PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Solution structures of the double-stranded RNA-binding domains from RNA helicase A

Takashi Nagata,[1,2,3,4] Kengo Tsuda,[4] Naohiro Kobayashi,[5] Mikako Shirouzu,[4] Takanori Kigawa,[4] Peter Güntert,[6,7,8,9] Shigeyuki Yokoyama,[4]* and Yutaka Muto[4]*

[1] Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

[2] Graduate School of Energy Science, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

[3] Department of Supramolecular Biology, Graduate School of Nanobioscience, Yokohama City University, Tsurumi-ku, Yokohama 230-0045, Japan

[4] RIKEN Systems and Structural Biology Center, Tsurumi-ku, Yokohama 230-0045, Japan

[5] Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan

[6] Institute of Biophysical Chemistry, Goethe University, Max-von-Laue-Strasse 9, 60438 Frankfurt/Main, Germany

[7] Center for Biomolecular Magnetic Resonance, Goethe University, Max-von-Laue-Strasse 9, 60438 Frankfurt/Main, Germany

[8] Frankfurt Institute for Advanced Studies, Goethe University, Max-von-Laue-Strasse 9, 60438 Frankfurt/Main, Germany

[9] Graduate School of Science and Technology, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan

ABSTRACT

RNA helicase A (RHA) is a highly conserved protein with multifaceted functions in the gene expression of cellular and viral mRNAs. RHA recognizes highly structured nucleotides and catalytically rearranges the various interactions between RNA, DNA, and protein molecules to provide a platform for the ribonucleoprotein complex. We present the first solution structures of the double-stranded RNA-binding domains (dsRBDs), dsRBD1 and dsRBD2, from mouse RHA. We discuss the binding mode of the dsRBDs of RHA, in comparison with the known dsRBD structures in their complexes. Our structural data provide important information for the elucidation of the molecular reassembly mediated by RHA.

Proteins 2012; 80:1699–1706.
© 2012 Wiley Periodicals, Inc.

Key words: RNA helicase A (RHA); DHX9; NDHII; double-stranded RNA (dsRNA); dsRBD; dsrm; NMR; structure.

INTRODUCTION

RNA helicase A (RHA or DHX9) is a highly conserved and ubiquitous protein that regulates many aspects of gene expression, such as transcription and translation of cellular and viral mRNAs.[1–4] Deregulation of RHA in humans greatly impacts cell growth, apoptosis, and innate response,[5] and RHA is reportedly involved in various

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: RIKEN Structural Genomics/Proteomics Initiative (RSGI), The National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), Yokohama Academic Foundation, Lichtenberg Program of the Volkswagen Foundation, Japan Society for the Promotion of Science (JSPS); Grant sponsor: Ministry of Education, Culture, Sports, Science and Technology of Japan; Grant numbers: 20570111, 23570146, and 21370047; Grant sponsor: Yokohama City University; Grant number: T2203; Grant sponsor: Advanced Medical Research Center of Yokohama City University [Research and Development Project II (10)]; Grant number: S2210.

*Correspondence to: Shigeyuki Yokoyama, RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. E-mail: yokoyama@biochem.s.u-tokyo.ac.jp or Yutaka Muto, RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. E-mail: ymuto@gsc.riken.jp.

Received 21 December 2011; Revised 3 February 2012; Accepted 7 February 2012
Published online 22 February 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24059

cancers, autoimmune diseases, and viral infections.[6] For example, in mice, the RHA knockout is embryonically lethal, due to defects in the differentiation of the embryonic ectoderm and gastrulation.[7] RHA recognizes highly structured nucleotides and catalytically rearranges the various interactions between RNA, DNA, and protein molecules to provide a platform for RNP (ribonucleoprotein complex).

Several RHA-responsive mRNAs contain highly structured untranslated regions (UTRs), which are critical for nuclear export and translation events.[4,8,9] Namely, constitutive transport element (CTE), which resides in the 3′ UTR of the type D retrovirus RNAs, is necessary for the nuclear export of the unspliced viral transcripts.[8,10] RHA binds directly to this CTE and recruits the nuclear export receptor Tap,[11] which may facilitate the consequent remodeling of the viral RNP for nuclear export and subsequent translation in the cytoplasm.[12] RHA also binds to a highly structured RNA element found in the 5′ UTR of several cellular and retroviral mRNAs,[4] referred to as PCE (post-transcriptional control element), as exemplified by the *junD* proto-oncogene and the HIV-1 *gag* gene.[4,13] RHA facilitates the nuclear export of these genes, and rather surprisingly, RHA also stimulates their translation in the cytoplasm.[4] Highly structured 5′ elements are generally believed to inhibit efficient translation. It has been proposed that RHA catalytically remodels the PCE and/or RNP, and consequently facilitates ribosome scanning, translation initiation, and polyribosome association.[4] Both CTE and PCE are >100 nucleotides in length, and contain a number of double-stranded, bulge, loop and stem-loop regions. RHA specifically recognizes the double-stranded RNA (dsRNA) structures.

Another line of evidence revealed that RHA functions in the human RNA interference (RNAi) pathway, as a small interfering RNA-loading factor of the RNA-induced silencing complex[9] (RISC). RISC is an RNP that uses a small single-stranded guide RNA, originating from small interfering RNA (siRNA), and micro RNA (miRNA) to find the complementary sequence in the target cellular and viral mRNAs. RISC then cleaves the target mRNA or suppresses its translation.[14,15] The core protein components of RISC are Dicer (an RNase III enzyme that generates a small dsRNA); trans-activation-responsive RNA binding protein (TRBP), protein activator of the interferon-induced kinase (PACT), and Argonaute 2 (Ago2), in which the guide RNA is loaded.[16–18] RHA facilitates the formation of the active RISC by interacting with siRNA, Ago2, Dicer, and TRBP, and then promotes the association of siRNA with Ago2.

These abovementioned biological functions of RHA are attributed to its three biological activities (helicase activity for nucleic acids, single- and double-stranded nucleic acid binding activities, and specific protein/protein interactions). Correspondingly, RHA comprises two double-stranded RNA-binding domains (dsRBDs), the DEIH box RNA helicase domain, a domain of unknown function (conserved in Ago2), the helicase-associated domain 2, a nuclear transport domain (recognized by importin-α), and the C-terminal RG domain (single-stranded nucleic acid binding). The dsRBDs of RHA are reportedly indispensable for the recognition of the above dsRNAs (CTE, PCE, and siRNA) and the related proteins.[9,11,19]

To elucidate the target specificity of RHA, structural studies for RHA have been performed. Recently, the crystal structure of the DEIH box RNA helicase domain, which is important especially for unwinding the dsRNAs, was solved and its unique features were revealed.[20] However, the structures of other domains in RHA have yet to be elucidated. In particular, the dsRBDs are often found in RNAi pathway-related proteins (for example, Dicer contains one, whereas TRBP and PACT contain three dsRBDs), and these dsRBDs not only function in dsRNA binding but also in protein binding (RHA, TRBP, and PACT use their dsRBDs to interact with one another[9,21]). Therefore, structural information about the two dsRBDs of RHA is important to elucidate the molecular reassembly mediated by RHA. In this article, we present the first solution structures of the dsRBDs, dsRBD1 and dsRBD2, from mouse RHA. The sequences of the mouse and human homologs of RHA share 90% identity; therefore, their structures and functions are expected to be nearly identical.[5] Our results revealed that the sequences, lengths, and structures of dsRBD1 and dsRBD2 are quite different. We discuss the binding mode of the dsRBDs of RHA, in comparison with the known dsRBD structures in their complexes.

MATERIALS AND METHODS

Sample preparation

The cDNA fragments encoding dsRBD1 (residues 4–89) and dsRBD2 (residues 163–262) of mouse RHA (UniProtKB/Swiss-Prot: O70133) were each cloned into pCR2.1 (Invitrogen). They were expressed with a His_6-tag, a TEV protease cleavage site, a $(Gly\text{-}Gly\text{-}Ser)_2$-Gly sequence at the N-terminus and a Ser-Gly-Pro-Ser-Ser-Gly sequence at the C-terminus. Uniformly ^{15}N, ^{13}C-labeled proteins were synthesized by the cell-free protein expression system.[22] After the reaction, the proteins were isolated by Ni affinity chromatography, and the His_6-tag was then removed by proteolysis. Subsequent cation-exchange chromatography yielded the highly purified proteins. The RHA dsRBD1 and dsRBD2 samples for NMR experiments were concentrated to 1.30 and 1.16 m*M*, respectively, in 20 m*M* sodium phosphate (pH 6.0), containing 100 m*M* NaCl, 1 m*M* DTT, and 0.02% NaN_3, in 90% H_2O/10% ^2H_2O.

NMR spectroscopy

All NMR data were collected at 298 K on Bruker AVANCE 600, 700, and 800 MHz NMR spectrometers,

Table I
Structural Statistics for the dsRBDl and dsRBD2 of RHA

NMR restraints	dsRBDl	dsRBD2
Distance restraints		
Total NOE	1635	2213
Intra-residue	375	561
Inter-residue		
Sequential ($\|i-j\|=1$)	427	550
Medium-range ($1<\|i-j\|<5$)	305	395
Long-range ($\|i-j\|\geq 5$)	524	707
Hydrogen bonds restraints[a]	40	33
Dihedral angle restraints		
φ and ψ	68/65	63/62
χ^1	22	7
Structure statistics (20 conformers)		
CYANA target function (Å^2)	0.41	0.07
Residual distance violations[b]		
Number > 0.1 Å	2	1
Maximum (Å)	0.34	0.16
Residual dihedral angle violations		
Number > 5°	0	0
Maximum (°)	2.08	1.19
AMBER energies (kcal/mol)		
Mean AMBER energy	−3024	−3694
Mean restraints violation energy	5.16	10.09
Ramachandran plot statistics (%)		
Residues in most favored regions	95.0	92.7
Residues in additionally allowed regions	4.9	7.3
Residues in generously allowed regions	0.0	0.1
Residues in disallowed regions	0.1	0.0
Average R.M.S.D. to mean structure (Å)[c]		
Protein backbone	0.38	0.30
Protein heavy atoms	1.26	1.01

[a]Four distance restraints between the amide and carbonyl group atoms (rNH−O = 1.7–2.2 Å, rN−O = 2.6–3.5 A, rNH−C = 2.6–3.3 Å, rN−C = 3.6–4.6 Å) were used for the hydrogen bond restraints, which were only employed in the CYANA calculation.
[b]The number of constraint violations was counted per model.
[c]For residues Lys5-Arg71 of dsRBDl and Lysl85-His255 of dsRBD2.

each equipped with a cryogenic probe. NMR spectra were processed with NMRPipe/NMRDraw.[23] Spectral analysis was performed with KUJIRA,[24] a program suite for interactive NMR analysis working with NMRView,[25] according to the methods described previously.[26] The backbone and side chain ^{1}H, ^{15}N, and ^{13}C resonances of the proteins were assigned by standard double- and triple-resonance NMR experiments,[27,28] and were deposited in the BioMagResDB (BMRB accession numbers 11456 and 11457). Distance restraints were derived from three-dimensional (3D) ^{15}N-edited and ^{13}C-edited nuclear Overhauser effect spectroscopy (NOESY)-HSQC spectra, each measured with a mixing time of 80 ms.

Structure calculations

Structure calculations of RHA dsRBD1 and dsRBD2 were performed using CYANA 2.0.17.[29–31] The structure calculations started from 200 randomized conformers, and used the standard CYANA simulated annealing schedule with 40,000 torsion angle dynamics steps per conformer. The 40 conformers with the lowest final CYANA target function values were further refined with AMBER9,[32] using the AMBER 2003 force field and a generalized Born model, as described previously.[26] The force constants for distance, torsion angle, and ω angle restraints were set to 32 kcal mol^{-1} Å^{-2}, 60 kcal mol^{-1} rad^{-1}, and 50 kcal mol^{-1} rad^{-2}, respectively. The 20 conformers that were most consistent with the experimental restraints were then used for further analyses. PROCHECK-NMR[33] and CHIMERA[34,35] were used to validate and to visualize the final structures, respectively. Detailed experimental data and structural statistics are summarized in Table I. The final ensembles of 20 conformers were deposited in the Protein Data Bank (PDB IDs 2RS6 and 2RS7). The electrostatic potential of the molecular surface was calculated using the online programs APBS and PDB2PQR (http://www.nbcr.net/pdb2pqr/).[36,37]

Molecular Modeling

Model structures of RHA dsRBD1 and dsRBD2, each in complex with RNA, were generated using the CHIMERA[34,35] and AMBER9[32] programs. The protein coordinates in the Xlrbpa:RNA complex (PDB ID: 1DI2) were replaced with those of the lowest energy-minimized structures of the RHA dsRBDs. These template structures were then refined using the AMBER 2003 force field. Initially, 1000 steps of energy minimization were applied *in vacuo* to eliminate the bad steric contacts, and then 2000 steps of energy minimization were performed using the generalized Born model.

RESULTS AND DISCUSSION

The solution structures of the RHA dsRBDs, which belong to the double-stranded RNA binding motif (dsrm) family (Pfam identifier: PF00035), were determined by NMR spectroscopy. The best fit superpositions of the backbone traces of the 20 lowest energy conformers for dsRBD1 (I4-V80) and dsRBD2 (L170-S262) are shown in Figure 1(a,b). These conformers exhibit good geometry, with excellent structure quality scores (Table I). The structure of dsRBD1 contains the common secondary structure elements αA(K5-R15), β1(A20-G27), β2(Q32-R39), β3(G47-S51), αB(K54-R71), and the extra helix-like segment, h-like C (S77-E79) [Fig. 1(a,c)]. The structure of the dsRBD2 begins with the extra helix-like segment, h-like o' (N171-H175), followed by αA′(L180-E194), β1′(K201-P207), β2′(N210-I221), β3′(R226-G234), and αB′(K237-L255) [Fig. 1(b,d)]. In both dsRBDs, αA (αA′) and αB (αB′) are packed against one face of a three-stranded anti-parallel β-sheet, a common feature of the core αββββα fold. The dsRBD1 has h-like C, while the dsRBD2 additionally has h-like o' and a C-terminal stretch (G256-S262). These terminal regions bring the N- and C-termini close together [Fig. 1(c,d)], and include extra hydrophobic cores (dsRBD1: F7, V66,

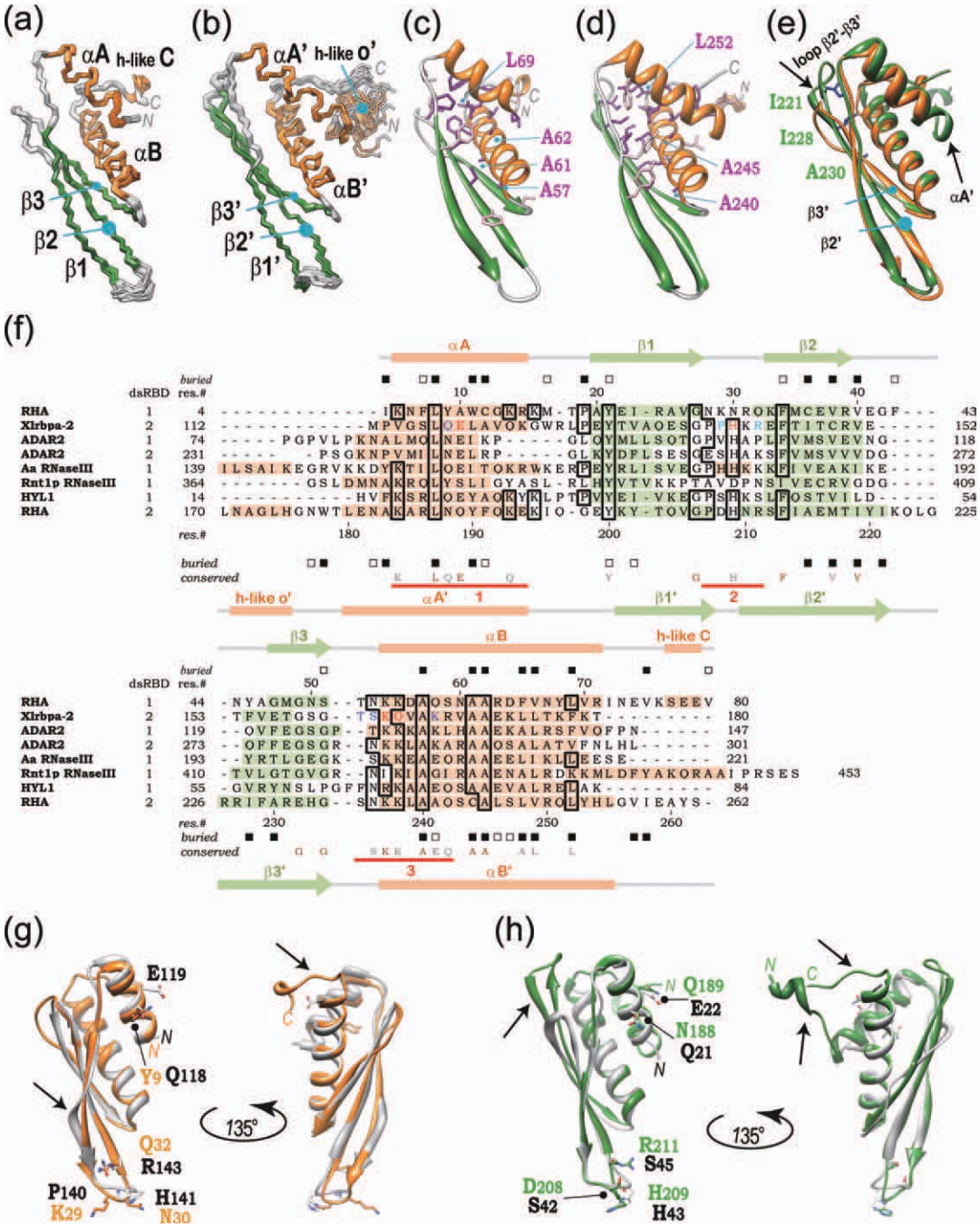

Figure 1

(**a,b**) Superimposition of 20 conformers of RHA dsRBD1$^{I4\text{-}V80}$ and dsRBD2$^{L170\text{-}S262}$. (**c,d**) Ribbon representations of the lowest energy structures of RHA dsRBD1 and dsRBD2. The residues that are important for hydrophobic core formation are displayed by sticks (highly conserved residues are labeled). (**e**) Superposition of the lowest energy structures of RHA dsRBD1 (orange) and dsRBD2 (green). The residues that are involved in the extension of the hydrophobic core are shown by sticks and labeled. (**f**) Structure-based sequence alignment of the RHA dsRBDs and selected dsRBDs. Xlrbpa-2 dsRBD2 and HYL1 dsRBD1 are structurally similar to RHA dsRBD1 and dsRBD2, respectively. Boxed residues are well conserved among the listed dsRBDs. Generally conserved dsRBD residues, obtained from the InterPro database (http://www.ebi.ac.uk/interpro), are highlighted in gray (50–70% conservation) and brown (> 70% conservation). Amino acid residues of Xlrbpa-2 dsRBD2 with backbones and/or side chains that reportedly form hydrogen bonds with dsRNA are highlighted for the backbones (cyan), side chains (blue), and both (red). The secondary structure elements are color-coded: α-helices (orange) and β-strands (green), and the structural elements are labeled. The accessible surface area (ASA) of the residues is presented: 0% < ASA ≤ 10% [purple in (c,d) and filled boxes in (f)], 10% < ASA ≤ 20% [pink in (c,d) and open boxes in (f)]. (**g,h**) Superposition of RHA dsRBD1 (orange) and Xlrbpa-2 dsRBD2 (gray); and that of RHA dsRBD2 (green) and HYL1 dsRBD1 (gray). The lowest energy structures were used for the RHA dsRBDs. Arrows indicate differences. Residues that interact with successive minor grooves of dsRNA in Xlrbpa-2 and HYL1 are shown by sticks with gray labels. The corresponding residues of RHA dsRBD1 and dsRBD2 are also shown (orange and green, respectively).

V75, and V80; dsRBD2: W178, V249, and I258 [Fig. 1(f)]), which seem to stabilize their locations. In the dsRBD2, Y261 stacks onto W178, while the plane of the Y253 ring approaches the ring of W178 from its side. These interactions seem to stabilize the extra hydrophobic core further. A search of the structures of the dsRBDs registered in the PDB revealed that the dsRBD2 of ADAR2 (PDB ID: 2L2K), and both the dsRBD1 and dsRBD2 of DGCR8 (PDB ID: 2YT4) also contain extra hydrophobic cores, and their N- and C-termini are located close together. In the Pfam database,[38] the corresponding dsRBDs of RHA (UniProt entry: DHX9_MOUSE) were predicted to be in the regions I4-L69 and N182-Q251, respectively, which are both shorter than those in our structures. Thus, our structures have included some terminal extensions and further defined the amino acid residues of the dsRBDs for RHA.

Despite the low-sequence identity (22%) and similarity (25%), the sequences of the RHA dsRBDs could be aligned over 76 amino acid residues (dsRBD1: I4-E79; dsRBD2: A183-S262), with 4-residue gaps. The structures of dsRBD1 and dsRBD2 could be superimposed over 56 C^{α}s, with an RMSD of 0.99 Å [Fig. 1(c–f)]. Structural differences were also found within the core αββββα fold. In dsRBD2, αA′ is elongated by one turn in the N-terminal region, as compared to dsRBD1 [Fig. 1(e,f)]. Additionally, β2′, the loop β2′-β3′, and β3′ are longer in dsRBD2 than the corresponding elements in dsRBD1, and dsRBD2 has extra residues in the loop β2′-β3′ (K222-G225), which are the origin of the aforementioned 4-residue gaps [Fig. 1(c–f)]. These augmented structures of dsRBD2 include additional hydrophobic amino acid residues, which extend the hydrophobic core [Fig. 1(c–f)].

A structural similarity search was performed using the PDBeFold server (www.ebi.ac.uk/msd-srv/ssm/). The overall structure of RHA dsRBD1 resembled that of Xlrbpa-2 dsRBD2[39] (PDB ID: 1DI2, *Z*-score = 8.8, RMSD = 1.71 Å, 19% sequence identity over 67 C^{α} with 3 gaps); while HYL1 dsRBD1[40] (PDB ID: 3ADI, *Z*-score = 3.9, RMSD = 1.74 Å, 30% sequence identity over 66 C^{α} with 9 gaps) was found to have the most similar overall structure to that of RHA dsRBD2. In both cases, there are distinct differences. RHA dsRBD1 has a longer C-terminus, while Xlrbpa-2 dsRBD2 has a bulge in the strand β1 [Fig. 1(g), arrows]. Meanwhile, RHA dsRBD2 has longer N- and C-termini, as well as a longer β2-β3 loop, than the corresponding elements in HYL1 dsRBD1 [Fig. 1(h), arrows]. The crystal structures of Xlrbpa-2 dsRBD2 and HYL1 dsRBD1 in complexes with dsRNA have been reported.[39,40] The locations of the residues that interact with the successive minor grooves of dsRNA in these two dsRBDs may be conserved in the dsRBDs of RHA [Fig. 1(g,h)]. The residues that may contribute to the dsRNA-binding in the RHA dsRBDs were predicted, based on these comparisons, and their binding modes will be discussed later.

The residues that are important for maintaining the overall structure are conserved, particularly in the C-terminal third of the dsRBDs.[41] These residues include A57, A61, A62, and L69 in RHA dsRBD1, and A240, A245, and L252 in dsRBD2 [Fig. 1(d,e)]. Other residues that are important for hydrophobic core formation are dispersed throughout the sequences of the dsRBDs. Although the locations of these residues are conserved, their residue types are more variable among the aliphatic and aromatic amino acids [Fig. 1(f)]. In general, the canonical dsRBD comprises 65–70 amino acids. However, insertions and deletions are often seen, particularly in the β-strands and the αA-β1, β1-β2, and β2-β3 loops.[41] In RHA, dsRBD2 has such an insertion in the loop β2-β3, and additionally, both dsRBDs contain extra structural elements in the N- and C-terminal regions [Fig. 1(c–f)]. It is tempting to speculate that these additional and/or novel structural features could form intramolecular and/or intermolecular binding sites that are relevant to the versatile functions of RHA.

A single canonical dsRBD, exemplified by Xlrbpa-2 RBD2,[39] binds on one side of a dsRNA, using an α-helical face composed of three regions: Region 1 (αA), Region 2 (loop β1-β2), and Region 3 (loop β3-αB and αB) [Fig. 1(f)]. Regions 1 and 2 interact with successive minor grooves, and Region 3 interacts with the mediating major groove. The structures of TRBP dsRBD2 and Staufen dsRBD3 have been reported in their RNA-free and RNA-bound forms.[20,40,42,43] These dsRBDs bind dsRNA without undergoing any particular conformational changes. Assuming that this is also the case with RHA, we built dsRNA docking models for each of the RHA dsRBDs, on the basis of the Xlrbpa-2 RBD2:dsRNA complex structure [Fig. 2(a,b)]. The amino acid residues that are within 10 Å of the RNA are shown in stick representations in Figure 2. All of these residues, accordingly, reside in the three conserved RNA-binding regions. These three regions are also highly conserved among the RHAs from different species (Supporting Information Fig. S1).[51]

The residues in Region 1 of RHA dsRBD1 (dsRBD2), K5 (K184), K14 (K193), and K16 (K195), are conserved in HYL1 [Fig. 1(f)]. In HYL1, however, only K17, equivalent to K5 (K184) in RHA, interacts with RNA.[40,44] The residues at the positions occupied by Q118 and E119 in Xlrbpa-2 are highly conserved among the dsRBDs, and are known to contact RNA [Fig. 1(f)] (this trend is also conserved in HYL1 and TRBP[20,40]). However, these positions are poorly conserved in RHA. In RHA dsRBD2, the residues at the corresponding positions are N188 and Q189 [Fig. 1(f)]. Their side chains are capable of forming similar contacts with RNA. In RHA dsRBD1, however, only Y9, which corresponds to Q118 in Xlrbpa-2, approaches the RNA [Figs. 1(f) and 2(a)]. The residue at the corresponding position in the dsRBD of Rnt1p RNase III is also tyrosine [Fig. 1(f)], and its phenol ring stacks

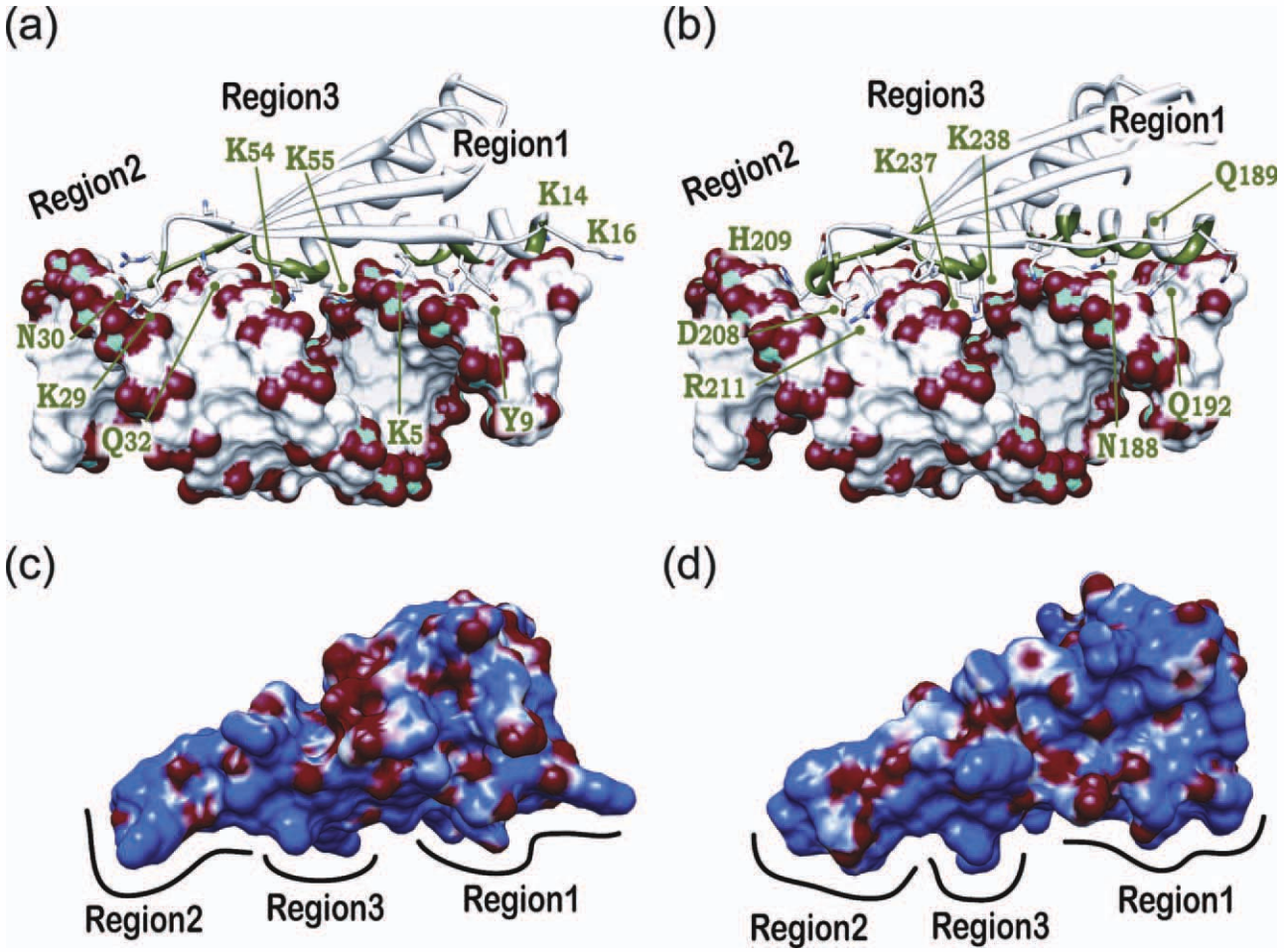

Figure 2

(a,b) Models of RHA dsRBD1$^{I4–V80}$ and dsRBD2$^{L170-S262}$, in their complexes with dsRNA. Proteins are shown by ribbons. Amino acid residues that are within 10 Å of the RNAs are colored green. Their side chains are displayed by sticks with colored heteroatoms: nitrogen (blue) and oxygen (red). The residues metioned in the text are labeled. RNA portions are illustrated in a surface representation: the backbone phosphorus and oxygen heteroatoms are colored cyan and red, respectively. **(c,d)** Electrostatic potential surface representations of the dsRBD1 and dsRBD2 of RHA. Negative (red) is set at -5 kT/e and positive (blue) is set at 5 kT/e. The structures are shown in the same orientation and scale as in (a) and (b), respectively.

next to the ribose and hydrogen bonds to the phosphate group.[45] Therefore, similar interactions may occur between Y9 in RHA and the RNA.

We noticed that the side chain of Q192 in Region 1 of RHA dsRBD2 closely approaches the RNA base [Figs. 1(f) and 2(b)]. The equivalent residue, Q161, in *Aquifex aeolicus* RNase III recognizes a guanine by two sequence-specific hydrogen bonds.[46] In the case of ADAR2 dsRBD1 and dsRBD2, the side chain methyl groups of M84 and M238, respectively, each contact an adenine base.[47] These findings were surprising at the time, since it was generally thought that dsRBDs are shape-specific, rather than sequence-specific, RNA binding domains.[48] The positions of these methionines and Q161 of *Aquifex aeolicus* RNase III are different in their equivalent αA helices (two-turns apart). This difference results in not only the recognition of distinct sequences but also the distinct register lengths between the two specific contacts on the RNA helix.[48] Note that the position corresponding to Q192 in RHA dsRBD2 is G13 in its dsRBD1, which lacks an RNA contact [Fig. 1(f)]. RHA dsRBD2 could be another example of a dsRBD that recognizes not only the shape of the RNA, but also its sequence.

In Region 2, the residue at the position occupied by H141 in Xlrbpa-2 is highly conserved among the dsRBDs [Fig. 1(f)]. Both the backbone and side chain of this residue are involved in RNA contacts.[39] The residue at the corresponding position in RHA dsRBD1 (dsRBD2) is N30 (H209) [Fig. 1(f)]. Histidine and asparagine are considered to be rather equivalent from a structural standpoint, because they have similar hydrogen bonding capabilities and space requirements. In Xlrbpa-2, P140

and R143 in Region 2 use their backbones to interact with RNA.[39] The corresponding residues in RHA dsRBD1 (dsRBD2) are K29 (D208) and Q32 (R211) [Fig. 1(f)]. Although the residues at these positions, except for R211, are different, not only their backbones but also their side chains may contact the RNA [Fig. 2(a,b)].

Region 3 reportedly forms a positively charged surface and contacts the major groove. This property is conserved in RHA dsRBD1 (dsRBD2) by K54 (K237) and K55 (K238) [Figs. 1(f) and 2(c,d)]. The importance of these residues in RNA binding was demonstrated with human RHA.[19] The deletion construct of human RHA containing two dsRBDs was subjected to binding assays with the PCE of spleen necrosis virus RNA and *junD* mRNA. In both cases, dissociation constants in the sub-nM range were observed.[19] Triple point-mutations were then introduced into this construct: K54, K55, and K236 (all of which are conserved in mouse) were replaced with alanine, alanine, and glutamate, respectively. This construct was also subjected to the binding assay, which revealed a severe reduction of RNA binding.[19] Since the sequences of the mouse and human homologs of RHA are highly conserved (90% identity), their structures and RNA-binding modes are expected to be quite similar, if not the same.

In this study, we determined the structures of the individual dsRBDs from RHA. In the native protein sequence, these two dsRBDs are connected by a linker. The structures of two tandemly linked dsRBDs have been determined for the protein kinase PKR and the microprocessor complex subunit DGCR8, each in the RNA-free form; and for adenosine deaminase ADAR2 in its free and RNA-bound forms. The linkers of PKR and ADAR2 were highly flexible, and no interdomain contacts were presented in all of the structures obtained.[47,49] On the other hand, the orientation of the two dsRBDs in DGCR8 was restricted.[50] In DGCR8, the two dsRBDs interact with each other by using the surface that is opposite from the RNA-binding sites. In addition, the extra α-helix that resides in the C-terminal region following the second dsRBD packs against both of the dsRBDs, and forms hydrophobic patches. The linker in DGCR8 contained two short helices; however, it did not contribute to the interdomain contact.

We compared the sequences of the RHAs from different species (mouse, human, bovine, and Xenopus), and searched for additionally conserved regions in the sequences around the dsRBDs, the linker between the dsRBDs (P81-D169), and the region following dsRBD2 (G263-V300). We found that two stretches (G111-E123 and W141-D169) within the linker and a stretch (G263-V300) in the C-terminal region are well conserved. The PSIPRED v3.0 server (bioinf.cs.ucl.ac.uk/psipred/) predicted two helices, D149-S165 (with a mixed confidence level) within the linker, and P282-E294 (with high confidence) in the C-terminal region, both with amphiphilic characteristics. We hypothesize that these extra conserved regions and predicted helices may support the RNA-binding, intraprotein and interprotein interactions by the dsRBDs. Further biochemical studies will clarify the target RNA sequences and define the binding sites of the target proteins. Understanding the detailed mechanisms of the versatile functions of RHA will require the structures of the complexes between RHA and its target molecules.

ACKNOWLEDGMENTS

The authors are grateful to Natsuko Matsuda, Yoko Motoda, Atsuo Kobayashi, Masaru Hanada, Masaomi Ikari, Fumiko Hiroyasu, Miyuki Sato, Satoko Yasuda, Yuri Tsuboi, Yasuko Tomo, Yukiko Fujikura, Takeshi Nagira, Hiroko Uda-Tochio, and Yukiko Kinoshita-Sakaguchi for sample preparation.

REFERENCES

1. Fuller-Pace FV. DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. Nucleic Acids Res 2006;34:4206–4215.
2. Huo L, Wang YN, Xia W, Hsu SC, Lai CC, Li LY, Chang WC, Wang Y, Hsu MC, Yu YL, Huang TH, Ding Q, Chen CH, Tsai CH, Hung MC. RNA helicase A is a DNA-binding partner for EGFR-mediated transcriptional activation in the nucleus. Proc Natl Acad Sci USA 2010;107:16125–16130.
3. Tettweiler G, Lasko P. A new model for translational regulation of specific mRNAs. Trends Biochem Sci 2006;31:607–610.
4. Hartman TR, Qian S, Bolinger C, Fernandez S, Schoenberg DR, Boris-Lawrie K. RNA helicase A is necessary for translation of selected messenger RNAs. Nat Struct Mol Biol 2006;13:509–516.
5. Ranji A, Boris-Lawrie K. RNA helicases: emerging roles in viral replication and the host innate response. RNA Biol 2010;7:775–787.
6. Bolinger C, Boris-Lawrie K. Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. Retrovirology 2009;6:8.
7. Lee CG, da Costa SoaresV, Newberger C, Manova K, Lacy E, Hurwitz J. RNA helicase A is essential for normal gastrulation. Proc Natl Acad Sci USA 1998;95:13709–13713.
8. Tang H, Gaietta GM, Fischer WH, Ellisman MH, Wong-Staal F. A cellular cofactor for the constitutive transport element of type D retrovirus. Science 1997;276:1412–1415.
9. Robb GB, Rana TM. RNA helicase A interacts with RISC in human cells and functions in RISC loading. Mol Cell 2007;26:523–537.
10. Gruter P, Tabernero C, von Kobbe C, Schmitt C, Saavedra C, Bachi A, Wilm M, Felber BK, Izaurralde E. TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. Mol Cell 1998;1:649–659.
11. Tang H, Wong-Staal F. Specific interaction between RNA helicase A and Tap, two cellular proteins that bind to the constitutive transport element of type D retrovirus. J Biol Chem 2000;275:32694–32700.
12. Hull S, Boris-Lawrie K. Analysis of synergy between divergent simple retrovirus posttranscriptional control elements. Virology 2003;317:146–154.
13. Bolinger C, Sharma A, Singh D, Yu L, Boris-Lawrie K. RNA helicase A modulates translation of HIV-1 and infectivity of progeny virions. Nucleic Acids Res 2010;38:1686–1696.
14. Siomi H, Siomi MC. On the road to reading the RNA-interference code. Nature 2009;457:396–404.

15. Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. Argonaute2, a link between genetic and biochemical analyses of RNAi. Science 2001;293:1146–1150.
16. Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, Shiekhattar R. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature 2005;436:740–744.
17. Haase AD, Jaskiewicz L, Zhang H, Laine S, Sack R, Gatignol A, Filipowicz W. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. EMBO Rep 2005;6:961–967.
18. Lee CG, Hurwitz J. A new RNA helicase isolated from HeLa cells that catalytically translocates in the 3′ to 5′ direction. J Biol Chem 1992;267:4398–4407.
19. Ranji A, Shkriabai N, Kvaratskhelia M, Musier-Forsyth K, Boris-Lawrie K. Features of double-stranded RNA-binding domains of RNA helicase A are necessary for selective recognition and translation of complex mRNAs. J Biol Chem 2011;286:5328–5337.
20. Yamashita S, Nagata T, Kawazoe M, Takemoto C, Kigawa T, Güntert P, Kobayashi N, Terada T, Shirouzu M, Wakiyama M, Muto Y, Yokoyama S. Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. Protein Sci 2011;20:118–130.
21. Kok KH, Ng MH, Ching YP, Jin DY. Human TRBP and PACT directly interact with each other and associate with dicer to facilitate the production of small interfering RNA. J Biol Chem 2007; 282:17649–17657.
22. Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A, Yokoyama S. Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. J Struct Funct Genomics 2004;5:63–68.
23. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
24. Kobayashi N, Iwahara J, Koshiba S, Tomizawa T, Tochio N, Güntert P, Kigawa T, Yokoyama S. KUJIRA, a package of integrated modules for systematic and interactive analysis of NMR data directed to high-throughput NMR structure studies. J Biomol NMR 2007;39:31–52.
25. Johnson BA. Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol 2004;278:313–352.
26. Nagata T, Suzuki S, Endo R, Shirouzu M, Terada T, Inoue M, Kigawa T, Kobayashi N, Güntert P, Tanaka A, Hayashizaki Y, Muto Y, Yokoyama S. The RRM domain of poly(A)-specific ribonuclease has a noncanonical binding site for mRNA cap analog recognition. Nucleic Acids Res 2008;36:4754–4767.
27. Clore GM, Gronenborn AM. Determining the structures of large proteins and protein complexes by NMR. Trends Biotechnol 1998;16:22–34.
28. Cavanagh J, Fairbrother WJ, Palmer AG, III, Skelton NJ. Protein NMR spectroscopy, principles and practice. Academic Press, Inc.: San Diego, CA; 1996.
29. Güntert P. Automated NMR structure determination from NMR spectra. Eur Biophys J 2009;38:129–143.
30. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 1997;273:283–298.
31. Herrmann T, Güntert P, Wüthrich 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.
32. Case DA, Cheatham TE, III, Darden T, Gohlke H, Luo R, Merz KM, Jr, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1668–1688.
33. Laskowski RA, Rullmann 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.
34. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 2004;25: 1605–1612.
35. Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE. Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinformatics 2006;7:339.
36. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci USA 2001;98:10037–10041.
37. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res 2004;32:W665–W667.
38. Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A. The Pfam protein families database. Nucleic Acids Res 2010;38:D211–D222.
39. 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.
40. 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.
41. Tian B, Bevilacqua PC, Diegelman-Parente A, Mathews MB. The double-stranded-RNA-binding motif: interference and much more. Nat Rev Mol Cell Biol 2004;5:1013–1023.
42. Bycroft M, Grunert S, Murzin AG, Proctor M, St Johnson 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.
43. Ramos A, Grunert 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.
44. Rasia RM, Mateos J, Bologna NG, Burdisso P, Imbert L, Palatnik JF, Boisbouvier J. Structure and RNA interactions of the plant MicroRNA processing-associated protein HYL1. Biochemistry 2010;49: 8237–8239.
45. 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.
46. Gan J, Shaw G, Tropea JE, Waugh DS, Court DL, Ji X. A stepwise model for double-stranded RNA processing by ribonuclease III. Mol Microbiol 2008;67:143–154.
47. Stefl R, Oberstrass FC, Hood JL, Jourdan M, Zimmermann M, Skrisovska L, Maris C, Peng L, Hofr C, Emeson RB, Allain FH. The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. Cell 2010;143: 225–237.
48. Stefl R, Skrisovska L, Allain FH. RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO Rep 2005;6:33–38.
49. Nanduri S, Carpick BW, Yang Y, Williams BR, Qin J. Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. EMBO J 1998;17:5458–5465.
50. Sohn SY, Bae WJ, Kim JJ, Yeom KH, Kim VN, Cho Y. Crystal structure of human DGCR8 core. Nat Struct Mol Biol 2007;14:847–853.
51. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 2010;38:W529–W533.