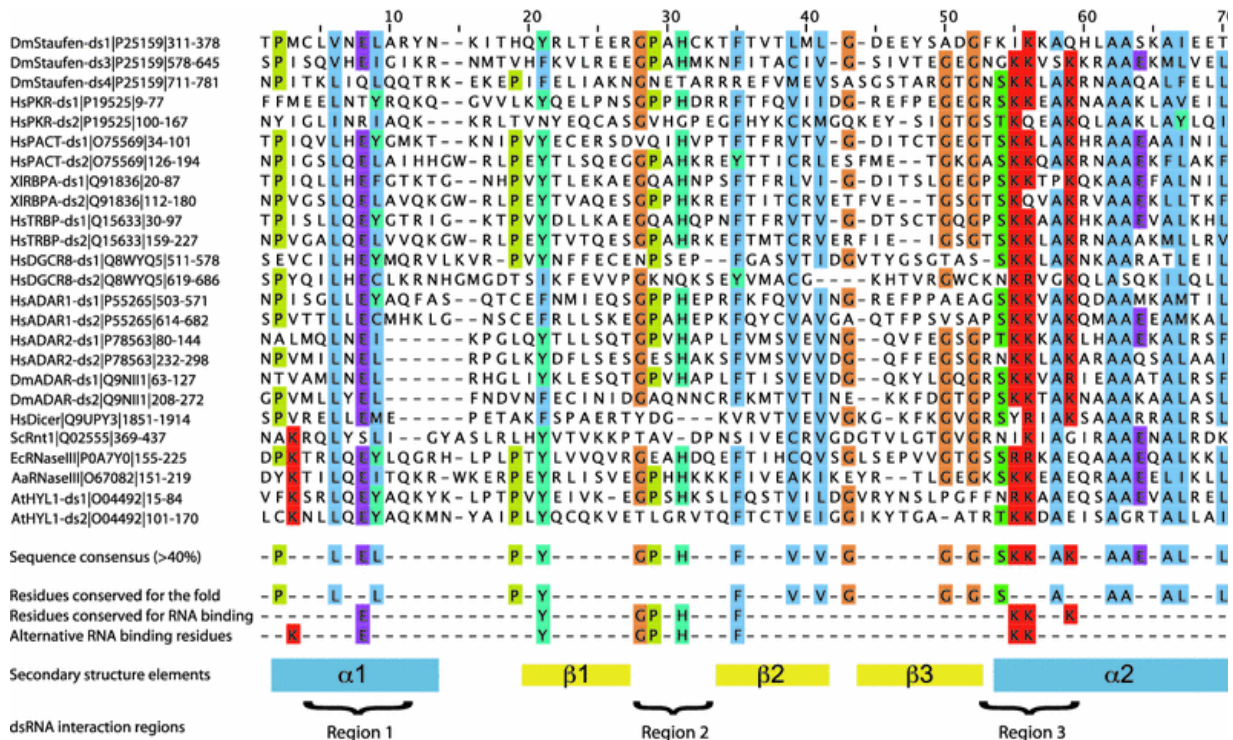


Figure 1.4 Multiple sequence alignment of various dsRBDs

Sequence alignment of dsRBDs from humans (*Homo sapiens*: Hs), fruit fly (*Drosophila melanogaster*: Dm), yeast (*Saccharomyces cerevisiae*: Sc), frogs (*Xenopus leavis*: Xl), Plants (*Arabidopsis thaliana*: At) and bacteria (*Escherichia coli*: Ec and *Aquifex aeolicus*: Aa) have been depicted in this figure. The alignment is colored by amino acid conservation >40% and properties. The conserved residues are color coded for conserved fold or secondary structure or canonical structured element. The three major dsRNA binding regions are also indicated. From Masliah, Barraud et al. 2013.

1.2.2 dsRBDs and RNAs

Table 1.1 indicates the number of high resolution structures of dsRBD in association with RNA substrates known till date. These complex structures have revealed three major dsRNA binding surfaces on dsRBD also shown in Figure 1.3. The three distinct regions of interaction include : (1) the N-terminal tip of the domain α_1 , (2) the β_1 - β_2 loop and (3) the N-terminal tip of helix α_2 (Ryter and Schultz, 1998a). dsRBDs span over a 15bp long dsRNA with two consecutive minor grooves with a major groove in between (Stefl et al., 2010).

Table 1.1 dsRBD in association with RNA substrates

Protein Name	Method	RNA substrate	References
<i>S.cerevisiae</i> Rnt1p	NMR	snoRNA snR47 capped by an AGAA or AAGU	178 180
<i>A.aeolicus</i> RNase III	X-ray	Artificial substrate such as: [GGCGCGCGCC] coaxially stacked duplex	18
<i>A.aeolicus</i> RNase III	X-ray	Artificial substrates such as: [AAAUUAUAUUUU] coaxially stacked duplex. 3 structures with 3 different substrates	57
<i>A.aeolicus</i> RNase III	X-ray	Coaxially stacked hairpin derived from R1.1 a natural substrate of E.coli RNaseIII	58
<i>A.aeolicus</i> RNase III	X-ray	Coaxially stacked hairpin derived from R1.1 a natural substrate of E.coli RNaseIII	56
<i>R.norvegicus</i> ADAR2	NMR	Hairpin dsRNA derived from the mammalian GluR2 R/G editing site. s	160
<i>D.melanogaster</i> Staufen	NMR	Artificial substrate with a UUCG tetraloop.	11
<i>X.leavis</i> RBPA	X-ray	Artificial substrate such as: [GGCGCGCGCC] coaxially stacked duplex	150
<i>H.sapiens</i> TRBPA	X-ray	Artificial substrate such as: [CGCGCGCGCG] coaxially stacked duplex	182
<i>A.thaliana</i> HYL1	X-ray	Artificial substrate such as: [CUCGAUAACC] coaxially stacked duplex	183
<i>A.thaliana</i> HEN1	X-ray	Coaxially stacked duplex derived from miR173/ miR173*, a canonical substrate of HEN1	80

Table adapted from Masliah, Barraud et al. 2013.

The molecular basis of substrate recognition by dsRBDs is achieved by virtue of the highly conserved regions of E8 in helix $\alpha 1$, the GPxH motif in the $\beta 1$ - $\beta 2$ loop and the positively charged KKxAK motif at N-terminal tip of the helix $\alpha 2$. These interactions are mostly achieved by the 2'-OH and/or $-\text{PO}_4$ groups and hence perfectly adapt to the shape of the RNA species. Despite similarities among the interacting surfaces, there are considerable variations in the recognition pattern of dsRBDs, especially those mediated by the loop of helix $\alpha 1$. The helix $\alpha 1$ of dsRBD2 of XlrbpA interacts non-specifically with the minor groove of RNA whereas the high resolution structures of dsRBD in complex with RNA stem loops also reveals specificity in interaction in case of Staufén dsRBD3, ADAR2 dsRBD1 and dsRBD2. Moreover the helix $\alpha 1$ of dsRBD of Rntp1 recognizes the specific shape of tetra loop in the minor groove of RNA. Another major way in which specificity is observed in case of proteins carrying multiple dsRBDs such as ADAR2 is through co-operation among multiple dsRBDs by virtue of which, when one dsRBD binds to the dsRNA, there is only restricted space for the second dsRBD to bind thus only a few binding sites are accessible as potential binding sites (Stefl et al., 2010). Pivoting on these findings, the least conserved secondary structural element the helix $\alpha 1$ seems to have a different spatial arrangement relative to the rest of the domain. In addition to it the nucleotide sequence in the RNA minor groove also modulates the specificity of a dsRBD (Błaszczuk et al., 2004; Gan et al., 2008; Gan et al., 2005; Huang et al., 2009; Ramos et al., 2000; Ryter and Schultz, 1998b; Wang et al., 2011; Wu et al., 2004b; Yamashita et al., 2011; Yang et al., 2010).

1.2.3 Additional functions of dsRBDs

dsRBD containing proteins have profound involvement in gene regulation. However not all of these regulatory functions are RNA dependent. Besides RNA-binding, additional functions have been suggested for a group of dsRBDs that tend to be less well conserved in their amino terminal ends. In fact, some of these less conserved dsRBDs (also referred to as type B dsRBDs) have been shown to be involved in protein localization, while others are required for homo- and/or hetero-dimerization (Doyle and Jantsch, 2002; Hitti et al., 2004; Micklem et al., 2000). For example, dsRBD2 of *Drosophila* Staufén is necessary for localization of *oskar*-mRNA by regulating microtubules. This is in contrast to dsRBD5 which is involved

in de-repression of translational activation. Similarly dsRBD3 of PACT (a protein activator of PKR) plays a pivotal role in activating PKR in the absence of RNA (Krovat and Jantsch, 1996b; Tian et al., 2004). Many proteins containing dsRBDs harbor a bipartite nuclear localization signal (NLS) and are hence localized into the nucleus (Saunders and Barber, 2003). Moreover it has also been reported that dsRBD alone can guide proteins to the nucleolus (Desterro et al., 2003a; Sansam et al., 2003; Tian and Mathews, 2001). Proteins such as PKR, NF90, and XlrpA are also associated with ribosomal rRNA. Moreover many dsRBD containing proteins such as ADAR1, NF90 are nucleo-cytoplasmic shuttling proteins. Exportin-5 mediates export of these two proteins from nucleus to cytoplasm. dsRBD2 of NF90 interacts with Exportin-5 and is regulated by binding of dsRNA (Brownawell and Macara, 2002b). dsRBDs can regulate the import and/or export of a protein by virtue of its interaction to an export or import factor such as Exportin-5 or Transportin-1 as in case of Staufen and ADAR1 (Fritz et al., 2009a; Macchi et al., 2004).

1.3. Nuclear targeting of proteins

Eukaryotic cells have an additional level of gene regulation by compartmentalizing transcription from translation by having a defined nucleus separating chromosomal material from the cytoplasm by nuclear membrane. Despite the physical barrier there is selective and efficient movement of macromolecules across the nuclear envelope via an elegant nucleocytoplasmic transport mechanism through complex pore structures, the nuclear pore complex (NPC).

NPCs are octagonal, rotation symmetric multi-protein channels embedded in the nuclear envelop (Reichelt et al., 1990). They form the largest known protein complexes in eukaryotic cells from 60MDa in yeast to 125 MDa in humans. NPCs consist of three different structural features based on location: 1) the cytoplasmic ring and eight cytoplasmic filaments 2) the membrane embedded central channel and the scaffold supporting it and 3) the nuclear ring and nuclear basket (Figure 1.5). Structurally NPCs are highly conserved among species from yeast to higher mammals however new evidence suggests that the protein composition of NPCs varies among cell types and tissues. At the molecular level they are composed of

several copies of more than 30 different nucleoporins (NUPs)(Raices and D'Angelo, 2012). NUPs contain the transmembrane domains, α -helices, β -propellers, WD domains and FG repeats (Chadrin et al., 2010; Hoelz et al., 2011). NUPs carrying 4 to 48 FG repeats filling the central channel of the NPC, thereby forming a meshwork that is responsible for controlling nucleocytoplasmic transport and that determines the pore permeability limit (D'Angelo and Hetzer, 2008).

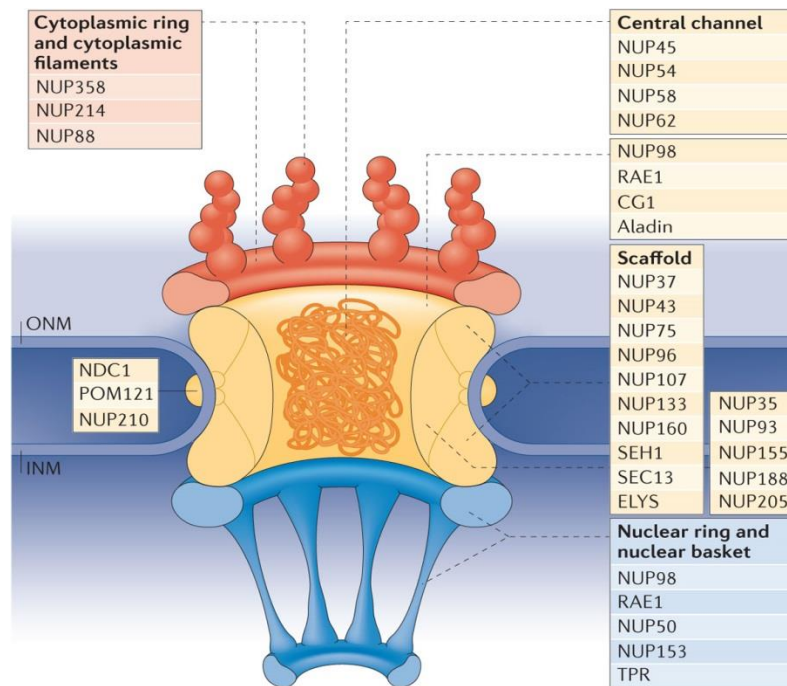


Figure1.5 Structural features of nuclear pore complex (NPC)

Schematic illustration of the nuclear pore complex (NPC) structure and composition. A cross-section normal to the plane of the nuclear envelope is illustrated. The cytoplasmic ring and cytoplasmic filaments are marked with *red* while the NPC scaffold and central transport channel are marked in *yellow* and *orange* respectively. The nuclear ring and nuclear basket are shown in *blue*. The nucleoporins (NUPs) comprising each of these structures are enclosed in respective boxes. Acronyms used in the figure are CG1: candidate gene 1; INM: inner nuclear membrane; NDC1: nuclear division cycle protein; ONM: outer nuclear membrane; POM121: pore membrane protein of 121 kDa; RAE1: RNA export 1; SEH1: SEC13 homologue protein 1.(Raices and D'Angelo, 2012)

In an unbound state the central pore is only 9nm in diameter, thereby allowing diffusion of small macromolecules of molecular mass 40-60 kDa (Paine et al., 1975). On the contrary, in the bound state with a soluble transport receptor, the nuclear pore can expand to 25nm in diameter thereby making it possible for protein

complexes as large as several million Daltons to pass through it (Vasu and Forbes, 2001). However, even small proteins and RNAs are actively transported resulting in a highly efficient and selective transport mechanism (Breeuwer and Goldfarb, 1990). A transport rate of 1000 transport events per second has been reported in eukaryotic cells (Ribbeck and Gorlich, 2001). Nuclear transport of most proteins and RNAs (cargoes) starts when an import or export receptor (karyopherin) binds to its cargo in the cytoplasm or the nucleus, respectively as illustrated in Figure 1.6. The complexes then move through the diffusion barrier of the pore by interacting with the FG repeats of NUPs to finally reach their destination compartment where cargoes are released (Raices and D'Angelo, 2012).

1.3.1 NLS and Import factors

Trafficking of macromolecules between the nucleus and cytoplasm through NPCs is a selective and an efficient process orchestrated by transport receptors that associate with their cargoes *via* cognate nuclear localization signals (NLSs) or nuclear export signals (NESs). Transport receptors of the Karyopherin- β (Kap β) family, also known as importins and exportins, account for the vast majority of the cargo flow through the NPC. Transport directionality is primarily driven by the RanGTPase nucleotide cycle, which produces an asymmetric distribution of RanGTP and RanGDP on both sides of the nuclear envelope (Figure 1.6)(Chook and Suel, 2011; Cook et al., 2007; Fried and Kutay, 2003; Gorlich and Kutay, 1999).

Over the past several years 14 Kap β s have been identified in yeast and more than 20 in mammals. Members of Kap β family share similar molecular weights (90-150k Da) and contain multiple tandem helical repeats known as HEAT repeats (explained in details in section 1.3.2). 13 of the yeast Kap β s and few of the mammalian Kap β s identified are known for involvement in import, export or bidirectional nuclear transport. Kap β s interact with both the cargo and the NUPs of NPC there by mediating nuclear import or export of the cargo protein (Chook and Blobel, 2001). Cargo- Kap β s interaction is regulated by Ran GTPases. The direction of transport is determined by the mode of interaction of interaction between the three proteins (RanGTPase, Kap β , and cargo) and their response to Ran nucleotide state (Figure 1.6). The nucleotide state of Ran is controlled by

RanGTPase activating protein (Ran GAP) and Ran guanine nucleotide exchange factor (RanGEF or RCC1) (Bischoff and Ponstingl, 1995; Fontoura et al., 2000; Pemberton et al., 1998).

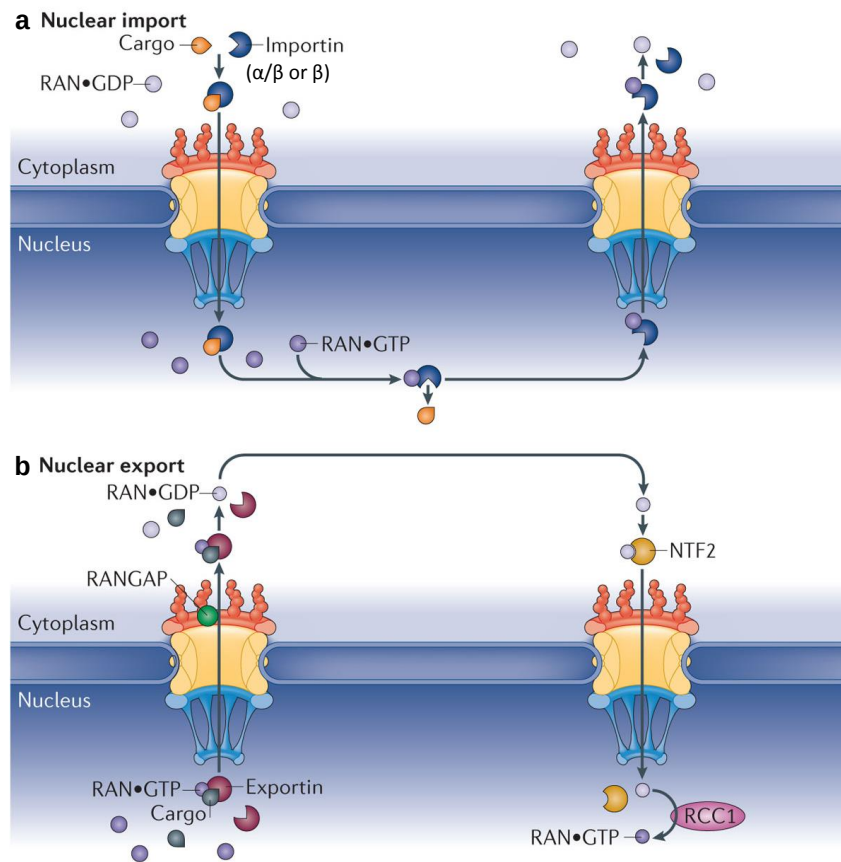


Figure 1.6 Nucleo-cytoplasmic transport of proteins

Mechanism of classical nuclear import-export cycle has been depicted here. Larger cargoes (such as proteins and some mRNAs) bind to nuclear transport receptors that recognize nuclear localization sequences (NLSs) and nuclear export sequences (NESs). **(A)** Nuclear import receptors (importins) after binding to their cargo in the cytoplasm are able to interact with the NPC and pass through the central transport channel. Once inside the nucleus they bind to Ran•GTP, which causes a conformational change that allows cargo release. The resulting complex of importin and Ran•GTP translocate back to the cytoplasm. **(B)** Nuclear export proteins (exportins) bind to their respective cargo and Ran•GTP in the nucleus and move through the NPC to the cytoplasm. The Ran•GTPase-activating protein (RANGAP) induces GTP hydrolysis by Ran in cooperation with two Ran•GTP-binding proteins (RANBP1 and RANBP2; not shown) at the cytoplasmic filaments of the NPC. The Ran•GTP concentration in the cytoplasm is therefore low. The resulting Ran•GDP complex interacts with the nuclear transport factor, NTF2 to translocate back to the nucleus where the Ran exchange factor RCC1 restores Ran•GTP levels. Figure from Raices and D'Angelo 2012.

RanGEF catalyzes nucleotide exchange thereby maintaining the nuclear Ran molecules in their GTP bound state while RanGAP catalyses nucleotide hydrolysis exhausting the cytoplasmic complexes of GTP-bound Ran. This is essential for re-cycling of the carrier protein (here Kap β), which in the absence of GTP-bound Ran retains back to the free-state thereby being ready for re-association with the cargo protein(Christophe et al., 2000).

Most of the cargo proteins bear the basic type canonical or classical NLS. They are a stretch of few basic amino acids essentially lysines or arginines that appear as single stretch known as monopartite NLSs (consensus sequence K/R₄₋₆) or appear as cluster of two basic stretches separated by a dozen amino acids known as bipartite NLS (consensus sequence (K/R)₂-X₁₀₋₁₂-(K/R)₃). The best studied example of a monopartite NLS is the SV40 large T-antigen and that of bipartite NLS is nucleoplasmin (Nigg, 1997). Based on the predominance of the lysines or arginines in the NLS of the proteins, nuclear import will be mediated either by heterodimer importins composed of importins- α and importins- β or by importins- β alone (Gorlich, 1998; Pollard et al., 1996b).

Besides the classical basic type NLSs, several other sequence motifs have been proposed to mediate nuclear import of cargo proteins. The M9 sequence seems to be a well-known example where a 38 amino acid long fragment serves to be the NLS. It overlaps in hnRNP A1 and A2. M9 sequence is recognized by Transportin-1, a close homologue of importin- β (Pollard et al., 1996b). Shorter NLS motifs that do not bear similarities with classical NLSs have also been discovered as seen in Sam68 (NLS sequence: PPXXR) or in Cdc6 (NLS sequence: S/TPXKRL/I) (Christophe et al., 2000). These NLS sequences are very heterogeneous in nature. This raises an obvious question about the diversity of recognition of NLSs by various import receptors.

1.3.2 Transportin-1 as import factor

Unlike Importin- β , which forms a heterodimer with the adaptor Importin- α to import substrates containing classical NLSs (cNLS) (Gorlich and Kutay, 1999; Lange et al., 2007), Transportin 1 (also known as Kap β 2; thereafter designated Trn1) binds cargoes in the cytoplasm without adaptors, targets them to the nucleus

through the NPC and releases them in the nucleus upon RanGTP binding (Bonifaci et al., 1997; Pollard et al., 1996a) as depicted in Figure 1.7.

Transportin1 belongs to Karyopherin β family (Kap β or Imp β). So far 30 crystal structures of Trn1 have been solved. Most of the structures available are either in complex with substrate NLSs or RanGTP. Full length wild type Trn1 has not yet been crystallized in a free form due to inherent flexibility of the protein. The available complex structures reveal that Trn1 is a superhelical S-like molecule composed of two overlapping arches (N- and C-terminal arches). Each of these arches is formed due to helical stacking of 20 HEAT repeats (H1-H20). Each of which is composed of two antiparallel α -helices (referred to as A and B helices). HEAT repeat was named after the protein huntingtin elongation factor 3, in which the motif was first identified (Andrade and Bork, 1995). There is a disordered loop connecting H8A and H8B helices. Trn1 has 24% sequence homology to Imp β but has a much longer H8 loop, important for substrate dissociation in the nucleus (Chook and Blobel, 2001; Imasaki et al., 2007; Pemberton et al., 1998; Siomi and Dreyfuss, 1995). Figures 1.7 and 1.8 illustrate both the structure and function of Trn1.

In contrast to the well-defined canonical NLSs recognized by Importin- α , Trn1 recognizes a diverse set of sequences with poor apparent conservation that cannot be appropriately described by a conventional consensus sequence. Nevertheless, structural and biochemical studies uncovered a set of loose principles common to some Trn1 signals (Cansizoglu et al., 2007; Imasaki et al., 2007; Lee et al., 2006; Suel et al., 2008). These principles can be described as follows: a peptide segment of 15-30 residues with intrinsic structural disorder, an overall basic character, and some weakly conserved sequence motifs including a relatively conserved proline-tyrosine dipeptide, leading to the name PY-NLSs for this class of signals (Lee et al., 2006).

However, many characterized Trn1 cargoes do not contain such a PY-NLS (Arnold et al., 2006; Chook and Suel, 2011; Jakel and Gorlich, 1998; Muhlhauser et al., 2001; Waldmann et al., 2007). Therefore, Trn1 not only interacts with the well-characterized PY-NLSs, but at least with another class of NLSs, the recognition of which is far less understood. Whether these other classes of NLSs

share some similarities with PY-NLSs remains to be determined. One Trn1 cargo that lacks a typical PY-NLS is the RNA-editing enzyme ADAR1.

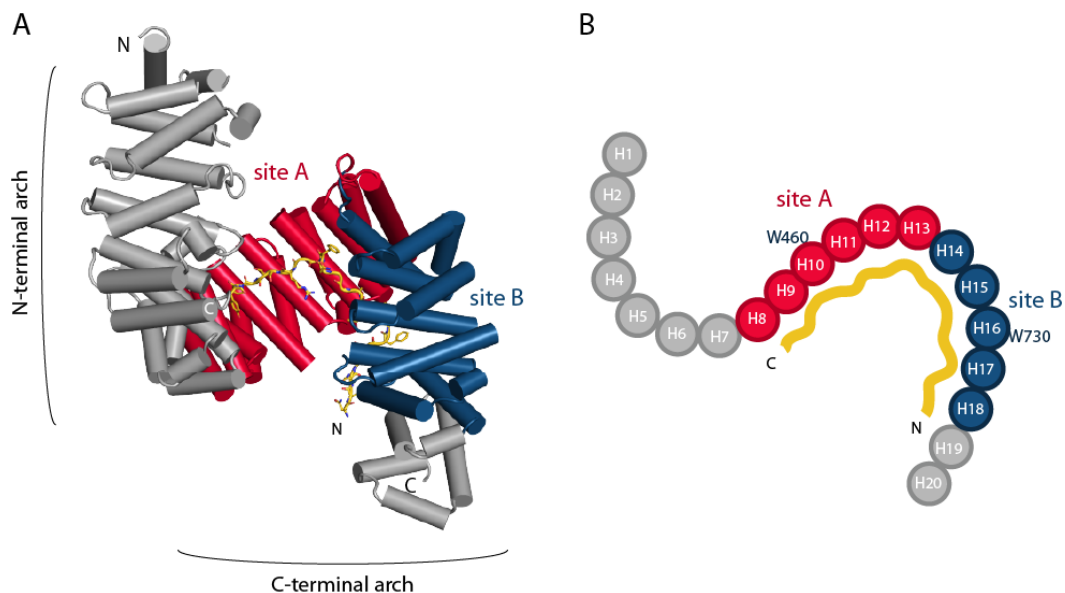


Figure 1.7 Structural basis of NLS recognition by Transportin 1

The structure of Trn1 in complex with hnRNP A1 NLS (PDB code 2H4M) has been chosen for illustration. **(A)** Cartoon representation of Trn1 with the two different sites identified for NLS recognition within the C-terminal arch, named site A and site B, colored *in red* and *in blue*, respectively. **(B)** Schematic representation of the 20 HEAT repeats of Trn1 with site A and site B of the C-terminal arch colored as in panel C, *i.e. in red* and *in blue*. Site A is composed of HEAT repeats 8-13 and site B of HEAT repeats 14-18. The NLS segment is schematically represented *in yellow*. Two tryptophans W460 in HEAT repeat H10 in site A and W730 in HEAT repeat H16 in site B have been reported to be particularly important for interaction with NLS such as hnRNP D, TAP and JKTBP (Imasaki et al., 2007). Figure has been adapted from (Barraud P. Banerjee S., 2014)

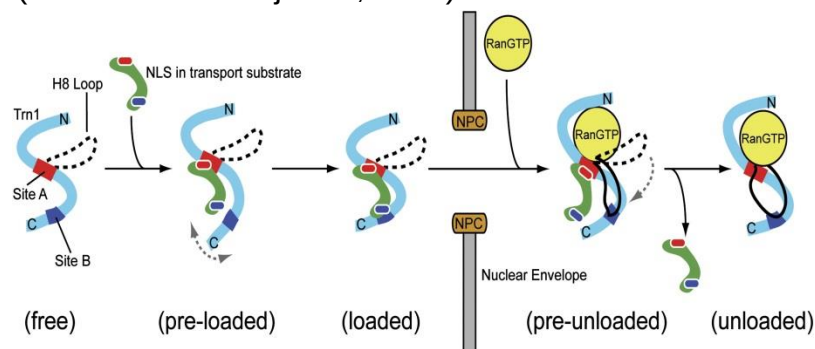


Figure 1.8 Nuclear Import Pathway Mediated by Trn1

Schematic representation highlighting the proposed the mechanism of nuclear import by Trn1. Trn1 is represented as an S-shaped ribbon (*cyan*). An NLS gets loaded on to the Trn1 using anchoring sites at the site A and site B represented in *red* and *blue*. The loading of the NLS makes Trn-1 compact such that it can pass

through the nuclear pore complex (NPC). Once the Trn1-NLS complex is in the cytoplasm, RAN-GTP displaces the NLS mediating nuclear import of the NLS harboring protein (Imasaki et al., 2007).

1.4. RBPs harboring NLSs

dsRBD-mediated nucleocytoplasmic trafficking has been reported for ADAR1 (Fritz et al., 2009b; Strehblow et al., 2002), but also for other dsRBD-containing proteins such as Elongation factor 1A (eEF1A), interleukin enhancer binding factor-3 (ILF3) (Calado et al., 2002; Gwizdek et al., 2004). For example, in human ILF3, a CRM-1 independent NES overlaps the dsRBD2 of the protein. Export of the protein to the cytoplasm is mediated by Exportin-5 (also belongs to Kap β family) which recognizes the NES in dsRBD2. Binding of dsRNA by ILF3 prevents binding of Exportin-5 and mediates nuclear retention (Brownawell and Macara, 2002a). Interestingly, it was also reported by the same group that dsRBD2 of ILF3 promotes export of a double-stranded RNA containing a minihelix by Exportin-5 (Gwizdek et al., 2004). This ambiguity suggests for varying mechanism for dsRBD mediated export by Exportin-5 which needs further detailed investigation. Moreover it has been also reported that cytoplasmic ILF3 in association with dsRNA is unable to enter the nucleus. Binding of dsRNA in the cytoplasmic ILF3 masks the bipartite NLS which is situated upstream of dsRBD1 preventing interaction with any import factor and aiding cytoplasmic retention (Brownawell and Macara, 2002a).

The dsRBD of human Dicer harbors an NLS mediating nuclear import of the protein via Imp β or Imp7 or Imp8. Solution structure of *Saccharomyces pombe* Dicer dsRBD revealed a C-terminal extension composed of a short α -helix together with zinc-finger binding motif (CHCC). Additional structural features of the dsRBD are essential for nuclear localization of the protein into the nucleus, although they did not directly serve as the NLS. Binding of dsRNA to dicer is dispensable for its processing function indicating nuclear import not being regulated by the binding of dsRNA. Instead, the helicase domain embedded in the protein modulates the nucleo-cytoplasmic shuttling of the protein by masking or exposing the NLS spanning in the dsRBD (Barraud et al., 2011; Doyle et al., 2013). Apart from the most important cytoplasmic role of Dicer in miRNA processing, Dicer has been shown to be involved in epigenetic regulation, including RNA-directed DNA methylation (RdDM) (Cernilogar et al., 2011; Doyle et al., 2012). Recently it has

been implicated that in mammalian cells, Dicer functions in the nucleus by involving in gene silencing of genes expressing convergent transcripts and in processing of dsRNA (Gullerova and Proudfoot, 2012). Hence NLS contained in the dsRBD of Dicer can modulate the nucleo-cytoplasmic shuttling of the protein and intern regulates its potential diverse functions.

Another interesting example of a RBP with dsRBD overlapping NLS is Staufen. Staufen localizes mRNAs and regulates certain mRNAs translationally in *Drosophila*. Staufen is conserved from *C.elegans* to humans and possess 4 or 5 dsRBDs based on the species. Both human Staufen1 and Staufen2 are nucleo-cytoplasmic shuttling proteins (Kiebler et al., 2005). The bipartite NLS of Staufen spans in between the dsRBD3 and dsRBD4 of the protein. However, binding of dsRNA does not mask the dsRBD3 and mediates nuclear import of the protein (Martel et al., 2006). The export of the shorter isoform of Staufen2 (Stau2₅₉) is mediated either by Crm-1 (Exportin-1) or Exportin-5 dependent pathway. The leucine rich NES mediating export of the protein is located at the N-terminal edge of dsRBD1 of the protein. However, the NES involved in the Exportin-5 export pathway overlaps the same dsRBD3. The nuclear export of the Stau2₅₉ and Stau2₆₂ via Exportin-5 is regulated in a RNA dependent manner (Macchi et al., 2004; Miki et al., 2005). This is similar to the Exportin-5 mediated nuclear export of ADAR1c where binding of RNA enhances export of the protein. However, the import of ADAR1c by Trn1 is regulated by dsRNA binding to the dsRBD3 which overlaps the NLS (Fritz et al., 2009b).

1.5. ADAR proteins: dsRBD containing protein harboring NLSs.

Adenosine deaminases acting on RNA (ADARs) family members convert adenosines (A) to inosines (I) by hydrolytic deamination in structured and double-stranded RNA substrates (Hogg et al., 2011; Nishikura, 2010b). Inosines formed by adenosine deamination is recognized as guanosine by the cellular machinery and hence, can alter the coding potential of mRNAs, but can also affect RNA stability, RNA structure or splicing, localization, and transport of cellular and viral RNAs (Hogg et al., 2011; Nishikura, 2010b). ADARs were first discovered in *Xenopus laevis* as double-stranded RNA unwinding proteins (Bass and Weintraub, 1988; Rebagliati and Melton, 1987). In the past years, ADARs have been found in almost

all metazoa from worms to humans. The search for A to –I deaminases in yeast led to the discovery of adenosine deaminases acting on tRNAs (ADATs). ADATs deaminate adenosine located adjacent to the anticodon or in the wobble position, to inosines in tRNAs, thereby indicating an evolutionary relationship to ADARs (Gerber et al., 1998; Gerber and Keller, 1999). ADARs might have evolved from ADATs by gaining dsRBDs since ADATs lack dsRBDs.

Three ADAR family members have been found in mammals, namely, ADAR1, ADAR2 and ADAR3. Many enzymatic substrates have been discovered for ADAR1 and ADAR2 both of which are ubiquitously expressed. However, since potential target substrates have not yet been identified for ADAR3 it is considered inactive although it is expressed in the central nervous system (Bass, 2002). ADARs have a common modular domain organization that includes one to three dsRBDs in their N-terminal region followed by a highly conserved C-terminal catalytic deaminase domain (Figure 1.9).

The expression pattern of human ADAR1 is highly regulated (Wagner et al., 1990). ADAR1 has two isoforms due to the presence of two transcription start sites and two promoters. The two promoters use alternate exons at the 5' end of the mRNA and express two isoforms ADAR1-i and ADAR1-c (George and Samuel, 1999). ADAR1-i is expressed from an interferon-inducible promoter resulting in a 150kDa protein which is predominantly cytoplasmic (Figure 1.9). In contrast to that, ADAR1-c is expressed constitutively from a downstream methionine (Met296) producing a 110kDa nucleo-cytoplasmic shuttling protein which is predominantly observed in the nucleus (Patterson and Samuel, 1995a) (Figure 1.9). The dsRBD3 overlaps an atypical NLS in ADAR1 (Eckmann et al., 2001). The structure and functions of dsRBDs of ADARs are explained later in detail in section 1.5.2 and 1.5.3.

ADAR1-i has two Z-DNA binding domains at the N-terminus, Z- α and Z- β instead of only one, Z- β in ADAR1-c. Only Z- α , but not Z- β can bind to left-handed form of DNA (Herbert et al., 1997). The crystal structure of Z- β shows the presence of an additional helix and lacks many crucial residues important for Z-DNA binding. Z- β domain has been suggested to be involved in protein-protein interaction during dimerization of ADAR1 (Athanasiadis et al., 2005). ADAR1-i has a CRM-1 dependent NES overlapping in the Z- α domain (Figure 1.9) (Strehlow et al., 2002).

Moreover, Z- α has a property to bind to Z-RNA as well (Peck et al., 1982). Hence, localizing ADAR1-i to actively transcribing DNA would allow it to efficiently edit RNA species prior to splicing. Infact, it was reported that A to-I editing by ADAR1-i increases in dsRNA substrates containing purine-pyrimidine repeats for example ones involved in the formation of Z-RNA which is favored by alternate guanosine-cytosine repeats (GC repeats)(Koeris et al., 2005; Peck and Wang, 1985). Hence, Z- α could also be involved in targeting ADAR1-i to Z-forming sequences such as the Z-DNA/Z-RNA hybrids formed during transcription or the ones formed during the trail of transcription in RNA viruses (Athanasiadis, 2012; Koeris et al., 2005).

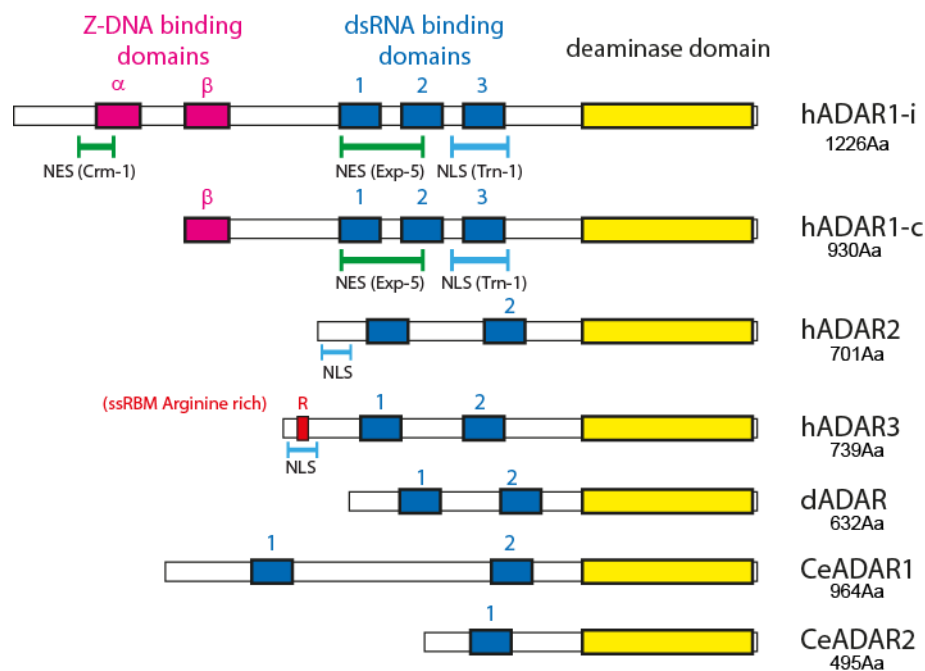


Figure 1.9 Domain organization of ADAR family members and NLS and NES repartition in human ADAR1 isoforms.

The ADAR family members are represented with their domain structure organization. Three ADARs are found in vertebrates (ADAR1, ADAR2 and ADAR3). One ADAR is found in *D. melanogaster* (dADAR) and two in *C. elegans* (CeADAR1-2). ADARs have a conserved C-terminal deaminase domain (*in yellow*) and diverse numbers of dsRNA binding domains (*in blue*). In addition ADAR1 has one or two copies of Z-DNA binding domains (*in pink*). The long isoform of ADAR1 is interferon-inducible (ADAR1-i), whereas the short isoform is constitutively expressed (ADAR1-c). ADAR1-i harbors a strong NES (*in green*) in its unique N-terminal region and is mostly cytoplasmic. ADAR1-c lacks the N-terminal NES and is primarily nuclear. Nuclear import of both protein versions is mediated by an atypical NLS (*in cyan*) that overlaps the third dsRBD present in either ADAR1 isoform. Figure adapted from (Barraud and Allain, 2012; Orlandi et al., 2012)

ADAR1 and ADAR2 are catalytically active when they form homodimers (Cho et al., 2003). Although the site for dimerization for vertebrate ADARs have not been yet reported the N-terminus of *Drosophila melanogaster* ADAR (dADAR) has been shown to be involved in dimerization (Gallo et al., 2003). Similarly the Z- β domain has been thought to be responsible for dimerization of ADAR1 and ADAR2 (Athanasiadis et al., 2005). However, ADAR3 does not dimerize and is suggested to be involved in repression of ADAR1 and ADAR2 activity by probably competing for common dsRNA substrates or by forming inactive heterodimers with them (Chen et al., 2000; Valente and Nishikura, 2007). ADAR3 has an arginine-rich, R-domain responsible for binding to ssRNA substrates (Figure 1.9). Interestingly, recently, an R-domain was also found in a minor splice variant of ADAR2 (Maas and Gommans, 2009). Alternative splicing events lead to several splice variants of ADAR1 and ADAR2 mediating altered enzyme activity (Jin et al., 2007; Rueter et al., 1999). Moreover, ADAR2 itself regulates its own activity by editing its own mRNA and there-by changing a splice site producing a truncated inactive protein (Feng et al., 2006; Palladino et al., 2000).

ADARs are indispensable for normal functioning of all animals tested so far. ADAR1 knockout mice are embryonically lethal. The mice die at prenatal E12 days and show defective hematopoiesis, wide spread apoptosis and liver disintegration (Hartner et al., 2009; Wang et al., 2000; Wang et al., 2004). The precise molecular mechanism leading to the death of the mice are not yet known. The under-lying mechanism for such severe phenotype of the mice has been proposed to be, due to the acceleration of interferon signaling pathway in the absence of ADAR1, leading to apoptosis by interferon stimulated genes (Hartner et al., 2004; Vitali and Scadden, 2010). Moreover, dyschromatosis symmetrica hereditaria (DHS), a disease associated with skin pigmentation is caused due to truncations of ADAR1 in humans. Point mutations leading to truncation of ADAR1 cause this mild heterozygous dominant phenotype with apparently normal physical and mental abilities (Miyamura et al., 2003; Zhang et al., 2004).

Knockout mice of ADAR2 are viable and have severe episodes of epileptic seizures, subsequently dying postnatal within 3 weeks. The seizures occur due to underediting of glutamate receptor subunit B (GluR-B) in the brain. The editing event leads to an amino acid change Q/R limiting the influx of calcium in the cell

(Seeburg et al., 2001; Sommer et al., 1991). The ADAR2 null mice can be rescued by a point mutation that encodes a pre-edited GluR-B mRNA (GluR-B^{R/R}) (Higuchi et al., 2000). Although these mice are viable they still display a range of subtle phenotypes ranging from a decreased acoustic startle response to decreased blood glucose level. The molecular mechanisms underlying these events still remain unknown. Interestingly, overexpression of ADAR2 in mice causes hyperplasia and obesity. Moreover, a catalytically inactive ADAR2 that retains the dsRNA binding ability can reproduce similar phenotypes. This suggests that editing independent phenotypes can be caused due to RNA binding of ADARs.

1.5.1 Substrates of ADAR

ADARs deaminate adenosine in cellular or viral RNA transcripts having either perfect or imperfect regions of dsRNAs (Bass, 2002; Gott and Emeson, 2000; Nishikura, 2010a) . A to -I editing by ADARs can be either be site specific or nonspecific. Nonspecific editing events occurs mostly within perfect dsRNA structures, deaminating up to 50% of all adenosines (hyper-editing) (Polson and Bass, 1994). Introns and 3' UTRs containing repetitive elements like SINEs (short interspersed nucleotide elements found in mouse), Alu elements or LINEs (long interspersed nucleotide elements found in primates) mainly contain hyper-editing sites. Since these elements exist in inverted orientation they can form extensive dsRNA structures (Levanon et al., 2004; Osenberg et al., 2010) . Despite such abundance of hyper-editing in repetitive elements the role of editing in such elements is still unknown. It has been reported that some mRNAs containing hyper-edited 3'UTR show nuclear localization (Zhang and Carmichael, 2001). Interestingly, other groups have shown that edited repetitive elements in the 3'UTR translocate to the cytoplasm onto the translating ribosomes and mostly do not affect the translation of the mRNA (Hundley et al., 2008) .

Site specific A to-I editing occurs in the coding mRNAs, pri-miRNAs, pre-miRNAs and mi-RNAs within imperfect dsRNA structures altering a single or a defined set of adenosines. In mRNAs, imperfect dsRNA structure is formed at the exon-intron junction complex by an intronic or exonic "editing complementary sequence "(ECS) and part of the edited exon (Nishikura, 2010a). Only a few dozen coding mRNA editing substrates have been identified so far, which is way limited

compared to the high frequency of editing events in repetitive elements. The majority of coding mRNA site-specific editing events were identified in the central nervous system, including ion channels and receptors of neurotransmitters. They have been nicely reviewed in (Penn et al., 2013; Pullirsch and Jantsch, 2010). For example, in GluR-B R/G site, genomically encoded AGA (coding for amino acid arginine or R) is altered to IGA (coding for amino acid glycine or G) there by causing a directed amino acid exchange R764G (Seeburg et al., 1998). The GluR-B R/G site is referred to as the R/G stem-loop and is the most widely used substrate for A-to-I editing studies since it forms a stem-loop containing three mismatches (Aruscavage and Bass, 2000). A few more examples are as follows: The serotonin receptor 2C (5-HT₂CR) is edited at five specific sites. Altered editing levels is associated with many human disorders such as schizophrenia, bipolar disorder, depression, anxiety and Prader-Willi syndrome (Iwamoto et al., 2009; Tariq and Jantsch, 2012). Modified receptor trafficking is mediated by editing of the gamma-aminobutyric acid type A (Gabra-3) (Daniel et al., 2011).

miRNAs and their precursors also form imperfect dsRNA stem-loops making them suitable substrates for site-specific A-to-I RNA editing, there by suggesting a relationship between the RNA editing and RNA interference mechanisms (Nishikura, 2006). Processing of pri-miRNA and pre-miRNA can be affected by deaminating adenosines at specific sites thereby regulating miRNA expression (Öhman, 2007). Gene targeting by miRNA is also modulated by editing events occurring at the seed sequence of miRNA leads to extension of the number of genes targeted by miRNAs. For example pri-miRNA 376-a1 can be edited at two sites, each by ADAR1 or ADAR2 thereby regulating miRNA mediated gene targeting (Kawahara et al., 2007).

A specific target recognition sequence is absent for ADARs. Despite that, ADARs are able to edit very efficiently at specific sites of perfect dsRNA structures. This indicates that both recognition of RNA shape and sequence are important for dsRBDs of ADARs to recognize, bind and position themselves on dsRNA substrates for the catalytic domain to mediate the deamination reaction. For ADAR2 dsRBDs have demonstrated that base-specific contacts are made between RNA and ADAR2. The nearest neighbors and the base opposite the editing site also influence the choice of the edited adenosine by ADARs (Stefl et al., 2010). U > A >

C > G is preferred at the 5' of the edited adenosine by both ADAR1 and ADAR2. While ADAR1 prefers a G > C ≈ A > U as the 3' base, ADAR2 has a similar preference for the 3' base (G > C > U ≈ A) (Eggington et al., 2011; Kawahara et al., 2008). Recently, a study identified that one conserved loop in the catalytic domain of ADAR was majorly responsible in specifying the next neighbor preferences demonstrating that not the direct recognition of neighboring bases but differential base flipping is important (Kuttan and Bass, 2012). U or a C is the most often base opposite the edited adenosine (Athanasiadis et al., 2004; Riedmann et al., 2008). For example, in case of a C, upon editing an A-C mismatch is converted into an I-C base pair. Additionally, ADARs also show selectivity towards RNA tertiary structure (bulges or loops) (Tay et al., 2011) and the number of editable adenosines in a dsRNA relative to its length (Bass, 1997).

1.5.2 dsRBDs of ADARs

dsRBDs of ADARs are involved in recognition and binding of dsRNA structure. However, the molecular basis explaining how dsRBDs preferentially choose certain edible sites specifically and most others non-specifically still remains poorly understood. Recent structures of ADAR2-dsRBDs bound to nearly natural coding substrate (R/G stem loop) have given some critical insights into the sequence-specific recognition of ADAR substrates (Steffl et al., 2010). On one hand the structures demonstrate conserved mode of recognition of the A-form RNA helix demonstrating three major point of contacts. (Helix α 1 and β 1- β 2 loop interact with the minor groove of the RNA-helix while the positively charged residues in the N-terminus end of the helix α 2 interact across the major groove phosphodiester backbone) (Gan et al., 2006; Ramos et al., 2000; Ryter and Schultz, 1998a). The structure also reveals two unexpected sequence-specific contacts of both dsRBDs to two consecutive minor grooves of the GluR-B RNA upper and lower stem loops (USL and LSL) at a register length of 10 and 9 base pair for dsRBD1 and dsRBD2 of ADAR2 respectively. The numbers of RNA base-pairs between two sequence-specific contacts are referred to as register length (Figure 1.10). This is similar to the *Aquifex aeolicus* RnasIII dsRBD which recognizes a dsRNA with G-X₁₀-G motif and *Drosophila melanogaster* Staufen, dsRBD3 which prefers a register length of 12 base pairs (Gan et al., 2006; Ramos et al., 2000). dsRBD1 of *Drosophila* ADAR

is less selective than mammalian ADARs since the contact made by $\alpha 1$ is different than those of mammalian ADARs (Keegan et al., 2011). This suggests that binding of certain dsRBDs might be sequence specific.

Perfect dsRNA structures found in cellular mRNAs or miRNAs are hairpin motifs (Woese et al., 1990). A hairpin motif can consist of internal loops formed due to mismatches in the region of dsRNA or apical loops, loop like structures capping the RNA hairpins. Apical loop-specific recognition of dsRNA by dsRBDs is not conserved (Varani, 1995). However, a few known structures of dsRNA-dsRBD reveal that dsRBDs make direct contact with a tetraloop via their helix $\alpha 1$ as in case of dsRBD3 of Staufen, dsRBD of Rnt1p and dsRBD1 of ADAR2 but not in dsRBD2 of ADAR2 (Ramos et al., 2000; Stefl et al., 2010; Wu et al., 2004a). This suggests that dsRNA substrates with apical loops are preferred for site-specific editing by ADARs. Interestingly it was reported that internal loops within an ADAR substrate could disassemble adjacent helices to convert a long, promiscuously deaminated substrate into a series of short, selectivity deaminated substrates (Lehmann and Bass, 1999). Larger internal loops ($\geq$ six nucleotides) acted as helix ends, whereas smaller loops ($\leq$ four nucleotides) did not which is contradicting to the recent structural studies where dsRBDs favor tetra-loops (Lehmann and Bass, 1999; Ramos et al., 2000; Stefl et al., 2010). Thus, the impact of internal loops and mismatches on dsRBD recognition is still needs further investigation.

In vitro studies have revealed that homo-dimerization of ADAR1, ADAR2 and of *Drosophila* dADAR is required for its catalytic activity of deamination (Cho et al., 2003). Although the site for dimerization for vertebrate ADARs have not been yet reported the N-terminus of *Drosophila* ADAR (dADAR) has been shown to be involved in dimerization (Gallo et al., 2003). However, ADAR3 does not dimerize and is suggested to be involved in repression of ADAR1 and ADAR2 activity by probably competing for common dsRNA substrates or by forming inactive heterodimers with them (Chen et al., 2000; Valente and Nishikura, 2007). The structure of ADAR2 dsRBDs with R/G stem-loops favor the dimerization model since both the dsRBDs bind one face of the RNA covering approximately 120 degrees of the space around the RNA helix. However, it is yet unknown how after dimerization the two catalytic domains would be positioned relative to each other and to the dsRBDs or whether dimerization is based on dsRNA substrate binding.

Moreover the interplay between multiple copies of the domain has not yet been addressed in detail for ADARs. They could aid ADARs in recognizing tertiary RNA structures or could increase their binding specificity. Structural data of full length protein in complex with dsRNA substrate is necessary to fully understand the mechanism of editing completely.

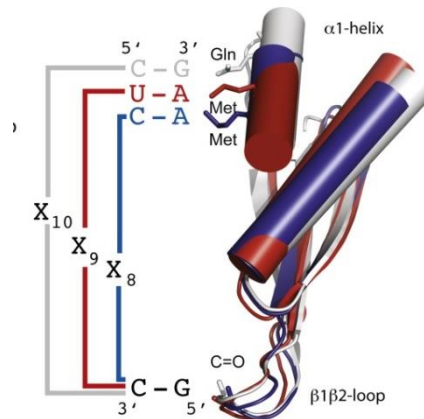


Figure 1.10: Register length recognition b dsRBDs of ADAR2

Schematic representation of the sequence specific contacts of dsRBD1 and dsRBD2 of ADAR2 with the GluR-B R/G stem loop. Overlay of the ADAR2 dsRBD1 (in blue), ADAR2 dsRBD2 (in red), and *Aquifex aeolicus* RNaseIII dsRBD (in gray) structures determines the register length between the two specific contacts on the RNA helix (Stefl et al., 2010).

1.5.3 An RNA regulated shuttling motif

As mentioned earlier, ADAR1 is expressed from two promoters giving rise to a long, interferon inducible version (ADAR1-i) and a constitutively expressed shorter version (ADAR1-c) (Patterson and Samuel, 1995b). ADAR1-i harbors a NES in its unique N-terminal region (Poulsen et al., 2001) and is mostly cytoplasmic, while ADAR1-c lacking the N-terminal NES is primarily nuclear (Figure 1.8) (Desterro et al., 2003b). Both isoforms of ADAR1 shuttle between the nucleus and cytoplasm (Fritz et al., 2009b; Strehblow et al., 2002)

Nuclear import of both ADAR1 versions is mediated by Trn1 via a NLS that overlaps the third dsRBD, and shows no similarity to a PY-NLS (Eckmann et al.,

2001; Fritz et al., 2009b; Poulsen et al., 2001; Strehblow et al., 2002). Importantly, RNA-binding by this dsRBD interferes with Trn1 binding and nuclear import thus constituting an RNA-regulated NLS (Fritz et al., 2009b). Like for ADAR1, interaction with dsRNA has often been described to regulate the trafficking function of dsRBDs. Importantly, the exact elements involved in the NLS-activity of ADAR1-dsRBD3 still remain elusive. Further, how ADAR1-NLS, which differs from PY-NLSs, interacts with Trn1 and how dsRNA binding competes with nuclear import remains unknown.

.

2. Specific Aims

In this thesis the overall aim was to globally characterize the NLS overlapping the dsRBD3 of ADAR1. The third dsRBD of ADAR1 not only binds RNA but is also required for nuclear import. ADAR1 is imported into the nucleus by Trn-1 binding to the dsRBD3. It was also shown that RNA-binding interferes with nuclear import and Trn-1 binding reference (Fritz et al., 2009b). However, the molecular mechanism underlying this special regulation of the dsRBD3-NLS was unknown. In this work I addressed the following points:

1.To decipher the structural features of dsRBD3 of hADAR1

Not all dsRBDs harbor NLS activity. It was therefore important to understand which structural features contribute to the NLS activity of dsRBD3. Using NMR analysis of the dsRBD3-hADAR1 purified domain, together with the help of our collaborators, Federic Allain's group at ETH Zurich, the hallmarks of this domain were identified. Structural signatures make this dsRBD an active NLS and hence an essential domain for nuclear ADAR1 RNA editing. This also led to the understanding how this NLS bearing dsRBD folded differently from other dsRBDs.

2.To identify the minimalistic NLS in dsRBD3 of hADAR1

Based on the NMR data, the interaction of the specific amino acids of dsRBD3 with Trn-1 and dsRNA were tested by in-vitro mutagenesis, import assays and pull down assays for their role in nuclear import. A minimal NLS could be designed by mimicking the changing the residues in non NLS bearing dsRBDs or known Trn-1 substrates.

3.To investigate how RNA binding competes with Transportin-1

RNA binding of dsRBD3 interferes with transportin-1(Trn-1) binding. This was tested by identifying the interacting surface of Trn-1 along with chemical shifts analyses, pull- down assays and molecular modelling.

3. Materials and Method

3.1 Cloning of cDNA for recombinant protein expression

The DNA sequence encoding the third dsRBD of human ADAR1 (residues 708-801) (Uniprot entry P55265), was sub-cloned by PCR amplification from FLAG-His-tagged full-length ADAR1-c plasmid (Fritz et al., 2009b) between *NdeI* and *XhoI* cloning sites in the *E. coli* expression vector pET28a. The constructs contain a N-terminal tag of 6x histidine residues used for protein purification.

The recombinant dsRBDs (wild-type and mutant) were expressed as GST fusion proteins. They were sub-cloned between *SmaI* and *Not I* cloning sites in *E.coli* GST expression vector pGEX-4T. For some GST proteins *BamHI* and *HindIII* restriction sites were used in the GST-pET24d expression vector. For a complete list of used primer pairs see Table 1.3 His-tagged transportin was cloned as described earlier (Fritz et al., 2009b).

3.2 Expression and Purification of recombinant proteins

Recombinant GST proteins were purified from *E. coli* CS41 using glutathione sepharose as described in (Fritz et al., 2009b). His-tagged transportin was purified from *E.coli* Top10 F' cells containing a Rep4 plasmid required for all pQE60 expression vectors. The cells were grown at 37 °C and induced at OD600 ~0.7 by adding isopropyl- β -D-thiogalactopyranoside to a final concentration of 1 mM. Cells were harvested 3 h after induction by centrifugation. Cell pellets were resuspended in Import buffer (110mM Kac, 2mM MgAc, 0,5mM EGTA, 20mM Hepes pH 7.3, 20mM NaCl, 0,1% Tween, 14mM β ME, 1mM PMSF and adjusted to a pH7,9) and lysed by sonication. Supernatant was loaded on a Ni-NTA column on a ÄKTA Prime purification system (Amersham Biosciences), and the protein of interest was eluted with an imidazole gradient in Import buffer. The fractions containing the protein were pooled, dialyzed against the 20mM Hepes pH7.3 5% glycerol, 14mM β ME.

His tagged dsRBDs proteins were overexpressed in BL21 (DE3) Codon-plus (RIL) cells in either LB media or M9 minimal media supplemented with $^{15}\text{NH}_4\text{Cl}$ and ^{13}C -labeled glucose by our collaborators at Federic Allain's lab in ETH Zurich as described below: The cells were grown at 37 °C to OD600 ~0.4, cooled down at 30

°C and induced at OD₆₀₀ ~0.6 by adding isopropyl-β-D-thiogalactopyranoside to a final concentration of 0.5 mM. Cells were harvested 16 h after induction by centrifugation. Cell pellets were resuspended in lysis buffer (Tris-HCl pH 8.0 50 mM, NaCl 1 M, imidazole 20 mM, DTT 1 mM) and lysed by sonication. Supernatant was loaded on a Ni-NTA column on a ÄKTA Prime purification system (Amersham Biosciences), and the protein of interest was eluted with an imidazole gradient. The fractions containing the protein were pooled, dialyzed against the NMR Buffer (NaPi pH 7.0 20 mM, KCl 100 mM, DTT 1 mM), and concentrated to ~0.9 mM using Vivaspin 5000 MWCO (Sartorius Stedim Biotech). All other ADAR1-dsRBD3-His-tagged constructs and mutants were expressed and purified similarly. All wild type and mutant proteins were checked to be properly folded by running (¹⁵N, ¹H)-HSQC spectra.

3.3 In vitro mutagenesis

All wild-type and mutant dsRBD constructs of interest were cloned into pcDNA3 into which an AUG codon was engineered upstream of a unique HindIII site, followed by a pyruvate kinase 6x myc fusion (Eckmann et al., 2001). Fragments were inserted into the HindIII site generating a translational fusion with pyruvate kinase-myc. For a complete list of used primer pairs see Table 1.3.

3.4 Transfection and immunofluorescence localization

PK-myc reporter plasmids were transfected into HeLa cells using Calcium phosphate or polyethyleneimine mediated transfection (Polyplus, France). Cells were fixed and stained 48 h after transfection using a monoclonal anti-myc antibody in combination with a secondary Alexa 488 secondary antibody and counterstained with DAPI as described (Eckmann et al., 2001). Microscopic confocal sections were taken on a Zeiss LSM710. Images were processed using Adobe Photoshop CS7.

3.5 Immunoprecipitation using recombinant proteins

Purified GST-dsRBD constructs and His-Trn1 constructs were incubated with glutathione beads (Biointex) at a concentration of 75 nM in 500 μL binding buffer (HEPES pH 7.9 20 mM, potassium acetate 110 mM, Mg(CH₃COO)₂ 2 mM, EGTA 1 mM, dithiothreitol 1 mM, PMSF 1 mM, NaCl 20 mM, glycerol 5% (v/v), Triton X-100

1% (v/v), 2-mercaptoethanol 14 mM) for 1 h at room temperature. Subsequently, the beads were washed three times in binding buffer and resuspended in SDS loading buffer. The co-precipitate was loaded on a 10% SDS-PAGE, blotted onto nitrocellulose. The blot was cut in half at the 70 kDa band. The upper part of the blot was detected with an anti-His antibody-rabbit (Rockland) 1:1000, detecting the Trn1 band while the lower part of the blot was detected with a GST antibody-goat (Sigma) 1:1000, detecting the dsRBD fusions. The secondary antibody of α rabbit AP 1:1000 and α goat AP 1:1000 were used for detection of the respective proteins.

The influence of RanGTP was to be checked, 400nM RanQ69L (Jena Bioscience) charged with GTP were added per reaction as well. RNA-dependence of the interaction was either investigated by digestion or by addition of dsRNA. In the first case the respective proteins were incubated with 0,1U RNase V1 (Ambion) per μ g protein at 37°C for 20min, whereby doubled-stranded RNA was degraded. Alternatively, dsRNA was added in a concentration of 0.1mg/ml dsRNA (Fritz et al., 2009b).

3.6 RNA samples for Immunoprecipitation

To assess the nucleic acid binding properties of dsRBD constructs, we produced dsRNA substrates by *in vitro* transcription with T7 polymerase, namely a dsRNA duplex of 30 bp as previously described in (Ramos et al., 2000). DNA oligos having an upstream T7 polymerase promoter signal were annealed at 68°C for 10min in the presence of 10mM MgCl₂. Later the annealed ds-oligo is cloned using SmaI and HindIII sites into puC18 plasmid which lacks T7 promoter. Subsequently dsRNA is produced by *in vitro* transcription with T7 polymerase using the MEGAscript® T7 kit (Life Technologies) following the manufacturer's protocol. Concentration of the dsRNA was measured using RiboGreen® RNA Quantitation Reagent by following manufacturer's instructions.

3.7 Import assay

For import assays on permeabilized cells Hela cells were grown on coverslips and treated with digitonin at a final concentration of 30 μ g/ml for 5 minutes on ice. Cells were then incubated with recombinantly expressed GST-tagged protein and His-tagged purified Trn1 protein both at 50 nM, GTP 1 mM, ATP 1 mM, creatine

phosphokinase 20 U/mL, creatine phosphate 5 mM, bovine serum albumin 0.1%, PMSF 1 mM, in import buffer (Adam et al., 1990). After 20 min at room temperature cells were washed in import buffer 3x 5min and immunostained using polyclonal anti GST antibody as mentioned below and imaged on a Zeiss LSM710.

3.8 NMR Spectroscopy and protein structure determination

All NMR experiments were carried out by our collaborators at Federic Allain's group at ETH Zurich. NMR spectra were recorded at 313 K on Bruker AVIII-500 MHz, AVIII-600 MHz, AVIII-700 MHz, AVIII-750 MHz and Avance-900 MHz spectrometers. The data were processed using TOPSPIN 3.0 (Bruker) and analyzed with Sparky (<http://www.cgl.ucsf.edu/home/sparky/>).

3.9 Protein structure determination

Protein structure determination was performed by our collaborators at Federic Allain's group at ETH Zurich. Automated NOE cross-peak assignments (Herrmann et al., 2002) and structure calculations with torsion-angle dynamics (Guntert et al., 1997) were performed using the macro noe assign of the software package CYANA 3.0 (Guntert, 2004).. The 20 best energy structures were analyzed with PROCHECK-NMR (Laskowski et al., 1996) and the iCING web server (Doreleijers et al., 2012) (<http://nmr.cmbi.ru.nl/icing/>). Structures were visualized and figures were prepared with program PYMOL (<http://www.pymol.org>).

3.10 RNA samples and isothermal titration calorimetry

To assess the nucleic acid binding properties of dsRBD constructs, we produced different RNA substrates. These dsRNA substrates were generated by our collaborators at Federic Allain's group at ETH Zurich, by *in vitro* transcription with T7 polymerase, namely a dsRNA duplex of 24 bp as previously described (Barraud et al., 2011). RNAs were purified by anion-exchange high-pressure liquid chromatography under denaturing conditions as previously described (Barraud et al., 2012).

ITC experiments were also performed by our collaborators at Federic Allain's group at ETH Zurich, on a VP-ITC instrument (MicroCal) calibrated according to the manufacturer's instructions.

3.11 Modelling the interaction with Transportin 1

The interaction between ADAR1-dsRBD3 and Trn1 was modeled by our collaborators at Federic Allain's group at ETH Zurich using the six available X-ray structures of Trn1 bound to various NLS sequences as initial structures (PDB codes 2H4M, 2OT8, 2Z5K, 2Z5N, 4FDD and 4JLQ)(Cansizoglu et al., 2007; Imasaki et al., 2007; Lee et al., 2006; Soniat et al., 2013; Zhang and Chook, 2012). The six structures were superimposed on the Trn1 proteins and the various NLS sequences were then aligned according to their relative position in the structures (Figure. 4.16A for the alignment). Then, the sequence of the N- and C-terminal fragments flanking ADAR1-dsRBD3 was manually aligned with the other NLSs to maximize sequence similarities while keeping a virtually extended peptide chain. This alignment together with the structures of Trn1 bound to the various NLSs was used to construct a set of highly ambiguous distance restraints in order to loosely constrain the N- and C-terminal tails to adopt an extended conformation along the path adopted by the NLS peptides in all crystal structures. 100 structural complexes were finally energy minimized and the 10 best energy structures were pooled and analyzed with PYMOL. A representative structure was then selected for figure preparation (Figure.4.16 and 4.17).

3.12 Accession Codes

The chemical shifts of ADAR1-dsRBD3 have been deposited in the BioMagResBank under accession number 19502. The coordinates of the structure have been deposited in the Protein Data Bank under accession code 2MDR.

4. Results

4.1 ADAR1-dsRBD3 with extended borders served ideal for NMR

The sequence representing the consensus of dsRBD3 spans Asn726-Gly794 in ADAR1. We have previously mapped the NLS in human ADAR1 to a region spanning Met690-Arg801 (Ryter and Schultz, 1998a). This region contains the third dsRBD of the protein and an N-terminal extension. Consequently GST tagged recombinant dsRBD3 (Met690-Arg801) of ADAR1 containing an internal -Tev protease cleavage site was purified using affinity chromatography. The dsRBD-NLS was released from the GST fusion part of the protein via an internal TEV-protease cleavage site. The purified domain was soluble in solution. However, purification of the protein was tedious and yield of the protein was very low. The NMR structural studies on the protein (performed by our collaborators at ETH Zurich) showed that the protein is well folded but there is a flexible region of around 20 amino acids at the N-terminal end of dsRBD3 that could not be resolved (Figure 4. 1).

Furthermore accurate examinations allowed us to map the minimal NLS to a region spanning Asn704-Arg801. However, when I tried to purify this protein, it was insoluble in solution. Additionally, protein sequence inspection through hydrophobic cluster analysis (performed by our collaborators) predicted the presence of potential secondary-structure elements outside the dsRBD (Callebaut et al., 1997) (Figure 4.2). This led us to extending the N-terminal of the domain Met704-Arg801. I purified several protein constructs with different cloning boundaries most of which were insoluble (data not shown). However, I was finally able to purify in solution an extended dsRBD with region of Met708-Arg801 which served to be an NLS (Figure 4.3). The solution structure of ADAR1 Met708-Arg801 that embeds dsRBD3 was therefore determined by our collaborators at ETH in Federic Allain's group (Figure 4.4).

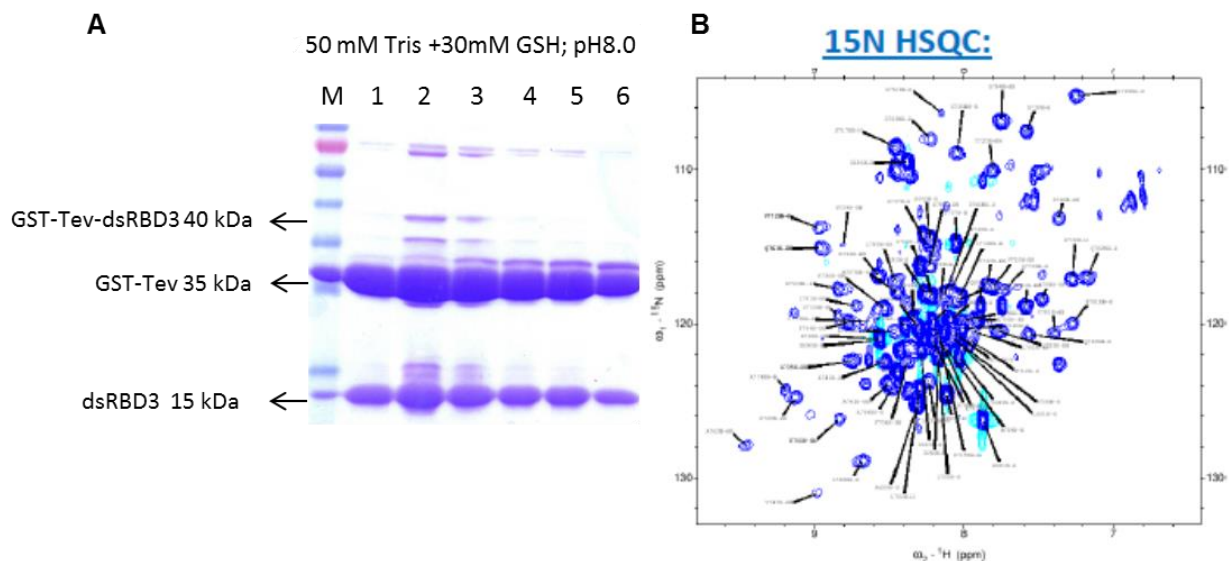


Figure 4.1 Extended dsRBD3 of ADAR1(Met690- Arg801)

(A) GST tagged dsRBD3 of ADAR1 was purified in soluble form. It had an internal TEV-cleavage site which could be easily cleaved using TEV protease at room temperature (methods section 3.2) leaving behind the GST-TEV tag of around 35kDa and the soluble dsRBD3 of 15kDa. However, downstream processing and recovery of the dsRBD was difficult and the yield was very poor. So a His tagged version of the same protein was purified which was later analyzed for NMR spectroscopy. **(B)** 2D HSQC generated by our collaborators at ETH Zurich (Frederic Allain's group) showed that the protein is well folded but there is a flexible region of around 20 amino acids at the N-terminal end of dsRBD3.

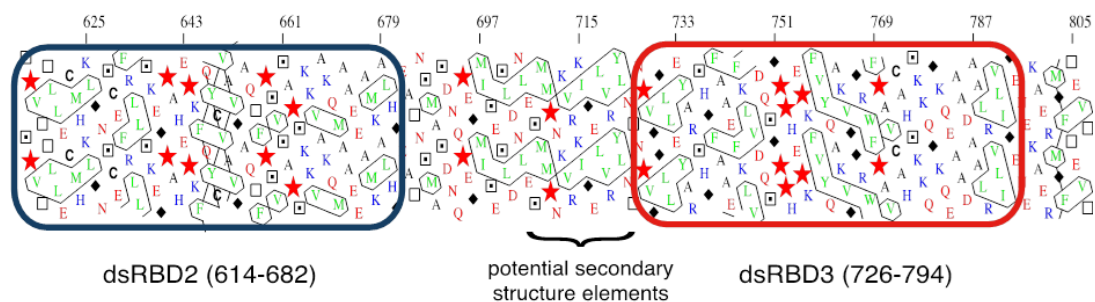


Figure 4.2 Residues preceding the canonical dsRBD fold are structured.

Hydrophobic cluster analysis (HCA) plot of the ADAR1 dsRBD2-dsRBD3 region. The HCA plot was created by our collaborators at ETH Zurich (Frederic Allain's group) with the online tool drawhca (<http://bioserv.impmc.upmc.fr/hca-form.html> – by Luc Canard). HCA uses a highly degenerate code for the sequences. Hydrophobic residues VLIF (strong); WMY (medium) are represented in *green* while AC (mimetic) represented in *black*. Hydrophilic residues DENQ are represented in *red* and HKR represented in *blue*. Proline is considered a chain breaker and is represented by ★. For detailed description of the special symbols refer (Gaboriaud et al., 1987; Lemesle-Varloot et al., 1990). Potential secondary structure elements can be inferred from the hydrophobic clusters present in the segment in between dsRBD2 and dsRBD3.

4.2 Nuclear localization of ADAR1 requires the region N-terminal to dsRBD3

The NLS in human ADAR1 covers a region spanning Met708-Arg801. This region contains the third dsRBD of the protein with 18 and 7 amino-acids long extensions at the N- and C-terminal ends, respectively (Figure 4.3 A). The region is necessary and sufficient for nuclear import of ADAR1 and of unrelated reporter constructs (Figure 4.3 B) (Eckmann et al., 2001). Deletion of the dsRBD3 alone Met708-Arg801 abolished NLS activity (Figure 4.3 B; also shown in Fritz, Strehblow et al. 2009). In order to verify that the extended N-terminal region was imparting NLS activity to this domain, I tested whether a construct encompassing only the canonical dsRBD3, but lacking the upstream stretch of amino-acids (Thr725-Arg801), would retain NLS activity. This shorter construct failed to show NLS activity (Figure 4.3 B). This resembles the extended dsRBD domain that has been observed in *S. pombe* Dicer dsRBD, where the C-terminal extension is essential to regulate the sub-cellular localization of the protein (Barraud et al., 2011). To strengthen the fact that the extended N-terminal region aided in the NLS activity and not misfolding of the protein, the domain Thr725-Arg801 was expressed, purified and subjected to NMR structural studies (Barraud P. Banerjee S., 2014).

4.3 ADAR1-dsRBD3 is an extended domain

The solution structure of human ADAR1-dsRBD3 Met708-Arg801 was determined by our collaborators at ETH Zurich (Frederic Allain's group). NMR data showed that ADAR1-dsRBD3 is extended by an N-terminal α -helix (thereafter called helix α_N) and two flexible basic patches at either end of the domain (there after called N-terminal module or NTM and C-terminal module or CTM). Apart from the additional features namely helix α_N , NTM and CTM the rest of the domain adopts a canonical dsRBD structure, with $-\alpha\beta\beta\beta\alpha-$ topology similar to Xlrbpa-dsRBD2 (Ryter and Schultz, 1998a). Pro727-Glu795 forms the dsRBD core domain maintaining the canonical features of a dsRBD (Bycroft et al., 1995; Kharrat et al., 1995). The helix α_N (Ile716-Asn726) is shorter in length than helix α_1 or α_2 . It is well structured and folds back remarkably to bring the NTM and CTM of the domain closer to each other and on the same surface of the protein (Figure 4.4 D-E).

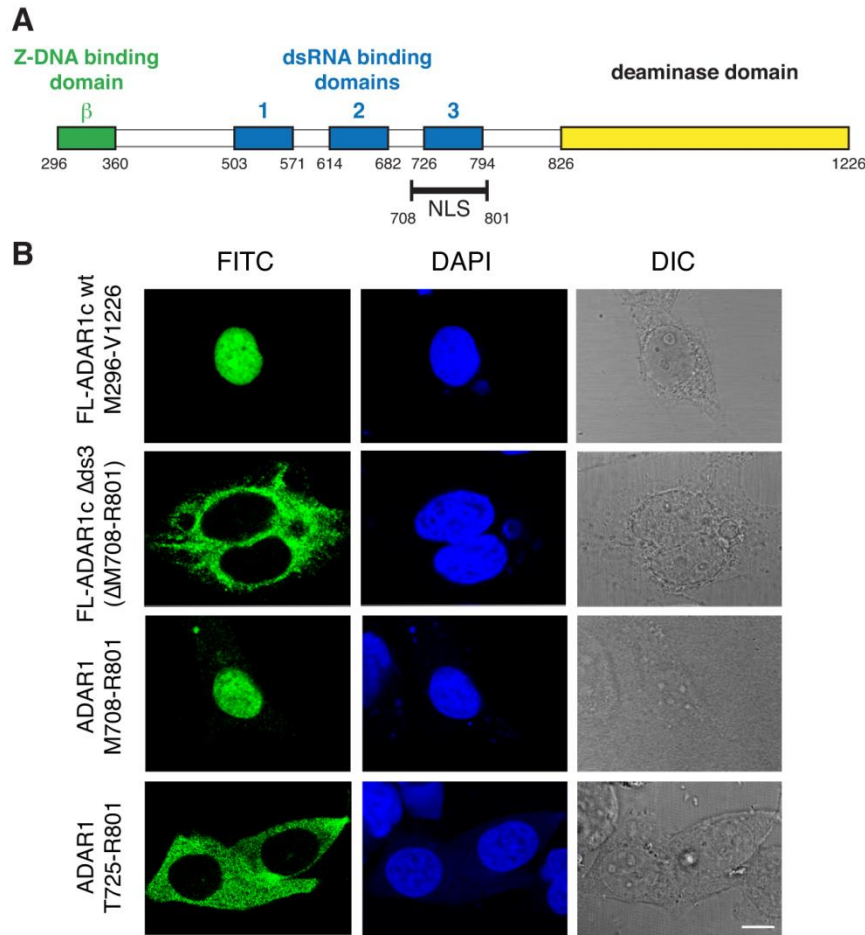


Figure 4.3 NLS activity is imparted by an N-terminal extension upstream dsRBD3

(A) Domain organization of human ADAR1c (Uniprot P55265 entry). The N-terminal Z-DNA binding domain is *in green*, the three dsRBDs are *in blue* and the C-terminal deaminase domain is *in yellow*. Annotated domain boundaries are reported below each domain. The minimal NLS fragment, as previously identified (23), is indicated (Met708-Arg801). **(B)** The constitutively expressed, full length ADAR1c spanning M296 to V1226 is localized to the nucleus. Removal of the third dsRBD Δ ds3 leads to cytoplasmic accumulation of myc-tagged ADAR1c. The minimal NLS spanning Met708 to Arg801 is sufficient to mediate nuclear transport of a myc-tagged pyruvate kinase (PK) reporter construct. A fragment starting at the beginning of the canonical dsRBD (Thr725-Arg801) lacks NLS activity. FITC channel shows transfected constructs in confocal sections. Nuclear DNA is stained with DAPI and cellular outlines are shown in the DIC channel. Scale bar = 10 μ m.

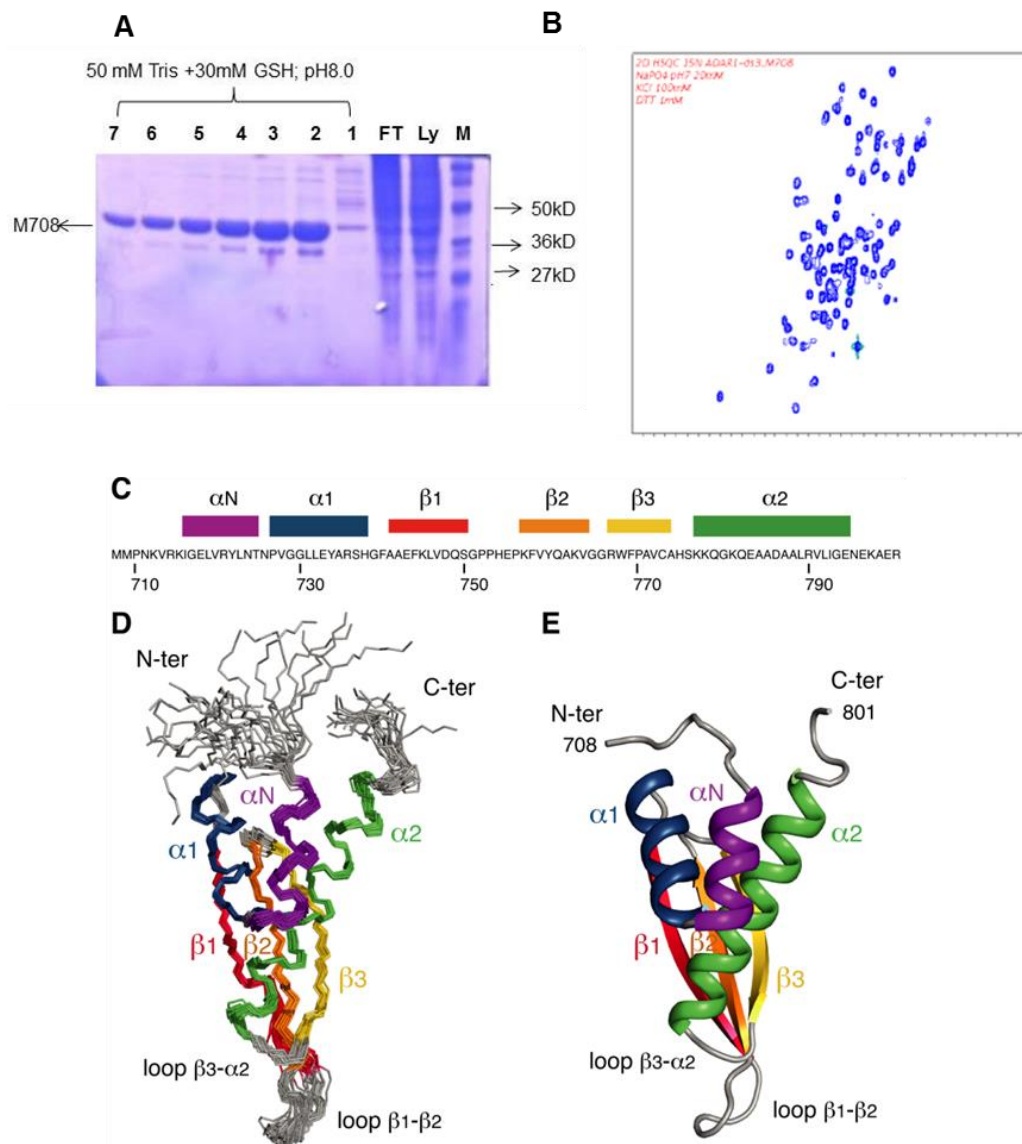


Figure 4.4 ADAR1-dsRBD3 (Met708-Arg801) is an extended domain

A) GST tagged dsRBD3 (Met708-Arg801) of ADAR1 was purified in soluble form. **(B)** 2D HSQC showed that the protein is well folded. The flexible region is removed and the HSQC spectrum was good so our collaborators at ETH Zurich proceeded with backbone and side chains assignments. **(C)** Sequence of ADAR1-dsRBD3 (Met708-Arg801) with the corresponding secondary structure elements indicated (color code is the same as in **D** and **E**). Amino-acid numbers refer to the full-length human ADAR1 protein (Uniprot P55265). **(D)** NMR ensemble. Overlay of the 20 final structures with color coded secondary structure elements: α1 in blue, β-strands 1–3 in red, orange and yellow, α2 in green and the additional N-terminal helix αN in purple. **(E)** Cartoon representation of the lowest energy structure. The same color code is used for secondary structure elements. Figures B, D and E were generated by our collaborators at ETH Zurich (Frederic Allain's group)(Barraud P. Banerjee S., 2014)

4.4 Extended features of ADAR1-dsRBD3 embody the nuclear localization elements

In order to seek for the molecular determinants responsible for the NLS activity in ADAR1, we focused our attention towards the unstructured peptide segments flanking the structured dsRBD3, *i.e.* Met708-Lys715 and Glu797-Arg801 in the N- and C-terminal parts, respectively (Figure 4.4 C and E). To test whether these flexible fragments are important for NLS activity, we mutated residues from the basic or hydrophobic portion of the N-terminal stretch into alanine (K712A, R714A, K715A and M708A, M709A, P710A, V713A for the basic and hydrophobic portions, respectively). Similarly, residues from the C-terminal stretch were mutated as well (K798A, E800A, R801A). Mutations were either introduced individually or in combinations. Their impact was evaluated *in vitro* by transfection of reporter constructs complemented by nuclear import assays on permeabilized HeLa (Figure 4.5 A, B, Fig 4.6, A, B also summarized in Table 1.2. Importantly, mutations in both the N- and C-terminal fragments strongly affect the NLS activity of the construct.

The most prominent effect was observed when changing Arg801 to alanine, which completely abolished NLS activity, both in the context of a pyruvate kinase reporter construct and in full length ADAR1c (Figure 4.7 A). Similarly, Trn1 was unable to transport recombinantly expressed R801A ADAR1-NLS to the nucleus in import assays on permeabilized cells (Figure 4.5B). Pull-down assays with R801A also support that direct interaction with Trn1 is lost in this mutant (Figure 4.7). In the N-terminal fragment, none of the single amino-acid substitutions showed any significant loss of nuclear accumulation (Figure 4.6 A, B). While a triple mutation of the basic amino acids K712A/R714A/K715A did not affect nuclear localization, the double mutation M708A/P710A showed both nuclear and cytoplasmic signals, suggesting that these residues are contributing to Trn1 interaction (Figure 4.5A). Furthermore, deletion of residues Met708-Lys715 clearly abolished nuclear accumulation (Figure 4.5A) and the deletion mutant (Δ M708-K715) failed to interact with Trn1. This was supported by both, *in vitro* import assays on permeabilized cells and pull down assays with Trn1 (Figure 4.5B, 4.7).

Having shown that both the N- and C-terminal fragments flanking the structured dsRBD3 are essential for NLS activity, we postulated that the extended dsRBD might act as a scaffolding element bringing closely together two distant stretches of

residues to form a functional bimodular NLS (Figure 4.9A). To test this hypothesis, I first checked whether the relative position of the N- and C-terminal fragments is important for NLS activity. Subsequently, I monitored the localization properties of a construct lacking helix α_N (deletion of residues Ile716-Asn724; Figure 4.9B). In this construct, the two modules are no longer on the same side of the domain.

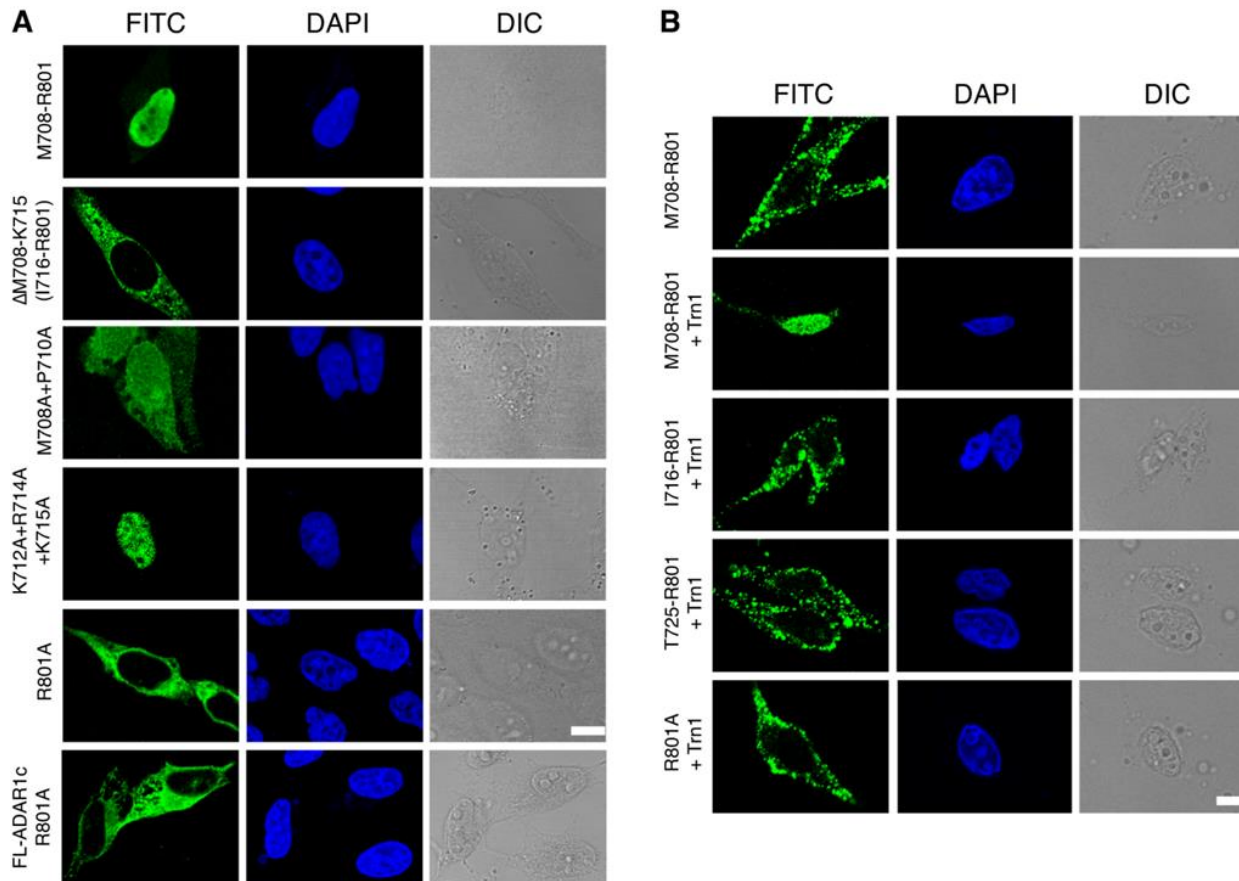


Figure 4.5 Elements flanking dsBD3 are important for NLS activity

(A) Confocal sections of indicated fragments transfected as myc-tagged PK fusions. Met708-Arg801 comprises the minimal NLS. Deletion of the N-terminal fragment Δ Met708-Lys715 abolishes NLS activity. Simultaneous mutation of M708A and P710A in this region strongly reduces NLS activity, while the basic amino acids are not important for NLS activity (K712A+R714A+K715A). The C-terminal Arg801 is essential for nuclear import, both in the context of the PK fusion construct and in full length ADAR1c. (B) Confocal sections of nuclear import assays on permeabilized cells using GST fusion constructs. Met708-Arg801 is imported into the nucleus in a Trn1 dependent fashion. A N-terminal deletion Δ Met708-Lys715 abolishes nuclear import. Also, the canonical dsRBD fold alone (Thr725-Arg801) fails to accumulate in the nucleus. The C-terminal R801A mutation abolishes nuclear import. Fusion proteins are visualized in the FITC channel. DAPI shows nuclear DNA; cellular outlines are showing the DIC channel. Scale bars = 10 μ m.

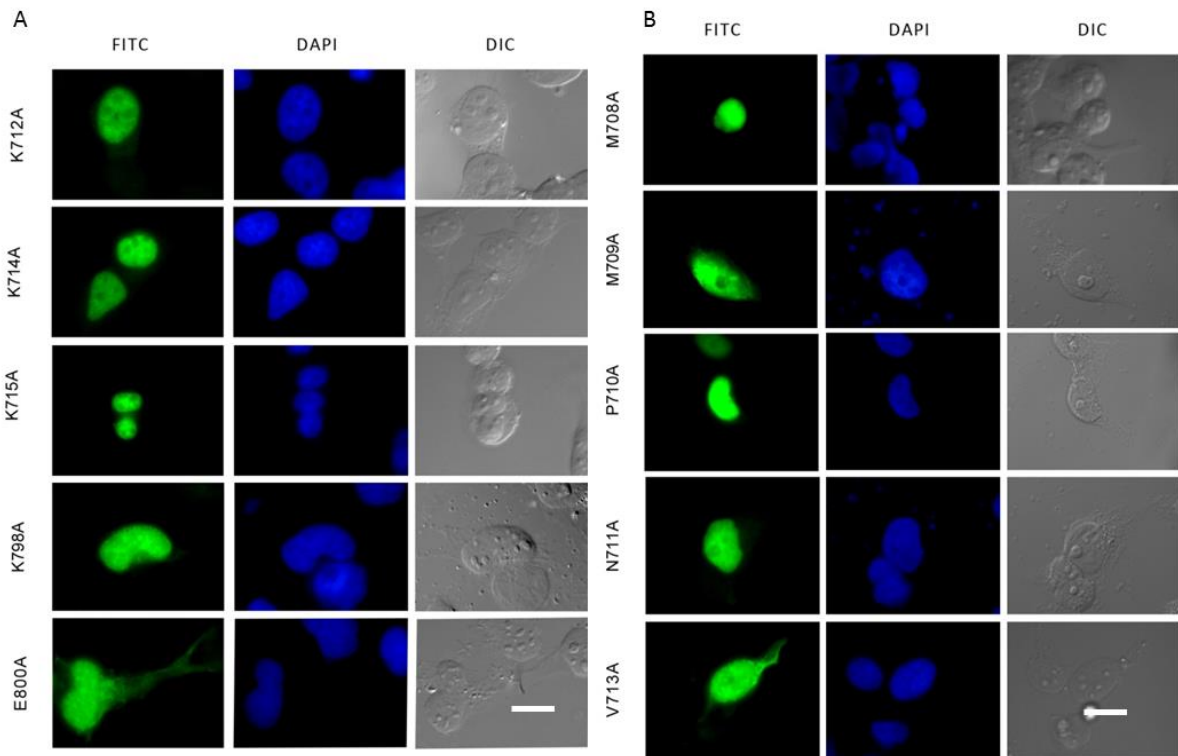


Figure 4.6 Basic elements flanking dsBD3 (except Arg801) are not important for NLS activity

A-B: Inverted Fluorescent microscopic images of indicated fragments transfected as myc-tagged PK fusions. **(A)** All basic amino acids except Arg801 in the extended unstructured elements are not important for NLS activity. Mutating each of these individually lysine residues (K712A; R714A; K715A; K798A) or the glutamic acid (E800A) individually to alanine do not affect the nuclear localization of the domain. **(B)** Mutation of the hydrophobic amino acids to alanine individually (M708A, M709A, P710A, and V713A) also does not abolish NLS activity. Fusion proteins are visualized in the FITC channel. DAPI shows nuclear DNA; cellular outlines are shown in the DIC channel. Scale bars = 20µm.

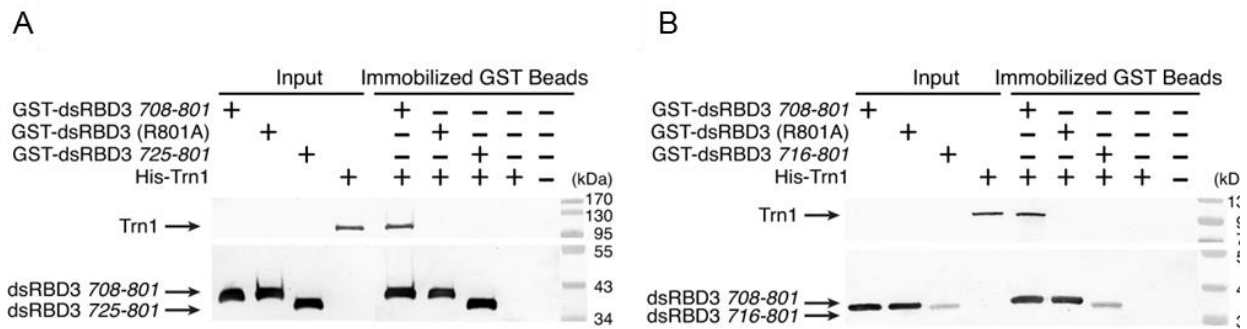


Figure 4.7 Elements flanking dsRBD3 interact with Trn-1

GST-fusion constructs of dsRBD constructs shown in panels **A** and **B** were tested for their ability to interact with Trn1 in pull down assays. Input lanes show the proteins used for the interaction studies. GST coupled wild-type construct (Met708-Arg801) can precipitate full length Trn1. Deletions of the dsRBD N-terminal fragment (716-801), the N-terminal fragment and helix α_N (725-801) or the R801A mutation abolish interaction with Trn1.

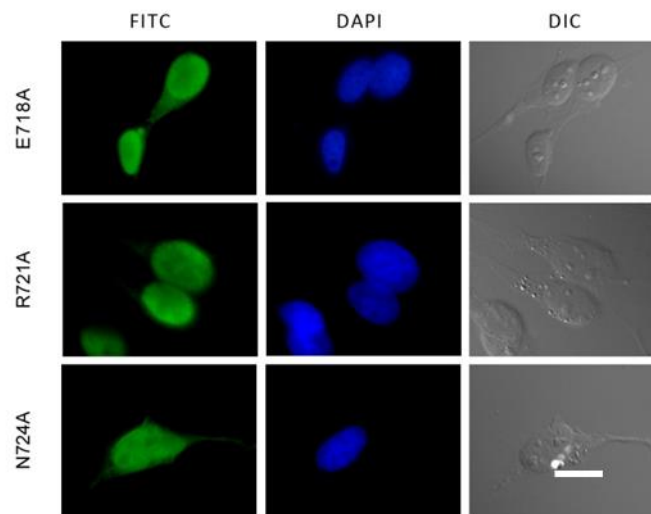


Figure 4.8 Exposed residues of helix α_N do not contribute to the NLS activity

Inverted Fluorescent microscopic images of indicated fragments transfected as myc-tagged PK fusions. The solvent exposed residues of the helix α_N individually when mutated to alanine do not abolish NLS activity. Fusion proteins are visualized in the FITC channel. DAPI shows nuclear DNA; cellular outlines are shown in the DIC channel. Scale bars = 20 μ m.

. Similarly, a triple mutant that disrupts helix α_N structure without deleting its sequence was generated. We mutated the three buried hydrophobics into polar residues, making contacts with α_1 and α_2 in order to impair the tight packing of α_N onto the rest of the domain (*i.e.* I716N/L719S/L723N). In both mutant proteins, the resulting constructs accumulated in the cytoplasm (Figure 4.9 B, C). However, a direct involvement of helix α_N in Trn1 binding and nuclear import cannot be truly excluded. To evaluate whether helix α_N could be directly involved in NLS activity, we mutated the solvent exposed residues of helix α_N into alanines (E718A, R721A and N724A) individually or in combination. Both the individual and triple mutated constructs fully retained NLS activity (Figure 4.8 and 4.9 also summarized in Table 1.2). These data show that the N- and C-terminal modules flanking ADAR1-dsRBD3 are essential elements of the NLS and their relative spatial positioning is critical for import. The extended dsRBD would act as a scaffolding element essential to position the two modules for a productive interaction with Trn1.

4.5 dsRBD3 acts as a scaffolding domain for the bimodular NLS

If the structured part of ADAR1-dsRBD3 is dispensable for NLS activity and only acts as a scaffolding domain, one could, in principle, replace ADAR1-dsRBD3 with an unrelated dsRBD lacking NLS activity while keeping the N- and C-terminal modules that form the NLS. Yet, helix α_N should be kept in order to appropriately position the modules of the bimodular NLS with respect to each other. I therefore constructed a chimeric protein containing a dsRBD of canonical structure that shows no NLS activity (*i.e.* Xlrbpa-dsRBD2; Figure 4.10 A, B), flanked by the N- and C-terminal fragments of ADAR1-dsRBD3, including helix α_N . In addition, I mutated three residues in α_1 and α_2 of Xlrbpa-dsRBD2 (*i.e.* S116G, T176R and K179I; Figure 4.10B) to allow the interaction of α_N with α_1 and α_2 . While Xlrbpa-dsRBD2 wild-type constructs show no NLS activity (4.10 A), the chimeric construct constitutes a fully functional NLS (Figure 4.11B) demonstrating that the sequence of the ADAR1-dsRBD3 is not important in itself but the overall structure is important for orienting the NTM & CTM for NLS activity.

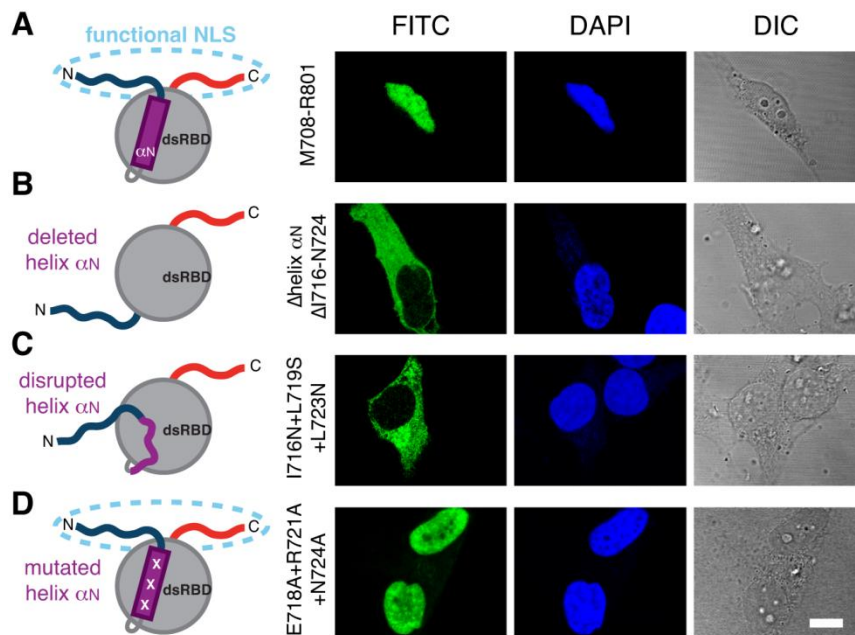


Figure 4.9 The relative orientation of the N- and C-terminal fragments flanking the dsRBD3 is essential in building a functional NLS.

Schematic representations (*left panels*) of ADAR1-dsRBD3 constructs tested for NLS activity. The dsRBD core domain is *in grey*; the N- and C-terminal fragments are *in blue* and *in red*, respectively, and the additional helix α N is *in purple*. Localization of the corresponding constructs (*right panels*). **(A)** In wild-type ADAR1-dsRBD3 (Met708-Arg801) a functional NLS is formed by the combination of the N- and C-terminal fragments. This functional NLS is highlighted *in cyan*. Consistently, a PK fusion harboring Met708-Arg801 accumulates in the nucleus as shown in the FITC channel. **(B)** Construct with a deleted helix α N (Met708-Lys715 fused to Thr725-Arg801) position the critical regions for Trn1 interaction on opposite phases of the domain and consistently lead to cytoplasmic accumulation of a reporter construct (Δ I716-N724).

(C) Construct with a disrupted helix α N carrying a triple mutation in the buried residues of helix α N (I716N, L719S and L723N) also fails to accumulate in the nucleus.

(D) Construct with helix α N carrying a triple mutation in the exposed residues of helix α N (E718A, R721A and N724A) can still properly position the N- and C terminal fragments and consistently accumulates in the nucleus. Transfected constructs are detected in confocal sections in the FITC channel. DNA is visualized by DAPI and cellular outlines are shown in the DIC channel. Scale bar = 10 μ m. See also Figure 4.6 and 4.10

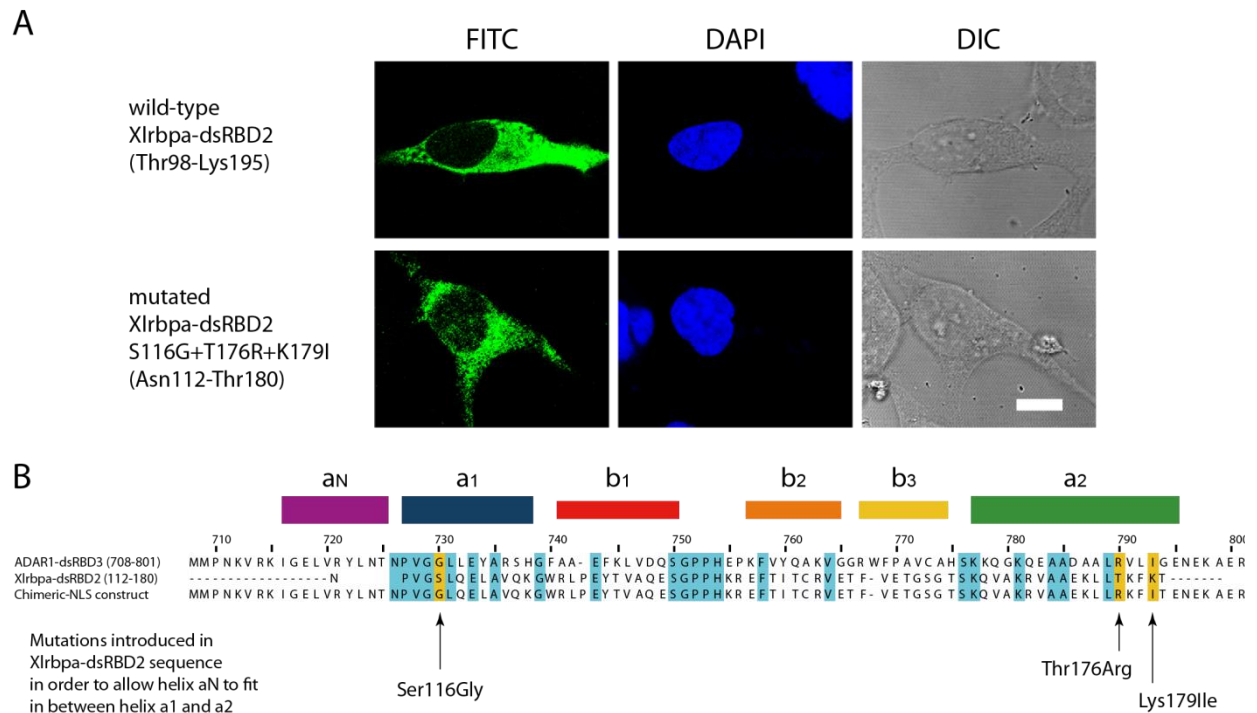


Figure 4.10 xIRBPA-dsRBD2 lacks the NLS activity

(A) The second dsRBD of XlrbpA Thr98-Lys195 does not confer NLS activity to a pyruvate kinase fusion construct (top row). Mutations in this XlrbpA construct (S116G, T176R, and K179I) do not alter the cytoplasmic localization. FITC shows localization of the fusion protein detected by Alexa488 antibody staining. DAPI shows nuclear DNA and DIC shows cellular outlines. Scale bar =10 μ m. (B) Sequence alignment of ADAR1-dsRBD3 (Met708-Arg801), Xlrbpa-dsRBD2 wild-type (Asn112-Thr180), and the chimeric construct resulting from the combination of ADAR1-dsRBD3 N- and C-terminal fragments flanking the Xlrbpa-dsRBD2 domain. Residue numbering and secondary structure elements of ADAR1-dsRBD3 are shown above the alignment. Identical residues between ADAR1-dsRBD3 and Xlrbpa-dsRBD2 are highlighted *in blue*. The three mutations introduced in Xlrbpa-dsRBD2 in order to adapt the surface of Xlrbpa-dsRBD2 for enabling the tight interaction of helix α_N in between helices α_1 and α_2 (*i.e.* S116G, T176R and K179I) are shown *in yellow*.

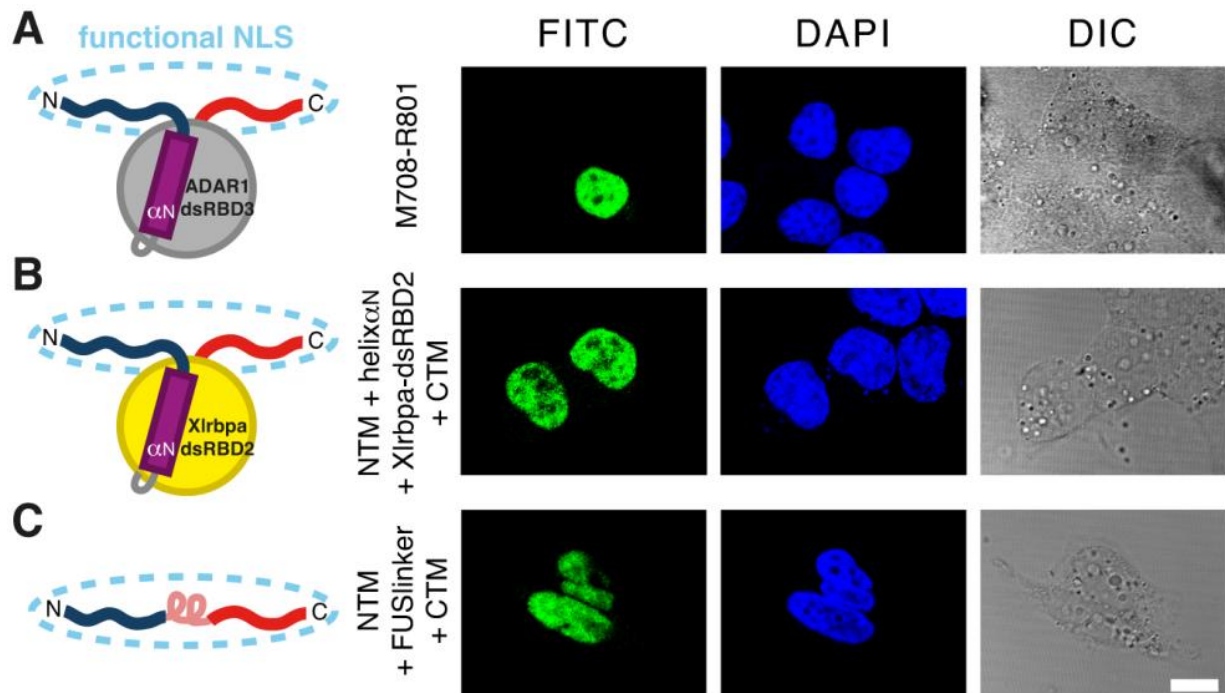


Figure 4.11 Structured domain of ADAR1-dsRBD3 only acts as a scaffold and is indispensable for NLS activity

Schematic representations of constructs tested (*left panels*). The additional helix α N is represented *in purple*. Localization of the corresponding fusion constructs upon transfection visualized in confocal sections (*right panels*). **(A)** Wild-type ADAR1-dsRBD3 construct (Met708-Arg801) represented *in grey* can transfer a PK fusion construct to the nucleus *via* the bimodular NLS. The functional NLS is highlighted *in cyan*. **(B)** Chimeric construct between ADAR1-dsRBD3 N- and C-terminal modules (NTM and CTM, the N- and C terminal fragments are *in blue* and *in red*, respectively) and Xlrpba-dsRBD2 core domain, represented *in grey* and *in yellow*, respectively can also mediate nuclear transport of a fusion protein. **(C)** Minimal NLS construct formed by adding a 7-residue linker of the FUS/ALS NLS (FUS linker is shown *in pink*) to join the N- and C-terminal modules of ADAR1-dsRBD3 (NTM and CTM, respectively). The exact sequence of the minimal peptide is **MMPNKVR** *GEHRQDR* **KAER** with the linker shown in italic and the N- and C-terminal modules from ADAR1-dsRBD3 in bold. This minimal NLS-construct is also able to transfer a PK fusion construct to the nucleus. Localization of the fusion construct is shown in the FITC channel, DNA is shown in the DAPI channel, and cellular outlines are shown in the DIC channel. Scale bar = 10 μ m. See also Figure 4.15A-D.

4.6 Designing a minimal putative NLS

Even though I could show that the NTM and CTM together with helix α_N are essential for NLS activity (Figure 4.11A, B), the actual molecular determinants controlling nuclear import of human ADAR1 *via* the interaction with Trn1 remained elusive. Moreover the chimeric construct of ADAR1-dsRBD3 and Xlrbpa-dsRBD2 (NTM+ α_N +Xlrbpa-dsRBD2+CTM) cannot exclude that the sequence of helix α_N plays a role (Figure 4.11A, B). To definitely prove that the entire ADAR1-dsRBD3, including helix α_N , is superfluous for the formation of a functional NLS, I wanted to design a minimal peptide substituting the dsRBD. Therefore, I made a fusion construct fusing bimodular NLS into a single peptide (NTM+ CTM). However this small peptide of 12 amino-acids accumulated in the cytoplasm indicating that the relative orientation of these two modules imparts the NLS activity (Figure 4.12A). So I needed to construct a peptide, where the N- and C-terminal modules are appropriately spaced to mimic their relative position in ADAR1-dsRBD3. I therefore used small linkers of alanines or glycines in multiple copies in between the NTM and CTM (Figure 4.12A-D). Moreover, as mentioned above, well-characterized Trn1 substrates consist of stretches of 15-30 residues with intrinsic structural disorder and poor apparent sequence conservation adopted a similar extended conformation upon Trn1 binding (Cansizoglu et al., 2007; Imasaki et al., 2007; Lee et al., 2006; Soniat et al., 2013; Zhang and Chook, 2012). But all these fusion constructs showed clear cytoplasmic localization (Figure 4.12).

This provided sufficient motivation for a more rigid linker in between these two modules. To achieve this challenging design, together with the help of our collaborators we used our structural model (Figure 4.14), and in particular the sequence and structural alignment of ADAR1-dsRBD3 with all the different NLSs (Figure 4.14A, B). Among all NLS peptides, we drew a particular attention to the FUS/ALS NLS, since its central region adopts an α -helical structure forming approximately two helical turns (Gly515-Arg521; PDB code 4FDD) (Zhang and Chook, 2012). The two helical turns could create a spacer that maintains the N- and C-terminal modules at the same distance as in the original dsRBD context (Figure 4.13B).

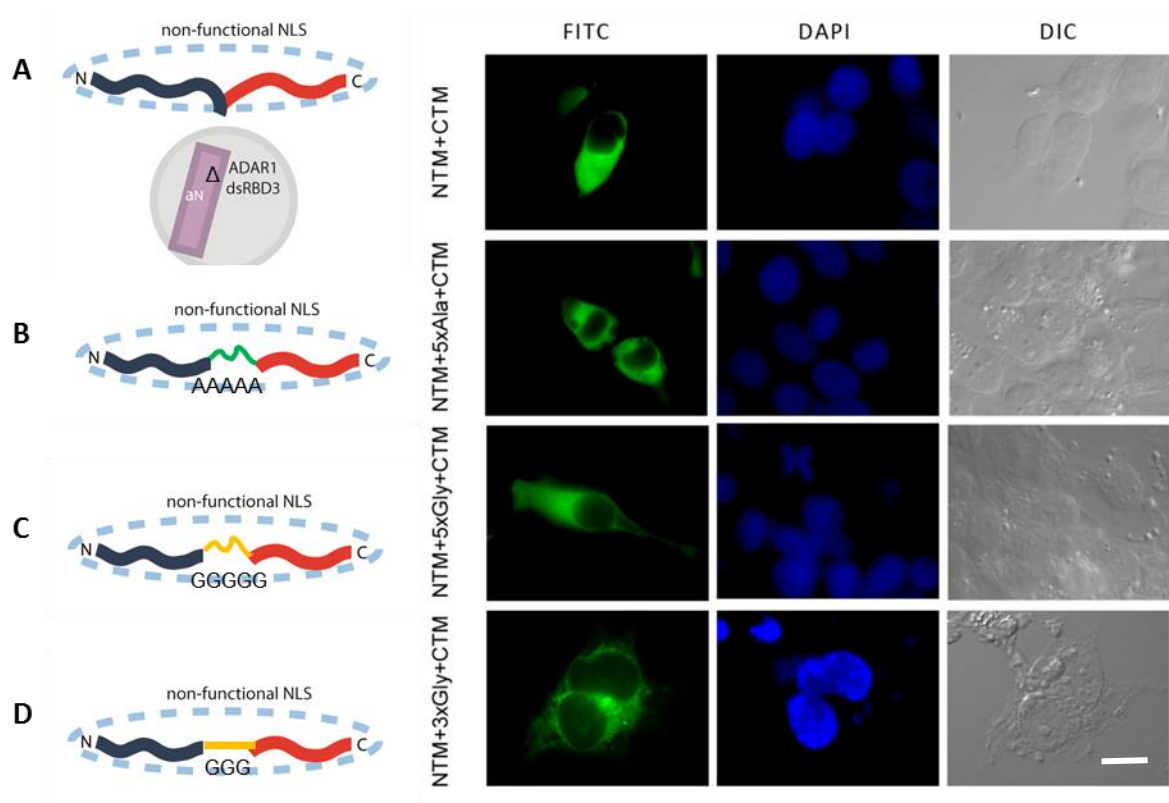


Figure 4.12 Rigid linker is required for forming an active NLS with NTM and CTM of ADAR1-dsRBD3

Schematic representations of constructs tested (*left panels*). The structured ADAR1-dsRBD3 is represented *in light shadow* since it has been deleted in all the constructs. Localization of the corresponding fusion constructs upon transfection visualized in inverted fluorescent microscope (*right panels*). (A) Minimal NLS construct formed by fusing the NTM and CTM (the N- and C terminal modules are *in blue* and *in red*, respectively), there by deleting the scaffolding wild-type ADAR1-dsRBD3 construct (Ile715-Lys798) cannot transfer the PK fusion construct to the nucleus. (B) Chimeric construct of ADAR1-dsRBD3 N- and C-terminal modules (NTM and CTM, the N- and C terminal fragments are *in blue* and *in red*, respectively), using a linker of five Alanine residues (shown in *green*) or (C) five glycine residues (shown in *yellow*) or (D) three glycine residues (shown in *yellow*) respectively cannot also mediate nuclear transport of the fusion protein. Localization of the fusion construct is shown in the FITC channel, DNA is shown in the DAPI channel, and cellular outlines are shown in the DIC channel. Scale bar = 20 μm.

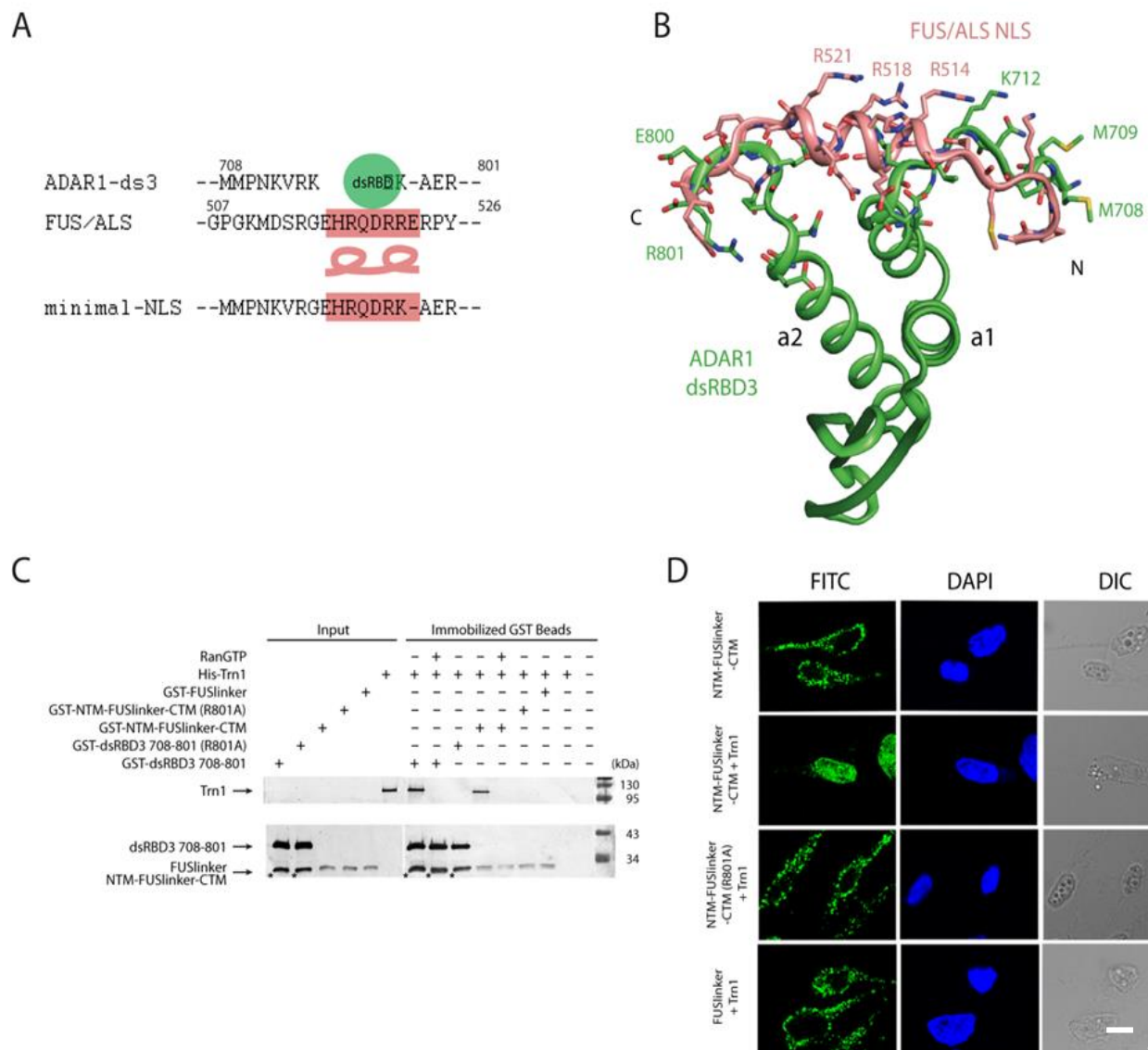


Figure 4.13 Design of a minimal NLS , the FUS-NLS

(A) Sequence alignment of ADAR1-dsRBD3 N- and C-terminal fragments (Met708-Arg801) and FUS/ALS PY-NLS. The sequence of the minimal NLS peptide designed with the N- and C-terminal modules of ADAR1-dsRBD3 linked by the two helical turns of the FUS/ALS PY-NLS is shown at the bottom. (B) Superposition of the structural model of the interaction between ADAR1-dsRBD3 and Trn1 (*in green*) and the FUS/ALS PY-NLS structure (*in pink*) (for clarity Trn1 proteins are not shown; refer to Figure 4.16 compiled by our collaborators at ETH Zurich (Frederic Allain's group)). (C) The N- and C-terminal modules of ADAR1-dsRBD3 (NTM and CTM for N-terminal module and C-terminal module, respectively) linked by the FUS/ALS linker (FUSlinker) can interact with Trn1 in a RanGTP sensitive manner. *Input* lanes show the material added to pull-down assays. *Immobilized GST beads* show the result of pull-downs of GST on glutathione sepharose. The FUS/ALS linker alone, fused to GST fails to interact with Trn1. Similarly, R801A mutations at the C-terminus of either ADAR1-dsRBD3 or the N- and C-terminal modules of dsRBD3 linked by the FUS/ALS linker (NTM-FUS linker-CTM) are sufficient to abolish the interaction with

Trn1. Asterisks marks breakdown products of GST-dsRBD3 708-801. **(D)** The NTM and CTM linked by the FUS/ALS linker (FUSlinker) are translocated to the nucleus in a Trn1-dependent manner in permeabilized cells. The R801A mutation in the CTM abolishes nuclear import. The FUS linker alone, lacking the NTM and CTM are not translocated to the nucleus by Trn1. The various FUS linker proteins are expressed as GST fusions, purified and used for import assays. GST fusions are visualized in the FITC channel. DAPI shows the outline of the nucleus whereas the whole cell can be seen in the DIC channel. Scale bar = 10 μ m. See also Figure 4.14.

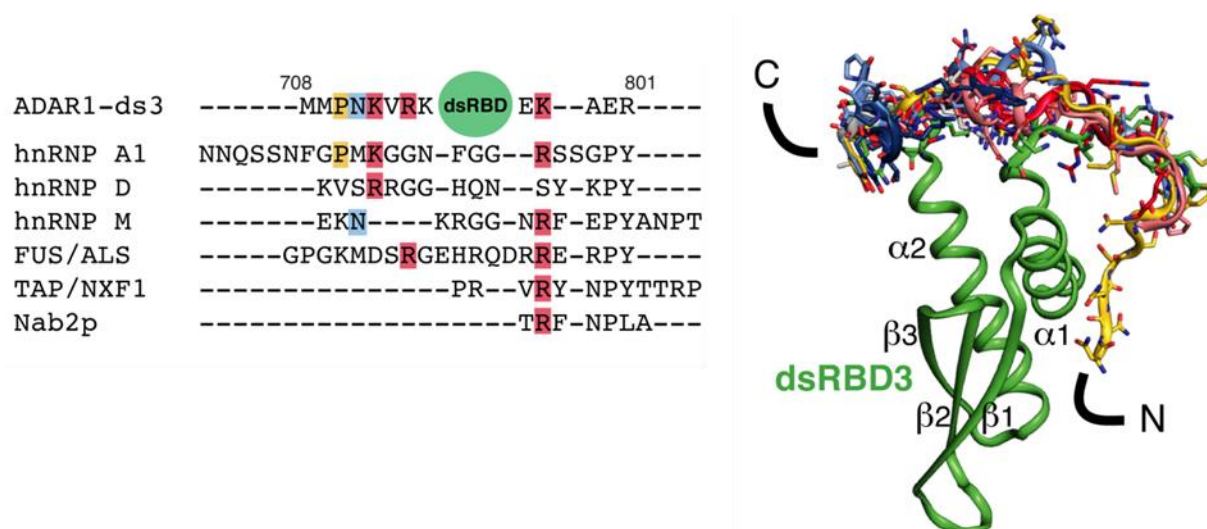


Figure 4.14 Superimposing known PY-NLSs on ADAR1-dsRBD3

(A) Structure-based alignment of the six available structures of Trn1 in complex with NLS peptides from the following proteins: hnRNP A1, hnRNP D, hnRNP M, FUS/ALS, TAP/NXF1 and Nab2p (PDB codes 2H4M, 2Z5N, 2OT8, 4FDD, 2Z5M and 4JLQ, respectively). The structures were superimposed on the Trn1 proteins (for clarity Trn1 proteins are not shown and only the NLS peptides are displayed as sticks) and the various NLS sequences were then aligned according to their relative position in the structures **(B)** Superposition of the structural model of the interaction between ADAR1-dsRBD3 and Trn1 with the different NLSs complexes (for clarity Trn1 proteins are not shown). These figures were generated by our collaborators at ETH Zurich (Frederic Allain's group).

So, I designed a minimal putative NLS consisting of the flanking fragments of ADAR1-dsRBD3 linked by the two helical turns of the FUS/ALS NLS (Figure 4.13A). While the FUS/ALS linker by itself lacks NLS activity (Figure 4.13C, D), addition of the N- and C-terminal fragments fully restores NLS activity (Figure 4.11C and Figure 4.13D), proving that ADAR1-dsRBD3 did not contribute to NLS activity *per se*, but instead provided the appropriate spacing between the N- and C- terminal modules.

This strongly validates our structural model and firmly establishes that ADAR1-NLS is a bimodular NLS formed by the combination of the N- and C-terminal fragments flanking ADAR1-dsRBD3.

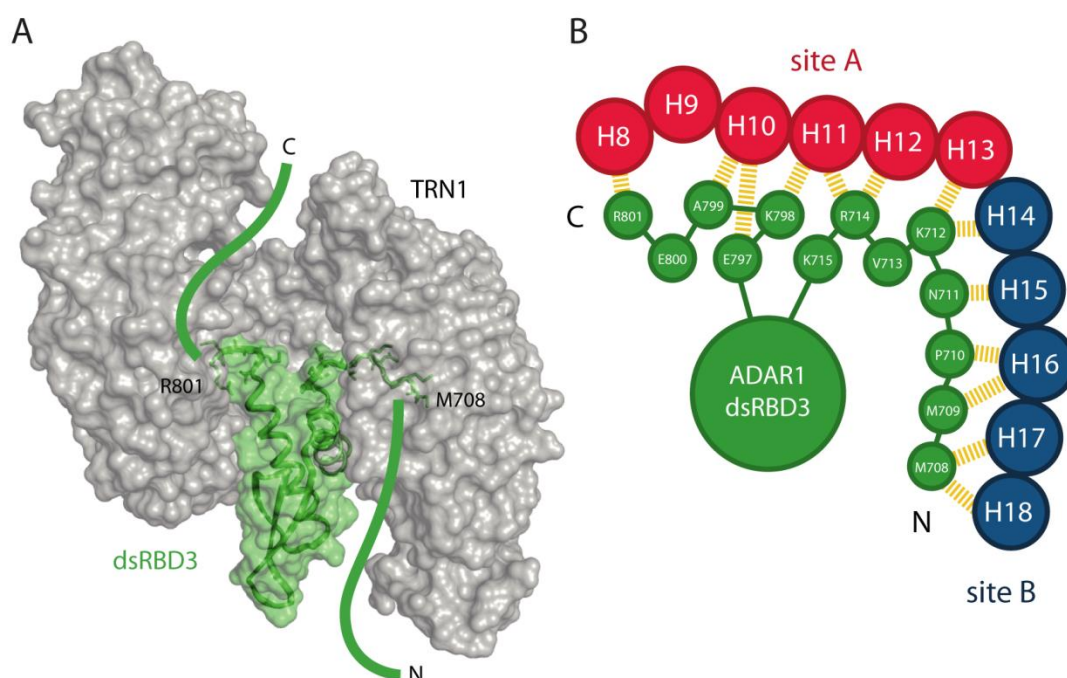


Figure 4.15 ADAR1-dsRBD3 fits perfectly into the groove of Transportin-1

(A) Surface representation of the structural model of the complex between ADAR1-dsRBD3 *in green* and Trn1 *in grey*. ADAR1-dsRBD3 perfectly fits without any atomic clashes into the Trn1 C-terminal arch formed by HEAT repeats 8 to 20. **(B)** Schematic representation of the molecular contacts seen in the structural model between the N- and C-terminal fragments of ADAR1- dsRBD3 *in green* and Trn1 HEAT repeats 8-18. Within the C-terminal arch, site A (H8-H13) and site B (H14-H18) are shown *in red* and *in blue*, respectively. Molecular contacts are shown as *yellow* dashed lines. (See also Figure 1.7 from introduction section 1.3.2). These figures were generated by our collaborators at ETH Zurich (Frederic Allain's group) (Barraud P. Banerjee S., 2014).

4.7 ADAR1-dsRBD3 uses the same binding surface as that of known Transportin-1 –PY NLSs.

The compatibility of the proposed molecular model of the ADAR1 bimodular NLS (Figure 4.9A) with the structures of both the extended dsRBD (contribution of this work) and Trn1 bound to well-characterized PY-NLSs needed further investigation (Cansizoglu et al., 2007; Imasaki et al., 2007; Lee et al., 2006; Soniat et al., 2013; Zhang and Chook, 2012). Rephrasing the problem statement, we wanted to explore if the extended dsRBD fits inside the C-terminal arch formed by HEAT repeats 8-20

(Chook and Blobel, 1999; Lee et al., 2006) (see also Figure 1.7 from section 1.3.2 for Trn1 nomenclature and geometry). We first asked if the same binding interface is used by Trn1 to bind the ADAR1-NLS and PY-NLSs. Particularly, two tryptophans (Trp460 and Trp730) in Trn1 have been proven crucial in binding to PY-NLSs (Imasaki et al., 2007; Lee et al., 2006). Trp460 belongs to repeat H10, in the N-terminal half of the arch, also called site A, whereas Trp730 is located in repeat H16, in the C-terminal half of the arch, called site B (Figure 4.15) (Imasaki et al., 2007). The binding of NLS peptides interacting at both the sites A and B is heavily impaired when the mentioned tryptophans are mutated to alanines (*i.e.* W460A and W730A). These Trn1 mutants have been observed to almost completely impair nuclear import of PY-NLSs (Imasaki et al., 2007). Furthermore, NLS peptides interacting only at site A (*e.g.* TAP NLS) are strongly impaired in their ability to bind to the W460A but not to the W730A mutant (Imasaki et al., 2007). This was our rationale to test the W460A and W730A mutants for the interaction with ADAR1-dsRBD3 by co-immunoprecipitation assays (Figure 4.16). Binding of ADAR1-dsRBD3 to Trn-1 is abolished by both the mutations, indicating that ADAR1-NLS is recognized by Trn1 using the same pockets as the ones used to bind PY-NLSs and that ADAR1-NLS is interacting with both site A and site B and therefore spans the entire C-terminal arch.

As an inception point, the interaction between ADAR1-dsRBD3 (Met708-Arg801) and Trn1 was modelled by our collaborators using the six available X-ray structures of Trn1 bound to various NLSs. Briefly highlighting the contribution of our collaborators, the structures were superimposed on the Trn1 proteins and the different NLSs were then aligned according to their relative position in the structures (Figure 4.14A). Subsequently, the sequences of the N- and C-terminal fragments flanking ADAR1-dsRBD3 were manually aligned with the aim to maximize the degree of sequence similarities while keeping a virtually extended peptide chain. The model was then constructed using docking protocols as described in the methods section in (Barraud P. Banerjee S., 2014). Although this model had limited empirical evidences, it is of great interest for the conceptual understanding of ADAR1-NLS recognition by Trn1. It exhibits that the extended ADAR1-dsRBD3 fits remarkably well into Trn1 C-terminal arch without any steric clashes (Figure 4.14A), even though ADAR1-dsRBD3 is closely surrounded by Trn1 and cannot be freely rotated in this cavity. Additionally, it also illustrates that the flexible N- and C-terminal modules, in

an extended conformation, can imitate canonical PY-NLSs and can interact with the C-terminal arch of Trn1 (Figure 4.14B). In this model, part of the N-terminal module (Met708-Lys712) interacts with site B at repeats H18-H14 and Lys712-Arg714 interact with repeats H13-H11 from site A (Figure 4.14B). The C-terminal module Glu797-Arg801 is exclusively interacting with site A at repeats H11-H8.

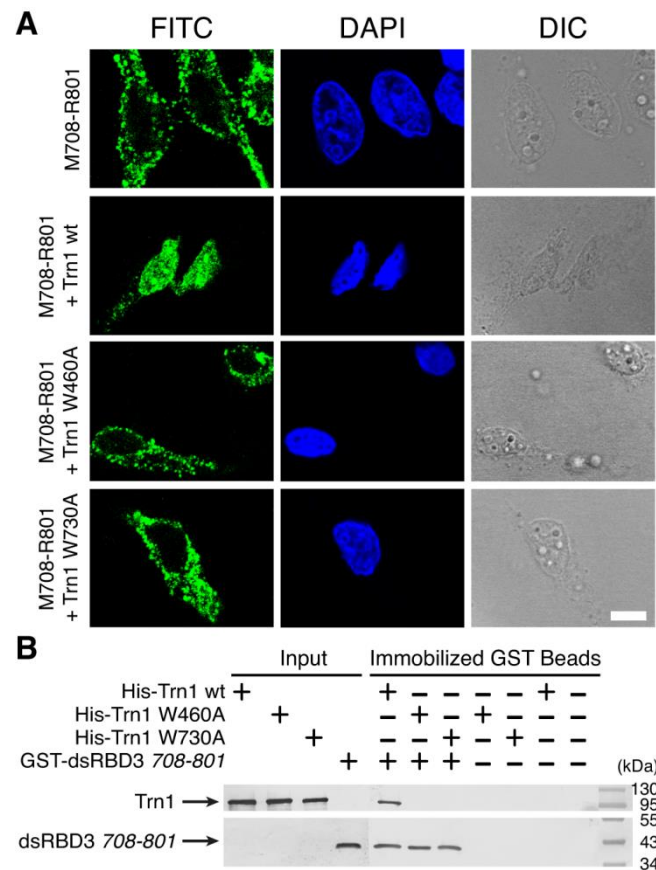


Figure 4.16 Binding pockets in Trn1 that are required for the interaction with PY-NLSs are also required for interaction with ADAR1-NLS

(A) GST-fusion protein harboring the ADAR1-NLS (Met708-Arg801) is transferred to the nucleus in permeabilized cells in the presence of wild-type Trn1. Mutation of the hydrophobics W460A or W730A that are essential for the interaction with PY-NLSs fail to transport the ADAR1-NLS to the nucleus. GST fusions are visualized in the FITC channel. DAPI shows the outline of the nucleus whereas the whole cell can be seen in the DIC channel. Scale bar = 10 μ m. (B) Mutations W460A and W730A in Trn1 fail to interact with ADAR1-NLS in pull down assays. Western blot shows His-tagged Trn1 and GST-tagged dsRBD input material. GST-dsRBD fusions coupled to GST-beads pull down wild-type Trn1 but none of the two mutant versions W460A or W730A.

Most importantly, the model strongly supports our hypothesis that the structured part of ADAR1-dsRBD3 (residues Ile716-Asn796) would not be necessary for the formation of the NLS *per se*, but rather simply acts as a scaffolding domain to form a functional NLS from distant parts of a polypeptide chain. We therefore challenged our structural model further.

4.8 Modulation of nuclear import by dsRNA binding

We have reported earlier and also re-confirmed that binding of ADAR1-dsRBD3 to Trn1 is strongly impaired in the presence of dsRNA which eventually leads to weaker nuclear import (Fritz et al., 2009b; Strehblow et al., 2002)(Figure 4.18). While analyzing the origin of the competition between Trn1 and dsRNA, our efforts to pick the exact reason was rather limited due to multiple possibilities spawning due to (a) a direct competition for the same residues of ADAR1-dsRBD3; (b) a structural rearrangement of ADAR1-dsRBD3 upon dsRNA binding impeding Trn1 binding; or (c) a mutually exclusive binding of either Trn1 or dsRNA due to steric hindrance. We therefore investigated dsRNA binding of ADAR1-dsRBD3.

First, we tested, using mutational analysis, that if any known dsRNA binding residues in canonical dsRBDs abolish NLS activity in our ADAR-dsRBD3 (Fritz et al., 2009b; Ramos et al., 2000). As mentioned, KKxAK motif at the N-terminal tip of helix $\alpha 2$ is well conserved for binding dsRNA substrates (Ryter and Schultz, 1998a). We therefore mutated these lysine residues either individually or in combination (K777A, K778A, K7881A and K777A+K778A) and tested their localization ability through our reporter construct (Figure 14.17). Lysine substitutions in the N-terminal tip of helix $\alpha 2$, do not affect nuclear localization of these constructs. Moreover, our collaborators at ETH Zurich could show that the mutant construct ⁷⁷⁷KK⁷⁷⁸ to ⁷⁷⁷AA⁷⁷⁸ also failed to bind to dsRNA by isothermal titration calorimetry (ITC), using a 24-base-pair duplex dsRNA substrate (Figure 14.19 A). This re-validated previously known facts. We extended our tests in order to find whether the positively charged residues from the flanking fragments are involved in dsRNA binding with ITC. The same 24-base-pair duplex dsRNA substrate was used for these tests by our collaborators. Mutants in the N- or C-terminal modules retain the capacity to bind dsRNA (e.g. R801A in the C-term, ⁷¹²KVRK⁷¹⁵ to ⁷¹²AVAA⁷¹⁵ or M708A+P710A or a deletion mutant in the N-term (Thr725-Arg801; Figure 14.19 B). This provides strong evidence that Trn1 and dsRNA are not competing for the same residues of ADAR1-dsRBD3.

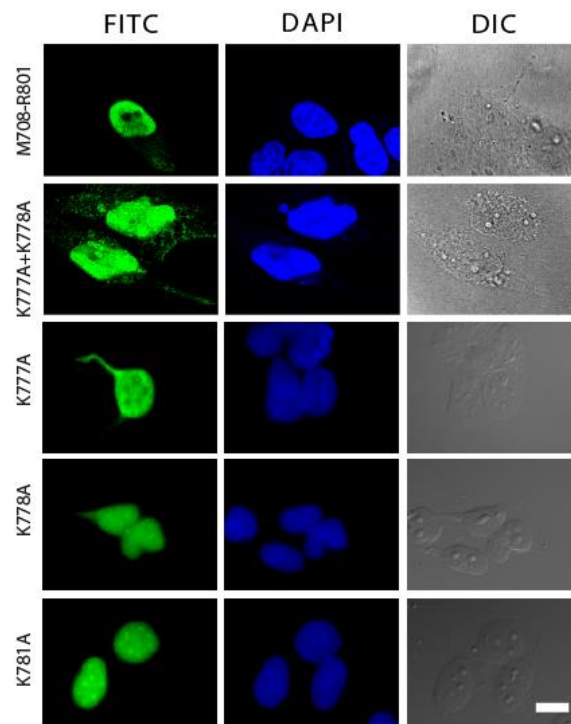


Figure 4.17 The dsRNA binding residues (KKxAK motif) of ADAR1-dsRBD3 are not involved in NLS activity

Mutations abolishing RNA binding by the dsRBD (K777A+K778A:K777A:K778A:K781) do not interfere with nuclear localization of a resulting fusion protein and localizes like the wild-type ADAR1-NLS Met708-Arg801. Transfected PK fusion proteins are visualized in the FITC channel, nuclei are stained with DAPI, and cellular outlines are shown in the DIC channel. Scale bar = 10 μ m.

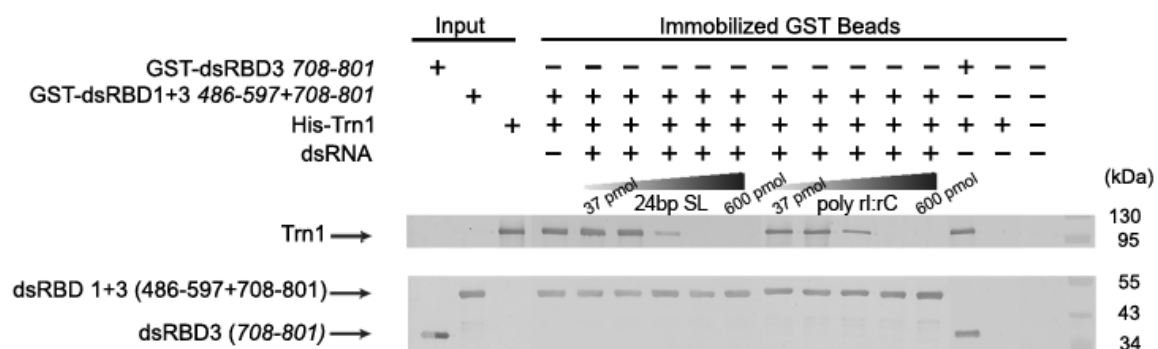


Figure 4.18 Double-stranded RNA interferes with Trn1 binding

37 pmol of a construct harboring dsRBD1 and dsRBD3 (ADAR1-dsRBD 1+3) fused to GST were coupled to GST-beads (where indicated). The coupled beads were incubated with 37 pmol of Trn1 in 500 μ l and with an increasing concentration of dsRNA (a 24bp stem loop, or poly rI: rC) ranging from 37 pmol to 600 pmol. Precipitated material was tested by western blotting for successful pull-down of His-Trn1. ADAR1-dsRBD 1+3 fusion protein nicely interacts with Trn1 but can be titrated off by a three-fold molar excess of double-stranded RNA over Trn1.

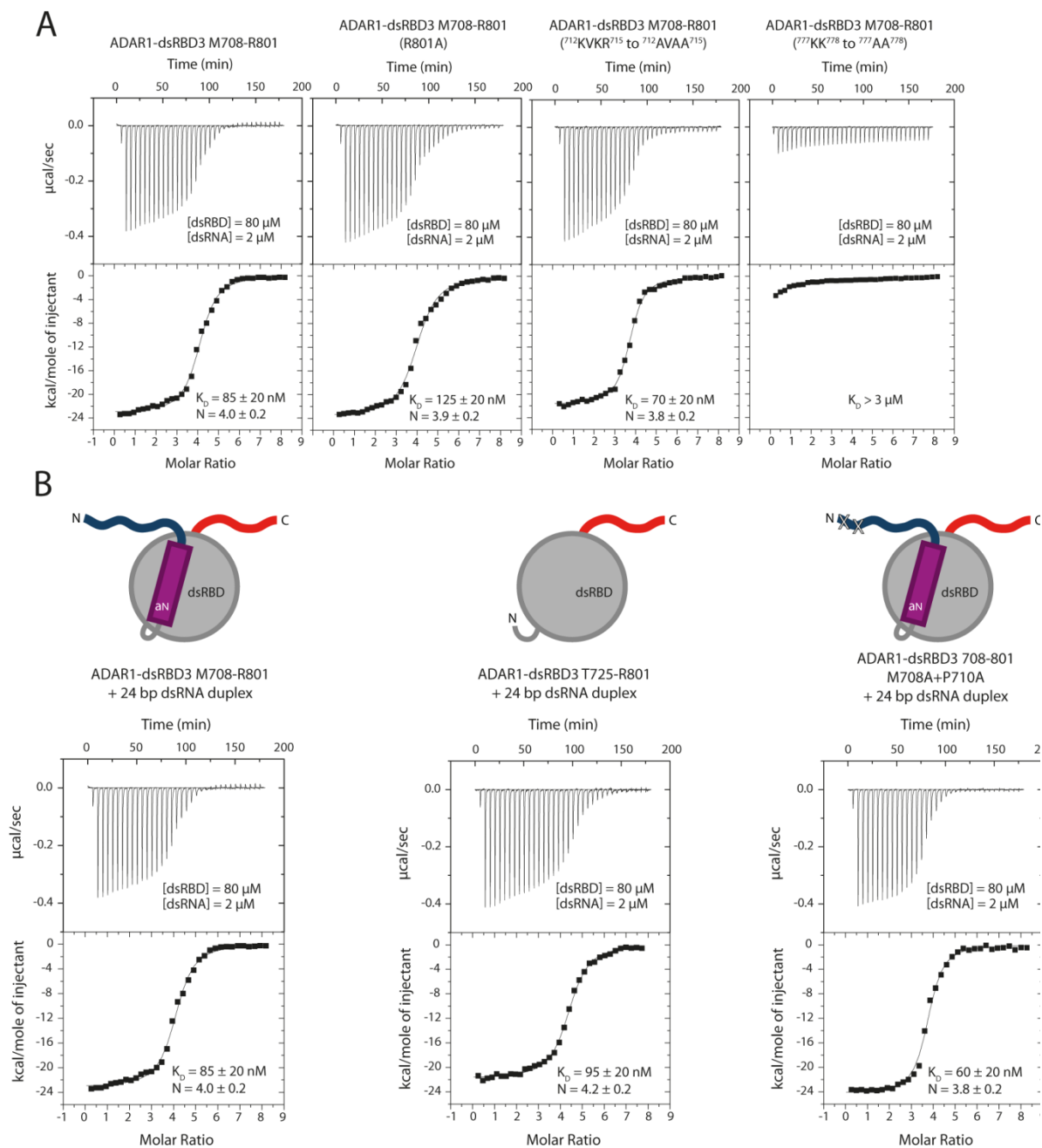


Figure 4.19 RNA binding properties of ADAR1-dsRBD3

(A) Isothermal titration calorimetry (ITC) of wild type ADAR1-dsRBD3 Met708-Arg801 construct and mutants showing that important residues for Trn1 interaction are not important for dsRNA binding. The name of the construct is indicated above each titration. **(B)** ITC monitoring the interaction between ADAR1-dsRBD3 Met708-Arg801 (*left panel*), Thr725-Arg801 (*middle panel*) and ADAR1-dsRBD3 708A+P710A mutant (*right panel*) and a 24 bp dsRNA duplex. Neither the N-terminal module nor the additional helix α_N is essential for dsRNA binding (Barraud P. Banerjee S., 2014).

Second, as a confirmatory step to validate that the N- and C-terminal fragments of ADAR1-dsRBD3 are not involved in dsRNA binding, ^{15}N -labeled ADAR1-dsRBD3 was titrated against a typical ADAR1 substrate, *GluR-2* LSL (Stefl et al., 2010) and *vice-versa* (Figure 4.20A). This part of the work was done by our collaborators at ETH Zurich. Hence, the regions of interaction could be mapped (Figure 4.20B). The NMR chemical shifts revealed a canonical mode of binding between ADAR1-dsRBD3 and dsRNA substrate (Masliah et al., 2013; Ryter and Schultz, 1998a). Structure work showed that the N- and C-terminal fragments do not bind RNA (Figure 4.20B). Furthermore, the regions in proximity of the N- and C-terminal modules are not involved in major structural re-arrangements that would prevent Trn1 binding. Overall, the data suggests that binding of Trn1 and dsRNA by ADAR1-dsRBD3 involves different binding interfaces and that the mutually exclusive binding of either Trn1 or dsRNA might result from steric hindrance.

Our collaborators also modeled ADAR1-dsRBD3 bound to *GluR-2* R/G stem-loop (Stefl et al., 2010) using the chemical shift perturbation, following the same procedure that has been employed to model the dsRNA-binding of *Drosophila* ADAR dsRBD1 (Barraud et al., 2012). The model was built by superimposing the ADAR1-dsRBD3 NMR structure onto the one of ADAR2-dsRBD2 bound to the *GluR-2* R/G stem-loop (PDB code 2L3J) (Stefl et al., 2010). The side chains of helices α_N and α_1 and of the $^{775}\text{HSKKxGK}^{781}$ motif at the N-terminal tip of helix α_2 , the entire β_1 - β_2 loop and the unstructured N- and C-terminal fragments were allowed to move freely whereas the rest of the protein was fixed. Finally, 20 lowest energy structures of the resulting protein-RNA complex were pooled and carefully analyzed (Figure 4.21A). Indeed, when this model is superposed onto the model of ADAR1-dsRBD3 bound to Trn1, binding of both the molecules simultaneously is prevented by the steric hindrance (Figure 4.21B, Figure 5.3), impairing nuclear import.

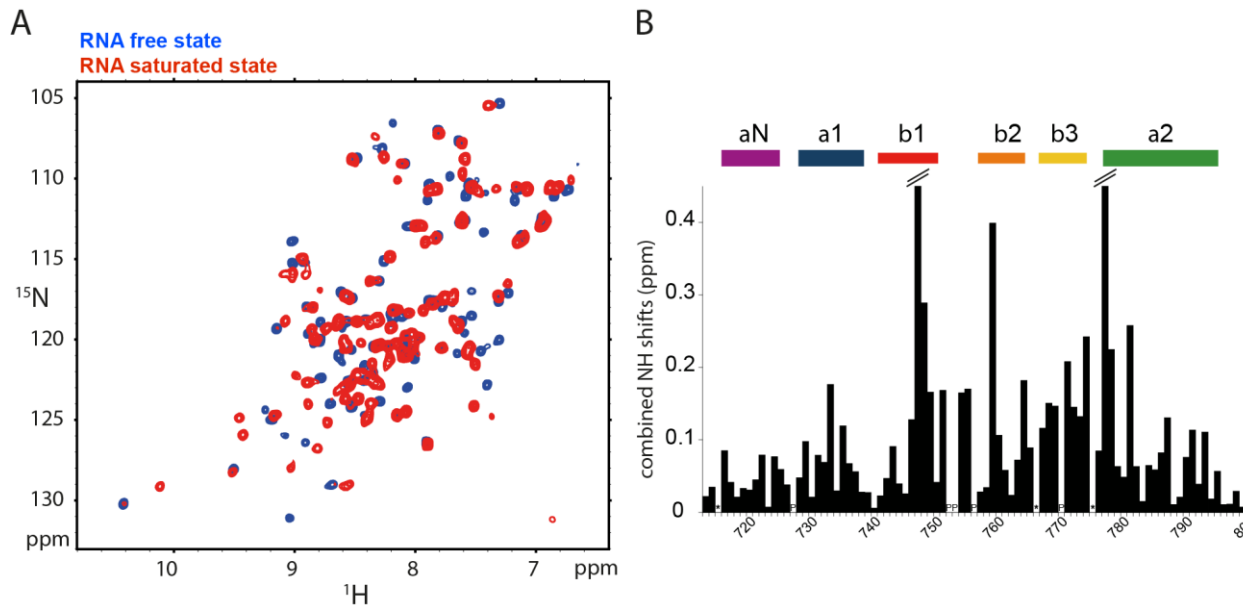


Figure 4.20 dsRNA binding to ADAR1-dsRBD3 does not alter the structure of the overall domain

(A) Superposition of (^{15}N , ^1H)-HSQC spectra representing NMR titration of the ^{15}N -labeled ADAR1-dsRBD3 Met708-Arg801 protein construct with increasing amount of unlabeled *GluR-2* R/G editing site LSL. Only the free state and the RNA saturated state of the protein (RNA:protein ratio of 5:1) are represented *in blue* and *red*, respectively. (B) NMR chemical shift mapping of combined NH shifts upon RNA binding plotted along ADAR1-dsRBD3 protein sequence. Residues marked with a P correspond to prolines. Residues marked with an asterisk have a missing amide assignment. These figures were generated by our collaborators at ETH Zurich (Frederic Allain's group) (Barraud P. Banerjee S., 2014).

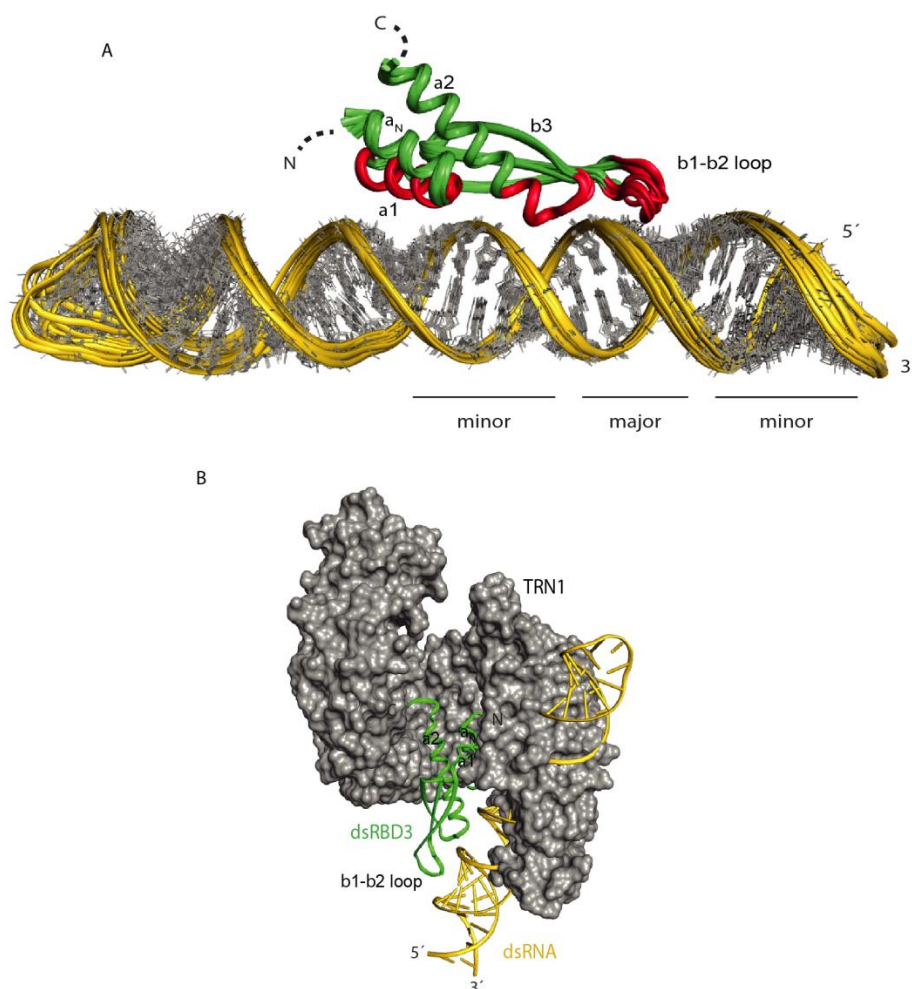


Figure 4.21 Steric hindrance prevents ADAR1-dsRBD3 from binding both dsRNA and transportin-1 simultaneously

(A) Overlay of the 20 lowest energy structures of the complex between ADAR1-dsRBD3 and *GluR-2* R/G stem-loop (a typical ADAR RNA substrate used for instance in Stefl *et al.* (2010)). The model was constructed by Barraud P of Frederic Allain's group. The RNA stem-loop is represented with a *yellow* backbone and *grey* lines. The dsRBD is represented as a cartoon *in green* and the three canonical regions of interaction (namely helix α_1 , the β_1 - β_2 loop, and the N-terminal tip of helix α_2) are highlighted *in red*. Helix α_1 and the β_1 - β_2 loop bind to two consecutive minor grooves, whereas the N-terminal tip of helix α_2 interacts with the phosphate backbone across the major groove. For clarity, the flexible N- and C-terminal fragments are not shown on the figure, although they have been included in the simulated annealing refinement protocol. **(B)** Superposition on ADAR1-dsRBD3 of the structures of the two complexes, namely ADAR1-dsRBD3 bound to dsRNA and the model of Trn1 bound to ADAR1-dsRBD3. The surface of Trn1 is shown *in grey*, ADAR1-dsRBD3 is shown as a cartoon *in green*, and the *GluR-2* R/G dsRNA stem-loop is shown as a cartoon *in yellow*. The dsRNA molecule extends across and strongly clashes with the C-terminal arch of Trn1 (Barraud P. Banerjee S., 2014).

5. Discussion

A-to-I editing by ADARs predominantly occurs in the nucleus of a cell due to the formation of double-stranded RNA structures that affects the subsequent stability of the targeted RNA. Since the I: U mismatch pairs generated due to editing events are less stable than A: U base pairs (Samuel, 2012). Moreover, I is recognized as G instead of A by ribosomes, during translation and by viral RNA-dependent RNA polymerases during RNA replication. ADARs can edit both cellular and viral RNAs. A detailed description of cellular substrates of ADARs has been discussed in section 1.5.1. A few examples of viral RNA substrates of ADARs include: hepatitis C virus (HCV), vesicular stomatitis virus (VSV), measles virus (MV), hepatitis delta virus (HDV), and human immunodeficiency virus type 1 (HIV-1). ADAR1 mediates both proviral (HDV, MV, VSV and HIV-1) and an antiviral effect (HCV) which is either editing-dependent (HDV and HCV), editing-independent (VSV and MV) or both (HIV-1) (Samuel, 2012). ADAR1-i predominantly localized in cytoplasm (see section 1.5 for details). ADAR-i impairs HCV replication via RNA editing at multiple sites in the nucleus (Taylor et al., 2005). Thus indicating that preferential movement of ADAR1 into different cellular compartments should be regulated. This is further supported by the fact that there is selectivity among the substrates for ADAR1. Moreover, we have previously showed that the nuclear import of ADAR1-c is regulated by RNA. Binding of dsRNA to ADAR1-c, prevents Trn-1 from importing the protein into the nucleus (Fritz et al., 2009b). Also many proteins shuttling between the cytoplasm and nucleus mediate RNA export (Lee et al., 1996). In addition, the binding of dsRNA to the dsRBD strengthens the export of the protein into the cytoplasm (Fritz et al., 2009b). Hence, studying the mechanism of nuclear import mediated by dsRBDs regulated by dsRNA was of primary goal of this thesis.

Our study shed light on a completely novel function of dsRBDs. While originally believed to be strictly involved in RNA-binding, dsRBDs have now been shown to act as interaction sites for nuclear transport factors, or to be homo- and hetero-dimerization domains (Doyle and Jantsch, 2002). We could interpret how the third dsRBD of ADAR1 interacts with the transport factor Trn-1 and how this process is regulated by dsRNA. This was achieved by looking at the structural features of the dsRBD3 of ADAR-1 by NMR spectroscopy performed by our collaborators at ETH Zurich (Federic Allain's group). Later this structural model was challenged by

biochemical and cell biological approach and we could precisely characterize the nuclear localization signal embedded in the third dsRBD of ADAR-1 crucial for nuclear import.

5.1 ADAR1-dsRBD3 is an extended dsRBD allowing a dual function

We have shown that ADAR1-dsRBD3 adopts an extended dsRBD structure. Similarly, other dsRBDs exhibit structural extensions (Masliah et al., 2013), such as the budding yeast RNase III Rnt1p which accommodates a long C-terminal helix α_3 (Leulliot et al., 2004; Wu et al., 2004c), and the fission yeast dicer Dicer1 which accommodates a short C-terminal helical turn and a -CHCC zinc-binding domain (Barraud et al., 2011). In contrast to previously reported dsRBD extensions, the one described here occurs upstream of the canonical dsRBD. The additional helix α_N is conserved in all ADAR1 proteins of vertebrates from fish to mammals, as evidenced by the absolute conservation of the pattern of hydrophobic and hydrophilic residues in this region (Figure 5.1).

Helix α_N is positioned between helices α_1 and α_2 which is similar to what was recently found in Staufen1-dsRBD5 for the Staufen-swapping motif (SSM) that drives dimerization of the protein (Gleghorn et al., 2013). Globally, the N-terminal helix of the SSM motif adopts a similar position and orientation to helix α_N in ADAR1-dsRBD3 which could be demonstrated by secondary structure predictions (figure 5.2). Importantly, it has been proposed that the dynamic nature of the 16-residue-long linker connecting the SSM sub-domain to Staufen1-dsRBD5 would allow the SSM to interact both in *cis* leading to a monomeric form and in *trans* leading to dimerization (Gleghorn et al., 2013). Although the structural feature is rather similar, the resulting functions are different, which underline the propensity of dsRBDs to acquire new functions *via* N- or C-terminal extensions (Barraud et al., 2011; Gleghorn et al., 2013).

Recently it was demonstrated that Staufen1 dsRBD3 bears a NLS mediating import of the protein into the nucleus. The amino acids constituting the bipartite NLS were basic lysine residues separated by 10 amino acids. The NLS is situated at the C-terminal edge of dsRBD3 of the protein (Martel et al., 2006). This led us to wonder if this domain behaved as ADAR1-dsRBD3 and had additional helix surrounding the

canonical dsRBD residues aiding the lysine residues in the formation of a bipartite NLS. However, amino acid sequence alignment and secondary structure prediction fail to show the presence of any structured element upstream or downstream of canonical Staufen1-dsRBD3 (Figure 5.2). Moreover, the import factor mediating the transport of Staufen-1 has still not been identified implying that the mechanism by which Staufen1-dsRBD3 NLS imports the protein is may be completely different from ADAR1-dsRBD3 (Doyle et al., 2013).

Moreover, we demonstrated here that the primary role of helix α_N is to act as a scaffolding element for a bimodular NLS to interact with Trn1. One might also ask if helix α_N might also be implicated in RNA binding similarly to the additional helix α_3 of Rnt1p (Leulliot et al., 2004; Wu et al., 2004c). The pull down assays and ITC titrations from our collaborators (Figure 4.18 and Figure 4.19) indicate that the contribution of helix α_N to RNA-binding is very limited. In the structural model of ADAR1-dsRBD3 bound to RNA (Figure 4.21A), only Asn726 in the short kink joining α_N to α_1 appears to be in the position to make direct contacts with RNA. Altogether, this argues against a strong implication of helix α_N in RNA binding.

5.2 Comparison of the ADAR1-NLS with classical bipartite NLSs

In the past two decades many proteins have been identified displaying nuclear import activity using bipartite NLSs. Proteins such as nucleoplasmin, N1, estrogen receptor, SWI5, p53, poly-ADPR polymerase, yeast ribosomal protein L29, Staufen1 etc. consists of bipartite NLSs having two basic clusters separated by a spacer of 10 random amino acids (Addison et al., 1990; Dingwall C, 1991; Moll et al., 1991; Picard et al., 1990; Underwood and Fried, 1990). Influenza virus polymerase basic protein 1(PBP1) was shown to have a spacer length of 16 amino acids between two basic clusters contrary to the adenovirus DNA binding protein which has 37 amino acid spacer (Morin et al., 1989; Nath and Nayak, 1990). These two unusual bipartite NLSs indicated that maintenance of certain distance between the two basic clusters was important for binding to the import factor and retaining NLS activity. Replacing the spacer with random flexible amino acids such as glycines or alanines could not mimic a function NLS activity (Figure 4.12). Replacing the entire globular domain with an short helix of 8 amino acids from FUS helix could create the appropriate space between the basic clusters (Figure 4.11C).

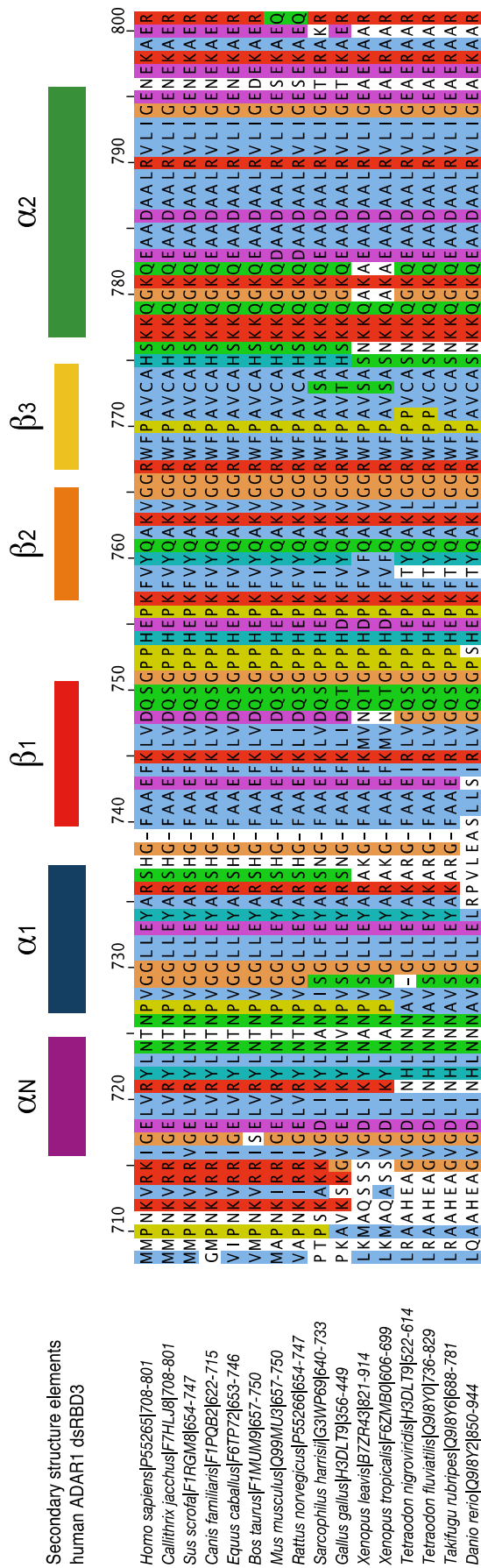


Figure 5.1: The N-terminal helix α_N is conserved in the ADAR1-dsRBD3 domains of vertebrates.

Sequence alignment of various ADAR1-dsRBD3 from *Homo sapiens*, *Callithrix jacchus*, *Sus scrofa*, *Canis familiaris*, *Equus caballus*, *Bos taurus*, *Mus musculus*, *Rattus norvegicus*, *Sarcophilus harrisii*, *Gallus gallus*, *Xenopus leavis*, *Xenopus tropicalis*, *Tetraodon nigroviridis*, *Tetraodon fluviatilis*, *Takifugu rubripes* and *Danio rerio*. Alignment was done with ClustalW. For each sequence, the name of the species is given in the first item, the accession number in the Uniprot database (<http://www.uniprot.org>) in the second item, and the range of amino-acid composing the dsRBD in the numbering of the full-length protein in the third item. The alignment is colored by amino-acid conservation and properties. Human ADAR1-dsRBD3 amino-acid numbering and secondary structure elements are shown on top of the alignment. The additional helix α_N is conserved in all ADAR1 proteins of vertebrates from fish to mammals, as evidenced by the absolute conservation of the pattern of hydrophobic and hydrophilic residues in this region (residues 716-725) (Barraud P. Banerjee S., 2014)

Figure 5.2: ADAR1-dsRBD3 domain has a unique extended N-terminal helix α_N

Secondary structure predictions for dsRBDs of two human nucleocytoplasmic shuttling proteins ADAR1 and Staufen1. There are no additional structural elements situated on the C-Terminal edge of dsRBD3 of Staufen1 which bears the bipartite NLS. Satufen1-dsRBD3 has similar functions as that of ADAR1-dsRBD3. However, there is no evidence of any extended helix also at the amino terminal of the domain. On the contrary the dsRBD5 of Staufen1 has two N-terminal short helices, SSM1 & SSM2. But there is no evidence of the domain involved in nuclear import of the protein. Alignment was done with PSIREN(Jones, 1999). The additional helix α_N is not conserved in all dsRBD containing proteins with similar functions. There is evidence of absolute conservation of the pattern of hydrophobic and hydrophilic residues in the canonical region (residues 726-795).(Gleghorn et al., 2013; Martel et al., 2006)

5.3 Comparison of the ADAR1-NLS with PY-NLSs

We have previously demonstrated that Trn1 acts as the unique nuclear import receptor of human ADAR1 *via* the segment Met708-Arg801, which encompasses the third dsRBD of the protein (Fritz et al., 2009b). In the present study, we identified the molecular determinants responsible for the interaction with Trn1 as the combination of the N- and C-terminal modules flanking the folded domain. These modules do not entirely comply with the collection of physical rules that describe the best-characterized substrates of Trn1, namely PY-NLSs (Chook and Suel, 2011; Xu et al., 2010). Their similarities and differences is thus of great interest.

First, PY-NLSs are disordered peptide segments, a feature that to some extent is fulfilled by ADAR1-NLS. Indeed, the modules flanking the folded dsRBD are structurally disordered in the free protein (Figure 4.4D).

Second, PY-NLSs have an overall basic character, and with five positive charges (Arg or Lys) within the two modules, ADAR1-NLS conforms perfectly to this rule. Finally, PY-NLSs display some weakly conserved sequence motifs with some structural conservation that were classified as three so-called epitopes (Cansizoglu et al., 2007). Epitope 1 is composed of a stretch of either hydrophobic or basic residues interacting with Trn1 site B (Figure 4.14 and Figure 4.15). Our structural modeling and *in vitro* data suggest that ADAR1-NLS holds a hydrophobic epitope 1 composed of residues ⁷⁰⁸MMPN⁷¹¹ (Figure 4.5A and Figure 4.15B). Epitope 2 is composed of a single and well-conserved positively charged residue (Arg/Lys/His) interacting with a negatively charged pocket formed by repeats H11-H12. In ADAR1-NLS, epitope 2 might be constituted by Lys798 since its position is very similar to the conserved Arg of other PY-NLSs (Figure 4.14A and Figure 4.15). Finally, epitope 3 is composed of a relatively well-conserved Pro-Tyr dipeptide interacting with Trn1 site A. There is no PY in the ADAR1 C-terminal module, but our structural model shows that residues ⁸⁰⁰ER⁸⁰¹ could occupy the corresponding Trn1 pocket and therefore represent a rather degenerated equivalent to epitope 3. Interestingly, although the PY energetic contribution is sometimes important and explains the sensitivity to However, the bimodular organization of ADAR1-NLS is unique with the two NLS modules separated by a small globular domain of about 9 kDa. The two modules are more than 80-residues apart in the primary structure (Figure 4.4C). mutations in this motif (*e.g.* hnRNP M, hnRNP D and FUS/ALS) (Cansizoglu et al., 2007; Imasaki et

al., 2007; Zhang and Chook, 2012), the contribution of epitope 3 to the overall binding affinity is sometimes relatively small and epitope 3 is then tolerant to mutations (Suel et al., 2008). In the case of ADAR1-NLS, however, Arg801 is essential for Trn1 binding and is thus an important constituent for nuclear import.

Therefore, considering the molecular determinants of ADAR1-NLS that we uncovered here, the number of currently predicted Trn1-interacting NLSs is most likely underestimated. Indeed, even though it is assumed that each epitope can accommodate substantial sequence diversity, the limits are still vague. The fact that some characterized Trn1 cargos lack a PY-NLS suggests that the actual variability is larger than currently conceived (Arnold et al., 2006; Jakel and Gorlich, 1998; Muhlhauser et al., 2001; Waldmann et al., 2007). We believe that more bimodular NLSs such as in ADAR1 will be found in the future and that our study will help identifying them.

5.4 Regulation of nuclear Import

The presence or absence of an NLS in a protein does not necessarily influence its efficiency of nuclear targeting. Transport receptors, cargo proteins and cell signaling molecules intricately interplay to highly regulate nucleocytoplasmic shuttling of a protein. Regulation of gene expression through transcription factors such as tumor suppressors or oncoproteins is critically governed by nuclear targeting of the cargo proteins (Macara, 1999; McLane and Corbett, 2009). The binding affinity of the import receptor to its cargo protein directly affects the rate of nuclear import. Hence modulating the receptor binding affinity can regulate the transport of an NLS containing cargo (Hodel et al., 2006; Timney et al., 2006). This can be achieved by physically blocking this interaction or by modifying the NLS either promoting or preventing interaction to the import receptor or by altering the nuclear transport machinery.

Alteration in the import receptors can cause deleterious effects to the cell. A truncated form of karyopherin- $\alpha 1$ (KPNA1) results in constitutive cytoplasmic retention of p53 in breast cancer cell lines, ZR-75-1, mediating decrease in apoptotic gene expression and increased expression of proliferating genes (Kim et al., 2000; Ko and Prives, 1996; Moll et al., 1992). Moreover, differential expressions of karyopherins have also been implicated in colon, breast and lung cancers (Dahl et al., 2006; Kau et al., 2004).

Changes within the NLS motif can also disrupt the nuclear import of a protein. In Swyer syndrome due to loss of NLS activity by virtue of mutation in the SRY transcription factor leads to developmental defects resulting in male-to-female sex reversal. The NLS of SRY is recognized by Imp β 1 and imported into the nucleus where SRY-dependent genes are up-regulated during normal development. In the case of Swyer syndrome individuals, the NLS of SRY is mutated resulting in decreased recognition by the import receptor and, consequently, less import of SRY into the nucleus decreasing transcription of target genes (Harley et al., 2003; McLane and Corbett, 2009). Intermolecular masking of an NLS by an upstream signaling molecule can also regulate nuclear targeting by either permitting or prohibiting masking of the NLS. Nuclear import of NF- κ B (nuclear-factor- κ B), a transcription factor, is mediated by phosphorylation and degradation of I κ B (inhibitor of κ B kinase) thereby unmasking the NLS of NF- κ B and facilitating binding of Imp- α/β receptors and hence up-regulating target genes (Beg et al., 1992). Recently panduratin A, a possible inhibitor of metastasized A549 cells has been discovered which results in a strong inhibition of NF- κ B activation, with a decrease in nuclear levels of NF- κ B/p65 and NF- κ B/p50 (Cheah et al., 2013).

The modulation of import activity has been reported for both Staufen1 and ADAR1, by masking of the NLSs which is embedded in dsRBD3 of both proteins. It is suggested that import of Staufen1 is regulated by a mechanism that involves cytoplasmic retention of the protein (Martel et al., 2006). RNA binding to dsRBD3 prevents nuclear import and enhances nuclear export (Fritz et al. 2009). For instance pri-miRNAs associated with ADAR1 are edited in the nucleus. By virtue of some of these editing events degradation of the RNA species occurs in the cytoplasm (O'Connell and Keegan, 2006; Yang et al., 2006). Only when the bound RNA is dissociated in the cytoplasm, ADAR1 will be accessible for interaction to Transportin1 and is ready for re-cycling into the nucleus.

5.5 Nuclear import regulation *via* an RNA-sensing NLS

A bimodular NLS where two active modules are interrupted by a small folded domain is an impressive molecular construction, but seems rather complicated to generate a functional NLS. However, this allows the corresponding NLS to be switched *on* and *off* depending on the presence of RNA associated with the intervening domain (in this case, a dsRBD; Figure 4.21B and Figure 5.3) (Fritz et al.,

2009b). Although the size of the intervening domain is crucial for this mechanism, since it should fit into the Trn1 C-terminal arch, it is conceivable that similar bimodular NLS interrupted by other domains might exist among Trn1 cargos.

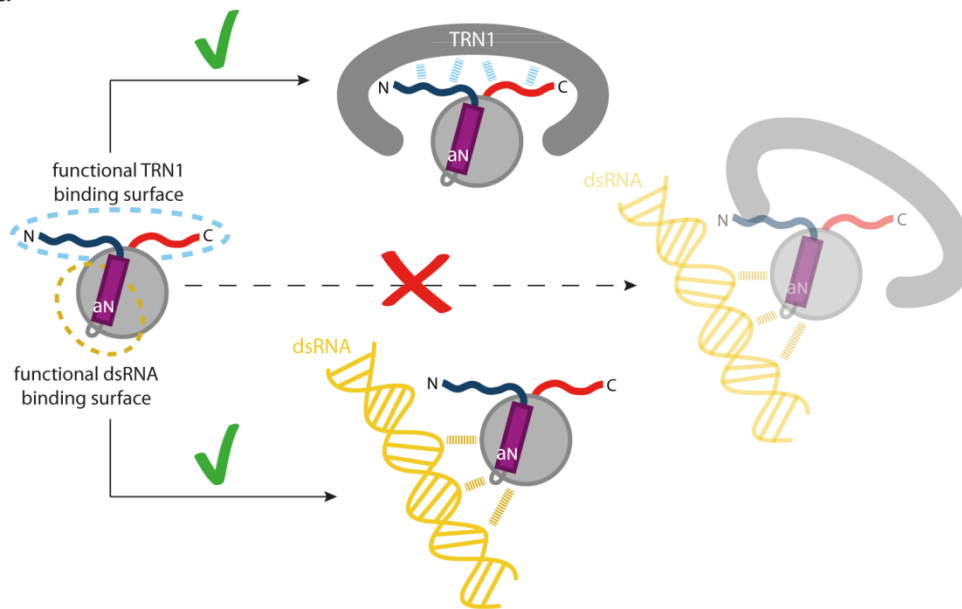


Figure 5.3 Dual role of ADAR1-dsRBD3

Schematic representation of the two functional binding interfaces of ADAR1-dsRBD3 (Met708-Arg801) and the association of these non-overlapping surfaces with their respective binding partner (*i.e.* Trn1 and dsRNA helices). These two binding sites cannot be occupied simultaneously due to the large dimensions of both Trn1 and dsRNA helices.

Interestingly, in addition to this RNA-sensing NLS, other mechanisms to regulate Trn1-mediated nuclear import have been described. For instance, post-translational modifications such as arginine methylation (Dormann et al., 2012) or cysteine oxidation *via* reactive oxygen species (Putker et al., 2013) have been recently reported to modulate Trn1-mediated nuclear import. In the case of ADAR1, the bimodular NLS is switched off upon binding to dsRNA substrates, which would prevent ADAR1 to carry RNAs back into the nucleus. The fact that dsRNA stimulates nuclear export of dsRBD-containing proteins (Brownawell and Macara, 2002a; Fritz et al., 2009b; Gwizdek et al., 2004) suggests that ADARs might leave the nucleus bound to substrate RNAs, for instance pre-miRNAs. The binding of the RNA-sensing NLS to Trn1 could then help the dissociation of ADAR1-associated RNAs in the cytoplasm and ensure an efficient reimport of ADAR1 free of RNAs to the nucleus

Thus binding of RNA to ADAR1-dsRBD3 is required to control the directionality of transport. However, it is unknown whether overexpression or binding of cytoplasmic or nuclear substrates can influence nucleo-cytoplasmic shuttling kinetics of wild-type protein but not the mutant. RNAs associated with ADAR1 either facilitating nuclear export by Exp-5 or nuclear import by Trn-1 still needs to be characterized to understand the modulation of import activity of ADAR1 more lucidly. In fact, for most dsRBD proteins that have been shown to act in the nucleus the NLS has not yet been identified experimentally. Given that many important regulators of RNA-metabolism such as Drosha, NF90, or TRBP are found in the nucleus, it is possible that nuclear transport of these proteins is also regulated by RNA-binding. The findings of this thesis have important implications for various research directions such as the understanding of NLS recognition by karyopherins, the nucleocytoplasmic trafficking regulation through RNA binding, and the regulation of ADAR proteins shuttling and localization.

6. References

1. Adam, S.A., Marr, R.S., and Gerace, L. (1990). Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *J Cell Biol* *111*, 807-816.
2. Addison, C., Jenkins, J.R., and Sturzbecher, H.W. (1990). The P53 Nuclear-Localization Signal Is Structurally Linked to a P34cdc2 Kinase Motif. *Oncogene* *5*, 423-426.
3. Andrade, M.A., and Bork, P. (1995). HEAT repeats in the Huntington's disease protein. *Nat Genet* *11*, 115-116.
4. Arnold, M., Nath, A., Hauber, J., and Kehlenbach, R.H. (2006). Multiple importins function as nuclear transport receptors for the Rev protein of human immunodeficiency virus type 1. *J Biol Chem* *281*, 20883-20890.
5. Aruscavage, P.J., and Bass, B.L. (2000). A phylogenetic analysis reveals an unusual sequence conservation within introns involved in RNA editing. *Rna* *6*, 257-269.
6. Athanasiadis, A. (2012). Zalpha-domains: At the intersection between RNA editing and innate immunity. *Seminars in Cell & Developmental Biology* *23*, 275-280.
7. Athanasiadis, A., Placido, D., Maas, S., Brown li, B.A., Lowenhaupt, K., and Rich, A. (2005). The Crystal Structure of the Z β Domain of the RNA-editing Enzyme ADAR1 Reveals Distinct Conserved Surfaces Among Z-domains. *J Mol Biol* *351*, 496-507.
8. Athanasiadis, A., Rich, A., and Maas, S. (2004). Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol* *2*, e391.
9. Barraud, P., and Allain, F.H. (2012). ADAR proteins: double-stranded RNA and Z-DNA binding domains. *Curr Top Microbiol Immunol* *353*, 35-60.

10. Barraud, P., Emmerth, S., Shimada, Y., Hotz, H.-R., Allain, F.H.-T., and Bühler, M. (2011). An extended dsRBD with a novel zinc-binding motif mediates nuclear retention of fission yeast Dicer. *Embo J* 30, 4223-4235.
11. Barraud, P., Heale, B.S., O'Connell, M.A., and Allain, F.H. (2012). Solution structure of the N-terminal dsRBD of Drosophila ADAR and interaction studies with RNA. *Biochimie* 94, 1499-1509.
12. Barraud P. Banerjee S., W.I.M., Michael F. Jantsch & Frédéric H.-T. Allain (2014). A bimodular nuclear localization signal assembled via an extended dsRBD acts as an RNA-sensing signal for Transportin 1 PNAS (*Submitted under revision*).
13. Bass, B.L. (1997). RNA editing and hypermutation by adenosine deamination. *Trends in Biochemical Sciences* 22, 157-162.
14. Bass, B.L. (2002). RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* 71, 817-846.
15. Bass, B.L., and Weintraub, H. (1988). An Unwinding Activity That Covalently Modifies Its Double-Stranded-Rna Substrate. *Cell* 55, 1089-1098.
16. Beg, A.A., Ruben, S.M., Scheinman, R.I., Haskill, S., Rosen, C.A., and Baldwin, A.S., Jr. (1992). I kappa B interacts with the nuclear localization sequences of the subunits of NF-kappa B: a mechanism for cytoplasmic retention. *Genes Dev* 6, 1899-1913.
17. Bischoff, F.R., and Ponstingl, H. (1995). Catalysis of Guanine-Nucleotide Exchange of Ran by Rcc1 and Stimulation of Hydrolysis of Ran-Bound Gtp by Ran-Gap1. *Method Enzymol* 257, 135-144.
18. 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, 457-466.
19. Bonifaci, N., Moroianu, J., Radu, A., and Blobel, G. (1997). Karyopherin beta2 mediates nuclear import of a mRNA binding protein. *Proc Natl Acad Sci U S A* 94, 5055-5060.

20. Breeuwer, M., and Goldfarb, D.S. (1990). Facilitated Nuclear Transport of Histone H1 and Other Small Nucleophilic Proteins. *Cell* 60, 999-1008.
21. Brownawell, A.M., and Macara, I.G. (2002a). Exportin-5, a novel karyopherin, mediates nuclear export of double-stranded RNA binding proteins. *J Cell Biol* 156, 53-64.
22. Bycroft, M., Grunert, S., Murzin, A.G., Proctor, M., and St Johnston, D. (1995). NMR solution structure of a dsRNA binding domain from *Drosophila* staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. *Embo J* 14, 3563-3571.
23. Calado, A., Treichel, N., Muller, E.C., Otto, A., and Kutay, U. (2002). Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *EMBO J* 21, 6216-6224.
24. Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J., Henrissat, B., and Mornon, J.P. (1997). Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cell Mol Life Sci* 53, 621-645.
25. Cansizoglu, A.E., Lee, B.J., Zhang, Z.C., Fontoura, B.M., and Chook, Y.M. (2007). Structure-based design of a pathway-specific nuclear import inhibitor. *Nat Struct Mol Biol* 14, 452-454.
26. Cernilogar, F.M., Onorati, M.C., Kothe, G.O., Burroughs, A.M., Parsi, K.M., Breiling, A., Lo Sardo, F., Saxena, A., Miyoshi, K., Siomi, H., *et al.* (2011). Chromatin-associated RNA interference components contribute to transcriptional regulation in *Drosophila*. *Nature* 480, 391-U151.
27. Chadrin, A., Hess, B., San Roman, M., Gatti, X., Lombard, B., Loew, D., Barral, Y., Palancade, B., and Doye, V. (2010). Pom33, a novel transmembrane nucleoporin required for proper nuclear pore complex distribution. *The Journal of cell biology* 189, 795-811.
28. Chang, K.-Y., and Ramos, A. (2005). The double-stranded RNA-binding motif, a versatile macromolecular docking platform. *FEBS Journal* 272, 2109-2117.

29. Cheah, S.-C., Lai, S.-L., Lee, S.-T., Hadi, A., and Mustafa, M. (2013). Panduratin A, a Possible Inhibitor in Metastasized A549 Cells through Inhibition of NF-Kappa B Translocation and Chemoinvasion. *Molecules* 18, 8764-8778.
30. Chen, C.X., Cho, D.S.C., Wang, Q.D., Lai, F., Carter, K.C., and Nishikura, K. (2000). A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *Rna* 6, 755-767.
31. Cho, D.S., Yang, W., Lee, J.T., Shiekhataar, R., Murray, J.M., and Nishikura, K. (2003). Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. *J Biol Chem* 278, 17093-17102.
32. Chook, Y., and Blobel, G. (2001). Karyopherins and nuclear import. *Current Opinion in Structural Biology* 11, 703-715.
33. Chook, Y.M., and Blobel, G. (1999). Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp. *Nature* 399, 230-237.
34. Chook, Y.M., and Suel, K.E. (2011). Nuclear import by karyopherin-betas: recognition and inhibition. *Biochim Biophys Acta* 1813, 1593-1606.
35. Christophe, D., Christophe-Hobertus, C., and Pichon, B. (2000). Nuclear targeting of proteins: how many different signals? *Cellular Signalling* 12, 337-341.
36. Cook, A., Bono, F., Jinek, M., and Conti, E. (2007). Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem* 76, 647-671.
37. D'Angelo, M.A., and Hetzer, M.W. (2008). Structure, dynamics and function of nuclear pore complexes. *Trends in Cell Biology* 18, 456-466.
38. Dahl, E., Kristiansen, G., Gottlob, K., Klamann, I., Ebner, E., Hinzmann, B., Hermann, K., Pilarsky, C., Dürst, M., Klinkhammer-Schalke, M., *et al.* (2006). Molecular Profiling of Laser-Microdissected Matched Tumor and Normal Breast Tissue Identifies Karyopherin $\alpha 2$ as a Potential Novel Prognostic Marker in Breast Cancer. *Clinical Cancer Research* 12, 3950-3960.

39. Daniel, C., Wahlstedt, H., Ohlson, J., Bjork, P., and Ohman, M. (2011). Adenosine-to-inosine RNA editing affects trafficking of the gamma-aminobutyric acid type A (GABA(A)) receptor. *J Biol Chem* 286, 2031-2040.
40. Desterro, J.M., Keegan, L.P., Lafarga, M., Berciano, M.T., O'Connell, M., and Carmo-Fonseca, M. (2003a). Dynamic association of RNA-editing enzymes with the nucleolus. *J Cell Sci* 116, 1805-1818.
41. Dingwall C, L.R. (1991). Nuclear targeting sequences--a consensus? *Trends Biochem Sci* 16, 478-481.
42. Doreleijers, J.F., Sousa da Silva, A.W., Krieger, E., Nabuurs, S.B., Spronk, C.A.E.M., Stevens, T.J., Vranken, W.F., Vriend, G., and Vuister, G.W. (2012). CING: an integrated residue-based structure validation program suite. *J Biomol NMR* 54, 267-283.
43. Dormann, D., Madl, T., Valori, C.F., Bentmann, E., Tahirovic, S., Abou-Ajram, C., Kremmer, E., Ansorge, O., Mackenzie, I.R., Neumann, M., *et al.* (2012). Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. *EMBO J* 31, 4258-4275.
44. Doyle, M., Badertscher, L., Jaskiewicz, L., Guttinger, S., Jurado, S., Hugenschmidt, T., Kutay, U., and Filipowicz, W. (2013). The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. *Rna* 19, 1238-1252.
45. Doyle, M., and Jantsch, M.F. (2002). New and old roles of the double-stranded RNA-binding domain. *J Struct Biol* 140, 147-153.
46. Doyle, M., Jaskiewicz, L., and Filipowicz, W. (2012). Chapter One - Dicer Proteins and Their Role in Gene Silencing Pathways. In *The Enzymes*, G. Feng, and T. Fuyuhiko, eds. (Academic Press), pp. 1-35.
47. Eckmann, C.R., Neunteufl, A., Pfaffstetter, L., and Jantsch, M.F. (2001). The human but not the *Xenopus* RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. *Mol Biol Cell* 12, 1911-1924.

48. Eggington, J.M., Greene, T., and Bass, B.L. (2011). Predicting sites of ADAR editing in double-stranded RNA. *Nat Commun* 2, 319.
49. Feng, Y., Sansam, C.L., Singh, M., and Emeson, R.B. (2006). Altered RNA editing in mice lacking ADAR2 autoregulation. *Mol Cell Biol* 26, 480-488.
50. Fierro-Monti, I., and Mathews, M.B. (2000). Proteins binding to duplexed RNA: one motif, multiple functions. *Trends in Biochemical Sciences* 25, 241.
51. Fontoura, B.M.A., Blobel, G., and Yaseen, N.R. (2000). The nucleoporin Nup98 is a site for GDP/GTP exchange on Ran and termination of karyopherin beta 2-mediated nuclear import. *Journal of Biological Chemistry* 275, 31289-31296.
52. Fried, H., and Kutay, U. (2003). Nucleocytoplasmic transport: taking an inventory. *Cell Mol Life Sci* 60, 1659-1688.
53. Fritz, J., Strehblow, A., Taschner, A., Schopoff, S., Pasierbek, P., and Jantsch, M.F. (2009a). RNA-Regulated Interaction of Transportin-1 and Exportin-5 with the Double-Stranded RNA-Binding Domain Regulates Nucleocytoplasmic Shuttling of ADAR1. *Mol Cell Biol* 29, 1487-1497.
54. Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J.P. (1987). Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett* 224, 149-155.
55. Gallo, A., Keegan, L.P., Ring, G.M., and O'Connell, M.A. (2003). An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. *Embo J* 22, 3421-3430.
56. Gan, J., Shaw, G., Tropea, J.E., Waugh, D.S., Court, D.L., and Ji, X. (2008). A stepwise model for double-stranded RNA processing by ribonuclease III. *Mol Microbiol* 67, 143-154.
57. Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2005). Intermediate states of ribonuclease III in complex with double-stranded RNA. *Structure* 13, 1435-1442.

58. Gan, J.H., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X.H. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. *Cell* 124, 355-366.
59. George, C.X., and Samuel, C.E. (1999). Human RNA-specific adenosine deaminase ADAR1 transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. *Proc Natl Acad Sci U S A* 96, 4621-4626.
60. Gerber, A., Grosjean, H., Melcher, T., and Keller, W. (1998). Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2. *Embo J* 17, 4780-4789.
61. Gerber, A.P., and Keller, W. (1999). An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* 286, 1146-1149.
62. Gleghorn, M.L., Gong, C., Kielkopf, C.L., and Maquat, L.E. (2013). Staufen1 dimerizes through a conserved motif and a degenerate dsRNA-binding domain to promote mRNA decay. *Nat Struct Mol Biol* 20, 515-524.
63. Gorlich, D. (1998). Transport into and out of the cell nucleus. *Embo J* 17, 2721-2727.
64. Gorlich, D., and Kutay, U. (1999). Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* 15, 607-660.
65. Gott, J.M., and Emeson, R.B. (2000). Functions and mechanisms of RNA editing. *Annu Rev Genet* 34, 499-531.
66. Gullerova, M., and Proudfoot, N.J. (2012). Convergent transcription induces transcriptional gene silencing in fission yeast and mammalian cells. *Nat Struct Mol Biol* 19, 1193-1201.
67. Guntert, P. (2004). Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278, 353--378.

68. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J Mol Biol* 273, 283--298.
69. Gwizdek, C., Ossareh-Nazari, B., Brownawell, A.M., Evers, S., Macara, I.G., and Dargemont, C. (2004). Minihelix-containing RNAs mediate exportin-5-dependent nuclear export of the double-stranded RNA-binding protein ILF3. *J Biol Chem* 279, 884-891.
70. Harley, V.R., Layfield, S., Mitchell, C.L., Forwood, J.K., John, A.P., Briggs, L.J., McDowall, S.G., and Jans, D.A. (2003). Defective importin beta recognition and nuclear import of the sex-determining factor SRY are associated with XY sex-reversing mutations. *Proc Natl Acad Sci U S A* 100, 7045-7050.
71. Hartner, J.C., Schmittwolf, C., Kispert, A., Muller, A.M., Higuchi, M., and Seeburg, P.H. (2004). Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. *J Biol Chem* 279, 4894-4902.
72. Hartner, J.C., Walkley, C.R., Lu, J., and Orkin, S.H. (2009). ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling (vol 10, pg 109, 2009). *Nat Immunol* 10, 551-551.
73. Herbert, A., Alfken, J., Kim, Y.G., Mian, I.S., Nishikura, K., and Rich, A. (1997). A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. *Proc Natl Acad Sci U S A* 94, 8421-8426.
74. Herrmann, T., Guntert, P., and Wuthrich, K. (2002). Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* 319, 209--227.
75. Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P.H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature* 406, 78-81.
76. 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, 25-30.

77. Hodel, A.E., Harreman, M.T., Pulliam, K.F., Harben, M.E., Holmes, J.S., Hodel, M.R., Berland, K.M., and Corbett, A.H. (2006). Nuclear localization signal receptor affinity correlates with in vivo localization in *Saccharomyces cerevisiae*. *J Biol Chem* 281, 23545-23556.
78. Hoelz, A., Debler, E.W., and Blobel, G. (2011). The structure of the nuclear pore complex. *Annu Rev Biochem* 80, 613-643.
79. Hogg, M., Paro, S., Keegan, L.P., and O'Connell, M.A. (2011). RNA editing by mammalian ADARs. *Adv Genet* 73, 87-120.
80. Huang, Y., Ji, L.J., Huang, Q.C., Vassilyev, D.G., Chen, X.M., and Ma, J.B. (2009). Structural insights into mechanisms of the small RNA methyltransferase HEN1. *Nature* 461, 823-U886.
81. Hundley, H.A., Krauchuk, A.A., and Bass, B.L. (2008). *C. elegans* and *H. sapiens* mRNAs with edited 3' UTRs are present on polysomes. *Rna* 14, 2050-2060.
82. Imasaki, T., Shimizu, T., Hashimoto, H., Hidaka, Y., Kose, S., Imamoto, N., Yamada, M., and Sato, M. (2007). Structural basis for substrate recognition and dissociation by human transportin 1. *Mol Cell* 28, 57-67.
83. Iwamoto, K., Bundo, M., and Kato, T. (2009). Serotonin receptor 2C and mental disorders: genetic, expression and RNA editing studies. *RNA Biology* 6, 248-253.
84. Jakel, S., and Gorlich, D. (1998). Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *Embo J* 17, 4491-4502.
85. Jantsch, M.F., and Gall, J.G. (1992). Assembly and localization of the U1-specific snRNP C protein in the amphibian oocyte. *The Journal of cell biology* 119, 1037-1046.
86. Jin, Y., Tian, N., Cao, J., Liang, J., Yang, Z., and Lv, J. (2007). RNA editing and alternative splicing of the insect nAChR subunit alpha6 transcript: evolutionary conservation, divergence and regulation. *BMC Evol Biol* 7, 98.

87. Jones, D.T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* 292, 195-202.
88. Kau, T.R., Way, J.C., and Silver, P.A. (2004). Nuclear transport and cancer: from mechanism to intervention. *Nat Rev Cancer* 4, 106-117.
89. Kawahara, Y., Megraw, M., Kreider, E., Iizasa, H., Valente, L., Hatzigeorgiou, A.G., and Nishikura, K. (2008). Frequency and fate of microRNA editing in human brain. *Nucleic Acids Research* 36, 5270-5280.
90. Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., and Nishikura, K. (2007). Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science* 315, 1137-1140.
91. Keegan, L.P., McGurk, L., Palavicini, J.P., Brindle, J., Paro, S., Li, X., Rosenthal, J.J., and O'Connell, M.A. (2011). Functional conservation in human and *Drosophila* of Metazoan ADAR2 involved in RNA editing: loss of ADAR1 in insects. *Nucleic Acids Research* 39, 7249-7262.
92. Kharraz, A., Macias, M.J., Gibson, T.J., Nilges, M., and Pastore, A. (1995). Structure of the dsRNA binding domain of *E. coli* RNase III. *Embo J* 14, 3572-3584.
93. Kiebler, M.A., Jansen, R.P., Dahm, R., and Macchi, P. (2005). A putative nuclear function for mammalian Staufen. *Trends in Biochemical Sciences* 30, 228-231.
94. Kim, I.S., Kim, D.H., Han, S.M., Chin, M.U., Nam, H.J., Cho, H.P., Choi, S.Y., Song, B.J., Kim, E.R., Bae, Y.S., *et al.* (2000). Truncated form of importin alpha identified in breast cancer cell inhibits nuclear import of p53. *J Biol Chem* 275, 23139-23145.
95. Ko, L.J., and Prives, C. (1996). p53: puzzle and paradigm. *Genes Dev* 10, 1054-1072.
96. Koeris, M., Funke, L., Shrestha, J., Rich, A., and Maas, S. (2005). Modulation of ADAR1 editing activity by Z-RNA in vitro. *Nucleic Acids Research* 33, 5362-5370.

97. Krovat, B.C., and Jantsch, M.F. (1996a). Comparative mutational analysis of the double-stranded RNA binding domains of *Xenopus laevis* RNA-binding protein A. *J Biol Chem* 271, 28112-28119.
98. Kuttan, A., and Bass, B.L. (2012). Mechanistic insights into editing-site specificity of ADARs. *Proceedings of the National Academy of Sciences* 109, E3295–E3304.
99. Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E., and Corbett, A.H. (2007). Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* 282, 5101-5105.
100. Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8, 477--486.
101. Lee, B.J., Cansizoglu, A.E., Suel, K.E., Louis, T.H., Zhang, Z., and Chook, Y.M. (2006). Rules for nuclear localization sequence recognition by karyopherin beta 2. *Cell* 126, 543-558.
102. Lee, M.S., Henry, M., and Silver, P.A. (1996). A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. *Genes Dev* 10, 1233-1246.
103. Lehmann, K.A., and Bass, B.L. (1999). The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol* 291, 1-13.
104. Lemesle-Varloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A., and Mornon, J.P. (1990). Hydrophobic cluster analysis: procedures to derive structural and functional information from 2-D-representation of protein sequences. *Biochimie* 72, 555-574.
105. Leulliot, N., Quevillon-Cheruel, S., Graille, M., van Tilbeurgh, H., Leeper, T.C., Godin, K.S., Edwards, T.E., Sigurdsson, S.T., Rozenkrants, N., Nagel, R.J., *et al.* (2004). A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNase III. *Embo J* 23, 2468-2477.

106. Levanon, E.Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z.Y., Shoshan, A., Pollock, S.R., Sztybel, D., *et al.* (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat Biotechnol* 22, 1001-1005.
107. Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* 8, 479-490.
108. Maas, S., and Gommans, W.M. (2009). Novel exon of mammalian ADAR2 extends open reading frame. *PLoS One* 4, e4225.
109. Macara, I.G. (1999). Nuclear transport: randy couples. *Curr Biol* 9, R436-439.
110. 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 Staufen2: nucleolar accumulation and isoform-specific exportin-5-dependent export. *J Biol Chem* 279, 31440-31444.
111. Martel, C., Macchi, P., Furic, L., Kiebler, M.A., and DesGroseillers, L. (2006). Staufen1 is imported into the nucleolus via a bipartite nuclear localization signal and several modulatory determinants. *Biochem J* 393, 245-254.
112. Martínez-Salas, E., Lozano, G., Fernandez-Chamorro, J., Francisco-Velilla, R., Galan, A., and Diaz, R. (2013). RNA-Binding Proteins Impacting on Internal Initiation of Translation. *International Journal of Molecular Sciences* 14, 21705-21726.
113. Masliah, G., Barraud, P., and Allain, F.H. (2013). RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. *Cell Mol Life Sci* 70, 1875-1895.
114. McLane, L.M., and Corbett, A.H. (2009). Nuclear localization signals and human disease. *IUBMB Life* 61, 697-706.
115. Micklem, D.R., Adams, J., Grunert, S., and St Johnston, D. (2000). Distinct roles of two conserved Staufen domains in oskar mRNA localization and translation. *EMBO J* 19, 1366-1377.

116. Miki, T., Takano, K., and Yoneda, Y. (2005). The Role of Mammalian Staufen on mRNA Traffic: a View from Its Nucleocytoplasmic Shuttling Function. *Cell Structure and Function* 30, 51-56.
117. Miyamura, Y., Suzuki, T., Kono, M., Inagaki, K., Ito, S., Suzuki, N., and Tomita, Y. (2003). Mutations of the RNA-Specific Adenosine Deaminase Gene (DSRAD) Are Involved in Dyschromatosis Symmetrica Hereditaria. *The American Journal of Human Genetics* 73, 693-699.
118. Moll, T., Tebb, G., Surana, U., Robitsch, H., and Nasmyth, K. (1991). The role of phosphorylation and the CDC28 protein kinase in cell cycle-regulated nuclear import of the *S. cerevisiae* transcription factor SWI5. *Cell* 66, 743-758.
119. Moll, U.M., Riou, G., and Levine, A.J. (1992). Two distinct mechanisms alter p53 in breast cancer: mutation and nuclear exclusion. *Proc Natl Acad Sci U S A* 89, 7262-7266.
120. Morin, N., Delsert, C., and Klessig, D.F. (1989). Nuclear localization of the adenovirus DNA-binding protein: requirement for two signals and complementation during viral infection. *Mol Cell Biol* 9, 4372-4380.
121. Muhlhauser, P., Muller, E.C., Otto, A., and Kutay, U. (2001). Multiple pathways contribute to nuclear import of core histones. *EMBO Rep* 2, 690-696.
122. Nath, S.T., and Nayak, D.P. (1990). Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1 N1). *Mol Cell Biol* 10, 4139-4145.
123. Nigg, E.A. (1997). Nucleocytoplasmic transport: signals, mechanisms and regulation. *Nature* 386, 779-787.
124. Nishikura, K. (2006). Editor meets silencer: crosstalk between RNA editing and RNA interference. *Nat Rev Mol Cell Biol* 7, 919-931.
125. Nishikura, K. (2010a). Functions and regulation of RNA editing by ADAR deaminases. *Annu Rev Biochem* 79, 321-349.

126. O'Connell, M.A., and Keegan, L.P. (2006). Drosha versus ADAR: wrangling over pri-miRNA. *Nat Struct Mol Biol* 13, 3-4.
127. Öhman, M. (2007). A-to-I editing challenger or ally to the microRNA process. *Biochimie* 89, 1171-1176.
128. Orlandi, C., Barbon, A., and Barlati, S. (2012). Activity regulation of adenosine deaminases acting on RNA (ADARs). *Mol Neurobiol* 45, 61-75.
129. Osenberg, S., Paz Yaacov, N., Safran, M., Moshkovitz, S., Shtrichman, R., Sherf, O., Jacob-Hirsch, J., Keshet, G., Amariglio, N., Itskovitz-Eldor, J., *et al.* (2010). Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. *PLoS One* 5, e11173.
130. Paine, P.L., Moore, L.C., and Horowitz, S.B. (1975). Nuclear-Envelope Permeability. *Nature* 254, 109-114.
131. Palladino, M.J., Keegan, L.P., O'Connell, M.A., and Reenan, R.A. (2000). dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *Rna* 6, 1004-1018.
132. Patterson, J.B., and Samuel, C.E. (1995a). Expression and Regulation by Interferon of a Double-Stranded-Rna-Specific Adenosine-Deaminase from Human-Cells - Evidence for 2 Forms of the Deaminase. *Mol Cell Biol* 15, 5376-5388.
133. Peck, L.J., Nordheim, A., Rich, A., and Wang, J.C. (1982). Flipping of cloned d(pCpG)n.d(pCpG)n DNA sequences from right- to left-handed helical structure by salt, Co(III), or negative supercoiling. *Proc Natl Acad Sci U S A* 79, 4560-4564.
134. Peck, L.J., and Wang, J.C. (1985). Transcriptional block caused by a negative supercoiling induced structural change in an alternating CG sequence. *Cell* 40, 129-137.
135. Pemberton, L.F., Blobel, G., and Rosenblum, J.S. (1998). Transport routes through the nuclear pore complex. *Curr Opin Cell Biol* 10, 392-399.

136. Penn, A.C., Balik, A., and Greger, I.H. (2013). Reciprocal regulation of A-to-I RNA editing and the vertebrate nervous system. *Front Neurosci* 7, 61.
137. Picard, D., Kumar, V., Chambon, P., and Yamamoto, K.R. (1990). Signal Transduction by Steroid-Hormones - Nuclear-Localization Is Differentially Regulated in Estrogen and Glucocorticoid Receptors. *Cell Regul* 1, 291-299.
138. Pollard, V.W., Michael, W.M., Nakielnny, S., Siomi, M.C., Wang, F., and Dreyfuss, G. (1996a). A novel receptor-mediated nuclear protein import pathway. *Cell* 86, 985-994.
139. Polson, A.G., and Bass, B.L. (1994). Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *The EMBO journal* 13, 5701-5711.
140. Poulsen, H., Nilsson, J., Damgaard, C.K., Egebjerg, J., and Kjems, J. (2001). CRM1 mediates the export of ADAR1 through a nuclear export signal within the Z-DNA binding domain. *Mol Cell Biol* 21, 7862-7871.
141. Pullirsch, D., and Jantsch, M.F. (2010). Proteome diversification by adenosine to inosine RNA editing. *RNA Biology* 7, 205-212.
142. Putker, M., Madl, T., Vos, H.R., de Ruiter, H., Visscher, M., van den Berg, M.C., Kaplan, M., Korswagen, H.C., Boelens, R., Vermeulen, M., *et al.* (2013). Redox-dependent control of FOXO/DAF-16 by transportin-1. *Mol Cell* 49, 730-742.
143. Raices, M., and D'Angelo, M.A. (2012). Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nat Rev Mol Cell Biol* 13, 687-699.
144. Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D., and Varani, G. (2000). RNA recognition by a Staufen double-stranded RNA-binding domain. *Embo J* 19, 997.
145. Rebagliati, M.R., and Melton, D.A. (1987). Antisense RNA injections in fertilized frog eggs reveal an RNA duplex unwinding activity. *Cell* 48, 599-605.

146. Reichelt, R., Holzenburg, A., Buhle, E.L., Jarnik, M., Engel, A., and Aebi, U. (1990). Correlation between Structure and Mass-Distribution of the Nuclear-Pore Complex and of Distinct Pore Complex Components. *J Cell Biol* 110, 883-894.
147. Ribbeck, K., and Gorlich, D. (2001). Kinetic analysis of translocation through nuclear pore complexes. *Embo J* 20, 1320-1330.
148. Riedmann, E.M., Schopoff, S., Hartner, J.C., and Jantsch, M.F. (2008). Specificity of ADAR-mediated RNA editing in newly identified targets. *Rna* 14, 1110-1118.
149. Rueter, S.M., Dawson, T.R., and Emeson, R.B. (1999). Regulation of alternative splicing by RNA editing. *Nature* 399, 75-80.
150. Ryter, J.M., and Schultz, S.C. (1998a). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *Embo J* 17, 7505-7513.
151. Samuel, C.E. (2012). ADARs: viruses and innate immunity. *Curr Top Microbiol Immunol* 353, 163-195.
152. Sansam, C.L., Wells, K.S., and Emeson, R.B. (2003). Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proceedings of the National Academy of Sciences of the United States of America* 100, 14018-14023.
153. Saunders, L.R., and Barber, G.N. (2003). The dsRNA binding protein family: critical roles, diverse cellular functions. *The FASEB Journal* 17, 961-983.
154. Seeburg, P.H., Higuchi, M., and Sprengel, R. (1998). RNA editing of brain glutamate receptor channels: mechanism and physiology. *Brain Res Brain Res Rev* 26, 217-229.
155. Seeburg, P.H., Single, F., Kuner, T., Higuchi, M., and Sprengel, R. (2001). Genetic manipulation of key determinants of ion flow in glutamate receptor channels in the mouse. *Brain Res* 907, 233-243.

156. Siomi, H., and Dreyfuss, G. (1995). A nuclear localization domain in the hnRNP A1 protein. *The Journal of cell biology* 129, 551-560.
157. Sommer, B., Kohler, M., Sprengel, R., and Seeburg, P.H. (1991). RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell* 67, 11-19.
158. Soniat, M., Sampathkumar, P., Collett, G., Gizzi, A.S., Banu, R.N., Bhosle, R.C., Chamala, S., Chowdhury, S., Fiser, A., Glenn, A.S., *et al.* (2013). Crystal structure of human Karyopherin beta2 bound to the PY-NLS of *Saccharomyces cerevisiae* Nab2. *J Struct Funct Genomics* 14, 31-35.
159. St Johnston, D., 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, 10979-10983.
160. Stefl, R., Oberstrass, F.C., Hood, J.L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R.B., *et al.* (2010). The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell* 143, 225-237.
161. Strehblow, A., Hallegger, M., and Jantsch, M.F. (2002). Nucleocytoplasmic distribution of human RNA-editing enzyme ADAR1 is modulated by double-stranded RNA-binding domains, a leucine-rich export signal, and a putative dimerization domain. *Mol Biol Cell* 13, 3822-3835.
162. Suel, K.E., Gu, H., and Chook, Y.M. (2008). Modular organization and combinatorial energetics of proline-tyrosine nuclear localization signals. *PLoS Biol* 6, e137.
163. Tariq, A., and Jantsch, M.F. (2012). Transcript diversification in the nervous system: a to I RNA editing in CNS function and disease development. *Front Neurosci* 6, 99.
164. Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S.M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Cunto, F., *et al.* (2011). Coding-independent regulation

of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell* 147, 344-357.

165. Taylor, D.R., Puig, M., Darnell, M.E.R., Mihalik, K., and Feinstone, S.M. (2005). New Antiviral Pathway That Mediates Hepatitis C Virus Replicon Interferon Sensitivity through ADAR1. *J Virol* 79, 6291-6298.

166. Tian, B., Bevilacqua, P.C., Diegelman-Parente, A., and Mathews, M.B. (2004). The double-stranded-RNA-binding motif: interference and much more. *Nat Rev Mol Cell Biol* 5, 1013-1023.

167. Tian, B., and Mathews, M.B. (2001). Functional Characterization of and Cooperation between the Double-stranded RNA-binding Motifs of the Protein Kinase PKR. *Journal of Biological Chemistry* 276, 9936-9944.

168. Timney, B.L., Tetenbaum-Novatt, J., Agate, D.S., Williams, R., Zhang, W., Chait, B.T., and Rout, M.P. (2006). Simple kinetic relationships and nonspecific competition govern nuclear import rates in vivo. *The Journal of cell biology* 175, 579-593.

169. Underwood, M.R., and Fried, H.M. (1990). Characterization of Nuclear Localizing Sequences Derived from Yeast Ribosomal Protein-L29. *Embo J* 9, 91-99.

170. 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, 16054-16061.

171. Varani, G. (1995). Exceptionally stable nucleic acid hairpins. *Annu Rev Biophys Biomol Struct* 24, 379-404.

172. Vasu, S.K., and Forbes, D.J. (2001). Nuclear pores and nuclear assembly. *Curr Opin Cell Biol* 13, 363-375.

173. Vitali, P., and Scadden, A.D. (2010). Double-stranded RNAs containing multiple IU pairs are sufficient to suppress interferon induction and apoptosis. *Nat Struct Mol Biol* 17, 1043-1050.

174. Wagner, R.W., Yoo, C., Wrabetz, L., Kamholz, J., Buchhalter, J., Hassan, N.F., Khalili, K., Kim, S.U., Perussia, B., McMorris, F.A., *et al.* (1990). Double-Stranded-Rna Unwinding and Modifying Activity Is Detected Ubiquitously in Primary Tissues and Cell-Lines. *Mol Cell Biol* 10, 5586-5590.
175. Waldmann, I., Walde, S., and Kehlenbach, R.H. (2007). Nuclear import of c-Jun is mediated by multiple transport receptors. *J Biol Chem* 282, 27685-27692.
176. Wang, Q., Khillan, J., Gadue, P., and Nishikura, K. (2000). Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science* 290, 1765-1768.
177. Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J., and Nishikura, K. (2004). Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. *J Biol Chem* 279, 4952-4961.
178. Wang, Z.H., Hartman, E., Roy, K., Chanfreau, G., and Feigon, J. (2011). Structure of a Yeast RNase III dsRBD Complex with a Noncanonical RNA Substrate Provides New Insights into Binding Specificity of dsRBDs. *Structure* 19, 999-1010.
179. Woese, C.R., Winker, S., and Gutell, R.R. (1990). Architecture of ribosomal RNA: constraints on the sequence of "tetra-loops". *Proc Natl Acad Sci U S A* 87, 8467-8471.
180. Wu, H., Henras, A., Chanfreau, G., and Feigon, J. (2004a). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc Natl Acad Sci U S A* 101, 8307-8312.
181. Xu, D., Farmer, A., and Chook, Y.M. (2010). Recognition of nuclear targeting signals by Karyopherin-beta proteins. *Curr Opin Struct Biol* 20, 782-790.
182. Yamashita, S., Nagata, T., Kawazoe, M., Takemoto, C., Kigawa, T., Guntert, P., Kobayashi, N., Terada, T., Shirouzu, M., Wakiyama, M., *et al.* (2011). Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. *Protein Sci* 20, 118-130.

183. Yang, S.W., Chen, H.Y., Yang, J., Machida, S., Chua, N.H., and Yuan, Y.A. (2010). Structure of Arabidopsis HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. *Structure* 18, 594-605.
184. Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R., and Nishikura, K. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. *Nat Struct Mol Biol* 13, 13-21.
185. Zhang, X.J., He, P.P., Li, M., He, C.D., Yan, K.L., Cui, Y., Yang, S., Zhang, K.Y., Gao, M., Chen, J.J., *et al.* (2004). Seven novel mutations of the ADAR gene in Chinese families and sporadic patients with dyschromatosis symmetrica hereditaria (DSH). *Hum Mutat* 23, 629-630.
186. Zhang, Z., and Carmichael, G.G. (2001). The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* 106, 465-475.
187. Zhang, Z.C., and Chook, Y.M. (2012). Structural and energetic basis of ALS-causing mutations in the atypical proline-tyrosine nuclear localization signal of the Fused in Sarcoma protein (FUS). *Proc Natl Acad Sci U S A* 109, 12017-12021.

Appendix

Table 1.2:

Mutations generated in this study and their nuclear localization activity.

A scheme of the structural elements of ADAR1-dsRBD3 is given above. Mutations in each of the structural elements are classified and presented with boxes in the corresponding color code. Point mutations are highlighted in bold. Deletions are indicated by dashes. The intracellular localization of the constructs tested as pyruvate kinase-myc fusion proteins is given in the right column. N= nuclear, C= cytoplasmic, N/C= impaired nuclear localization.

	N-terminal module (NTM) αN α1 β1 β2 β3 α2 C-terminal module (CTM)											
		708	720	730	740	750	760	770	780	790	800	
M708-801		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
M708A		AAMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
M709A		MAPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
P710A		MMANKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
N711A		MMPAKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
V713A		MMPNKARKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
M708A+P710A		AMANKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										NC
M709A+N711A		MAPAKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
K712A		MMPNAVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
R714A		MMPNKVAKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
K715A		MMPNKVRAIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
K712A+R714A+K715A		MMPNAVAAIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
K798A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENEAAER										N
E800A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAAR										NC
R801A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAEA										C
K712A+R714A+K715A+R801A		MMPNAVAAIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAEA										C
ΔM708-K714		-----IGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										C
E718A		MMPNKVRKIGALVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
R721A		MMPNKVRKIGELVAYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
N724A		MMPNKVRKIGELVRYLATNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
E718A+R721A+N724A		MMPNKVRKIGALVAYLATNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										N
I716N+L719S+L723N		MMPNKVRKNGESVRYNNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										C
ΔI716-N724		MMPNKVRK-----TNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										C
T725-801A		-----TNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKKQKGQEAADAALRVLIGENKAER										C
K777A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSAKQKGQEAADAALRVLIGENKAER										N
K778A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSAQKGQEAADAALRVLIGENKAER										N
K781A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSSKQGAQEAADAALRVLIGENKAER										N
K777A+K778A		MMPNKVRKIGELVRYLNTNPVGGLLEYARSHGFAAEFKLVDQSGPPHEPKFVYQAKVGGRWFPVAVCAHSAAQKGQEAADAALRVLIGENKAER										N
NTM+CTM x1RBPads2 (T98-K195)		MMPNKVRK-----KAER										C
x1RBPads2 S116G+T176R+K179I		----TLRDPKPPNQMQENPVGSLQELAVQKGWRLPEYTVAQESGPPHKREFTITCRVETFVETSGTSKQVAKRVAAEKLLTKFTISTDNIPLNKLIGNK										C
x1RBP-NLS		-----NPVGGLQELAVQKGWRLPEYTVAQESGPPHKREFTITCRVETFVETSGTSKQVAKRVAAEKLLRKFIT-----										C
FUS-NLS		MMPNKVRK-----GEHRQDR-----KAER										N
FUS-NLS+R801A		MMPNKVRK-----GEHRQDR-----KAEA										C
FUS linker		-----GEHRQDR-----										C

Table 1.3:

List of primers used for cloning of various constructs.

Sequences of forward and reverse primers used to clone the pyruvate kinase-myc fusion constructs used for nuclear localization studies and for protein expression.

PK-Myc-pcDNA3.1+constructs				
Deletion/Mutations	Forward/Primer(5'-3')	Number	Reverse/Primer(5'-3')	Number
ΔNBPΔ716-801	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3464	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
ΔHelixΔNBPΔ725-801	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3330	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
ΔHelixΔNBPΔ708-715+725-801	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3410; MJ3332	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
N-Terminal Basic Patch(NBP)ΔC-Terminal Basic Patch(CBP)ΔMutations Basic Residues				
K712A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3411	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
R714A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3412	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
K715A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3413	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
K712A+Δ714A+Δ715A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3409	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
R801A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3326	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3329
K712A+Δ714A+Δ715A+R801A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3412	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3329
K788A	CTCCGCGAAGCTTATGCGTTCTGCGGCTCGTTCTCCCAATCAAGAC	MJ3326	CTCCGCGAAGCTTATGCGTTCTGCGGCTCGTTCTCCCAATCAAGAC	MJ3497
E800A	CTCCGCGAAGCTTATGCGTTCTGCGGCTCGTTCTCCCC	MJ3326	CTCCGCGAAGCTTATGCGTTCTGCGGCTCGTTCTCCCC	MJ3498
Hydrophobic Residues				
M708A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3582	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
M709A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3583	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
P710A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3584	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
N711A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3585	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
V713A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3586	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
M708A+P710A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3587	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
M709A+N711A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3588	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
Alpha-Helix-0 Mutations Exposed Residues				
E718A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3423	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
R721A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3424	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
N724A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3425	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
E718A+Δ721A+ΔN724A	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3426	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
Hydrophobic Residues				
E716N+L719S+L723N	CTAGAAGGCCACCCACAGGGTTGGTATTGTATCTACGCTTTCGCCATTCTCTGACCTTGTGGGCATCATTAA	MJ821	CTAGAAGGCCACCCACAGGGTTGGTATTGTATCTACGCTTTCGCCATTCTCTGACCTTGTGGGCATCATTAA	MJ3822
Alpha-Helix-1 Mutations Exposed Residues				
E733A	SCAAAGCCATGGAGCGGGCGTACGCCAAAAGGCCACCCACAGG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3592; MJ3326	SCAAAGCCATGGAGCGGGCGTACGCCAAAAGGCCACCCACAGG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3589; MJ3327
R736A	SCAAAGCCATGGAGCGGGCGTACGCCAAAAGGCCACCC;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3592; MJ3326	SCAAAGCCATGGAGCGGGCGTACGCCAAAAGGCCACCC;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3590; MJ3327
E733A+R736A	SCAAAGCCATGGAGCGGGCGTACGCCAAAAGGCCACCCACAGG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3592; MJ3326	SCAAAGCCATGGAGCGGGCGTACGCCAAAAGGCCACCCACAGG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3591; MJ3327
Alpha-Helix-2 Mutations RNA-binding Residues				
K777A	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3326; MJ3493	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3492; MJ3327
K778A	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3326; MJ3494	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3492; MJ3327
K781A	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3326; MJ3496	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3492; MJ3327
K777A+K778A+K781A	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3326; MJ3495	CTGTGTGCGCAGACGGCTGGGAACAG;CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3492; MJ3327
Fusion Constructs dsFusions Constructs				
NBP+CBP	TGCGTTCTGCGTTCTCTGACCTTGTGGGATCATAAGCT	MJ3325	TGCGTTCTGCGTTCTCTGACCTTGTGGGATCATAAGCT	MJ3324
NBP+AAAA+CBP	AGCTTATGCGTTCTGCGGCGGGCGGCTCTGACCTTGTGGGATCATTAA	MJ3554	AGCTTATGCGTTCTGCGGCGGGCGGCTCTGACCTTGTGGGATCATTAA	MJ3555
NBP+GGGG+CBP	AGCTTATGCGTTCTGCAACCAACCAACACCTTCTGACCTTGTGGGATCATTAA	MJ3552	AGCTTATGCGTTCTGCAACCAACCAACACCTTCTGACCTTGTGGGATCATTAA	MJ3553
NBP+GGG+CBP	AGCTTATGCGTTCTGCAACCAACCTTCTGACCTTGTGGGATCATTAA	MJ3550	AGCTTATGCGTTCTGCAACCAACCTTCTGACCTTGTGGGATCATTAA	MJ3551
xIRBPΔ constructs				
xIRBPΔ-dsRBD2-Δ98-195(Δlex)	CGTGAAGCTTATTTTATTCATACAGCTTG	MJ1004	CGTGAAGCTTATTTTATTCATACAGCTTG	MJ1005
xIRBPΔ-dsRBD2Δadapted;ΔN112-180	GATCAAGCTTAAATCGTTGGCTCACTACAGG	MJ3891	GATCAAGCTTATGCTTAAATTTGGTAAGGAGTTTC	MJ3890
xIRBPΔ-dsRBD2Δadapted;ΔN112-180 withΔ116G+T176R+K179I	GATCAAGCTTATGCTTAAATTTGTAAGGAGTTTC	MJ3889	GATCAAGCTTATGCTTAAATTTGTAAGGAGTTTC	MJ3890
xIRBPΔ-NLSΔ2+NBP+CBP	GATCAAGCTTATGCTTAAATTTGGTAAGGAGTTTC;GATCAAGCTTATGCGTTCTGCGTTGATAAATTTGGGAAGGATTTCTCTGCACAGACCC	MJ3889; MJ3892	GATCAAGCTTATGCTTAAATTTGGTAAGGAGTTTC;GATCAAGCTTATGCGTTCTGCGTTGATAAATTTGGGAAGGATTTCTCTGCACAGACCC	MJ3890; MJ3893
FUS constructs				
NBP+FUS+CBP	GATCAAGCTTATGCGTTCTGCTTGGCATCTGCTGTGCTCACCCCT	MJ3901	GATCAAGCTTATGCGTTCTGCTTGGCATCTGCTGTGCTCACCCCT	MJ3902
NBP+FUS+CBP(R801A)	GATCAAGCTTATGCTTCTGCTTGGCATCTGCTGTGCTGTC	MJ3326	GATCAAGCTTATGCTTCTGCTTGGCATCTGCTGTGCTGTC	MJ4004
FUSΔlone	AGCTTATGCGATCTGTCTGCTCACCATA	MJ4005	AGCTTATGCGATCTGTCTGCTCACCATA	MJ4006
ΔNBPΔ constructs				
ΔNBPΔΔ708-801	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3326	CTGAAGCTTATGCGTTCTGCGTTCTCGTTCTCCCC	MJ3327
ΔAR1ΔTΔR801A	PO ₂ JTCTCCCAATCAAGACAGGAGAG	MJ4019	PO ₂ JTCTCCCAATCAAGACAGGAGAG	MJ4020
Recombinant Protein Constructs				
GST fusions				
GST-dsRBD3ΔM708-801	CATCCCGGGATGATGCCCAAGAGTCAAGGAG	MJ3322	CGATGCGGGCGCGGCTTCTGCTTCTGCTTCTCC	MJ3252
GST-dsRBD3Δ725-801	CATCCCGGGATGATGCCCAAGAGTCAAGGAG	MJ3323	CGATGCGGGCGCGGCTTCTGCTTCTGCTTCTCC	MJ3252
GST-dsRBD3ΔM708-801	GCTACGGGATCATGATGCCCAAGAGTCAAGGAG	MJ3894	GGCGGCAAGCTTACAGCTTCTGCTTCTGCTTCTCC	MJ3659
GST-dsRBD3ΔR801A	GCTACGGGATCATGATGCCCAAGAGTCAAGGAG	MJ3656	GGCGGCAAGCTTACAGCTTCTGCTTCTGCTTCTCC	MJ3657
GST-FUSΔNLS	GCTACGGGATCATGATGCCCAAGAGTCAAGGAG	MJ3656	CCGCAAGCTTACAGCTTCTGCTTCTGCTTCTCC	MJ4033
GST-FUSΔNLS(R801A)	GCTACGGGATCATGATGCCCAAGAGTCAAGGAG	MJ3656	CCGCAAGCTTACAGCTTCTGCTTCTGCTTCTCC	MJ4033
GST-FUSΔlinker	GATCCGGTGAGCAGACAGGATCGCTGAA	MJ4034	AGCTTACAGGATCTGCTGTGCTGCTCAGG	MJ4035
His fusions				
HisΔW730ΔM1	CACAAAGCTTCTTCTGACGATGTAAGCCCTTG	MJ3921	GACCAAGCTTATGATGGTGGTGGTGGATGAGATCTGAGTCCAAACACCAATAA	MJ3922

Acknowledgements

I would like to convey my sincere gratitude to my supervisor Prof. Michael Jantsch for his ideas and extended help discussing my results every Tuesday. I shall be ever thankful to him for the patience he showed in communication and most importantly for having faith in me. I would like to express my sincere thankfulness to Prof. Federic Allain from ETH Zurich, for his constant support and valuable comments. I am extremely thankful to Dr. Pierre Barraud my collaborator post-doc, for his enormous patience in dealing all NMR related queries of mine, critical appreciation and always keeping my moral high. My project was funded by FWF Funding agency. I also want to thank the DK RNA biology program for their constant support.

I have greatly benefited from the many people that have graced the benches of the Jantsch lab over the years. I would like to thank Conny, for her extreme patience for dealing with me almost every day since we were neighbors, and thank you for being my Deutsch Übersetzerin. I am thankful to Aamira and Mansoureh for having a special connection with me both in lab and personally and for our late night ventures from lab to Landstrasse for some spicy food. I am ever grateful to my lab drinking buddy, Maja to help me survive through all the failed protein expressions. It was my immense pleasure to be able to share pipettes (red one) with senior postdoc, Konstantin and (cute coverslips) with junior postdoc, Mamta and borrowed postdoc, Jacki for all the hearty laughter sessions. Mamta needs special mention for her help through the long confocal sessions and for giving me a crash Photoshop course as well. I cannot but forget to thank all Jantsch lab technicians from Hans, Dani, Peter and Eric who have always tried to help me in their own way. I want to thank especially Hans for his constant selfless support.

I would like to thank Ana Krolo for helping me with the French Press and expression plasmids and Susanne Zich for antibody and enzyme supplies when in great need. I had an extended Indian family in Vienna; a big thanks to you all: Prabha, Madhura, Raji, Vinoth and Kullu for sharing and listening both happy and sad stories.

Special thanks goes to my family back home in India (my father, my parent in-laws and my brother) for their continued love, support and encouragement. Thanks for learning to use skype, whatsapp and viber and making it possible to never feel like I was so far away. Finally, I would like to thank Arnab, my loving husband for his timeless support and for instilling the ambition to pursue a scientific career. If not for you, my world would fall apart.

Curriculum Vitae

Personal Information

Name	Silpi Banerjee
Date of Birth	4th December 1982
Nationality	Indian
Work Address	Dr. Bohrgasse 9, 5.122, 1030 Wien
Telephone	+43 1 4277 56231 2325365
Email	silpi.banerjee@univie.ac.at

Education

- 2010 –2014 **PhD** in Molecular Biology at the University of Vienna
Supervisor: Prof Michael Jantsch
Thesis: Nuclear import by an ADAR1 double stranded RNA binding domain. Student of the International Doctoral Program in RNA Biology.
- 2009 - 2010 **Master in Research (MRes)** in Medical Molecular BioSciences (Cancer studies) at Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK.
Supervisor: Prof Olaf Heidenrich
Thesis: Calpains in Acute myeloid Leukemia. Awarded NUPIS scholarship for supporting education in Newcastle.
- 2005 - 2007 **MSc.** in Biotechnology, Bharathiar University, Coimbatore, India.
Supervisor: Dr. R. Sathishkumar
Thesis: Development of a Derivative Spectroscopic method to quantify α -tocopherol in plant tissues of *Lycopersicon Esculentum*.
- 2002 - 2005 **BSc.** in Botany (Honours), Dalmia College, Rajgangpur, Sambalpur University, India.

Professional Activities

- 2008– 2009 Business Development Executive, Mott MacDonald Private Ltd, Mumbai, India, for a Biotech Project titles “ Blood Plasma Fractionation for manufacturing of Human Plasma proteins”
- 2006 Summer internship at Saha Institute of Nuclear Physics, Kolkata, India. Project Title: “Role of RNF 36 in Apoptosis”.

Publications

A bimodular nuclear localization signal assembled via an extended dsRBD acts as an RNA-sensing signal for Transportin 1

Pierre Barraud^{1*}, Silpi Banerjee^{2*}, Weam Mohamed¹, Michael F. Jantsch² & Frederic H. Allain¹.

* These authors contributed equally to this work.

¹Institute of Molecular Biology, ETH Zurich. ²Dept. of Chromosome Biology, MFPL, Vienna. Manuscript under Revision process in PNAS

Additional Achievements

- At the 18th Annual Meeting of the RNA Society, Davos, Switzerland, June 11-16, 2013 presented a poster “An emerging role for double-stranded RNA binding domains: regulating the sub-cellular localization of proteins”
- At the Gordon Research Conference on RNA Editing, January 2013 , Galveston, Houston, Texas presented a poster Title “Nuclear import by an ADAR1 double-stranded RNA binding domain”.
- At the Embo Meeting, September 10th -13th 2012, Vienna presented a poster Title “dsRBDs of ADAR1: RNA binding & Nuclear Localization Signal”.
- Selected among 50 participating candidates for International DK-RNA Biology PhD Program in 2010.
- Selected as one of the top 50 participants among a total of 550 applicants to attend the International Talent Program for Life Sciences & Health going to be held at Amsterdam, Netherlands in May 2010, organized by Hyphen Projects in cooperation with the Ministry of Economic Affairs
- Awarded the NUIPS Bursary to support Post-graduate studies at Newcastle University, UK in 2009.
- Awarded Fellowship for Summer Project at Saha Institute of Nuclear Physics, Kolkata by SINP in 2006.
- Attended the North East Postgraduate Research conference in Oct 2009, at Newcastle University.
- Participant at 3rd and 4th National Biotechnology Conference, conducted by J.N.U, New Delhi, highlighting the “Commercialization of Products”, in 2005 and by CCMB at Hyderabad highlighting on “Over-expression of Proteins” in 2006, respectively.