Solution structure of the N-terminal dsRBD of Drosophila ADAR and interaction studies with RNA

Published in final edited form as:
Biochimie. 2012 July ; 94(7): 1499–1509. doi:10.1016/j.biochi.2011.12.017.

Solution structure of the N-terminal dsRBD of *Drosophila* ADAR and interaction studies with RNA

Pierre Barraud[1], **Bret S.E. Heale**[2,*], **Mary A. O'Connell**[2], and **Frédéric H.-T. Allain**[1,†]

[1]Institute of Molecular Biology and Biophysics, ETH Zurich, CH-8093 Zürich, Switzerland [2]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine at Edinburgh University, Crewe Road, Edinburgh EH4 2XU, UK

Abstract

Adenosine deaminases that act on RNA (ADAR) catalyze adenosine to inosine (A-to-I) editing in double-stranded RNA (dsRNA) substrates. Inosine is read as guanosine by the translation machinery; therefore A-to-I editing events in coding sequences may result in recoding genetic information. Whereas vertebrates have two catalytically active enzymes, namely ADAR1 and ADAR2, *Drosophila* has a single ADAR protein (dADAR) related to ADAR2. The structural determinants controlling substrate recognition and editing of a specific adenosine within dsRNA substrates are only partially understood. Here, we report the solution structure of the N-terminal dsRNA binding domain (dsRBD) of dADAR and use NMR chemical shift perturbations to identify the protein surface involved in RNA binding. Additionally, we show that *Drosophila* ADAR edits the R/G site in the mammalian *GluR-2* pre-mRNA which is naturally modified by both ADAR1 and ADAR2. We then constructed a model showing how dADAR dsRBD1 binds to the *GluR-2* R/G stem-loop. This model revealed that most side chains interacting with the RNA sugar-phosphate backbone need only small displacement to adapt for dsRNA binding and are thus ready to bind to their dsRNA target. It also predicts that dADAR dsRBD1 would bind to dsRNA with less sequence specificity than dsRBDs of ADAR2. Altogether, this study gives new insights into dsRNA substrate recognition by *Drosophila* ADAR.

Keywords

RNA editing; ADAR; Drosophila; dsRBD; dsRBM; NMR; *GluR-2* R/G site

1. Introduction

Adenosine-to-inosine (A-to-I) RNA editing occurs via hydrolytic deamination of adenosine and is catalyzed by members of the adenosine deaminases that act on RNA (ADARs)

†Corresponding author: FH-T Allain, Institute of Molecular Biology and Biophysics, ETH Zurich, Schafmattstrasse 20, CH-8093 Zürich, Switzerland. Tel.: +41 44 633 3940, Fax: +41 44 633 1294. allain@mol.biol.ethz.ch.
*Present address: University of Utah School of Medicine, Division of Gastroenterology, Salt Lake City, Utah 84132, USA
Author contributions: P.B., M.A.O'C. and F.H.-T.A. designed the project; P.B. prepared protein and RNA samples for structural studies and ITC, measured and analyzed NMR data, performed structure calculations, structural modelling, and ITC measurements. P.B. analyzed structure and model. B.H. performed *in vivo* editing assays. P.B., M.A.O'C. and F.H.-T.A. wrote the manuscript; all authors discussed the results and approved the manuscript.

Conflict of interest; The authors declare no conflict of interest.

Accession numbers; The chemical shifts of dADAR dsRBD1 have been deposited in the BioMagResBank under accession number 17936. The coordinates of the structure have been deposited in the Protein Data Bank under accession code 2LJH.

Supplementary data; Supplementary data associated with this article can be found in the online version of the article.

Europe PMC Funders Author Manuscripts

protein family (for reviews see [1-4]). Because inosine base-pairs with cytosine, it is read as a guanine by most cellular processes, and A-to-I editing can therefore create a codon for a different amino acid, a stop codon or even a new splice-site contributing to the diversification of protein created from a single gene [1-3]. Whereas A-to-I modification is mostly non-specific within perfect dsRNA substrate, specific deamination of a single or limited set of adenosine residues often occurs in imperfect double-stranded RNA (dsRNA) regions containing bulges, loops and mismatches. Those specifically modified substrates are typically formed by base pairing between edited exons and complementary sequences in flanking introns in pre-mRNA. Importantly, the structural determinants controlling this type of site-specific RNA editing are still not completely understood [5-8].

Many editing events in both *Drosophila* and humans occur in neuronal tissues and have been shown to be important for proper functioning of the nervous system, where A-to-I editing produces functionally important isoforms of proteins involved in synaptic neurotransmission [9-17]. In vertebrates, one of the best studied ADAR substrate is the pre-mRNA encoding for subunit 2 of the α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionic acid (AMPA) subtype of glutamate receptor (*GluR-2*; also referred to as *GluR-B*) which is edited at two specific sites, *i.e.* the R/G and Q/R editing sites [9, 18, 19]. These two sites are edited with a high specificity and reduction of editing efficiency at these sites leads to dramatic effects on central nervous system [20, 21]. They are therefore often used as model system for studies of editing by ADARs.

ADARs from all organisms share a similar domain organization that includes from one to three copies of a dsRNA binding domain (dsRBD; also referred to as dsRBM for dsRNA binding motif) in their N-terminal region followed by a C-terminal adenosine deaminase catalytic domain. The role of the dsRBDs in ADAR proteins is to recognize and bind to dsRNA thereby bringing the deaminase catalytic domain to its substrate adenosine at specific editing sites. Whereas two functional enzymes are present in vertebrates (ADAR1 and ADAR2), *Drosophila* has a single ADAR protein (dADAR) related to vertebrates ADAR2 [22, 23]. The structure of mammalian ADAR2 deaminase domain [24] and ADAR2 dsRBDs have been determined in their free state [25]. Additionally, structures of ADAR2 dsRBDs have been determined in complex with the *GluR-2* R/G site RNA substrate [7]. Importantly, structural studies on ADAR dsRBDs have been limited so far to mammalian ADAR2.

The dsRBD is a ~65-75 amino acids domain with specific binding capacity for dsRNA, which is found in many eukaryotic and prokaryotic proteins presenting a large variety of functions [26-29]. The structures of different dsRBDs have been determined uncovering a mixed α/β fold with a conserved αβββα topology in which the two α-helices are packed against the three-stranded anti-parallel β-sheet [30, 31]. In addition, structures of dsRBDs have been determined in complex with dsRNA, predominantly with non-natural RNA duplexes [32-38], revealing the canonical mode of dsRNA recognition by dsRBDs. Molecular recognition is made via three regions of interaction: helix α1 and the loop between β1 and β2 contact dsRNA minor grooves at one turn of interval whereas the short loop between β3 and α2 together with the N-terminal part of helix α2 contact the dsRNA phosphate backbone across the major groove. These studies also indicated that a subclass of dsRBDs prefers stem-loops over A-form RNA helices [33, 35]. In some structures, a sequence-specific contact has been observed between a main-chain carbonyl of the β1-β2 loop and an amino group of a guanosine in the minor groove [32, 34, 36, 37]. In addition, a large variety of contacts have been described between helix α1 and both regular minor-grooves [32, 34, 36, 37] and apical loop structures [33, 35, 38]. Even if most of these contacts are not sequence-specific, it has been proposed earlier that helix α1 could help achieve substrate specific recognition via the modulation of its contacts with apical RNA

loop structures [33]. Nevertheless, the most common idea about dsRBDs is that they recognize the A-form helix of dsRNA in a sequence-independent manner [26, 29, 32]. Indeed, the majority of dsRBD-RNA interaction involve contacts with the 2′-hydroxyl groups of the ribose sugar rings and with non-bridging oxygen of the phosphodiester backbone.

However, the recent structures of ADAR2 dsRBDs bound to the *GluR-2* R/G site have shown that the binding is achieved via a direct readout of the RNA sequence in the minor groove of the dsRNA substrate [7], giving some critical insights to reconsider our current understanding of the sequence-specificity of dsRBDs. ADAR2 dsRBDs use helix α1 and the β1–β2 loop as molecular rulers to find their binding register in the RNA minor groove. Interestingly, dsRBD1 recognize a G-X9-A motif whereas dsRBD2 binds a slightly different motif: G-X8-A [7]. The length and the positioning of helix α1 relative to the β1–β2 loop appear to be the key structural determinant that control the register length of the two dsRBDs [7]. These structures provide key elements for our understanding of editing site specificity as they reveal that dsRBD2 specifically recognizes the base pair 3′ adjacent to the editing site, therefore bringing the deaminase domain close to its substrate adenosine.

The determination of the register of binding of other dsRBDs from other ADAR proteins is very attractive as it could help to understand the molecular determinant explaining the selectivity of editing, and eventually could help to predict binding sites of ADAR dsRBDs and sites of editing. However, the subtle differences in the mode of binding by different dsRBDs [7] render the prediction of the register length for a particular dsRBD only on the basis of its sequence difficult if not impossible [4]. For this reason, the determination of precise structures of other ADAR dsRBDs are essential to extend and improve our understanding of editing specificity.

In the present study, we have determined the solution structure of the first dsRBD of *Drosophila* ADAR and characterized its interaction with RNA. Additionally, we show that *Drosophila* ADAR edits the R/G site in the mammalian *GluR-2* pre-mRNA and we could build a model showing how dADAR dsRBD1 binds to the *GluR-2* R/G stem-loop. This model reveals that most side chains interacting with the RNA sugar-phosphate backbone are already oriented and ready to bind to the dsRNA partner. The model also predicts that dADAR dsRBD1 would bind to dsRNA with less sequence specificity than dsRBDs of ADAR2. Altogether, this study gives new insights into dsRNA substrate recognition by the unique ADAR protein of *Drosophila*.

2. Material and methods

2.1 Cloning, expression and purification of dADAR dsRBD1

The DNA sequence encoding the first dsRBD of dADAR (residues 48-140) (Uniprot entry Q9NII1) was cloned in *E. coli* expression vector pET28a+ between NdeI and XhoI cloning sites. The constructs contain a N-terminal tag whose sequence MGSSHHHHHHSSGLVPRGSHM includes a 6 histidine stretch used for protein purification. Proteins were overexpressed in BL21(DE3) Codon-plus (RIL) cells in either LB media or M9 minimal media supplemented with $^{15}NH_4Cl$ and ^{13}C-labelled glucose. 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 15 h after induction by centrifugation. Cell pellets were resuspended in lysis buffer (Tris-HCl pH 8.0 50 mM, NaCl 1 M, EDTA 1mM, DTT 1 mM) and lysed by sonication. Cell lysates were centrifuged 40 min at 45,000 g. 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 6.5 25 mM, KCl 25 mM, DTT 1 mM), and concentrated to ~1.2 mM with a Vivaspin 5000 MWCO (Sartorius Stedim Biotech).

2.2 NMR Spectroscopy

All NMR spectra were recorded at 298 K on Bruker AVIII-500 MHz, AVIII-600 MHz, AVIII-700 MHz and Avance-900 MHz spectrometers (all equipped with a cryoprobe). The data were processed using TOPSPIN 2.1 (Bruker) and analyzed with Sparky [39]. Protein resonances were assigned with 2D (^{1}H-^{15}N)-HSQC, 2D (^{1}H-^{13}C)-HSQC, 3D HNCA, 3D HNCACB, 3D CBCA(CO)NH, 3D (H)CCH-TOCSY, 3D [^{13}C; ^{15}N; ^{1}H] HCC(CO)NH-TOCSY, 3D [^{1}H; ^{15}N; ^{1}H] HCC(CO)NH-TOCSY, 3D NOESY-(^{1}H-^{15}N)-HSQC and two 3D NOESY-(^{1}H-^{13}C)-HSQC optimized for the observation of protons attached to aliphatic carbons and to aromatic carbons, respectively. We recorded all 3D NOESY spectra with a mixing time of 120 ms.

2.3 Protein structure determination

Automated NOE cross-peak assignments [40] and structure calculations with torsion-angle dynamics [41] were performed with the macro noeassign of the software package CYANA 2.1 [42]. Peak lists of the three NOESY spectra were generated as input with the program ATNOS [43] and manually cleaned to remove artefact peaks. The input also contained 32 hydrogen-bond restraints. Hydrogen bonded amides were identified as slowly exchanging protons in presence of D_2O. Their bonding partner was identified from preliminary structures as well as from analysis of the characteristic NOE pattern found in α-helices and β-sheets. We calculated 100 independent structures that we refined in a water shell with the program CNS 1.21 [44, 45] as previously described [46]. The 20 best energy structures were analyzed with PROCHECK-NMR [47]. Structures were visualized and figures were prepared with program PYMOL [48].

2.4 RNA sample and isothermal titration calorimetry

The modified *GluR-2* upper stem-loop RNA (GluR-2 USL) [7] was produced by *in vitro* transcription with T7 polymerase (RNA sequence 5′ GGUAGUAUAACAAUAUCCGUGUUGUUAUAGUACC 3′). It corresponds to the *GluR-2* upper stem capped with the *GluR-3* tetraloop, as previously described [7]. RNA was purified by anionexchange high-pressure liquid chromatography under denaturing conditions, as previously described [49]. The RNA was annealed by heating at 95 °C for 5 min and snap cooling on ice, to favour a stem-loop conformation.

ITC experiments were performed on a VP-ITC instrument (MicroCal) calibrated according to the manufacturer's instructions. The samples of protein and nucleic acids were prepared in and dialyzed against the ITC buffer (NaPi pH 6.5 25 mM, KCl 75 mM, 2-mercaptoethanol 2 mM). The concentration of protein and nucleic acid was determined by OD absorbance at 280 and 260 nm, respectively. The sample cell (1.4 mL) was loaded with 6 μM of *GluR-2* USL RNA; dADAR dsRBD1 concentration in the syringe was 120 μM. Titration experiments were done at 25 °C and consisted of 34 injections, each of 8 μL volume with a 5-min interval between additions. Stirring rate was 307 r.p.m. Raw data were integrated, corrected for nonspecific heats and normalized for the molar concentration. Three parameters were fitted (the association constant Ka, the binding enthalpy ΔH and the number of site N) using the equation for 1:1 binding model. Measurements were repeated three times and mean value and standard deviation were calculated for each parameter.

Europe PMC Funders Author Manuscripts

2.5 Editing Assay for GluR-2 mRNA R/G site by *Drosophila* ADAR

A plasmid encoding the rat *GluR-2* pre-mRNA encompassing the R/G editing site pcDNA3 FLAG_GRG SS [50] was used for subcloning into pAc5.1/V5-HisA vector (Invitrogen) so that it could be transfected into *Drosophila* S2 cells. Different *Drosophila* ADAR isoforms were expressed in the pRMHa3 vector which has an inducible metallothionein promoter.

S2 cells were cultured under standard conditions and transfected with Cellfectin according to the manufactures instructions. The total amount of DNA transfected was 1.25 μg. After induction of *Drosophila* ADAR with copper sulfate, cells were collected and RNA prepared using Trizol reagent (Invitrogen). Random hexamers were used to generate cDNA. PCR amplification was performed with sequence specific primers that were also used for sequencing; CAGCCAGCAGTCGTCTAATC (sense) and CGGTACGCGTAGAATCGAGA (antisense).

2.6 Modelling the interaction with dsRNA

The refined ensemble of structures of dADAR dsRBD1 was superimposed on the backbone atoms with the 20 refined structure of ADAR2 dsRBD1 in complex with *GluR-2* upper stem-loop RNA (pdb code 2L3C) [7]. This initial superposition allowed us to generate 400 starting structures of protein/RNA complexes between dADAR dsRBD1 and *GluR-2* USL RNA, each of them built as a unique pair of conformers. Each individual structure was subjected to a refinement protocol in CNS 1.21 with no experimental energy terms. First, the structures were energy minimized with a conjugate gradient minimization and subsequently a rigid body minimization with two rigid groups defined as one for the protein and one for the RNA stem-loop. Second, these minimized structures were subjected to a restrained simulated annealing protocol in implicit water. It consisted of 6 ps of dynamics at 1000 K followed by cooling to 25 K over 26 ps. Different type of restraints were applied for the interface and for the rest of the molecules; (i) the side chains of helix α1, the side chains of the $_{108}$RSKKVAR$_{114}$ motif at the N-terminal tip of helix α2 as well as the side chains and backbone of loop β1-β2 were set to unrestrained atoms; (ii) the backbone of helix α1 and the $_{108}$RSKKVAR$_{114}$ motif as well as the nucleotides in the RNA loop were harmonically restrained to their initial position, allowing small motions for these parts; (iii) all the rest of the protein and the RNA stem were set to fixed atoms. The resulting complexes were finally energy minimized and the 10 best energy structures were pooled as the refined ensemble and analyzed with PYMOL [48].

3. Results

3.1 *Drosophila* ADAR dsRBD1 adopts a canonical dsRBD fold

To get structural insights into the molecular determinants leading to *Drosophila* ADAR (dADAR) specificity, we first determined the solution structure of the first dsRBD of dADAR with 2208 NOE-derived constraints. The structure is very precise with a backbone r.m.s.d. over the entire domain (residues 64-126) of 0.30 ± 0.06 Å for the ensemble of 20 conformers (Figure 1A). Constraints also include 32 hydrogen-bond restraints to backbone amides that were established as slowly exchanging protons in presence of D_2O. Their hydrogen-bond acceptors were identified from preliminary structures. Assignment and structure calculation procedures are described in the material and method section. NMR experimental constraints and refinement statistics are presented in Table I.

The domain adopts the canonical αββββα topology with the two α-helices packed against the 3-stranded anti-parallel β-sheet (Figure 1B) [30, 31]. Hydrophobic side chains from all six secondary structure elements (α1, β1-3 and α2) contribute to the hydrophobic core of the domain (Figure 1B and 1C). Even though dADAR dsRBD1 displays the canonical

αββββα topology, slight differences with reported dsRBD structures are present. First, helix α1 is one helical turn shorter at its C-terminal tip than canonical dsRBDs like dsRBD2 of Xlrbpa [32] (Figure 2). This type of shorter helix α1 have been reported in the dsRBD structures of ADAR2 [25] and yeast RNAse III Rnt1p [35, 51]. Remarkably, helix α1 of dADAR dsRBD1 does not present an additional turn at its N-terminal extremity like it has been reported in the case of ADAR2 dsRBD1 [25] (Figure 2). The length of helix α1 is therefore similar to the one of ADAR2 dsRBD2 (Figure 2) even though its amino acid sequence is more similar to ADAR2 dsRBD1 (Figure 3E). This point will be further discussed later. Secondly, a β-bulge is present at the end of β1 with two consecutive residues facing up towards the hydrophilic side of the β-sheet, namely E81 and S82 (Figure 1C). Additionally, the β1-β2 loop which is sometimes conformationally heterogenous [25, 33, 46] presents here a well defined conformation (Figure 1A).

3.2 *Drosophila* ADAR dsRBD1 interacts with GluR-2 R/G upper stem loop RNA

So far, the only structural data of ADAR dsRBDs bound to an RNA substrate are the structures of ADAR2 dsRBDs bound to the *GluR-2* R/G stem-loop [7]. Therefore, we decided to characterize the RNA binding properties of dADAR dsRBD1 with the same RNA, namely the *GluR-2* R/G stem-loop, in order to benefit from the available structural information. In ADAR2, dsRBD1 has been shown to bind to the upper part of the *GluR-2* R/G stem-loop (in short the upper stem-loop or USL), whereas dsRBD2 binds to the lower part of the stem-loop [25]. We first tested whether dADAR dsRBD1 would also interact with the *GluR-2* USL. We therefore used the same modified *GluR-2* stem-loop used in the structural study of ADAR2 [7], namely a *GluR-2* stem-loop in which the apical pentaloop has been replaced by the tetraloop found in the *GluR-3* R/G editing site (Figure 3A). This construct gave NMR data of better quality [7]. Isothermal titration calorimetry (ITC) confirmed the strong protein-RNA interaction (Kd = 0.40 ± 0.05 μM) with two dsRBDs bound per RNA USL (Figure 3D). This protein:RNA ratio of 2:1 is interesting as ADAR proteins are known to be active as dimers and thus two pairs of dsRBDs are certainly binding together on ADAR RNA substrates [52, 53].

3.3 Mapping the RNA binding surface of *Drosophila* ADAR dsRBD1

To get a more detailed view of the interaction between dADAR dsRBD1 and *GluR-2* USL, we performed NMR titrations by adding step by step either ^{15}N-labelled dsRBD1 into a *GluR-2* USL RNA sample or by adding RNA into a ^{15}N-protein sample (Figure 3B). In both situations, the exchange regime corresponds to a fast to intermediate exchange regime on the NMR chemical shift time-scale, as revealed by broadening and disappearance of most resonances of the protein amides for substoichiometric amount of RNA and their reappearance for a RNA:protein ratio of 2:1 and above. However, some specific amide resonances seem not to come back even in excess of RNA. This severe broadening leading to unobservable signals is sometimes observed in fast to intermediate exchange regime for resonances that experience large chemical shift variation between free and bound state [54, 55], meaning for residues of the interface. We cannot exclude other sources of broadening to be responsible for the disappearance of these signals, such as residual dynamics at the interface or multiple binding conformations. Indeed, the ITC titration showed that two protein sites exist on the USL RNA (Figure 3D), suggesting that an exchange between the two different bound states could possibly contribute to the resonances broadening. In any case these signals must belong to residues of the protein-RNA interface. The analysis of these NMR titrations enables us to determine the regions of the protein involved in RNA binding (Figure 3B). These residues forming the protein-RNA interface are mapped on the surface of the protein (Figure 3C) as well as on the dsRBD amino acid sequence (Figure 3E). These residues belong to the three canonical regions known to participate in dsRNA binding, namely the helix α1 and the β1-β2 loop which interact in two consecutive minor

grooves of a RNA helix, and the β3-α2 loop together with the N-terminal tip of helix α2 which interact across the RNA major groove [7, 32, 33, 35]. It is important to note that this NMR titration does not permit to directly conclude about the participation in dsRNA binding of residues V87 and H88 located in the β1-β2 loop because in one case (V87) we miss the amide assignment in the free form and because in the other case (H88) the amide signal lies in the middle and overlapped region of the ^{15}N-^{1}H HSQC spectrum (121.9 – 8.35 ppm), making impossible the assignment of its bound state on the only basis of the RNA titration. Therefore, the β1-β2 loop does not appear as drastically affected upon RNA binding on Figure 3C. But the disappearance of A89 amide (Figure 3B) is a clear indication that the β1-β2 loop is definitely involved in RNA binding. Altogether, this shows that dADAR dsRBD1 binds to *GluR-2* USL RNA using a canonical mode of interaction.

3.4 *Drosophila* ADAR edits mammalian GluR-2 R/G site

Having shown that dADAR dsRBD1 binds to *GluR-2* R/G USL RNA, we were interested to know whether this mammalian site of editing, which is naturally modified by both ADAR1 and ADAR2 could be edited by the *Drosophila* enzyme. This question was of primary interest in order to support our structural NMR characterization, and has of course important implications for the understanding of the molecular determinant at the origin of the distinct selectivity of the different ADAR enzymes. Indeed, we recently demonstrated that *Drosophila* ADAR edits other sites in mammalian *GluR-2* pre-mRNA, namely the Q/R site which is a natural substrate of ADAR2 and an intronic hotspot 1 site which is preferentially modified by ADAR1 [22], showing that dADAR is less selective than mammalian ADARs. Additionally, mammalian ADAR2 can modify highly specific sites of editing in *Drosophila* substrates [22], indicating an overlap in the editing specificity of dADAR and ADAR2.

We thus investigated with an *ex vivo* editing assay the capacity of dADAR to edit the *GluR-2* R/G site. We co-transfected plasmids encoding the *Drosophila* ADAR enzyme together with a minigene encoding the *GluR-2* R/G site into *Drosophila* S2 cells, and subsequently determine the editing levels by sequencing as described in material and methods. Interestingly, dADAR efficiently modifies the mammalian *GluR-2* R/G site, with approximately 90% editing efficiency for the highest dADAR expression level (Figure 4 and Supplementary Figure S1). The editing efficiency of two different isoforms of dADAR, namely the S isoform and the G isoform, has been investigated. However, dADAR can edit its own pre-mRNA at the so-called S/G site [23], therefore when dADAR S isoform is transfected into cells a mixture of unedited and edited isoforms is obtained, refers to as dADAR S/G. We noticed that this mixture of isoform is more active at editing mammalian *GluR-2* R/G site than the completely edited isoform dADAR G (Figure 4) which is consistent with previous observations [14]. Altogether, this further confirms the overlapping specificity of editing between *Drosophila* and mammalian ADARs and suggests that RNA substrate recognition by different ADAR proteins must share common structural features.

3.5 Modelling the interaction of *Drosophila* ADAR dsRBD1 with RNA

In order to characterize the molecular determinants leading to dsRNA binding by dsRBD1 of *Drosophila* ADAR, we decided to take advantage of the available structure of the complex between ADAR2 dsRBD1 and the *GluR-2* USL, and to build a structural model of dADAR dsRBD1 bound to the same RNA stem-loop. The ITC titration showed that two dsRBDs are bound per *GluR-2* USL RNA (Figure 3D). Our goal was not to build a model that would describe how these two dsRBDs would arrange on a single stem-loop, but to compare the molecular determinants at the origin of RNA recognition in the case of dADAR and ADAR2. Ultimately, the structural model could help us to understand the selectivity of interaction and particularly the length of the binding register of *Drosophila* dsRBD1. The model was therefore build by superimposing the dADAR dsRBD1 structure onto the one of

ADAR2 dsRBD1. The resulting bundle of structures were refined in implicit water with a restrained simulated annealing protocol as described in material and methods. Different type of restraints were applied for the interface as determined by NMR titration experiments (Figure 3) and for the rest of the molecule. Shortly, the side chains of helix α1 and of the $_{108}$RSKKxAR$_{114}$ motif at the N-terminal tip of helix α2, as well as the entire β1-β2 loop were allowed to move freely; whereas the rest of the protein was fixed during the simulated annealing protocol. Initially, the protein and the RNA were also allowed to adjust their relative orientation as two rigid blocks in a rigid body minimization. Finally, the 10 lowest energy structures of the protein-RNA complex (Figure 5B) were compared to the initial protein bundle and protein-RNA contacts were carefully analyzed.

The superimposition of the initial structures of free dsRBD1 with the refined model bound to dsRNA revealed that most side chains interacting with the sugar-phosphate backbone are already correctly oriented and ready to bind to the dsRNA. In particular the long side chains of K110, K111 and R114 are only slightly displaced to adapt to the RNA phosphate backbone over the major-groove (Figure 6). This situation is not unique to dADAR dsRBD1, as the stacking of positively charged residues equivalent to K110 and R114 (dADAR dsRBD1 residue numbering – Figure 3E) over conserved aromatic rings, namely Y78 for K110 and F92 for R114, allows these long side chains to adopt an elongated conformation (Figure 6). It has been proposed earlier that these aromatic side chains would therefore be essential for proper orientation of the long positively charged residues within the KKxAK consensus motif at the N-terminal tip of helix α2 [32], and their importance for the dsRNA binding capacity of dsRBDs have been also demonstrated [30, 33, 56]. Surprisingly, the side chain of K111 also presents an elongated conformation in the free form which in this case cannot be explained by a stacking onto an aromatic residue. In addition, the β1-β2 loop is in our structure also properly structured in the free form as it needs only small amplitude movements to adapt to its RNA partner (Figure 6). This situation does not seem to be shared by all dsRBDs as the β1-β2 loop has been sometimes reported to be structurally heterogeneous in the free form, therefore requiring larger structural rearrangement upon dsRNA binding [25, 30, 31, 33].

Non-sequence specific contacts are found between the $_{108}$RSKKVAR$_{114}$ motif and the sugar phosphate backbone across the major groove of the RNA USL. Side chains of K110, K111 and R114 contact non-bridging oxygen of the phosphate backbone (Figure 5A). Those electrostatic interactions form the typical pattern of dsRNA recognition via dsRBDs. Additionally, R108 contacts the 2′-hydroxyl of nucleotide G22 (Figure 3A and 5A). Contacts are also observed between the upper part of helix α1 and the RNA tetraloop (Figure 5A) and are consistent with the chemical shift changes observed in the loop region of the RNA (U34, C35 and C36) upon interaction with dADAR dsRBD1 (Supplementary Figure S2). Interestingly, a sequence specific contact is found between the main-chain carbonyl of V87 in the β1-β2 loop and the amino group of G22 in the minor groove (Figure 5A and D), as previously seen in the ADAR2 complex [7]. This contact is observed in 4 of the 10 structures of the bundle. Remarkably, in the ADAR2 dsRBD1 USL structure, an other sequence specific contact is formed in helix α1 between nucleotide A32 and a methionine side chain [7], but no similar contact is observed in our model (Figure 5A and C). Indeed, the residue equivalent to the methionine contacting adenine 32 in the minor groove is an alanine, namely A67 (Figure 3E and 5A), the side chain of which is too short to contact RNA bases in the minor groove. Chemical shift perturbation between free and bound form of the protein confirms the participation of A67 methyl group in the protein RNA interface and excludes the adjacent methionine M68 (Figure 3E and 5A) to be part of the interaction as its methyl group is not affected upon protein binding (Supplementary Figure S2). Overall, the model suggests that dADAR dsRBD1 which makes only one sequence specific contact

via its β1-β2 loop would bind to dsRNA with less specificity than ADAR2 dsRBD1 and dsRBD2.

4. Discussion

4.1 Structural comparison of dADAR and ADAR2 dsRBD1 structures

In this study, we determined the solution structure of the N-terminal dsRBD of *Drosophila* ADAR. This dsRBD shares 60 % identity and 76 % similarity with ADAR2 dsRBD1 (Figure 3E), whose structure was also determined by solution NMR [25]. However, in the present structure, more NOE-derived constraints were extracted from the NOESY spectra, 2208 compared to 1754, and more long range NOE distances were used (|i-j| ≥ 5 residues), 726 compared to 596 (Table I and [25]). As a result, the present structure is more precise with an heavy atom r.m.s.d. of 0.56 ± 0.07 Å over the entire domain (ADAR2 dsRBD1 heavy atom r.m.s.d. = 1.28 ± 0.23 Å).

The sequences at the N-terminal tip of helix α1 are very similar (Figure 3E), but nonetheless the two structures differ substantially, with helix α1 of ADAR2 forming an additional turn at its N-terminus (Figure 2). This difference may have important implications for RNA recognition as the length and the positioning of helix α1 appear to be the key structural element that determine the register length of different dsRBDs [7]. In order to understand the molecular basis of this important difference, we compared the two sets of NOE-derived constraints that define the structure of helix α1. It appeared that the additional turn in ADAR2 dsRBD1 is not supported by NOE-derived constraints, but is defined by hydrogen-bonds and torsion angle restraints. These hydrogen-bonds and torsion angle restraints were applied based on secondary chemical shift values (R. Stefl, personal communication). In dADAR dsRBD1, the residues preceding N64 (Figure 3E) were not visible in the NMR spectra and we could thus not obtain the assignment of their Cα and Cβ chemical shifts. We can therefore not evaluate if real differences exist between the two helix α1 in dADAR dsRBD1 and ADAR2 dsRBD1, or if the observed differences are only coming from the different protocols that were used for structure calculation.

4.2 Specificity of editing

One important question with dsRBDs of ADAR proteins, is to understand the molecular basis of their specificity that may result in the modification of a single adenosine in an entire dsRNA structure, like in the *GluR-2* R/G stem-loop. Also in *Drosophila* there is extensive site-specific editing events as 972 editing sites were identified within 597 genes that lead to 630 codon changes [17]. Therefore the question arises how can an enzyme remain specific yet recognize so many different substrates. Recently, structures of ADAR2 dsRBDs in complex with natural dsRNA substrates revealed that this sequence specific substrate recognition is achieve via a direct readout of the dsRNA minor groove [7]. In the present study, we built a model reporting the interaction between dADAR dsRBD1 and a dsRNA substrate modified by this enzyme. This model suggests that a main-chain carbonyl group in the β1-β2 loop is contacting the amino group of a guanosine in the minor groove. This type of contact has been observed in both dsRBDs of ADAR2 [7], but also in various crystal structures of dsRBDs in complex with dsRNA [32, 34, 36, 37], and therefore seems to be a common feature of dsRBD-dsRNA sequence-specific recognition.

However, our model predicts no other sequence-specific contact that could be formed via helix α1. Interestingly, sequence specific contacts involving helix α1 seem to be more diverse, as different type of residue for base recognition, namely methionine and glutamine, have been observed in the reported structures [7, 34, 36]. Therefore, we believe that dADAR dsRBD1 binds to dsRNA with less selectivity than other dsRBDs that use two sequence-

specific contacts to find a single binding register in the RNA minor groove. First, this could explain why in NMR titrations, we do not observe all the protein signals to come back even in excess of RNA (Figure 3B). The domain would fail to be fixed at a single position and might slide back and forth between different sites, resulting in a broadening of the resonances at the protein-RNA interface. Additionally, this might also account for the looser specificity of dADAR as compared with ADAR1 and ADAR2 [22]. Nevertheless, the affinity of binding between dADAR dsRBD1 and *GluR-2* USL is strong (Kd = 0.40 ± 0.05 μM) and is very close to the affinity that have been determined for ADAR2 dsRBD1 binding to the same RNA substrate (Kd = 0.33 ± 0.03 μM) [7]. This high affinity of binding can account for the high editing efficiency *Drosophila* ADAR has for the human *GluR-2* R/G site (Figure 4).

The lack of a second sequence-specific contact from helix α1 results from a single residue difference where a methionine in ADAR2 is changed to an alanine in dADAR (A67 - Figure 3E). Interestingly, such a mutation M67A (dADAR residue numbering) has been introduced in ADAR2 and results in a five fold decrease in editing efficiency for the *GluR-2* R/G site [7]. It would be interesting to test whether such a mutant with less editing efficiency on the R/G site would also start editing other adenosines with a higher efficiency, and especially some ADAR1 specific sites that are efficiently edited by the *Drosophila* enzyme [22].

4.3 Dimerization of ADAR proteins on RNA substrates

Intrigued by the presence of two binding sites of high affinity (Figure 3D) for dADAR dsRBD1 on the *GluR-2* USL, where only one high affinity binding site was observed in our previous study of ADAR2 dsRBD1 [7], we decided to examine the structural aspects of how two dsRBDs could bind together to the *GluR-2* USL. This is of particular interest in the context of ADAR dimerization. ADAR proteins are indeed known to be active as dimers [52, 53] and therefore two pairs of dsRBDs must be accommodated onto ADAR RNA substrates. Moreover, even if the footprint of the RNA in its bound form (RNA:protein ratio of 1:2 – Supplementary Figure S2) is consistent with the binding of one protein in the loop region of the *GluR-2* RNA (as represented in Figure 5A), the disappearance of the signal of U20 cannot be explained by this binding event and comes probably from the binding of a second protein molecule onto the *GluR-2* USL RNA. The model we built for a single dsRBD bound to the *GluR-2* USL revealed that only one sequence specific contact towards an amino group of a guanosine in the minor groove would be sufficient for recognition and binding (Figure 5). Therefore, we analyzed all the possibilities of binding for two dsRBDs contacting a guanosine in the minor groove via their β1–β2 loop. There were in total three possibilities, but in one of those (binding to G41 in the minor groove) the two proteins would clash over the major groove, the binding sites being too close to each other. The two remaining possibilities are shown in Figure 7. In one of those, the two dsRBDs have the same orientation and in the other adopt an anti-parallel orientation. Even if it might not reflect dimerization *in vivo*, we consider that the first model is more likely to describe the binding of the second dsRBD that we observed in our *in vitro* system, and this for two reasons: (i) the position of the second dsRBD with contact to G19 in the first model would explain the disappearance of the resonances of U20 in the 2D TOCSY of the complex (Supplementary Figure S2); (ii) the minor groove at the bottom of the RNA stem would not be long enough for an optimal interaction with helix α1 in the second model (Figure 7C). In the context of the full length *GluR-2* RNA, the situation is likely to be different. Indeed, on the one hand the stem would be longer, enabling an optimal accommodation of helix α1 in the second model (Figure 7C), and on the other hand, G19 which is contacted in the first model, is not present in the natural *GluR-2* sequence. It has been introduced here together with G18 as starting nucleotides for an optimal *in vitro* RNA transcription (Figure 3A).

Therefore, even if our *in vitro* data are consistent with the first model (Figure 7A and B), we do not pretend here to describe the *in vivo* dimerization of the protein. Note that no change in the resonances of U20 was observed in the case of ADAR2 dsRBD1, consistent with the fact that only one high affinity binding site exists for ADAR2 dsRBD1 on the *GluR-2* USL [7].

Altogether, this suggests that the formation of an active dimer complex can take advantage of two different situations, either involving highly specific dsRBDs which recognize their substrate via two sequence-specific contacts (like the ones of ADAR2), or involving less specific dsRBDs which recognize their substrate with only one sequence specific contact (like dsRBD1 of *Drosophila* ADAR). In the first situation, dimerization would occur via a first strong binding of two dsRBDs to highly specific sites and followed by weaker binding of the same two dsRBDs to suboptimal binding sites. In the second case, dimerization will occur by simultaneous binding of four dsRBDs with medium specificity allowing equally efficient dimerization on the RNA substrate but looser specificity. This looser specificity of dADAR dsRBD1 could explain the broader range of substrates efficiently edited by the unique ADAR protein of *Drosophila* that is known to be able to modify substrates specific to each human enzyme ADAR1 or ADAR2.

5. Conclusion

In summary, the results presented in this paper report the structure of dADAR dsRBD1 and the characterisation of its binding properties to dsRNA. We also report that the *Drosophila* enzyme can efficiently modify the mammalian *GluR-2* R/G site showing an overlap in the editing specificity of mammalian and *Drosophila* ADAR. Finally, we constructed a model accounting for the interaction between dADAR dsRBD1 and RNA. This model reinforces our current understanding of sequence-specific RNA recognition by dsRBDs. Indeed, it suggests that the interaction of a carbonyl group in the β1-β2 loop with an amino group of a guanosine in the minor groove of the RNA would be a widespread feature of dsRBD-dsRNA sequence-specific recognition, whereas the various type of sequence-specific contacts involving helix α1 would allow the recognition of different types of RNA sequences. More structures of dsRBDs in their free form or in complex with natural dsRNA target will certainly help to extend and refine our appreciation of sequence-specific RNA recognition by dsRBDs and eventually to allow the prediction of binding sites of dsRBD with an increased precision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to Grégoire Masliah for helpful discussions on simulated annealing protocols, to Julien Boudet for technical assistance and invaluable advices on ITC measurements and to Christophe Maris for his remarkable and very useful dedication to the HPLC maintenance. We thank all the members of the Allain laboratory for fruitful discussions. Research in the Allain laboratory is supported by the Swiss National Science Foundation (Nr 31003ab-133134) and the SNF-NCCR structural biology. P.B. was supported by the Postdoctoral ETH Fellowship Program. M.O.C. was supported by funding from the Medical Research Council (U1275.1.5.1.1). B.H. was supported by a Fellowship from the Marie Curie Foundation.

Abbreviations

ADAR Adenosine deaminase that act on RNA

dADAR	*Drosophila* ADAR
dsRBD	double-stranded RNA binding domain
r.m.s.d.	root mean square deviation

References

[1]. Bass BL. RNA editing by adenosine deaminases that act on RNA. Annu Rev Biochem. 2002; 71:817–846. [PubMed: 12045112]

[2]. Valente L, Nishikura K. ADAR gene family and A-to-I RNA editing: diverse roles in posttranscriptional gene regulation. Prog Nucleic Acid Res Mol Biol. 2005; 79:299–338. [PubMed: 16096031]

[3]. Nishikura K. Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochemm. 2010; 79:321–349. [PubMed: 20192758]

[4]. Barraud P, Allain FH. ADAR Proteins: Double-stranded RNA and Z-DNA Binding Domains. Curr Top Microbiol Immunol. 2012; 353:35–60. [PubMed: 21728134]

[5]. Lehmann KA, Bass BL. The importance of internal loops within RNA substrates of ADAR1. J Mol Biol. 1999; 291:1–13. [PubMed: 10438602]

[6]. Stephens OM, Haudenschild BL, Beal PA. The binding selectivity of ADAR2's dsRBMs contributes to RNA-editing selectivity. Chem Biol. 2004; 11:1239–1250. [PubMed: 15380184]

[7]. Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB, Allain FH-T. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. Cell. 2010; 143:225–237. [PubMed: 20946981]

[8]. Eggington JM, Greene T, Bass BL. Predicting sites of ADAR editing in double-stranded RNA. Nat Commun. 2011; 2:319. [PubMed: 21587236]

[9]. Sommer B, Köhler M, Sprengel R, Seeburg PH. RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. Cell. 1991; 67:11–19. [PubMed: 1717158]

[10]. Burns CM, Chu H, Rueter SM, Hutchinson LK, Canton H, Sanders-Bush E, Emeson RB. Regulation of serotonin-2C receptor G-protein coupling by RNA editing. Nature. 1997; 387:303–308. [PubMed: 9153397]

[11]. Hoopengardner B, Bhalla T, Staber C, Reenan R. Nervous system targets of RNA editing identified by comparative genomics. Science. 2003; 301:832–836. [PubMed: 12907802]

[12]. Bhalla T, Rosenthal JJC, Holmgren M, Reenan R. Control of human potassium channel inactivation by editing of a small mRNA hairpin. Nat Struct Mol Biol. 2004; 11:950–956. [PubMed: 15361858]

[13]. Palladino MJ, Keegan LP, O'Connell MA, Reenan RA. A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. Cell. 2000; 102:437–449. [PubMed: 10966106]

[14]. Keegan LP, Brindle J, Gallo A, Leroy A, Reenan RA, O'Connell MA. Tuning of RNA editing by ADAR is required in Drosophila. EMBO J. 2005; 24:2183–2193. [PubMed: 15920480]

[15]. Paro S, Li X, O'Connell MA, Keegan LP. Regulation and Functions of ADAR in Drosophila. Curr Top Microbiol Immunol. 2012; 353:221–236. [PubMed: 21761288]

[16]. Stapleton M, Carlson JW, Celniker SE. RNA editing in Drosophila melanogaster: New targets and functional consequences. RNA. 2006; 12:1922–1932. [PubMed: 17018572]

[17]. Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, et al. The developmental transcriptome of Drosophila melanogaster. Nature. 2011; 471:473–479. [PubMed: 21179090]

[18]. Higuchi M, Single FN, Köhler M, Sommer B, Sprengel R, Seeburg PH. RNA editing of AMPA receptor subunit GluR-B: a base-paired intron-exon structure determines position and efficiency. Cell. 1993; 75:1361–1370. [PubMed: 8269514]

Europe PMC Funders Author Manuscripts

[19]. Lomeli H, Mosbacher J, Melcher T, Höger T, Geiger JR, Kuner T, Monyer H, Higuchi M, Bach A, Seeburg PH. Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. Science. 1994; 266:1709–1713. [PubMed: 7992055]

[20]. Maas S, Kawahara Y, Tamburro KM, Nishikura K. A-to-I RNA editing and human disease. RNA Biol. 2006; 3:1–9. [PubMed: 17114938]

[21]. Hood JL, Emeson RB. Editing of Neurotransmitter Receptor and Ion Channel RNAs in the Nervous System. Curr Top Microbiol Immunol. 2012; 353:61–90. [PubMed: 21796513]

[22]. Keegan LP, McGurk L, Palavicini JP, Brindle J, Paro S, Li X, Rosenthal JJC, O'Connell MA. Functional conservation in human and Drosophila of Metazoan ADAR2 involved in RNA editing: loss of ADAR1 in insects. Nucleic Acids Res. 2011; 39:7249–7262. [PubMed: 21622951]

[23]. Palladino MJ, Keegan LP, O'Connell MA, Reenan RA. dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. RNA. 2000; 6:1004–1018. [PubMed: 10917596]

[24]. Macbeth MR, Schubert HL, Vandemark AP, Lingam AT, Hill CP, Bass BL. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. Science. 2005; 309:1534–1539. [PubMed: 16141067]

[25]. Stefl R, Xu M, Skrisovska L, Emeson RB, Allain FH-T. Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. Structure. 2006; 14:345–355. [PubMed: 16472753]

[26]. Fierro-Monti I, Mathews MB. Proteins binding to duplexed RNA: one motif, multiple functions. Trends Biochem Sci. 2000; 25:241–246. [PubMed: 10782096]

[27]. Saunders LR, Barber GN. The dsRNA binding protein family: critical roles, diverse cellular functions. FASEB J. 2003; 17:961–983. [PubMed: 12773480]

[28]. Chang K-Y, Ramos A. The double-stranded RNA-binding motif, a versatile macromolecular docking platform. FEBS J. 2005; 272:2109–2117. [PubMed: 15853796]

[29]. Stefl R, Skrisovska L, Allain FH-T. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO Rep. 2005; 6:33–38. [PubMed: 15643449]

[30]. Bycroft M, Grunert S, Murzin AG, Proctor M, St Johnston D. 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. 1995; 14:3563–3571. [PubMed: 7628456]

[31]. Kharrat A, Macias MJ, Gibson TJ, Nilges M, Pastore A. Structure of the dsRNA binding domain of E. coli RNase III. EMBO J. 1995; 14:3572–3584. [PubMed: 7628457]

[32]. Ryter JM, Schultz SC. Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. EMBO J. 1998; 17:7505–7513. [PubMed: 9857205]

[33]. Ramos A, Grünert S, Adams J, Micklem DR, Proctor MR, Freund S, Bycroft M, St Johnston D, Varani G. RNA recognition by a Staufen double-stranded RNA-binding domain. EMBO J. 2000; 19:997–1009. [PubMed: 10698941]

[34]. Blaszczyk J, Gan J, Tropea JE, Court DL, Waugh DS, Ji X. Noncatalytic assembly of ribonuclease III with double-stranded RNA. Structure. 2004; 12:457–466. [PubMed: 15016361]

[35]. Wu H, Henras A, Chanfreau G, Feigon J. 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 USA. 2004; 101:8307–8312. [PubMed: 15150409]

[36]. Gan J, Tropea JE, Austin BP, Court DL, Waugh DS, Ji X. Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. Cell. 2006; 124:355–366. [PubMed: 16439209]

[37]. Yang SW, Chen HY, Yang J, Machida S, Chua NH, Yuan YA. Structure of Arabidopsis HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. Structure. 2010; 18:594–605. [PubMed: 20462493]

[38]. Wang Z, Hartman E, Roy K, Chanfreau G, Feigon J. Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. Structure. 2011; 19:999–1010. [PubMed: 21742266]

[39]. Goddard, T.; Kneller, D. SPARKY 3. University of California; San Francisco: 2006.

[40]. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol. 2002; 319:209–227. [PubMed: 12051947]

[41]. Guntert P, Mumenthaler C, Wuthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol. 1997; 273:283–298. [PubMed: 9367762]

[42]. Guntert P. Automated NMR structure calculation with CYANA. Methods Mol Biol. 2004; 278:353–378. [PubMed: 15318003]

[43]. Herrmann T, Guntert P, Wuthrich K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. J Biomol NMR. 2002; 24:171–189. [PubMed: 12522306]

[44]. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr. 1998; 54:905–921. [PubMed: 9757107]

[45]. Brunger AT. Version 1.2 of the Crystallography and NMR system. Nat Protoc. 2007; 2:2728–2733. [PubMed: 18007608]

[46]. Barraud P, Emmerth S, Shimada Y, Hotz H-R, Allain FH-T, Bühler M. An extended dsRBD with a novel zinc-binding motif mediates nuclear retention of fission yeast Dicer. EMBO J. 2011; 30:4223–4235. [PubMed: 21847092]

[47]. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996; 8:477–486. [PubMed: 9008363]

[48]. DeLano, WL. The PyMOL molecular graphics system. 2002.

[49]. Duss O, Maris C, von Schroetter C, Allain FH-T. A fast, efficient and sequence-independent method for flexible multiple segmental isotope labeling of RNA using ribozyme and RNase H cleavage. Nucleic Acids Res. 2010; 38:e188. [PubMed: 20798173]

[50]. Bratt E, Ohman M. Coordination of editing and splicing of glutamate receptor pre-mRNA. RNA. 2003; 9:309–318. [PubMed: 12592005]

[51]. Leulliot N, Quevillon-Cheruel S, Graille M, van Tilbeurgh H, Leeper TC, Godin KS, Edwards TE, Sigurdsson STL, Rozenkrants N, Nagel RJ, Ares M, Varani G. A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. EMBO J. 2004; 23:2468–2477. [PubMed: 15192703]

[52]. Cho D-SC, Yang W, Lee JT, Shiekhattar R, Murray JM, Nishikura K. Requirement of dimerization for RNA editing activity of adenosine deaminases acting on RNA. J Biol Chem. 2003; 278:17093–17102. [PubMed: 12618436]

[53]. Gallo A, Keegan LP, Ring GM, O'Connell MA. An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. EMBO J. 2003; 22:3421–3430. [PubMed: 12840004]

[54]. Dominguez C, Schubert M, Duss O, Ravindranathan S, Allain FH-T. Structure determination and dynamics of protein-RNA complexes by NMR spectroscopy. Prog Nucl Magn Reson Spectrosc. 2011; 58:1–61. [PubMed: 21241883]

[55]. Barraud P, Dardel F, Tisné C. Optimizing HSQC experiment for the observation of exchange broadened signals in RNA-protein complexes. Comptes Rendus Chimie. 2008; 11:474–479.

[56]. Krovat BC, Jantsch MF. Comparative mutational analysis of the double-stranded RNA binding domains of Xenopus laevis RNA-binding protein A. J Biol Chem. 1996; 271:28112–28119. [PubMed: 8910425]

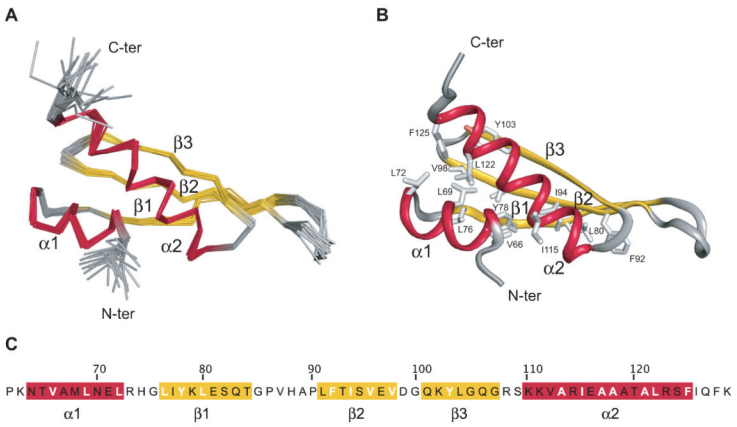

Figure 1. NMR solution structure of the N-terminal dsRBD of *Drosophila* ADAR

(**A**) NMR ensemble. Overlay of the 20 final structures with α-helices 1 and 2 *in red* and β-strands 1-3 *in yellow*. (**B**) Cartoon representation of the lowest energy structure with the same colour for the secondary structure elements. The residues forming the hydrophobic core of the domain are depicted as white sticks. (**C**) Sequence of the dsRBD1 of dADAR with the corresponding secondary structure elements indicated with the same colors as in panels A and B. The residues forming the hydrophobic core of the domain are displayed in white bold letters. Amino-acid number refer to the *Drosophila* ADAR protein (Uniprot entry Q9NII1).

Europe PMC Funders Author Manuscripts

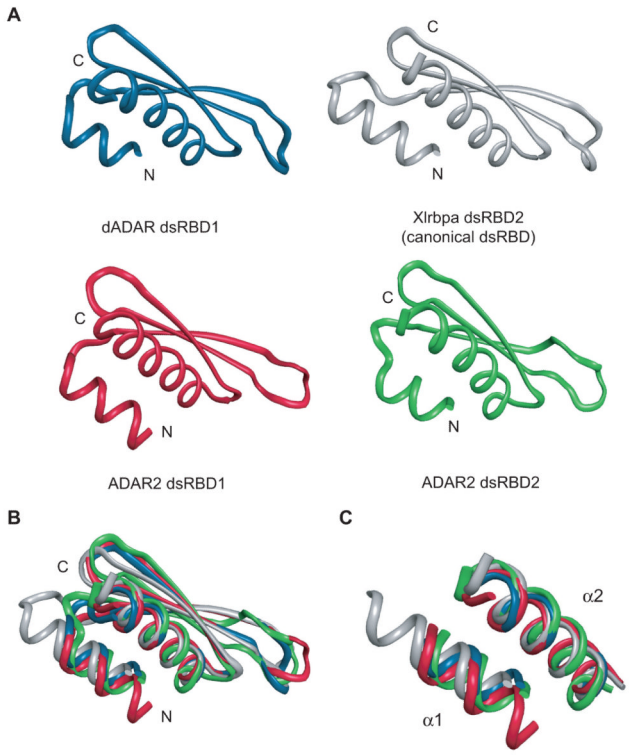

Figure 2. Structural comparison with other dsRBDs

(**A**) Structure of dADAR dsRBD1 *in blue* (PDB code 2LJH), Xlrbpa dsRBD2 *in grey* (PDB code 1DI2), ADAR2 dsRBD1 *in red* (PDB code 2B7T) and ADAR2 dsRBD2 *in green* (PDB code 2B7V). (**B**) Superimposition of the four dsRBDs with the same colour-code as in panels **A**. The structures were superimposed over Cα atoms over the entire domain. (**C**) Comparison of the position and length of helix 1 and helix 2.

Europe PMC Funders Author Manuscripts

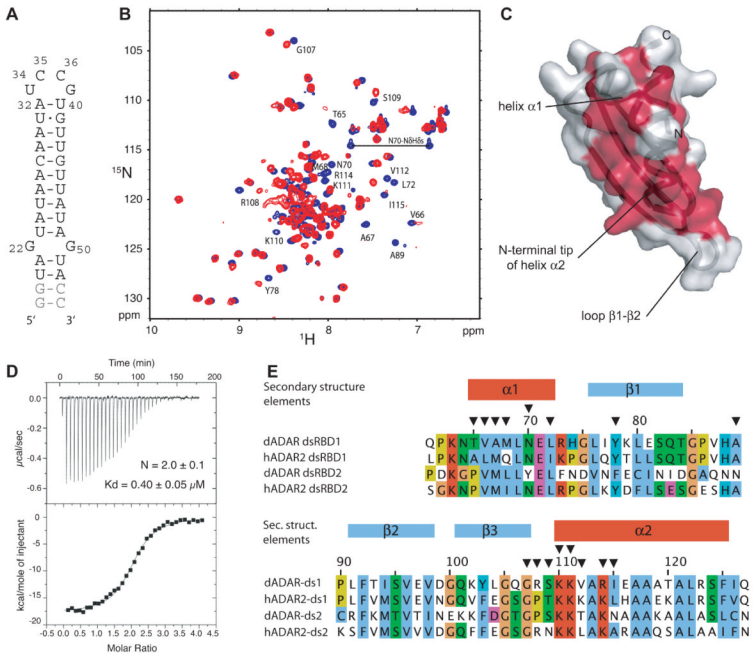

Figure 3. *Drosophila* ADAR dsRBD1 interacts with the *GluR-2* R/G site upper stem-loop

(**A**) Sequence and secondary structure of the *GluR-2* R/G upper stem-loop (USL). Numbering correspond to the entire *GluR-2* R/G stem-loop [7]. (**B**) Superimposition of ^{1}H-^{15}N HSQC spectra representing NMR titration of the ^{15}N-labelled dADAR dsRBD1 protein with increasing amount of unlabelled *GluR-2* R/G USL. For simplification, only the free state and the RNA saturated state of the protein (RNA:protein ratio of 4:1) are represented in blue and red, respectively. Resonances that do not reappear at a RNA:protein ratio > 2:1 or that present chemical shift perturbation > 0.1 ppm are labelled with their amino-acid number. These residues forming the interface of interaction with the RNA are reported in red on the surface of the dsRBD structure of panel **C** and marked with black arrows on the dsRBD sequence of panel **E**. (**C**) Chemical shift perturbations upon RNA binding mapped to the surface of the protein identifies the RNA binding surface. The three canonical region of interaction are involved in RNA binding [32]. (**D**) Affinity of dADAR dsRBD1 for *GluR-2* R/G USL as determined by ITC. (**E**) Sequence alignment of dADAR dsRBD1 and dsRBD2 and human ADAR2 dsRBD1 and dsRBD2. The alignment is coloured by amino acid conservation and properties. dADAR dsRBD1 secondary structure elements are shown on top of the alignment. Residues involved in RNA binding, as determined by NMR titration, are marked with black arrows above the alignment.

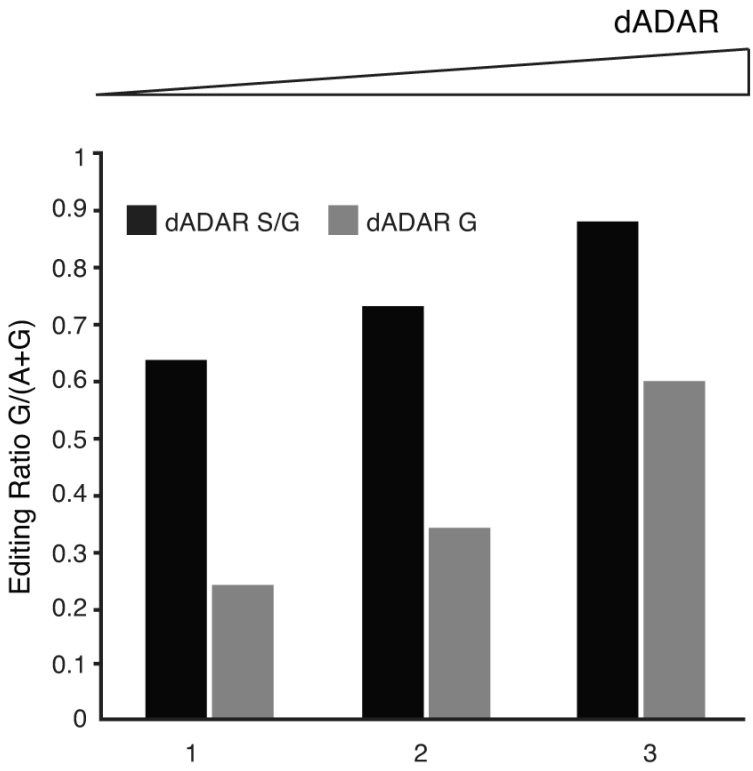

Figure 4. *Drosophila* ADAR can edit the mammalian *GluR-2* pre-mRNA encoded by a minigene

Editing ratio at the *GluR-2* R/G site by *Drosophila* ADAR. The ratio of the Guanosine signal to the combined signal of Adenosine and Guanosine is shown. A ratio of 1 implies 100% editing whereas a ratio of 0 indicates no detectable editing. Increasing amounts of DNA encoding *Drosophila* ADAR were co-transfected with a minigene encoding the *GluR-2* R/G site. The DNA concentration of the minigene was titrated so as to keep the total amount of DNA constant at 1.250 μg. After RT-PCR, sequencing was performed to ascertain editing levels. *Drosophila* ADAR can edit its own pre-mRNA, therefore when dADAR S isoform is transfected into cells a mixture of unedited and edited isoforms is obtained; dADAR S/G (*Black columns*). However this mixture of isoform is more active than the completely edited isoform dADAR G (*Grey columns*) which is consistent with previous observations [14]. See also Supplementary Figure S1.

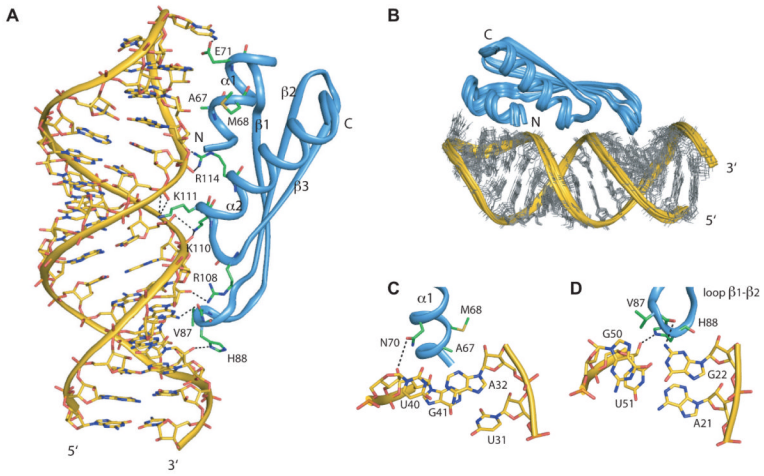

Figure 5. Modelling the interaction of *Drosophila* ADAR dsRBD1 with RNA

(**A**) Representative structure from the ensemble of models of the complex between dADAR dsRBD1 and *GluR-2* R/G USL RNA. The RNA is represented as a *yellow* stick model and the protein is shown as a cartoon *in blue* with important residues shown *in green*. Intermolecular hydrogen bonds are indicated by black dotted lines. (**B**) Overlay of the 10 lowest energy structures of the dsRBD1-USL model. (**C-D**) Close-up view of interactions in the RNA minor groove mediated by helix α1 (**C**) and the β1-β2 loop (**D**).

Europe PMC Funders Author Manuscripts

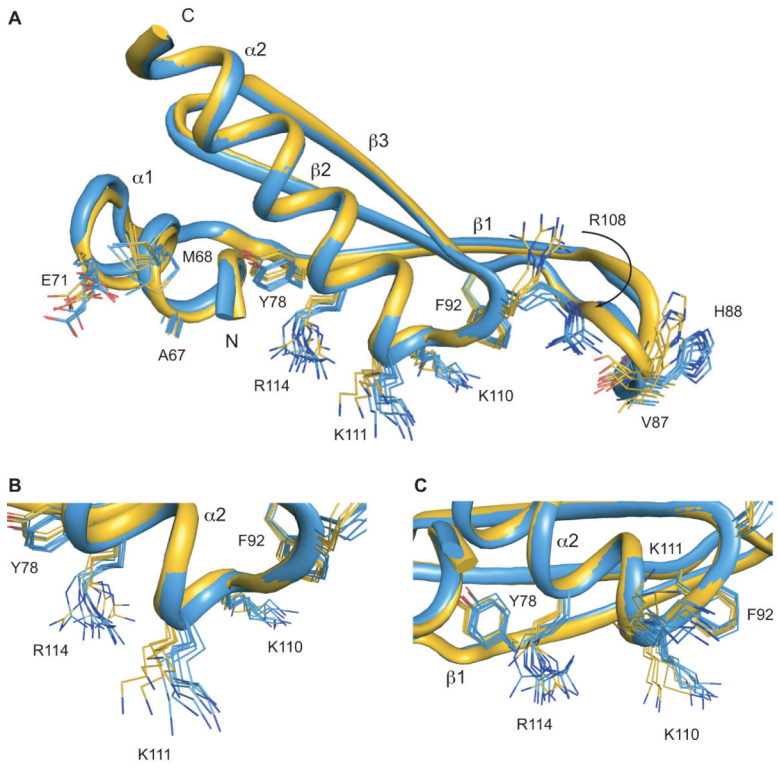

Figure 6. Most side chains need only small displacement to adapt for dsRNA binding

Comparison of free dsRBD1 *in yellow* and dsRBD1 bound to dsRNA *in blue*. The structures are superimposed over Cα atoms over the entire domain. (**A**) Overall view. Only the side chain of R108 experiences large displacement. (**B-C**) Close-up views. The three side chains that bind to the phosphate backbone across the major groove, namely K110, K111 and R114, are only slightly displaced to adapt for RNA binding.

Europe PMC Funders Author Manuscripts

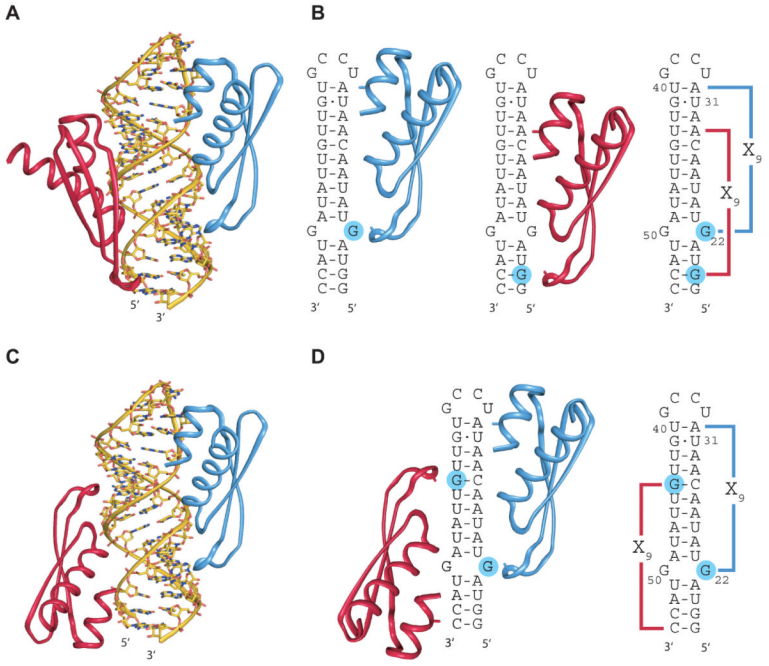

Figure 7. Modelling the interaction of two dsRBDs on the *GluR-2* USL

(**A-B**) One possibility of binding two dsRBDs onto *GluR-2* USL. The two dsRBDs recognize guanosines 22 and 19 leading to a parallel orientation of the dsRBDs. (**A**) Spatial organization of the two dsRBDs. (**B**) Schematic representation of the binding sites onto the *GluR-2* USL sequence. (**C-D**) Second possibility of binding two dsRBDs onto *GluR-2* USL. The two dsRBDs recognize guanosines 22 and 44 leading to an anti-parallel orientation of the dsRBDs. (**C**) Spatial organization of the two dsRBDs. (**D**) Schematic representation of the binding sites onto the *GluR-2* USL sequence.

Table I

NMR experimental restraints and structural statistics

Distance restraints	
Total NOE	2208
Intraresidue	435
Sequential	563
Medium range (|i-j| < 5 residues)	484
Longe range (|i-j| ≥ 5 residues)	726
Hydrogen bond	32
Structure statistics	
NOE violations (mean ± s.d.)	
Number of NOE violations > 0.2 Å	0.6 ± 0.8
Maximum NOE violation (Å)	0.21 ± 0.05
R.m.s.d. from average structure [a]	
Backbone	0.30 ± 0.06 Å
Heavy atoms	0.56 ± 0.07 Å
Mean deviation from ideal covalent geometry	
Bond length (Å)	0.0036
Bond angles (°)	0.47
Ramachandran analysis	
Most favored region	86.9 %
Allowed region	12.9 %
Disallowed region	0.3 %

[a]Protein r.m.s.d. was calculated using residues 64-126 for the ensemble of 20 refined structures