Solution Structure of the Second RRM Domain of RBM5 and Its Unusual Binding Characters for Different RNA Targets

ARTICLE *in* BIOCHEMISTRY · JULY 2012

Impact Factor: 3.02 · DOI: 10.1021/bi300539t · Source: PubMed

7 AUTHORS, INCLUDING:

Jiahai Zhang
University of Science and Technology of China
86 PUBLICATIONS **1,033** CITATIONS
SEE PROFILE

Qingguo Gong
University of Science and Technology of China
31 PUBLICATIONS **247** CITATIONS
SEE PROFILE

Jihui Wu
University of Science and Technology of China
98 PUBLICATIONS **1,414** CITATIONS
SEE PROFILE

Yunyu Shi
University of Science and Technology of China
126 PUBLICATIONS **2,207** CITATIONS
SEE PROFILE

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

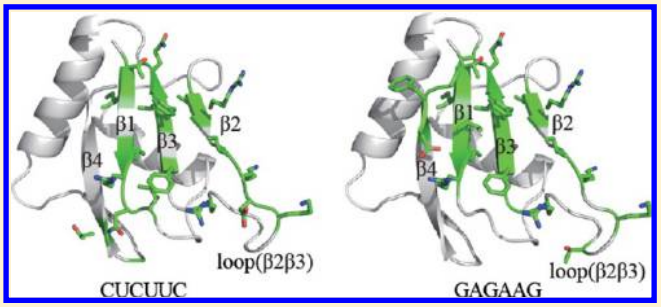
pubs.acs.org/biochemistry

Solution Structure of the Second RRM Domain of RBM5 and Its Unusual Binding Characters for Different RNA Targets

Zhenwei Song, Peiwen Wu, Peng Ji, Jiahai Zhang, Qingguo Gong, Jihui Wu,* and Yunyu Shi*

Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230027, People's Republic of China

ABSTRACT: The RNA binding motif protein 5 (RBM5), also known as LUCA15 or H37, containing two RNA recognition motifs, is a component of the spliceosome A complex. Previously, it has been reported that RBM5 bound to a U/C-rich sequence upstream of the In100 element at intron 9 of caspase2 pre-mRNA that enhanced the formation of proapoptotic caspase-2L isoform. In the present study, we solved the solution structure of the RBM5 RRM2 core domain and characterized its unusual binding capability for different RNA sequences. We found that the RBM5 RRM2 could preferentially bind to both CU rich and GA rich sequences with affinity in 10^{-5} molar range. Further NMR experiments revealed that the dual RNA molecules could be accommodated on almost the same region of the protein's β-sheet surface and that both the N- and C-terminal regions of the protein were involved in the recognition. Our studies provide evidence for the RBM5 sequence specific interaction with the cis-acting element in pre-mRNA regulating alternative splicing.



The RBM5 protein, also known as LUCA-15 or H37, is a RNA binding protein widely found in the nuclei of mammalian cells.[1] The RBM5 gene, as a putative tumor suppressor, maps to the human chromosomal locus 3p21.3 that is frequently deleted in lung cancers. The expression of the RBM5 gene is decreased in 70−80% of lung cancers. Down-regulation of RBM5 gene expression is also found in various solid tumors.[2] By contrast, RBM5 is consistently overexpressed in breast cancer.[3] Collectively, altered levels of RBM5 causes changes in the transcript levels of about 35 genes involved in cell proliferation and apoptosis.[4] RBM5 also appears to be required for normal p53 transcription activity.[5] However, the underlying mechanisms of these regulations are still not well understood. RBM5 is associated with prespliceosomal complex.[6−8] It is an interaction partner of U2AF65. The biochemical functional studies of RBM5 were focused on the regulation for the alternative splicing of apoptosis-related genes, including caspase-2, FAS, and c-FLIP at the post-transcriptional level.[9,10]

The full-length RBM5 is a multidomain protein with 815 amino acid residues that includes two RNA recognition motifs (RRM), two zinc fingers, one arginine-serine domain, and one glycine patch (Figure 1A). The protein also contains one octamer repeat (OCRE) motif, which interacts with U5 snRNP proteins and plays a key role in alternative splicing regulation of Fas gene.[9] Recently, it was found that the RBM5 RanBP2-type zinc finger can recognize single-stranded RNA containing a GGU motif.[11] It was also reported that RBM5 can bind to a U/C-rich sequence upstream of the In100 element in intron 9 of caspase2 pre-mRNA and regulate its alternative splicing, and that this may induce apoptosis.[10] However, which part of RBM5 is involved in this binding is still unknown.

RBM5 belongs to the RBM family protein. It has been noted that RBM5, along with the proteins RBM6 and RBM10, comprises a small family with high similarity and overlapping functional activities.[9,12] The similarity between RBM5 and the proteins RBM6 and RBM10 are 30% and 50% identity, respectively. In addition, the RRM2 motif is the most conserved motif of this family. The similarity between RRM5 RRM2 and RBM10 RRM2 is 69% identity (Figure 1A and B). RBM5 RRM2 contains two well conserved amino-acid sequences of the RRM family,[13,14] which are referred to as RNP1 (consensus sequence as [R/K]-G-[F/Y]-[G/A]-[F/Y]-[I/L/V]-X-[F/Y], where X can be any amino acid) and RNP2 (consensus sequence as [I/L/V]-[F/Y]-[I/L/V]-X-N-L), respectively (Figure 1C). In RBM5 RRM2, there are corresponding RNP2 regions from Ile231 to Ile236 and RNP1 regions from Arg274 to Leu281.

In this study, we solved the solution structure of the RBM5 RRM2 core domain (231−316) and characterized the dynamic properties of the RRM2 core domain (224−325) with the N- and C-terminal extension. Our data showed that both the N- and C-terminal regions of the protein were involved in the recognition. In the presence of a CUCUUC sequence, the C-terminal extension enhanced the RNA binding affinity, while part of the N-terminal extension became more rigid than the RNA-free state. Surprisingly, we found that the RBM5 RRM2

Received: April 26, 2012
Revised: July 19, 2012
Published: July 27, 2012

dx.doi.org/10.1021/bi300539t | *Biochemistry* 2012, 51, 6667−6678

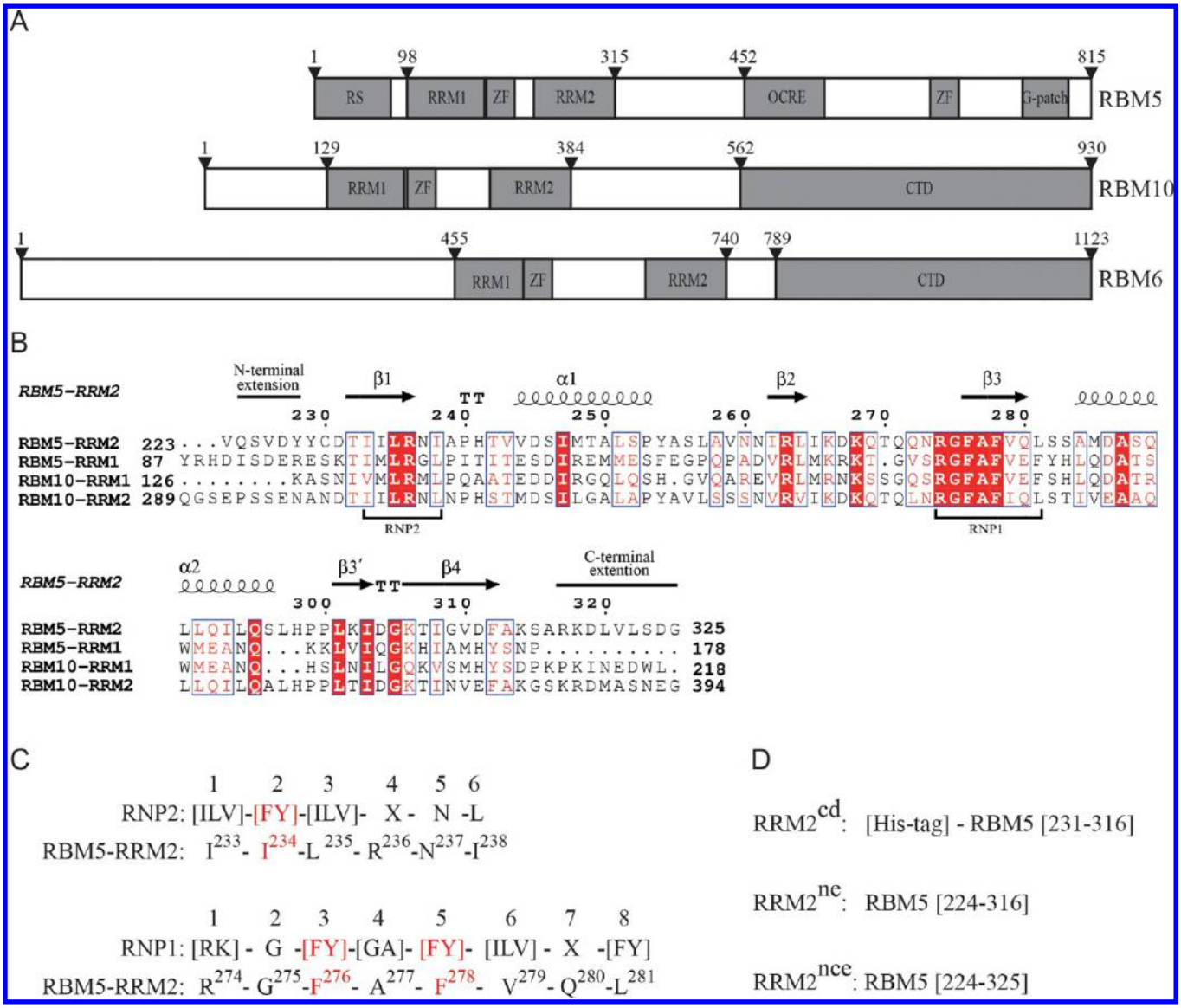


Figure 1. Primary structure of the RBM5 RRM2. (A) Schematic diagram of protein motifs in the primary structure of RBM5. RBM5 possesses one arginine−serine (RS) domain, two RNA recognition motifs (RRM), two Zinc Finger (ZnF) domains, an OCtamer REpeat (OCRE) domain, and one glycine patch. (B) Multiple sequence alignment of the RBM5 RRM2. RBM5 sequences from *Homo sapiens* (SWISSPROT P52756), RBM10 sequence from *Homo sapiens* (SWISSPROT P98175), and RBM6 sequence from *Homo sapiens* (SWISSPROT P78332) were aligned using the Tea-Coffee online service and the ESPript program (Easy Sequencing in PostScript).[35,36] Residues identical in all sequences are highlighted in red columns. Conserved residues are colored red in blue frames. Secondary structure elements for RBM5 RRM2 were shown above the sequence alignment with β strands as arrows, helices as squiggles, and turns as TT. The signature sequences of RNP1 and RNP2 are indicated by the black brackets. (C) The details of RNP1 and RNP2. (D) The list of recombinant protein fragments used in this study, with superscripts cd, ne, and nce corresponding to core domain (cd), N-terminal extension (ne), N- and C-terminal extension (nce), respectively.

could preferentially bind both the [5′-(CUCUUC)-3′] and [5′-(GAGAAG)-3′] sequences. Through comparisons to other complex structures of RRM with target RNA molecules, we depicted at least two base-specific binding pockets on the β-sheet surface. The one located at the β4-β1 strands can recognize and discriminate a guanine base group, and the other located at the β3-β2 strands can accommodate a pyrimidine base group. However, a careful comparison the results of the chemical shift perturbation in the presence of each RNA molecule showed that the RBM5 RRM2's RNA binding surface may have the flexibility to recognize RNA. These results partially explained the RBM5 RRM2's RNA sequence preference. Taken together, our study provided a structural basis for understanding the RNA recognition manner of RBM5 RRM2.

■ MATERIALS AND METHODS

Protein Expression and Purification. The DNA fragments of RRM2^{cd}, RRM2^{ne}, and RRM2^{nce} (Figure 1 D) were amplified from a plasmid containing RBM5 full length cDNA sequence by PCR. The DNA fragment of RRM2^{cd} was ligated into modified pET28a vectors without the thrombin cleavage site using NdeI and XhoI restriction sites. The DNA fragments of RRM2^{ne} and RRM2^{nce} were ligated into other modified pET28a vectors with substitution of the thrombin cleavage site to a TEV cleavage site. RRM2^{nce} mutants were generated by QuikChange Mutagenesis Kit (Takara) and confirmed by DNA

dx.doi.org/10.1021/bi300539t | *Biochemistry* 2012, 51, 6667−6678

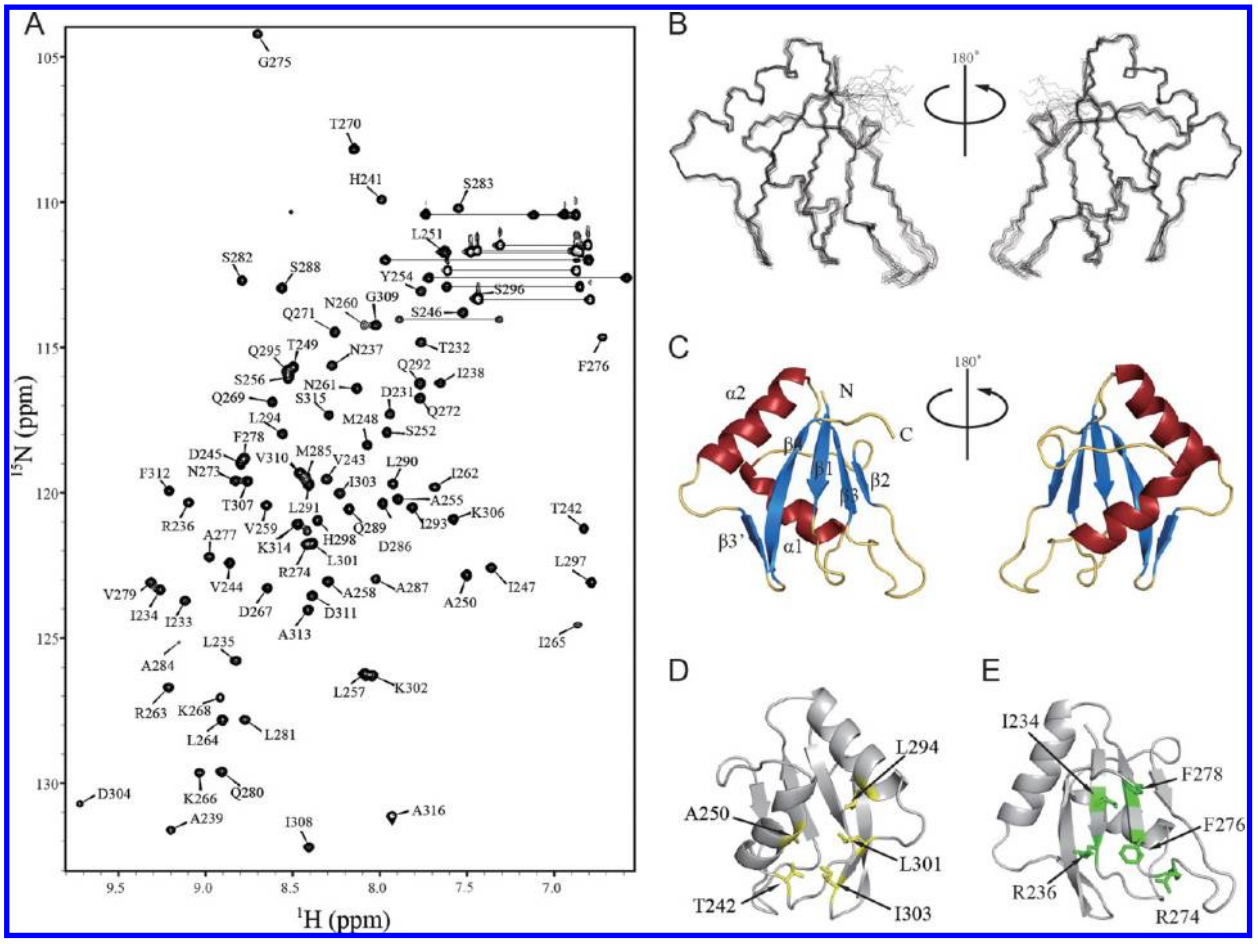


Figure 2. Solution structure of the RBM5 RRM2. (A) Two-dimensional ^{15}N−^{1}H HSQC spectrum of RBM5 RRM2. The side chain resonances of asparagine and glutamine are connected by horizontal lines. (B) Stereo view of the 20 lowest energy structures of RBM5 RRM2 superimposed for best fit of the N, CO, and CA atoms of residues 231−313. The protein backbone is shown as a wire model. The orientations of B and C are the same. (C) Ribbon representation of the energy-minimized structure with secondary structure elements highlighted. (D) Hydrophobic core formation by the $\beta3'$ β-hairpin. The side chains of these hydrophobic residues are colored yellow. (E) Solvent-exposed residues of RNP1 and RNP2 are presented in green.

sequencing. All of the recombinant RBM5 RRM2 proteins were produced in *E. coli* BL21 (DE3) Gold cells induced with 0.5 mM isopropyl-1-β-D-thiogalactoside at 16 °C overnight. Uniformly, ^{15}N or ^{15}N/^{13}C-labeled fusion proteins were produced by growing the bacteria in LeMaster and Richards minimal medium using $^{15}NH_4Cl$ (0.5 g/L) and/or $^{13}C_6$-glucose (2.5 g/L) as the sole nitrogen and carbon sources. The harvested bacteria were lysed in buffer A (20 mM Tris and 500 mM NaCl at pH 8.0). Then the fusion proteins were loaded on a Ni-chelating column (Qiagen) following the protocol. Before eluting from the column, an additional wash with high salt buffer B (20 mM Tris and 1 M NaCl, pH 8.0) was performed. For RRM2^{ne} and RRM2^{nce}, the fused TEV cleavage site was removed by TEV digestion in buffer C (50 mM Tris-HCl, 50 mM NaCl, 2.5 mM EDTA, and 5 mM DTT, pH 8.0) overnight at 16 °C. Three kinds of RBM5 RRM2 proteins were further purified by size exclusion chromatography on a Superdex 75 column (GE Healthcare) equilibrated with buffer D (20 mM Tris, 500 mM NaCl, 2.5 mM EDTA, and 1 mM NaN_3, pH 8.0). Before use, the purified proteins were exchanged into NMR buffer (40 mM phosphate sodium, 50 mM NaCl, 1 mM NaN_3, and 2.5 mM EDTA, pH 6.4) and concentrated to 10−20 mg/mL.

NMR Spectroscopy and Data Processing. All NMR spectra of 1.5 mM uniformly ^{15}N, ^{13}C-labeled RBM5 RRM2^{cd} in NMR buffer with 10% D_2O or 100% D_2O were recorded at 298 K on a Bruker DMX600 spectrometer with a cryoprobe. The proton chemical shift was referenced relative to the frequency of the ^{1}H resonance of water, which was further calibrated by DSS. The ^{13}C and ^{15}N chemical shifts were referenced indirectly based on the ratio of the proton frequency to the carbon/nitrogen frequency.[15] Backbone and side-chain assignments were obtained by using standard NMR methodology, containing the following spectra: 2D ^{1}H−^{15}N HSQC and ^{1}H−^{13}C HSQC, 3D HNCO, HN(CA)CO, CBCA(CO)NH, CBCANH, (H)C(CO)NH-TOCSY, HBHA(CBCA)(CO)NH, H(C)(CO)NH-TOCSY, HCCH−COSY, and HCCH-TOCSY. Assignments were checked for consistency with 3D ^{1}H−^{15}N-NOESY-HSQC and ^{1}H−^{13}C-NOESY-HSQC. NOESY spectra were recorded with mixing times of 130 ms. All NMR spectra were processed with NMRPipe and NMRDraw.[16] The spectra assignment and analysis were performed using Sparky software.[17]

For the RNA titration experiments, CUCUUC RNA (purchased form Takara Bio., Inc.) was dissolved in NMR buffer to make a 32 mM solution. NMR titration experiments were performed by adding increasing amounts of unlabeled

dx.doi.org/10.1021/bi300539t | *Biochemistry* 2012, 51, 6667−6678

synthetic CUCUUC RNA to ^{15}N-labeled RRM2^{nce} protein. 2D ^{1}H−^{15}N HSQC spectra were recorded at RNA/Protein mole ratios of 0.3:1, 0.5:1, 0.8:1, and 1:1 to follow the resonances perturbed by RNA binding to the protein. The weighted average of ^{15}N and ^{1}H$_N$ perturbation values were calculated as follows: $\Delta\delta = [(\Delta\delta_{HN})^2 + (\Delta\delta_N/6.5)^2]^{1/2}$.

Structure Calculations. The solution structure of the RBM5 RRM2 was calculated using Xplor-NIH2.24[18,19] with 1474 NOE distance restraints, 128 dihedral angles, and 58 hydrogen bonds. NOE distance restraints were obtained from 3D ^{1}H−^{15}N-NOESY-HSQC and ^{1}H−^{13}C-NOESY-HSQC spectra. Dihedral angle (Φ and Ψ) restraints were obtained from the backbone and $^{13}C^{\beta}$ chemical-shift values using the program TALOS+.[20] Hydrogen bond restraints were defined from the slow exchange amide protons mainly in the regular secondary structures. A standard script was modified to anneal the structures from extended conformations. Structural quality was assessed using PROCHECK.[21] Molecular graphics were generated using MOLMOL[22] and PyMOL (http://www.pymol.org).

Fluorescence Polarization. Lyophilized 5,6-carboxyfluorescein (FAM) labeled RNA fragments were purchased from Takara Bio., Inc., dissolved in diethypyrocarbonate (DEPC) treated NMR buffer with 20 μM, and stocked in −80 °C. For the fluorescence polarization titration experiments, the FAM-labeled RNA fragments were diluted to a final concentration of 200 nM/400 nM in corresponding buffers (as described in Figure 3). Eighty microliters of this diluted RNA solution was mixed with the same volume of a gradient protein stock and incubated for 15 min. All of the fluorescence titrations were carried out in a SpectraMax M5 (Molecular Devices) plate reader at 20 °C. The fluorophore was excited at 495 nm, and both vertical and horizontal emissions were recorded at 525 nm. The data were analyzed with Microcal ORIGIN software, using a single-site binding model for curve fitting. The data were normalized for visualization purposes.

RESULT

Solution Structure of the RBM5 RRM2. The solution structure of the recombinant protein RBM5 RRM2 was determined by multidimensional heteronuclear NMR spectroscopy. The recombinant protein RRM2^{cd} constitutes a 6-His tag and RBM5 RRM2 (231−316). The ^{15}N−^{1}H HSQC experiment shows a well dispersed spectrum of amide groups (Figure 2A), indicating a well folded domain. All the expected resonances of the backbone residues were assigned with the missing peaks corresponding to the first 8 residues (MGHHHHHH- at the N-terminus as an artifact from the vector). The structure of this recombinant protein was calculated using Xplor-NIH2.24 with a total of 1674 restraints, including 1474 NOE derived distance restrains, 58 hydrogen bond restrains, and 128 dihedral angle restraints. An ensemble of the 20 lowest energy structure sorted from 200 calculated structures were well determined and highly converged (Figure 2 B). These structures have a backbone root-mean-square deviation of 0.418 Å for the well-defined region. A full summary of the structural statistics is given in Table 1. The coordinates' 20 NMR structures have been deposited into the Protein Data Bank (code ID: 2LKZ).

The tertiary structure of RBM5 RRM2 displays a canonic fold of RRM family with a β1-α1-β2-β3-α2-β4 topology. The fold comprises two α-helices packed against a four-stranded antiparallel β-sheet. The core structure consists of residues D231-A313 representing a ribbon structure with the lowest energy (Figure 2 C). The four strands arranged in an antiparallel β4-β1-β3-β2 fashion correspond to the residues Lys306-Phe313 (β4), Thr232-Arg236 (β1), Asn273-Gln280 (β3), and Ile262-Lys267 (β2), respectively. α-Helix 1 lies almost perpendicular to α-helix 2, corresponding to the residues Val244-Ala250 (α1) and Met285-Gln295 (α2), respectively. An additional short β-hairpin was identified between α2 and β4 (β3′, Leu301-Ile303). This β-hairpin closely contacts with the two α-helices and forms a large hydrophobic core region involving residues Thr242, Ala250, Leu294, Leu301, and Ile303 (Figure 2 D). Thus, the linker connecting α2 and β4 is ordered in the 20 lowest energy structures. The residues Leu301 and Ile303 were highly conserved in four RRMs (Figure 1B) indicating that the β-hairpin block was a common character in the protein folding process.

Generally, the canonical RRM fold has two highly conserved sequences, termed RNP2 and RNP1 (Figure 1 C), which correspond to the β1 and β3 strands, respectively. In the case of RBM5 RRM2, position 2 of RNP2 is occupied by an isoleucine residue (Ile234), instead of an aromatic amino acid. In RNP1 of the RBM5 RRM2, all of the solvent-exposed amino acids are well-matched with the consensus sequence. Thus, the residue pattern on the β-sheet surface indicates a potential role of RBM5 RRM2 in RNA binding (Figure 2E).

The N- and C-terminal regions of RRM (outside of the core domain) often contribute to RNA binding by increasing the

Table 1. Summary of Conformational Restraints and Structural Statistics for the 20 Energy-Refined Conformers of the RBM5 RRM2 Structure

Experimental Restraints	
NOE distance restrains	1474
intraresidue	306
short range ($I - j = 1$)	453
medium range ($2 \le I - j \le 4$)	282
long range ($I - j \ge 5$)	433
hydrogen bonds	58
TALOS dihedral angel constraints	128
X-PLOR Energies (kcal·mol^{-1})	
E_{total}	110.75 ± 1.54
E_{bond}	3.51 ± 0.11
E_{angle}	69.67 ± 0.64
E_{cdih}	0.35 ± 0.07
$E_{improper}$	6.89 ± 0.25
E_{vdw}	9.36 ± 0.45
E_{noe}	20.98 ± 1.32
Experimetal Restraints Violation Analysis	
NOE > 0.5 A	0
dihedral angel >5	0
Ramachandran Plot Analysis (%)[a]	
residues in most favored regions	84.6
additionally allowed regions	14.4
generously allowed regions	1.0
residues in disallowed regions	0
rmsd from the Mean Structure (Å)[b]	
backbone (N, CA, CO) atoms	0.418
heavy atoms	0.777

[a]Ramachandran plot of ordered residues, including residues 10−92, was analyzed by Procheck. [b]Root mean square deviation was calculated for the ordered residues, including residues 10−92.

dx.doi.org/10.1021/bi300539t | Biochemistry 2012, 51, 6667−6678

protein−RNA interaction network.[14] To investigate the characteristics of RBM5 RRM2 interaction with RNA, we prepared other construct proteins with the N- or/and C-terminal extension, RRM^{ne} (224−316) and RRM^{nce} (224−325), for this study. The different protein fragments are shown in Figure 1D.

RNA Sequence Preference of the RBM5 RRM2. It has been reported that the full length RBM5 interacted with a CU rich sequence [5′-(CUCUUUUCCUAAGAACUUGGCU-CUUCUCU)-3′] upstream of an In100 element in intron 9 of Casp-2 pre-mRNA.[10] On the basis of this, we selected a number of hexanucleotide RNAs and measured the equilibrium dissociation constants (K_ds) of RBM5 RRM2^{nce} with these RNA fragments (Figure 3A). Fluorescence titrations for 5′-FAM labeled RNA fragments were performed in 10 mM HEPES buffer (pH 7.4) with 150 mM NaCl. The K_ds were determined by fitting fluorescence polarization (FP) data of RNA fragments upon RRM2^{nce} binding with 1:1 model. As a result, the protein RRM2^{nce} bound to the sequence [5′-(CUCUUC)-3′] with a K_d value of 57 μM, 2-fold stronger than the sequence [5′-(AACACA)-3′] and also stronger than the sequence [5′-(ACUUUA)-3′]. The differences of binding affinities for the sequence [5′-(UCUUCU)-3′] (K_d = 77 μM), the sequence [5′-(CUUCUC)-3′] (K_d = 86 μM), and the sequence [5′-(UCCUCU)-3′] (K_d = 94 μM) are relatively small.

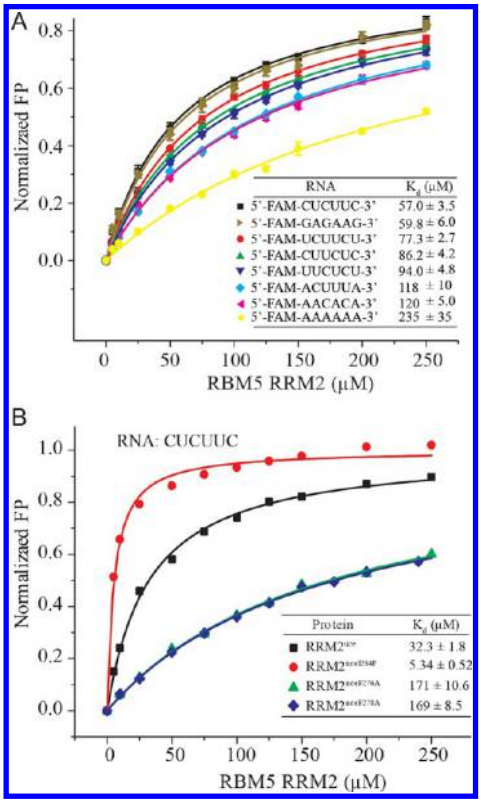


Figure 3. (A) Properties of RNA binding by FP titration. Fluorescence polarization of 5′-FAM labeled RNA fragments upon RRM2^{nce} titration in 10 mM HEPES buffer (pH 7.4) containing 150 mM NaCl, 4 mM DTT, and 1 mM NaN_3. The RNA concentration was 200 nM for each point. Equilibrium dissociation constants (K_d) of RNA fragments are obtained by fitting to the 1:1 binding model. Fitted curves are colored respectively. The K_d values are indicated in the graphs. (B) Point mutant analyses for RNA binding affinity. Fluorescence polarization of 5′-FAM labeled CUCUUC RNA upon RRM2^{nce}, RRM2^{nceI234F}, RRM2^{nceF276A}, and RRM2^{nceF287A} titration in NMR buffer. The RNA concentration was 200 nM for the RRM2^{nce} and RRM2^{nceI234F} assay and increased to 400 nM for the RRM2^{nceF276A} and RRM2^{nceF287A} assay.

Since the sequence [5′-(CUCUUUUCCUAAGAACUUGG-CUCUUCUCU)-3′], upstream of an In100 element in intron 9 of Casp-2 pre-mRNA, contains the AGAA sequence, we also investigated the interaction between RBM5 RRM2^{nce} with the RNA fragment [5′-(GAGAAG)-3′] and found that RBM5 RRM2^{nce} could also bind the sequence [5′-(GAGAAG)-3′] with a K_d value of 59.8 μM, almost the same affinity as that for binding with the CUCUUC RNA but much stronger than that for the sequence [5′-(AAAAAA)-3′] (K_d = 235 μM).

Taken together, these data showed that the sequences [5′-(CUCUUC)-3′] and [5′-(GAGAAG)-3′] are both preferred target RNA for the RBM5 RRM2, among the sequences examined.

RNA Binding Surface. To further clarify the RNA sequence preference of the RBM5 RRM2, we carried out an NMR chemical shift perturbation study with both of the [5′-(CUCUUC)-3′] and [5′-(GAGAAG)-3′] RNAs. We observed that, in both cases, some of the protein resonances shifted in a stepwise directional manner (Figure 4C and D), indicating that upon RNA binding, the exchange between the RNA-bound and RNA-free forms is fast on the NMR time scale. Next, we quantified the results of perturbation and calculated the perturbation values using the following equation: $\Delta\delta = [(\Delta\delta_{HN})^2 + (\Delta\delta_N/6.5)^2]^{1/2}$ (see Figure 5A and B). As a result, 20 residues (Thr232, Ile233, Leu235, Arg236, Ile262, Arg263, Ile265, Lys266, Lys268, Arg274, Phe276, Phe278, Val279, Gln280, Arg317, Asp319, Leu320, Val321, Leu322, and Ser323) were commonly affected by both of these RNAs. In addition, five residues (Asn237, Ile238, Asp267, Gly305, and Thr307) were specifically affected by [5′-(CUCUUC)-3′] RNA. Eight residues (Ile234, Thr270, Gly275, Ala277, Leu281, Asp311, Phe312, and Ala313) were specifically affected by [5′-(GAGAAG)-3′] RNA. Mapping the perturbed residues on the tertiary structure of RRM2^{cd} depicted that the RBM5 RRM2 directly makes contact with both of these RNAs through its β-sheet surface (Figure 5D and E).

Sequence-specific proteins make contact with single-stranded nucleotides commonly via the interaction between aromatic side chains of protein with nucleotide bases.[23] From chemical shift perturbation data, we found that Phe276, which is located at β3 (position 3 of RNP1), is involved in the interaction with either [5′-(CUCUUC)-3′] RNA or [5′-(GAGAAG)-3′] RNA. We also noticed that instead of a tyrosine at β1 (position 2 of RNP2) for most of the RRM domain, there is an Ile234 at position 2 of RNP2 for RBM5 RRM2. The chemical shift of residue Ile234 changed very little during the titration of both RNAs. The chemical shift changed a lot at Ile238 in the titration of [5′-(CUCUUC)-3′] RNA or at Leu235 in the titration of [5′-(GAGAAG)-3′] RNA.

It is interesting that not only the two central β strands (β1 and β3) are involved in RNA binding but also β2 strand is involved in RNA binding in both [5′-(CUCUUC)-3′] and [5′-(GAGAAG)-3′], especially in residue Ile265. When [5′-(GAGAAG)-3′] was titrated, the chemical shift of Phe312 at β4 was dramatically changed, but it did not happen in [5′-(CUCUUC)-3′] titration.

Most of the above residues involved in the interaction with RNA are hydrophobic residues which may form aromatic

dx.doi.org/10.1021/bi300539t | Biochemistry 2012, 51, 6667−6678

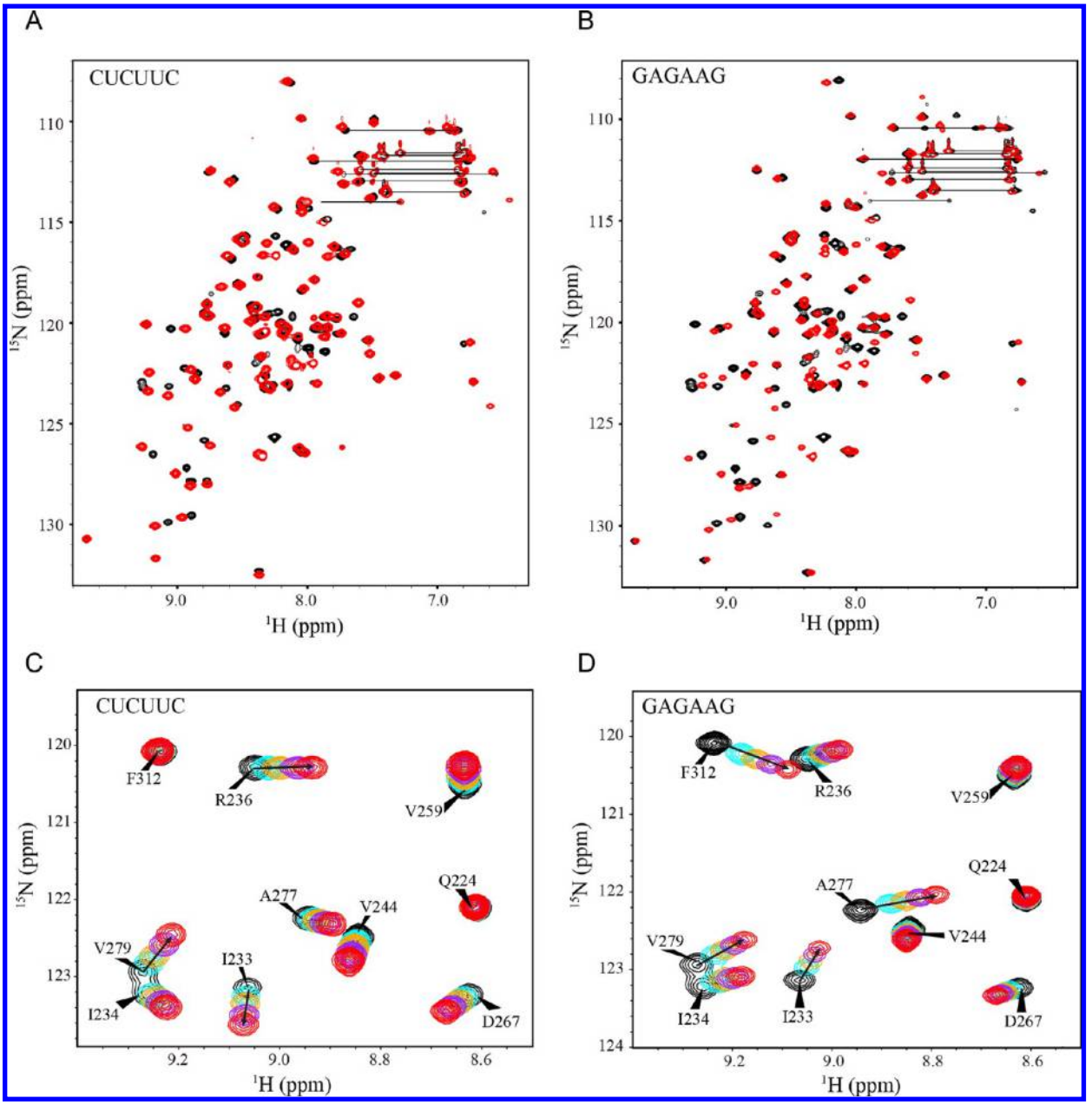


Figure 4. NMR chemical shift perturbation of RRM2^{nce} upon RNA binding. (A,B) Superposition of ^{15}N−^{1}H HSQC spectra of ^{15}N-labeled RRM2^{nce} (1 mM), free and in complex with CUCUUC/GAGAAG. The protein/RNA ratio is approximately 1:1. The cross-peaks of free RRM2^{nce} are shown in black, and those in complex are shown in red. (C,D) Close-up views of the ^{15}N−^{1}H HSQC spectra. Black arrows indicate chemical shift perturbations observed upon RNA binding. The molar ratio of protein/RNA was 1:0 (black), 1:0.25 (cyan), 1:0.5 (yellow), 1:0.8 (magenta), and 1:1 (red).

interactions with the RNA base, such as Phe276, Phe312, or just form the hydrophobic pocket to accommodate the RNA base (Figure 5F).

C-terminal Extension Enhanced CUCUUC Binding. As described above, several RRMs with the N- and C-terminal extensions facilitate RNA binding. In our study, we found that the longer RBM5 RRM2 construct (RRM2^{nce}) which includes the C-terminal extension could bind the sequence [5′-(CUCUUC)-3′] more tightly than the RRM2^{ne} construct (Figure 6C). Accordingly, from the perturbation data, we found that the chemical shifts of five residues including Arg317, Asp319, Leu320, Val321, and Leu322 located at the C-terminal extension moved dramatically indicating that the C-terminal region of RRM2^{nce} was involved in the interaction with both the CUCUUC and GAGAAG RNAs (Figures 5A and B and 6D).

The T2 and heteronuclear NOE relaxation data showed that without the RNA molecule, the backbone amides of Asp319, Leu320, and Val321, which are located at the central region of C-terminal extension, were only slightly more flexible than residues in the α1-β2 and β2-β3 loops region (Figure 6A). Moreover, we observed NOE correlations between Val321Hγ and the protons in the aromatic ring of Phe276 and Phe278, corresponding to positions 3 and 5 of RNP1, respectively. The same situations were observed for Leu322Hδ (Figure 6B). These data indicate that without RNA the C-terminal extension may locally make contact with the β3 strand at β-sheet surface of the core RRM2 by hydrophobic interactions. In contrast, we

dx.doi.org/10.1021/bi300539t | *Biochemistry* 2012, 51, 6667−6678

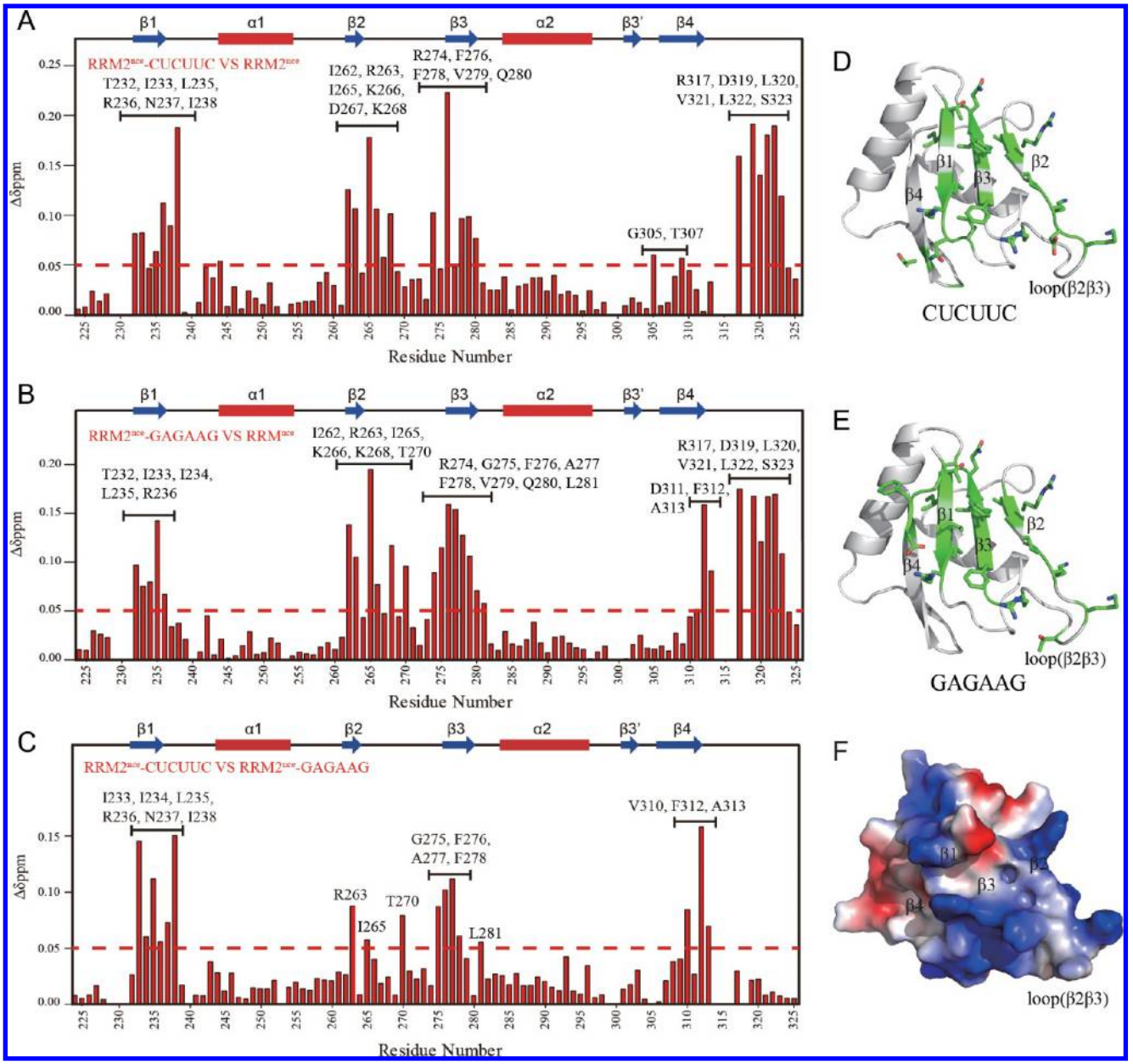


Figure 5. Quantification of chemical shift perturbations. (A,B) Histogram plot presents chemical shift perturbations upon the addition of CUCUUC/GAGAAG. The molar ratio of protein/RNA was 1:1. The perturbation values were calculated as follows: $\Delta\delta = [(\Delta\delta_{HN})^2 + (\Delta\delta_N/6.5)^2]^{1/2}$. The mean value is denoted as a dotted red line. The perturbed residues were denoted on the graph. The secondary structural elements were depicted on the histogram plot. (C) Chemical shift comparison between the RRM2^{nce}-CUCUUC state and the RRM2^{nce}-GAGAAG state. (D,E) The perturbed residues were mapped on the tertiary structure of RRM2^{cd}. The perturbed residues with side chains were colored in green. (F) The electronic potential surface of RRM2^{cd}. The orientations of D, E, and F are the same.

observed that when the CUCUUC RNA bound to the RRM2^{nce} protein, the T2 and heteronuclear NOE relaxation data showed that the protein C-terminal region (Arg317-Gly325) was much more flexible than that in RNA-free RRM2^{nce} (Figure 6A). These observations more likely indicated that during RNA binding, the C-terminal extension was forced to stay away from the core RRM body and then exposed the aromatic ring of Phe278, which was stacked with the base group of the RNA molecule.

N-Terminal Extension Becomes Rigid upon CUCUUC Binding. We cannot observe the peaks of the residues (Tyr229, Cys230, Asp231, Lys314, Asp315, and Leu316) in the spectrum of the RNA-free state. Upon CUCUUC RNA binding, the residues Tyr229, Cys330, and Asp231 appeared on the spectrum. Interestingly, from the beginning of the titration the amide resonances of these three residues also moved in a fast exchange manner along with increasing amounts of RNA. Moreover, in the presence of RNA, the T2 and heteronuclear NOE relaxation data indicate that the N-terminal extension root region is as rigid as the RRM core (Figure 6A). These results demonstrate that the N-terminal extension of RBM5 RRM2 is also involved in RNA recognition.

■ DISCUSSION

Our NMR data analysis of RBM5 RRM2 reveals several features of protein−RNA interfaces. The most striking

dx.doi.org/10.1021/bi300539t | Biochemistry 2012, 51, 6667−6678

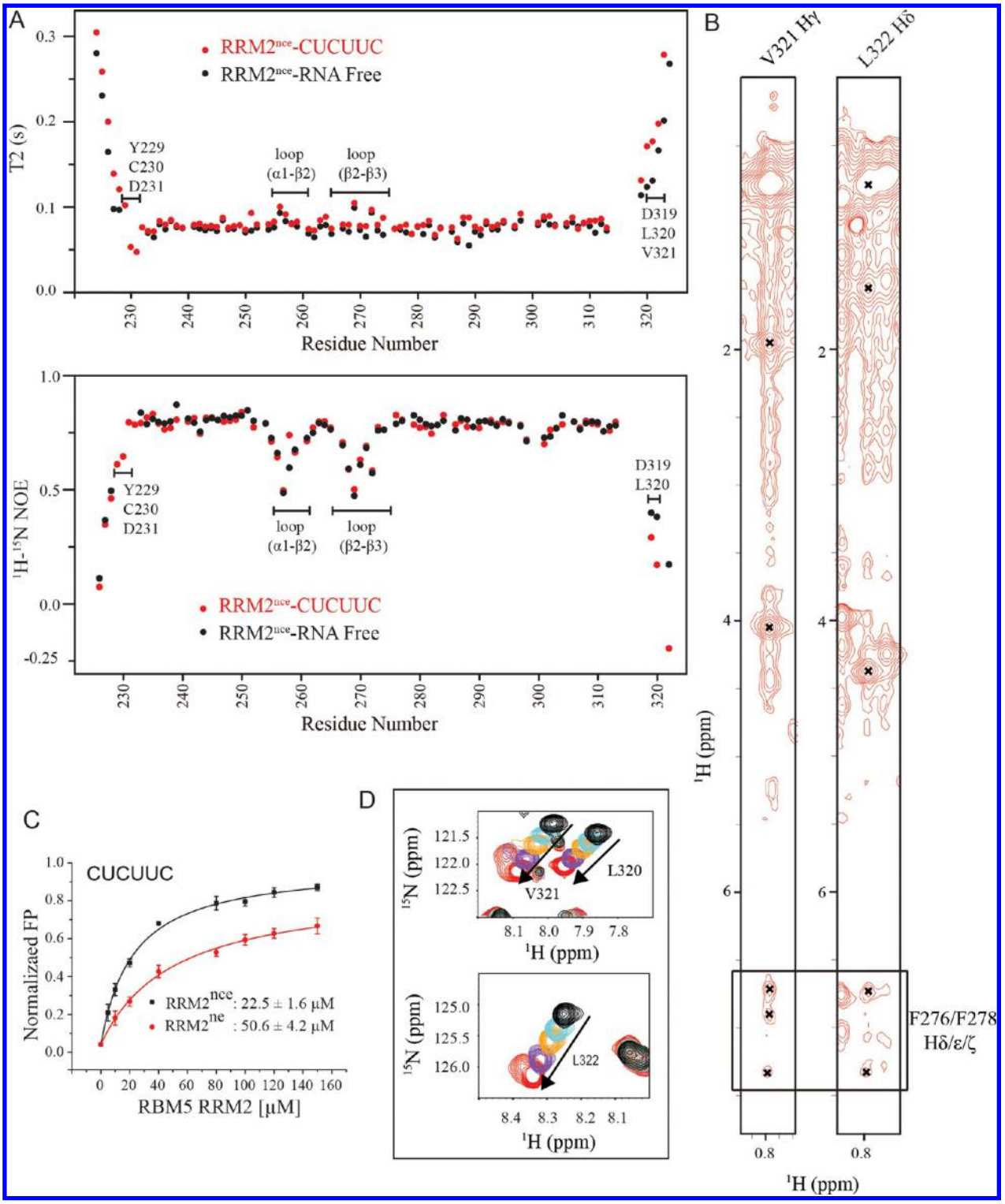


Figure 6. Dynamics of RBM5 RRM2^{nce}. (A) Superposition of the RNA-bound (red) onto the RNA-free (black) RBM5 RRM2^{nce} of T2 and heteronuclear NOEs of backbone resonances. Residues Asp324 and Gly325 (extremely fast motions) for which T2 and heteronuclear NOEs values outside the scale are not shown. (B) Section of 3D ^{13}C-NOESY-HSQC spectra of RNA-free RRM2^{nce}. The cross-peaks of V321Hγ/L322Hδ to the protons derived from the aromatic ring of Phe276 and Phe278 are shown in the black box below. (C) C-terminal extension enhanced CUCUUC binding. Fluorescence polarization of 5′-FAM labeled CUCUUC RNA upon RRM2^{nce} and RRM2^{ne} titration, respectively, in NMR buffer. (D) Close-up views of the ^{15}N−^{1}H HSQC spectra. The aliphatic residues (Leu320, Val321, and Leu322) on the C-terminal region are shown.

observation is the presence of two preferred target RNA sequences, the sequences [5′-(CUCUUC)-3′] and [5′-(GAGAAG)-3′]. For insights into the base-specific recognition of the RBM5 RRM2, we searched for the complex structures of the RRMs that recognize a polypurine or polypyrimidine RNA fragment. We found that the mechanisms of RNA recognition by Tra2-β RRM and SXL RBD1 gave clues about the RBM5 RRM2 RNA binding characteristics.

dx.doi.org/10.1021/bi300539t | *Biochemistry* 2012, 51, 6667−6678

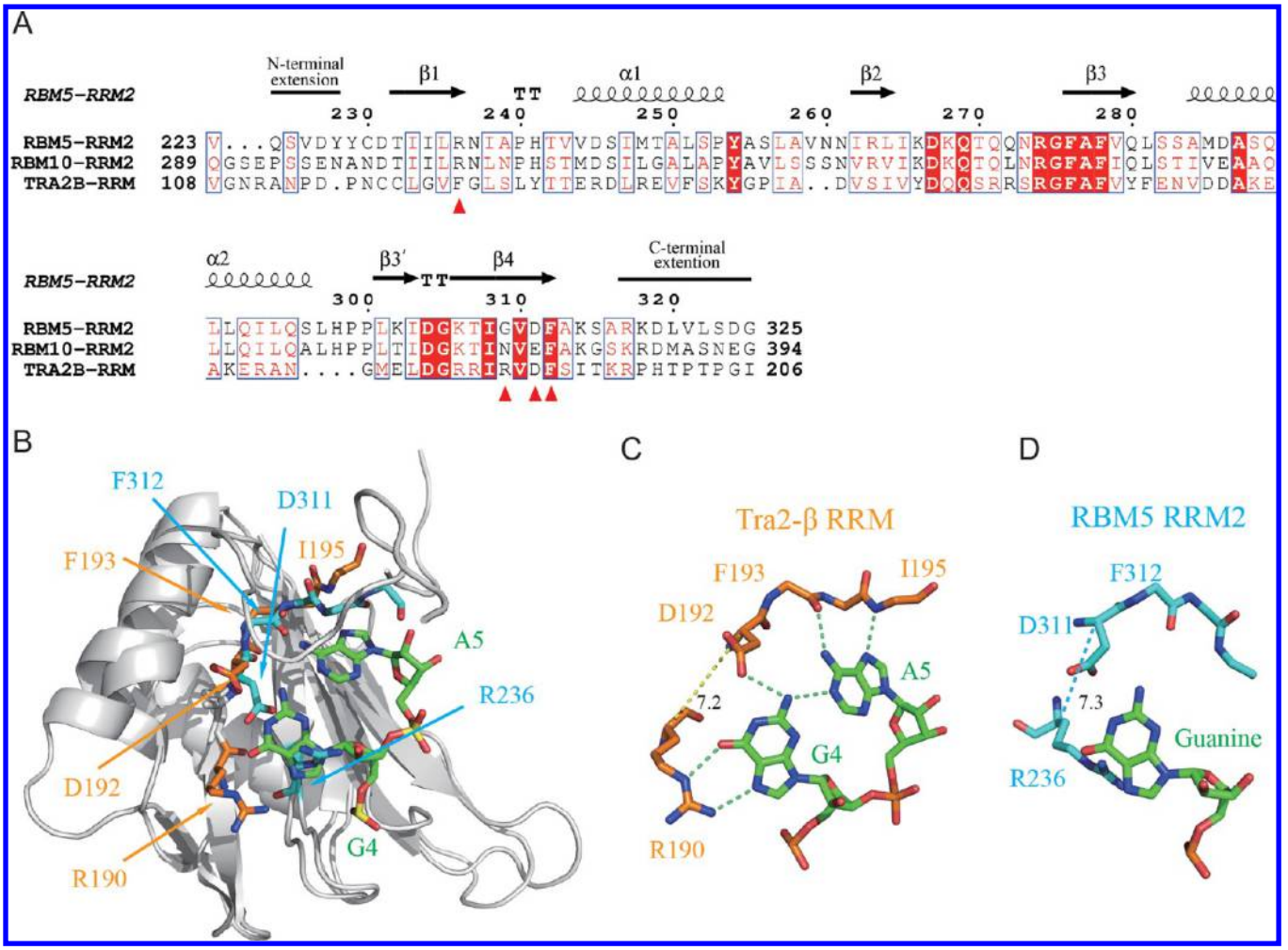


Figure 7. Comparison of the RNA recognition modes between the RBM5 RRM2 and Tra2-β RRM. (A) Multiple sequence alignment of the RBM5 RRM2 and other RRMs. The RRMs from RBM5 (UniProt P52756), RBM10 (UniProt P98175), and Tra2-β (UniProt Q8N1H4) were aligned as in Figure 1A. The red triangles indicate the conserved residues that are important for base-specific recognition. (B) Structural superposition of the RBM5 RRM2 onto the TRA2B RRM with the G4A5 nucleotides (green). The side chains of the residues from the RBM5 were shown and denoted in blue. Those from the Tra2-β were shown and denoted in orange. (C) Close-up views of the base-specific hydrogen-bond network of GA nucleotides. The green dotted lines represent the hydrogen bonds, and the orange dotted lines show the distances between Cα atoms of the denoted residues. (D) The model of guanine base binding pocket of RBM5 RRM2. The blue dotted lines show the distances between Cα atoms of the Asp311 and Arg236 residues.

End of the β4 Strand Region Contributes to Preferential Recognition for the Guanine Base. When we analyzed the NMR perturbation data, we observed a dramatic difference of perturbation values of Phe312 between GAGAAG and CUCUUC fragment titration (Figure 5A and B). Thus, the end of the β4 strand region of RBM5 RRM2 is only affected by GAGAAG fragment binding, but not by CUCUUC fragment, revealed that a base-specific binding pocket may exist. A multiple sequence alignment of the RBM5 RRM2 with the Tra2-β RRM (Figure 7A) showed a high similarity around the β4 strand region in the primary sequence. Thus, we compared the 3D structure of RBM5-RRM2 with Tra2-β RRM in complex with the GAAGAA fragment (PDB ID: 2RRA) to investigate how RBM5-RRM2 recognizes the RNA target (Figure 7B). In the case of Tra2-β, the side chains of residues Arg190 and Asp192 formed a delicate base-specific binding pocket (Figure 7C and D). The N2 amino, O6, and N7 groups of G4 formed hydrogen bonds with D192 Oδ, Arg192 Hε, and Hη.[24] In addition, it was often found that the guanine O6 and N7 are simultaneously in contact with an arginine guanidinium group in protein−RNA complex structures.[25,26] However, the side chain of aspartate can only act as a hydrogen-bond acceptor which expels the O2 group of pyrimidine.[23] In the case of RBM5 RRM2, we found that both solvent-exposed side chains of Asp311 and Arg236 composed a similar base-specific binding pocket on the β-sheet surface. Furthermore, the distance between Cα atoms of these residues is 7.3 Å, almost the same as the case of Tra2-β (Figure 7C and D). The residue composition pattern clarifies the preference for the guanine base moiety by the RBM5 RRM2, which is in good agreement with our fluorescence polarization data that the protein RRM2^{nce} could bind the sequence [5′-(GAGAAG)-3′] about 4 times stronger than the sequence [5′-(AAAAAA)-3′] (Figure 3).

Uracil Base Specific Binding Pocket. When we focus on the RNA binding character of the β3 and β2 strands region, we found a RNA binding pocket which well accommodates a pyrimidine base group. Multiple sequence alignment and structural superposition of the RBM5 RRM2 with the SXL RBD1 revealed the mechanisms of this binding site. In the case

dx.doi.org/10.1021/bi300539t | Biochemistry 2012, 51, 6667−6678

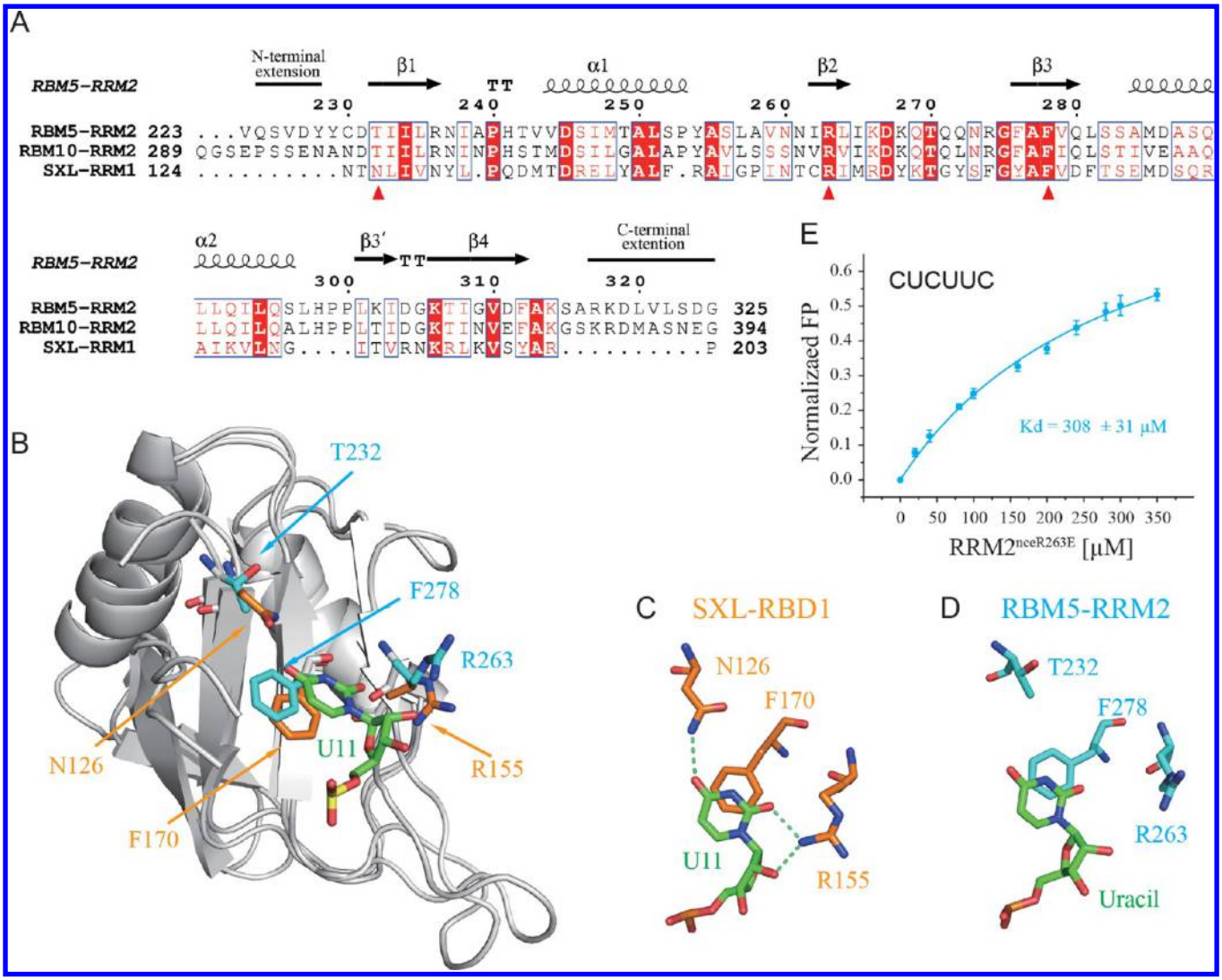


Figure 8. Comparison of the RNA recognition modes between RBM5-RRM2 and SXL RMM. (A) Multiple sequence alignment of the RBM5 RRM2 and other RRMs. The RRMs from RBM5 (UniProt P52756), RBM10 (UniProt P98175) and SXL (UniProt P19339) were aligned as in Figure 1A. The red triangles indicate the conserved residues that formed the binding pocket for a uracil recognition. (B) Structural superposition of RBM5 RRM2 onto SXL-RBD1 with the U11 nucleotide (green). The side chains of the residues from RBM5 are shown and denoted in blue. Those from the SXL are shown and denoted in orange. (C) Close-up views of the binding pocket of the base moiety of uracil. The green dotted lines represent the hydrogen bonds. (D) The model of uracil base binding pocket of RBM5 RRM2. (E) Point mutant analysis on the uracil binding pocket. Fluorescence polarization of 5′-FAM labeled CUCUUC RNA upon $RRM2^{nceR263E}$ titration in 40 mM phosphate buffer (pH 6.4) containing 50 mM NaCl, 4 mM DTT, and 1 mM NaN_3.

of SXL, the side chains of Phe170, Arg 155, and Asn126 formed a uracil base-specific binding pocket (Figure 8B). The uracil base of U11 stacks on the aromatic ring of Phe170. The O2′, O2, and O4 groups of U11 form hydrogen bonds to the Arg155 Hη1, Arg155 Hη2, and Asn126Hδ.[27] In the case of RBM5, the similar base-specific binding pocket was formed by the residues Phe278, Arg263, and Thr232 (Figure 8C and D). The position corresponding to the recognition of the O4 group is occupied by a Thr residue, instead of an Asn residue. The smaller side chain of Thr would accommodate both the uracil O4 group and the cytosine N4 amino group, while the −OH group of Thr could act as both the hydrogen-bond donor and acceptor. In addition, both uracil and cytosine containing an O2 group could form hydrogen bonds to the Arg263 side chain. This may explain the one-point mutant protein $RRM2^{nceR263E}$ bound to the [5′-(CUCUUC)-3′] sequence with a K_d value of 308 μM in NMR buffer condition, about 10-fold weaker than the wild type $RRM2^{nce}$ protein (Figure 8E).

Loss of Stacking Interaction and Low Binding Affinity. In the canonical RRM, three highly conserved aromatic residues (positions 3 and 5 of RNP1 and position 2 of RNP2) are present in the center of the β-sheet surface and normally formed stacking interactions.[14] A survey of the RRM and RNA complex structures revealed that a nucleotide stacked with the aromatic side chain always provides a higher affinity than the replacement by nonaromatic side chain.[23] In the case of the FOX-1 complex, mutation of a phenylalanine on position 5 of RNP1, to alanine, leads to a ∼30 000-fold loss in affinity.[28] In the case of RBM5 RRM2, one conserved aromatic residue on position 2 of RNP2 is replaced by residue Ile234. Furthermore, on the side of the β-sheet surface, the α1-β2 and β2-β3 loop regions also lack aromatic residues. These features may cause the RBM5 RRM2 to recognize the RNA

dx.doi.org/10.1021/bi300539t | Biochemistry 2012, 51, 6667−6678

sequence with affinity only in the 10^{-5} molar range. Compared to other single RRM complexes with single stranded RNA, for example, the FOX-1 with 1 nM, the Tra2-β1 with 2 μM, CUG-BP1 RRM3 with 1.9 μM, and SRSF2 with 0.27 μM, the affinity of the RBM5 RRM2 binding the sequence [5′-(CUCUUC)-3′] with 57 μM is quite low. Our fluorescence polarization data show that the one-point mutant protein RRM2^{nceI234F} (the position 2 of RNP2) is 6-fold stronger than the wild type when binding to the [5′-(CUCUUC)-3′] sequence with a K_d value of 5 μM in NMR buffer condition. We also tested the binding affinity of RRM2^{nceF276A} (position 3 of RNP1) and RRM2^{nceF278A} (position 5 of RNP1); in each case, the mutation causes a 6-fold loss in affinity (Figure 3B).

Flexibility of RNA Recognition. In this study, it is much unexpected that the complementary RNA sequences [5′-CUCUUC-3′] and [5′-GAGAAG-3′] are both preferred target RNAs for RBM5 RRM2. Further, the chemical shift perturbation data showed that the RBM5 RRM2 could make contact with both RNAs through highly similar binding surfaces, especially at the β1, β2, and β3 regions (Figure 5D and E). It seems that the protein pockets on the β-sheet surface of RBM5 RRM2 can accommodate different bases equally well. It has been reported that RRMs identified can bind diverse sequences with similar affinity, even in sequence-specific mode. It has been found that the nucleotide binding pocket U3 of HuD can accommodate the A3 base without any conformational change in the protein pocket. However, two different complex structures showed that the U3 adopts an anti conformation, and at the same position, A3 adopts a syn conformation.[29,30] Very recently, Allain's group revealed that the SRSF2 RRM could also recognize either G or C nucleotides at the same binding pockets without protein conformational change. Similarly, two guanines adopt a syn conformation to expose their Hoogsteen edge that resembles the Watson−Crick edge of cytosine with an anti conformation.[30] The above mentioned examples imply that some of the RRM RNA binding pockets have a capability to recognize the purine and pyrimidine exchange bases equally well. In the case of RBM5 RRM2, the RNA binding surfaces depicted by mapping the perturbed residues were highly similar at the β1, β2, and β3 regions in the presence two different RNAs. However, considering the directions of chemical shift movement, at the same binding surface regions, the perturbation pattern of some residues (Ile233, Leu235, and Arg236 in the β1 strand region; Arg263 and Ile265 in the β2 strand region; Phe276 and Phe278 in the β3 strand region) was quite different (Figure 5 C), which may be due to the different ring current effect of purine compared with that of pyrimidine. Our NMR studies in combination with the exhausted FP measurements for K_d values provide strong evidence that the RBM5 RRM2 has a capability to accommodate both the purine and pyrimidine sequences properly.

RNA-binding proteins specifically bind to target pre-mRNAs to regulate alternative splicing, in which there is a canonical interaction between cis and trans elements in the level of gene control.[31] These sequence-specific interactions are not always with high binding affinity, but the specificities are usually biologically critical.[14,32] For instance, the four RRM domains of PTB, each bind their RNA targets with the dissociation constant in the micromolar range. However, any mutations of binding sites to poly(C) or poly(U) alter PTB function in splicing considerably.[33,34] It has been reported that mutating the RBM5-binding sequence (CU rich sequence) obviously changes the RBM5 activity in regulating casp-2 alternative splicing.[10] Accordingly, our studies provide evidence for RBM5 sequence-specific interaction with this cis element in intron 9 of casp-2 pre-mRNA.

The present structural study revealed the flexibility of RNA recognition of RBM5 RRM2. The complex structures of the RBM5 RRM2 and the CU/GA RNA will precisely clarify the RNA recognition mode.

■ ASSOCIATED CONTENT

Accession Codes

Structure coordinate of RBM5 RRM2 has been deposited in the Protein Data Bank (accession code 2LKZ).

■ AUTHOR INFORMATION

Corresponding Author

*(J.W.) Tel: 86-551- 3600394. Fax: 86-551- 3600374. E-mail: wujihui@ustc.edu.cn. (Y.S.) Tel: 86-551-3607464. Fax: 86-551-3600441. E-mail: yyshi@ustc.edu.cn.

Funding

This work was supported by the National Basic Research Program of China [2012CB917201, 2011CB966302, and 2011CB911104] and the Chinese National Natural Science Foundation [30830031 and 31170693].

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Professor Xu Song (Sichuan University) for supplying the full-length RBM5 plasmid. We thank Dr. Weiwei Wang, Dr. Qin Su, and Dr. Jianping Liu for helpful advice and discussions. We thank Dr. F. Delaglio and Professor A. Bax for providing the software NMRPipe, Professor T. D. Goddard and Professor D. G. Kneller for providing Sparky, Professor A. T. Brünger for providing the program CNS, Dr. R. Koradi and Professor K. Wüthrich for providing MOLMOL, and Warren L. DeLano for providing PyMOL.

■ ABBREVIATIONS USED

RRM, RNA recognition motif; HSQC, heteronuclear single quantum coherence; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TOCSY, total correlation spectroscopy; DEPC, diethypyrocarbonate; K_d, equilibrium dissociation constants; RNP, ribonucleoprotein; c-FLIP, cellular FLICE inhibitory protein; Tra2-β, transformer2-β; PTB, polypyrimidine tract binding protein; SXL, sex-lethal protein

■ REFERENCES

(1) Drabkin, H. A., West, J. D., Hotfilder, M., Heng, Y. M., Erickson, P., Calvo, R., Dalmau, J., Gemmill, R. M., and Sablitzky, F. (1999) DEF-3(g16/NY-LU-12), an RNA binding protein from the 3p21.3 homozygous deletion region in SCLC. *Oncogene 18*, 2589−2597.

(2) Ramaswamy, S., Ross, K. N., Lander, E. S., and Golub, T. R. (2003) A molecular signature of metastasis in primary solid tumors. *Nat. Genet. 33*, 49−54.

(3) Rintala-Maki, N. D., Goard, C. A., Langdon, C. E., Wall, V. E., Traulsen, K. E., Morin, C. D., Bonin, M., and Sutherland, L. C. (2007) Expression of RBM5-related factors in primary breast tissue. *J. Cell Biochem. 100*, 1440−1458.

(4) Mourtada-Maarabouni, M., Keen, J., Clark, J., Cooper, C. S., and Williams, G. T. (2006) Candidate tumor suppressor LUCA-15/

RBM5/H37 modulates expression of apoptosis and cell cycle genes. *Exp. Cell Res. 312*, 1745−1752.
(5) Kobayashi, T., Ishida, J., Musashi, M., Ota, S., Yoshida, T., Shimizu, Y., Chuma, M., Kawakami, H., Asaka, M., Tanaka, J., Imamura, M., Kobayashi, M., Itoh, H., Edamatsu, H., Sutherland, L. C., and Brachmann, R. K. (2011) p53 transactivation is involved in the antiproliferative activity of the putative tumor suppressor RBM5. *Int. J. Cancer 128*, 304−318.
(6) Hartmuth, K., Urlaub, H., Vornlocher, H. P., Will, C. L., Gentzel, M., Wilm, M., and Luhrmann, R. (2002) Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16719−16724.
(7) Behzadnia, N., Golas, M. M., Hartmuth, K., Sander, B., Kastner, B., Deckert, J., Dube, P., Will, C. L., Urlaub, H., Stark, H., and Luhrmann, R. (2007) Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *EMBO J. 26*, 1737−1748.
(8) Chen, Y. I., Moore, R. E., Ge, H. Y., Young, M. K., Lee, T. D., and Stevens, S. W. (2007) Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors. *Nucleic Acids Res. 35*, 3928−3944.
(9) Bonnal, S., Martinez, C., Forch, P., Bachi, A., Wilm, M., and Valcarcel, J. (2008) RBM5/Luca-15/H37 regulates Fas alternative splice site pairing after exon definition. *Mol. Cell 32*, 81−95.
(10) Fushimi, K., Ray, P., Kar, A., Wang, L., Sutherland, L. C., and Wu, J. Y. (2008) Up-regulation of the proapoptotic caspase 2 splicing isoform by a candidate tumor suppressor, RBM5. *Proc. Natl. Acad. Sci. U.S.A. 105*, 15708−15713.
(11) Nguyen, C. D., Mansfield, R. E., Leung, W., Vaz, P. M., Loughlin, F. E., Grant, R. P., and Mackay, J. P. (2011) Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA. *J. Mol. Biol. 407*, 273−283.
(12) Sutherland, L. C., Rintala-Maki, N. D., White, R. D., and Morin, C. D. (2005) RNA binding motif (RBM) proteins: a novel family of apoptosis modulators? *J. Cell Biochem. 94*, 5−24.
(13) Varani, G., and Nagai, K. (1998) RNA recognition by RNP proteins during RNA processing. *Annu. Rev. Biophys. Biomol. Struct. 27*, 407−445.
(14) Maris, C., Dominguez, C., and Allain, F. H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J. 272*, 2118−2131.
(15) Rule, G. S., and Hitchens, T. K. (2006) *Fundamentals of Protein NMR Spectroscopy*, Vol. *5*, Springer, New York.
(16) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(17) Goddard, T. D., and Kneller, D. G. *SPARKY 3*, University of California, San Francisco, CA.
(18) Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*, 65−73.
(19) Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc. 48*, 47−62.
(20) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR 44*, 213−223.
(21) Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR 8*, 477−486.
(22) Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Biomol. Graphics 14* (51−55), 29−32.
(23) Auweter, S. D., Oberstrass, F. C., and Allain, F. H. (2006) Sequence-specific binding of single-stranded RNA: is there a code for recognition? *Nucleic Acids Res. 34*, 4943−4959.
(24) Tsuda, K., Someya, T., Kuwasako, K., Takahashi, M., He, F., Unzai, S., Inoue, M., Harada, T., Watanabe, S., Terada, T., Kobayashi, N., Shirouzu, M., Kigawa, T., Tanaka, A., Sugano, S., Guntert, P., Yokoyama, S., and Muto, Y. (2011) Structural basis for the dual RNA-recognition modes of human Tra2-beta RRM. *Nucleic Acids Res. 39*, 1538−1553.
(25) Morozova, N., Allers, J., Myers, J., and Shamoo, Y. (2006) Protein-RNA interactions: exploring binding patterns with a three-dimensional superposition analysis of high resolution structures. *Bioinformatics 22*, 2746−2752.
(26) Allers, J., and Shamoo, Y. (2001) Structure-based analysis of protein-RNA interactions using the program ENTANGLE. *J. Mol. Biol. 311*, 75−86.
(27) Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature 398*, 579−585.
(28) Auweter, S. D., Fasan, R., Reymond, L., Underwood, J. G., Black, D. L., Pitsch, S., and Allain, F. H. (2006) Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *EMBO J. 25*, 163−173.
(29) Wang, X., and Tanaka Hall, T. M. (2001) Structural basis for recognition of AU-rich element RNA by the HuD protein. *Nat. Struct. Biol. 8*, 141−145.
(30) Daubner, G. M., Clery, A., Jayne, S., Stevenin, J., and Allain, F. H. T. (2012) A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. *EMBO J. 31*, 162−174.
(31) Matlin, A. J., Clark, F., and Smith, C. W. (2005) Understanding alternative splicing: towards a cellular code. *Nat. Rev. Mol. Cell Biol. 6*, 386−398.
(32) Oberstrass, F. C., Auweter, S. D., Erat, M., Hargous, Y., Henning, A., Wenter, P., Reymond, L., Amir-Ahmady, B., Pitsch, S., Black, D. L., and Allain, F. H. (2005) Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science 309*, 2054−2057.
(33) Auweter, S. D., Oberstrass, F. C., and Allain, F. H. (2007) Solving the structure of PTB in complex with pyrimidine tracts: an NMR study of protein-RNA complexes of weak affinities. *J. Mol. Biol. 367*, 174−186.
(34) Perez, I., Lin, C. H., McAfee, J. G., and Patton, J. G. (1997) Mutation of PTB binding sites causes misregulation of alternative 3′ splice site selection in vivo. *RNA 3*, 764−778.
(35) Notredame, C., Higgins, D. G., and Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol. 302*, 205−217.
(36) Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics 15*, 305−308.