PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Novel RNA recognition motif domain in the cytoplasmic polyadenylation element binding protein 3

Kengo Tsuda,[1,2] Kanako Kuwasako,[1,2,3] Takashi Nagata,[1,4,5] Mari Takahashi,[1,2] Takanori Kigawa,[1] Naohiro Kobayashi,[1,6] Peter Güntert,[7,8,9,10] Mikako Shirouzu,[1,2] Shigeyuki Yokoyama,[1,11]* and Yutaka Muto[1,2,3]*

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

[2] Division of Structural and Synthetic Biology, RIKEN Center for Life Science Technologies, Tsurumi-ku, Yokohama 230-0045, Japan

[3] Faculty of Pharmacy and Research Institute of Pharmaceutical Sciences, Musashino University, Tokyo 202-8585, Japan

[4] Structural Energy Bioscience Research Section, Institute of Advanced Energy, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

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

[6] Laboratory of Molecular Biophysics, Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, Japan

[7] Tatsuo Miyazawa Memorial Program, RIKEN Genomic Sciences Center, Yokohama 230-0045, Japan

[8] Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, Goethe University Frankfurt, Frankfurt am Main, Germany

[9] Frankfurt Institute of Advanced Studies, Goethe University Frankfurt, Frankfurt am Main, Germany

[10] Department of Chemistry, Graduate School of Science and Engineering, Tokyo Metropolitan University, Hachioji, Japan

[11] RIKEN Structural Biology Laboratory, Tsurumi-ku, Yokohama 230-0045, Japan

ABSTRACT

The family of cytoplasmic polyadenylation element binding proteins CPEB1, CPEB2, CPEB3, and CPEB4 binds to the 3′-untranslated region (3′-UTR) of mRNA, and plays significant roles in mRNA metabolism and translation regulation. They have a common domain organization, involving two consecutive RNA recognition motif (RRM) domains followed by a zinc finger domain in the C-terminal region. We solved the solution structure of the first RRM domain (RRM1) of human CPEB3, which revealed that CPEB3 RRM1 exhibits structural features distinct from those of the canonical RRM domain. Our structural data provide important information about the RNA binding ability of CPEB3 RRM1.

Proteins 2014; 82:2879–2886.
© 2014 Wiley Periodicals, Inc.

Key words: 3′-UTR; CPEB3; CPE; GluR2; Mrna; RRM; NMR; structure.

Additional Supporting Information may be found in the online version of this article.
Grant sponsor(s): RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, the Ministry of Education, Culture, Sports, Science and Technology of Japan; Grant sponsor: Human Frontier Science Program (HFSP to Y.M.); Grant sponsor: Scientific Research of the Japan Society for the Promotion of Science (JSPS; Grant-in-Aid to P.G.) and by the Volkswagen Foundation.
Takanori Kigawa's current address is RIKEN Quantitative Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan.
Kengo Tsuda and Kanako Kuwasako contributed equally to this work.
*Correspondence to: Shigeyuki Yokoyama, RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan. E-mail: yokoyama@riken.jp or Yutaka Muto, Faculty of Pharmacy, Department of Pharmaceutical Sciences, Musashino University, 1-1-20 Shinmachi Nishitokyo-shi, Tokyo 202-8585, Japan. E-mail: ymuto@musashino-u.ac.jp
Received 28 April 2014; Revised 3 July 2014; Accepted 15 July 2014
Published online 26 July 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24651

INTRODUCTION

The family of human cytoplasmic polyadenylation element binding proteins CPEB1, CPEB2, CPEB3, and CPEB4 has a common domain organization, involving two consecutive RNA recognition motif (RRM) domains followed by a zinc finger domain (ZF) in the C-terminal region [Fig. 1(A)].[1] They bind to the 3′-untranslated region (3′-UTR) of mRNA and play significant roles in mRNA metabolism and translation regulation.[1]

Among the members of the CPEB family, CPEB1 was first identified in the translational control of several maternal mRNAs during oocyte maturation. Before fertilization, CPEB1 binds to the consensus 5′-UUUUUAU-3′ sequence [the cytoplasmic polyadenylation element (CPE)] in the 3′-UTR of several maternal mRNAs in the oocytes, and maintains their translational dormancy.[1] On the phosphorylation of CPEB1 in response to fertilization, it regulates the cytoplasmic polyadenylation of these mRNAs, along with other factors such as PAPD4, leading to translation initiation in the process of oocyte maturation.[2] Three other homologous proteins, CPEB2, CPEB3, and CPEB4, were successively identified based on their sequence similarity to CPEB1. CPEB2, and CPEB4 also bind to the CPE sequence, but with lower affinity as compared to CPEB1.[1] Conversely, CPEB3 reportedly binds to a specific RNA sequence in the 3′-UTR of the glutamate receptor (GluR2) mRNA,[3] and negatively regulates GluR2 translation by accelerating the deadenylation and decay of its mRNA, through interactions with the Tob protein bound to Caf1 deadenylase.[4] The target RNA sequence of CPEB3 is considered to adopt a specific secondary structure, while the CPE sequence is in the single-stranded form.[3] Huang *et al.* reported that CPEB1 RRM1 plays a substantial role in facilitating CPEB1 binding to the CPE sequence. Conversely, for strong binding of CPEB3 to the target structured RNA, the CPEB3 RRM1 is not sufficient, and the cooperativity between RRM1 and RRM2 is necessary.[3] Thus, although CPEB1 and CPEB3, as well as CPEB2 and CPEB4, are in the same protein family, their preferences for target RNA molecules and the involvement of RRM1 of CPEB protein family in the target RNA binding are distinct from each other.

The RRM domain is present in numerous eukaryotic RNA binding proteins, and mainly plays important roles in sequence-specific RNA binding.[5–7] The canonical RRM consists of a four-stranded antiparallel β-sheet packed against two α-helices, with a typical β1-α1-β2-β3-α2-β4 topology [the amino-acid sequence of the first RRM of U1A is shown in Fig. 1(B), as a canonical RRM]. Two well-conserved consensus sequences, RNP2 and RNP1, correspond to the β1 and β3 strands, respectively [Fig. 1(B)], and the RRM typically interacts with RNA bases on its β-sheet surface, through stacking interactions mediated by the conserved and exposed aromatic rings located on the β1 and β3 strands.[5–8] However, several RRMs with unique structural features, such as PARN RRM and hnRNP-F RRM, reportedly do not utilize the β-sheet surface to interact with RNA. In addition, some RRM domains function as protein–protein interaction modules through the β-sheet surface or the opposite α-helical surface.[5–7] The elucidation of novel structures of these widespread protein modules has thus revealed an increasing number of RNA/protein recognition modes by the RRM domain.

Comparisons of the primary sequences between the CPEB family members and the canonical RRMs revealed that the first RRM (RRM1) of the CPEB family possesses a longer β2-β3 loop region [Fig. 1(B)]. In addition, a long amino-acid segment is inserted between the α2 and β4 secondary structural elements. Therefore, structural information about the first RRM of the CPEB family is needed to clarify the multifunctional properties of the CPEB protein family.

In this study, we solved the solution structure of CPEB3 RRM1, which is the first structure determined for a functional domain of any member of the CPEB family. We discovered that CPEB3 RRM1 possesses an additional β-hairpin structure immediately following the second α-helix, and adopts a unique structure, as compared to the canonical RRM fold. Moreover, we used NMR to examine the interactions of CPEB3 RRM1 with the reported target RNA molecules, and provide evidence that the novel structural insertion is involved in the RNA binding activity of CPEB3 RRM1.

MATERIALS AND METHODS

Protein expression and purification

The DNA fragment encoding the CPEB3 RRM1 domain (residues Ser440-Asp540) (SwissProt accession no. Q8NE35) was subcloned by PCR from the human full-length cDNA clone into the expression vector pCR2.1 (Invitrogen), as a fusion with an N-terminal 6-His affinity tag and a TEV protease cleavage site. The $^{13}C/^{15}N$-labeled fusion protein was synthesized by a cell-free protein expression system.[9] The reaction mixture was first adsorbed to a HiTrap chelating column (Amersham Biosciences) with buffer A (50 m*M* sodium phosphate buffer, pH 8.0, containing 500 m*M* sodium chloride and 20 m*M* imidazole). The protein was eluted with buffer B (50 m*M* sodium phosphate buffer, pH 8.0, containing 500 m*M* sodium chloride and 500 m*M* imidazole). To remove the His-tag, the eluted protein was incubated with TEV protease at 30°C for 1 h. The tag-free CPEB3 RRM1 was further purified by HiTrap Q and HiTrap SP column chromatography (GE Healthcare).

Nuclear magnetic resonance spectroscopy

For nuclear magnetic resonance (NMR) measurements, the samples were concentrated to 1.0 m*M* in 20 m*M*

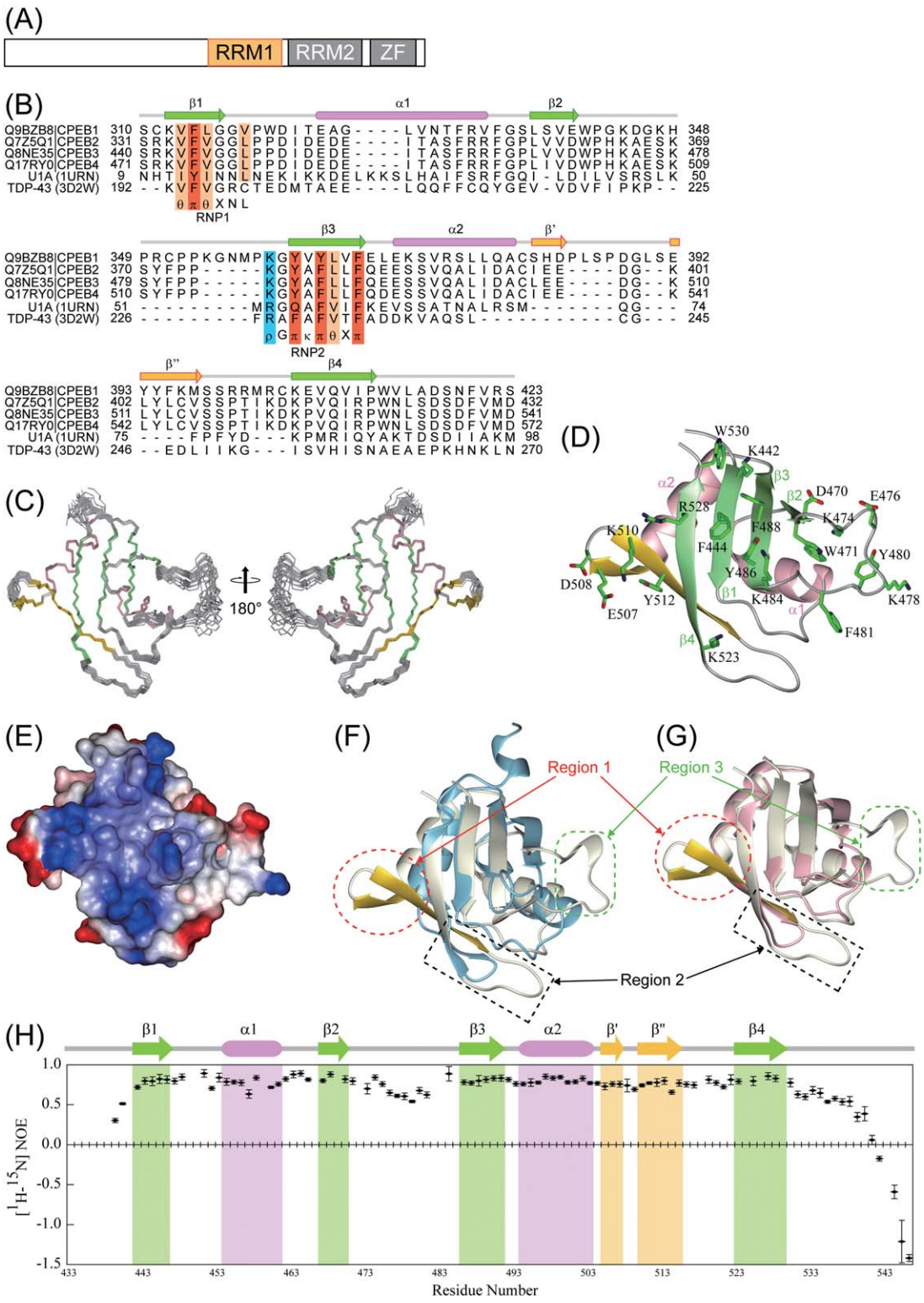

Figure 1

Structures and dynamics of CPEB3 RRM1. (**A**) Schematic representation of the CPEB3 protein. (**B**) Sequence alignment of the RRM1s of the human CPEB family and the RRMs of two previously reported structures: CPEB1 (Q9BZB8), CPEB2 (Q7Z5Q1), CPEB3 (Q8NE35), CPEB4 (Q17RY0), U1A, and TDP-43 RRM2. The secondary structure elements of CPEB3 RRM1 are displayed above the amino-acid sequences. The sequences of RNP1 and RNP2 are displayed below the sequences, where θ, π, ρ, κ, and X indicate I/L/V, F/Y, K/R, G/A, and any amino acid residues, respectively. (**C**) Superposition of the 20 conformers of the CPEB3 RRM1 NMR structure (Gly439-Asn531). The secondary elements are colored as follows: α-helix, pink; β1–4 strands, light green; and β′, β″ strands, light orange. (**D**) Ribbon representation of the lowest-energy conformer of the CPEB3 RRM1 in view of the β-sheet surface. The side chains of aromatic, positively charged and negatively charged residues are showed in neon green. Heteroatoms are colored blue (nitrogen) and red (oxygen), respectively. (**E**) Electrostatic potential surface representation of CPEB3 RRM1 in view of the β-sheet surface. Negative and positive charges are colored blue and red, respectively. (**F** and **G**) Superpositions of the NMR structures of CPEB3 RRM1 with U1A RRM (F) and TDP-43 RRM2 (G), respectively. U1A RRM and TDP-43 RRM2 are colored light blue and pink, respectively. CPEB3 RRM1 is colored light yellow, and the characteristic β-hairpin structure within CPEB3 RRM1 is colored gold. (**H**) Heteronuclear NOE values of CPEB3 RRM1. The secondary structure elements are displayed at the top of the diagram.

Tris-d11-HCl buffer (pH 7.0), containing 100 m*M* NaCl, 1 m*M* 1,4-*DL*-dithiothreitol-d_{10} (*d*-DTT) and 0.02% NaN_3 (in 90% H_2O/10% D_2O), using an Amicon Ultra-15 filter (3000 MWCO, Millipore). NMR experiments were performed at 298 K on Bruker AV600 and AV800 spectrometers. The ^1H, ^{15}N, and ^{13}C chemical shifts were referenced relative to the frequency of the ^2H lock resonance of water. Backbone and side-chain assignments of CPEB3 RRM1 were obtained by using a combination of standard triple resonance experiments.[10,11] 2D [^1H, ^{15}N]-HSQC and 3D HNCO, HN(CA)CO, HNCACB, and CBCA(CO)NH spectra were used to obtain the ^1H, ^{15}N, and ^{13}C assignments of the protein backbone. The ^1H and ^{13}C assignments of the nonaromatic side-chains, including all prolines, were obtained using 2D [^1H, ^{13}C]-HSQC, and 3D HBHA(CO)NH, H(CCCO)NH, (H)CC(CO)NH, HCCH-COSY, and HCCH-TOCSY spectra. Assignments were checked for consistency with 3D ^{15}N-edited [^1H, ^1H]-NOESY and ^{13}C-edited [^1H, ^1H]-NOESY spectra. The ^1H and ^{13}C spin systems of the aromatic rings of Phe, Tyr, Trp, and His were identified using 3D HCCH-COSY and HCCH-TOCSY experiments, and 3D ^{13}C-edited [^1H, ^1H]-NOESY was used for the sequence-specific resonance assignment of the aromatic side-chains. A 3D HNHA spectrum was measured to determine the φ dihedral angle restraints. 3D ^{15}N- and ^{13}C-edited [^1H, ^1H]-NOESY spectra were recorded with mixing times of 80 ms. Analyses of the processed data and manual chemical-shift assignments were performed with the programs NMRView and KUJIRA. The chemical shift information about the backbone and side chain ^1H, ^{13}C, and ^{15}N resonances of CPEB3 RRM1 was deposited in the BioMagResDB (BMRB accession number: 11563).

The measurements of the ^1H-^{15}N heteronuclear nuclear overhauser effect (NOE) values were performed at 298 K on a Bruker AV600 spectrometer equipped with a cryo-probe, using the ^{15}N, ^{13}C-labeled CPEB3 RRM1 at a concentration of 200 μ*M*. The ^1H-^{15}N heteronuclear NOE values were calculated as the ratio between the cross-peak intensities with and without ^1H saturation.[12] The errors were estimated from the root-mean-square values of the baseline noise in the two spectra.

Structure calculations

The three-dimensional structure of CPEB3 RRM1 was determined with the program CYANA 2.1, which implements automated NOE assignments and structure calculations with torsion angle dynamics. Restraints for the φ and χ^1 dihedral angles were obtained by analyzing the 3D NOESY and 3D HNHA spectra,[13] and the patterns of the inter- and intraresidual NOE intensities, respectively.[14]

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. Further refinements by restrained molecular dynamics followed by restrained energy minimization were performed for the 40 conformers with the lowest final CYANA target function values, using the AMBER9 program with the AMBER 2003 force field and a generalized Born model, as described previously.[15] The 20 conformers that were most consistent with the experimental restraints were then used for further analyses. PROCHECK-NMR and MOLMOL were used to validate and to visualize the final structures, respectively. The final ensemble of 20 conformers was deposited in the Protein Data Bank (PDB ID: 2RUG).

Table I
Structural Statistics for the Solution Structure of CPEB3 RRM1

	CPEB3 RRM1
NMR restraints	
Distance restraints	
Total NOE	1646
Intraresidue	473
Inter-residue	
Sequential ($\|i-j\|=1$)	383
Medium-range ($1<\|i-j\|<5$)	242
Long-range ($\|i-j\|\geq 5$)	548
Hydrogen bond restraints[a]	16
Dihedral angle restraints	
φ angle	12
χ^1 angle	42
Structure statistics (20 structures)	
AMBER energies (kcal/mol)	
Mean AMBER energy	−3017.8
Mean restraints violation energy	5.64
Ramachandran plot statistics (%)	
Residues in most favored regions	85.3
Residues in additionally allowed regions	14.1
Residues in generously allowed regions	0.5
Residues in disallowed regions	0.1
Average R.M.S.D. to mean structure (Å)	
Protein backbone[b]	0.39
Protein heavy atoms[b]	0.96

[a]Only used in the CYANA calculation.
[b]For residues 442–472, 486–529.

RESULT AND DISCUSSION

Overall structure of the first RRM domain of CPEB3

We solved the solution structure of CPEB3 RRM1 (residues Ser440-Asp540) with excellent structure quality scores [Table I and Fig. 1(C,D)]. In total, 1646 NOE distance restraints derived from 3D ^{15}N-edited [^1H, ^1H]-NOESY and ^{13}C-edited [^1H, ^1H]-NOESY spectra were assigned and used in the structure calculation, together with 54 dihedral angle restraints (Table I).

The solution structure of CPEB3 RRM1 reveals that the Lys442-Trp530 region is the core RRM body, consisting of a six-stranded antiparallel β-sheet and two α-helices with a β1-α1-β2-β3-α2-β′-β″-β4 topology [the

numbered β strands are related to the typical four β-strands of the canonical RRMs, and the β′ and β″ strands are the unique ones in CPEB3 RRM1; Fig. 1(B)]. The β1 strand (Lys442-Gly446) is followed by the α1 helix (Glu454-Arg462). The β2 and β3 strands comprise Leu467-Asp470 and Tyr486-Phe491, respectively. Intriguingly, the unique β′ (Leu505-Glu507) and β″ (Lys510-Val515) strands immediately follow the α2 helix (Glu494-Ala503) [Fig. 1(C,D)] and form an additional β-hairpin structure, which is located beside the following β4 strand (Lys523-Pro529) and is involved in the formation of a positively charged six-stranded antiparallel β-sheet structure [Fig. 1(D,E)].

Structural comparison between CPEB3 RRM1 and canonical RRMs

We compare the structure of CPEB3 RRM1 with those of two previously reported RRM domains: U1A [PDB ID: 1URN; Fig. 1(F)], which is commonly referred to as the canonical RRM, and the second RRM domain of TDP-43 [PDB ID: 3D2W; Fig. 1(G)], with an overall structure very similar to that of CPEB3 RRM1 (*Z*-score: 7.4, RMSD: 1.76 Å for the C^{α} atoms of matched residues at the best 3D superimposition), as searched by the PDBeFold server (www.ebi.ac.uk/msd-srv/ssm/).

The structure superpositions of CPEB3 RRM1 with the U1A and TDP-43 RRMs reveal distinct structural differences in three regions [Regions 1, 2, and 3; Fig. 1(F,G]. Generally, the canonical RRM has a short β-turn structure immediately following the α2 helix in Region 1, and a β-hairpin structure just before the β4 strand in Region 2 [Fig. 1(F)]. In the case of CPEB3 RRM1, the amino-acid segment in Region 1 is longer than those of the canonical RRMs, and adopts the β′-β″-hairpin structure. In addition, the amino-acid segment in Region 2 of CPEB3 RRM1 is longer than those in the canonical RRMs, and forms a rather loose β-hairpin structure. These structural features contribute to the formation of the novel antiparallel β′-β″ sheet structure [Fig. 1(C,D,F)].

In addition, the β2-β3 loop (Trp471-Gly485) of CPEB3 RRM1, which corresponds to Region 3, is longer than those of the canonical RRMs, and exhibits a unique structural feature [Fig. 1(F,G)]. Three aromatic amino acids (Trp471, Tyr480, and Phe481) are located in the β2-β3 loop. NOEs were observed between the Hζ2 atom of Trp471 and the Hε atoms of Lys474, and between the Hη2 atom of Trp471 and the Hε atoms of Lys474, respectively, indicating that the aromatic ring of Trp471 stacks with the side chain of Lys474 and stabilizes the loop structure [Fig. 1(D) and Supporting Information Fig. 1(A)]. Conversely, the aromatic rings of Tyr480 and Phe481 are exposed to the solvent, and are surrounded by positively and negatively charged amino acid residues [Fig. 1(E)]. In the canonical RRM, this loop region is important for the RNA/RRM interaction, through electrostatic and hydrophobic interactions. Therefore, this loop region in CPEB3 RRM1, which includes exposed aromatic amino acids with an abundance of charged residues, could be advantageous for the interaction of CPEB3 RRM1 with RNA [Fig. 1(E)].

An amino-acid sequence comparison among all CPEB family members revealed that these characteristic structural features seem to be commonly shared within CPEB2-4 [Fig. 1(B)]. Therefore, we suggest that the first RRMs of CPEB2-4 form an extensive six-stranded β-sheet surface, including the characteristic β′-β″-hairpin structures. In the case of CPEB1 RRM1, larger numbers of amino acid residues compose the additional β-hairpin structure (Regions 1 and 2) and the loop region between the β2 and β3 strands (Region 3), as compared to those of the other CPEB2-4 proteins [Fig. 1(B)]. These findings suggest that the structure of CPEB1 RRM1 is slightly different from those of the CPEB2-4 proteins, and this could be the underlying reason for the differences in the target RNA sequences between CPEB1 and the other CPEB proteins.

Aromatic amino-acid residues on the extended β-sheet structure

In the canonical RRMs, the aromatic amino-acid residues on RNP1 and RNP2 play important roles in the protein/RNA interactions. The corresponding aromatic amino acids [Phe444 (β1), and Tyr486 and Phe488 (β3)] were identified in CPEB3 RRM1 [Fig. 1(D)] as well as the other members of the CPEB family [Fig. 1(B)]. In addition to the canonical aromatic amino acids, CPEB3 RRM1 has another aromatic residue (Tyr512) on the β″ strand, which is the characteristic structure of CPEB3 RRM1 [Fig. 1(D)]. Positively charged residues (β1: Lys442, β2-β3 loop: Lys478 and Lys484, β″: Lys510, and β4: Lys523 and Arg528) and negatively charged residues (β2: Asp470, β2-β3 loop: Glu476, β′: Glu507, and β′-β″ loop: Asp508) are located on the extensive β-sheet surface and surround these hydrophobic residues [Fig. 1(D,F)]. These findings suggest that CPEB3 RRM1 could interact with a target molecule through its extensive β-sheet surface.

Conversely, the aromatic ring of Trp530 interacts with those of Phe488 (β3) and Lys442 (β1) by cation-π interactions [Supporting Information Fig. 1(B)], and seems to fill the space where the RNA bases could be located [Fig. 1(D)]. In the canonical RRMs, the position of Trp530 (at the end of the β4 strand) is frequently occupied by a small amino-acid residue (Ala or Ser), which forms a van der Waals contact with the RNA base, which is stacked with the aromatic amino-acid residue on the β1 strand (corresponding to Phe444 in CPEB3 RRM1). Therefore, it could be considered that unless the interaction among Lys442, Phe488, and Trp530 changes such

that the space just above Phe488 opens, the canonical RNA recognition surface (β1 and β3 strands) of CPEB3 RRM1 could not be utilized for interactions with RNA molecules.

Dynamics of CPEB3 RRM1

To examine the stability of the characteristic structural elements of CPEB3 RRM1, we performed NMR experiments on a 600 MHz instrument to measure the ^{1}H-^{15}N heteronuclear NOEs, as shown in Figure 1(H). The ^{1}H-^{15}N heteronuclear NOE values of the last 5 residues in the C-terminal region were either very small or negative, indicating that the C-terminal region is highly flexible. Within the core RRM body (Lys442-Trp530), the NOE values for the β1, β2, β3, β4, β′, β″, α1, and α2 secondary elements were positive and high, indicating that these secondary elements form a rigid structure. These results reveal that this β′-β″-strand structure is as rigid as the other secondary structure regions, and contributes to the formation of the unique structure of CPEB3 RRM1.

The length of the β2-β3 loop of CPEB3 RRM1 (Region 3; Trp471-Gly485) is longer than those of the canonical RRMs. As described above, however, the interaction between Trp471 and Lys474 stabilizes the loop structure. Correspondingly, the lower heteronuclear NOE values are confined to the segment spanning residues Asp475-Phe481, suggesting the slightly higher flexible for this region.

Molecular interactions of CPEB3 RRM1

The CPEB3 protein reportedly binds to the 3′-UTR in the GluR2 mRNA. Previous SELEX experiments indicated two binding-sequence candidates in the target structured RNA sequences in the GluR2 mRNA, 5′-GAGGAU-3′ and a poly-U stretch,[3] while the CPE sequence contains only a poly-U stretch. Therefore, we examined if CPEB3 RRM1 could interact with the two RNA segments (the GAGGAU and U_5 sequences) by NMR titration experiments, and discovered that the GAGGAU sequence more effectively interacts with CPEB3 RRM1 than the U_5 sequence [Fig. 2(A)].

On addition of the GAGGAU sequence, two regions [the charged residues near Tyr512 and those near Tyr480 and Phe481 (including the N-terminus of the α1-helix)] were mainly perturbed. These two regions are located on the unique structural elements of CPEB3 RRM1 (the β′-β″ sheet and the β2-β3 loop), respectively. Thus, these results suggest that the two unique structural elements are involved in the recognition of the purine-rich RNA sequence, although the single GAGGAU sequence does not seem to cover both of the regions simultaneously.

NOE measurements revealed that Asp453 on the N-terminus of the α1-helix, which was the most effectively perturbed amino-acid residue, contacts Tyr480 and Phe481. Thus, it is considered that the chemical shift perturbation on the N-terminus of the α1-helix was triggered by the interaction between the aromatic amino-acid residues (Tyr480 and Phe481) and the target RNA molecules (when the aromatic amino-acid residues interact with the target RNA molecule, the ^{1}H-^{15}N resonances of the amino-acid residues in the proximity of the aromatic rings are also strongly affected[16]). As the aromatic side-chains of Tyr480 and Phe481 protrude from beneath the β2-β3 loop, it is assumed that the RNA molecule is recognized underneath the β2-β3 loop, which is also observed in the structured RNA recognition by the DbpA RRM.[17]

The curve fitting of the titrations of Asp453 and Glu456 indicated K_d values of 0.34 and 0.19 m*M*, respectively, indicating that the CPEB3 RRM1 does not exhibit any biologically meaningful RNA binding activity by itself. Normally, the aromatic amino-acid residues on RNP2 and RNP1 play substantial roles in RNA binding by the canonical RRM. However, they are not affected by the interaction between the CPEB3 RRM and the GAGGAU sequence. This weak binding of CPEB3 RRM1 alone seems to be caused by the lack of participation of the canonical aromatic amino-acid residues in RNA binding.

As described above, the U_5 sequence caused a very slight change in the ^{1}H-^{15}N resonances of CPEB3 RRM1 [Fig. 2(A) bottom]. However, the mapping of the expanded chemical shift change [Fig. 2(A) bottom] onto the CPEB3 RRM1 structure reveals that the U_5 sequence could affect the resonances on the bottom edge of the β-sheet surface (we could not estimate the K_d value for the U_5 sequence, because of the very weak interaction with CPEB3 RRM1). Although the U_5 sequence could not effectively interact with CPEB3 RRM1 by itself, the present NMR data suggest that the β-sheet surface of CPEB3 RRM1 is the potential binding site for the poly-U stretch.

Recent mutational experiments by Chen *et al.* have revealed that the aromatic amino-acid residues on the β1 and β3 strands of CPEB3 RRM1 are necessary for CPEB3 RRM1-RRM2-ZF to interact with the aforementioned SELEX-selected structured RNA fragment.[18] In our structural study of CPEB3 RRM1, Trp530, at the end of the β4-strand, seems to fill the space on RNP1 and inhibit RNA binding on RNP2 and RNP1 [Fig. 1(D) and Supporting Information Fig. 1(B)]. Thus, it is conceivable that the cooperative RNA recognition of CPEB3 RRM1-RRM2-ZF changes the environment around Trp530 of CPEB3 RRM1, to relieve the interaction of Trp530 with Phe488 (β3) and Lys442 (β1) such that CPEB3 RRM1 could interact with RNA via RNP1 and RNP2.

Furthermore, an amino-acid sequence comparison between the CPEB1 and CPEB3 RRM1s revealed that

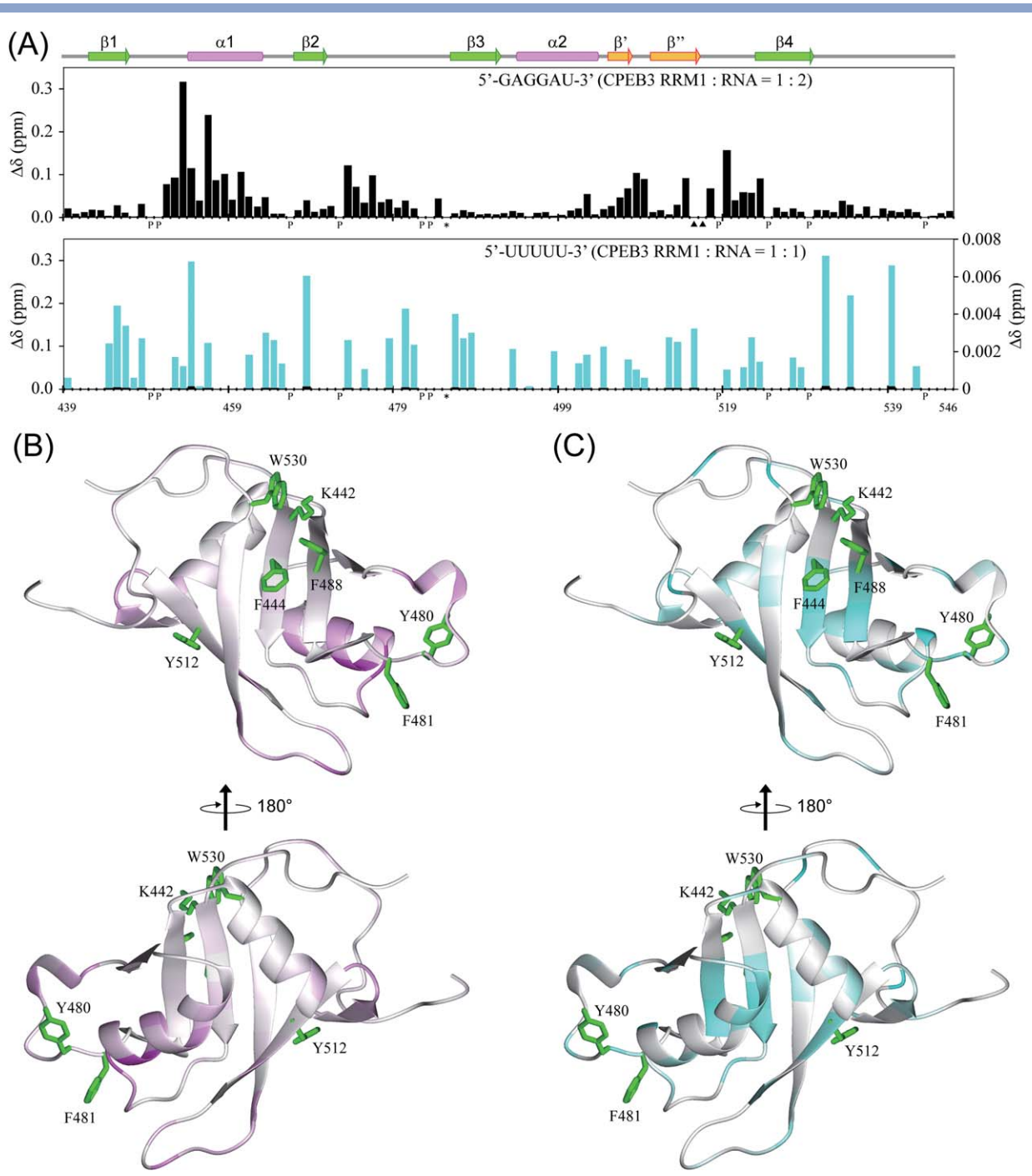

Figure 2

Molecular interactions of CPEB3 RRM1. (**A**) NMR chemical shift perturbations of the CPEB3 RRM1 upon RNA binding. (top) [5′-(GAGGAU)-3′] (protein:RNA = 1:2), and (bottom) [5′-(UUUUU)-3′] (protein:RNA = 1:1). In both histograms, the amino acid residue numbers are on the horizontal axis and the magnitudes of the chemical shift changes, indicated by a scale on the left vertical axis, are plotted as black bars. In the case of [5′-(UUUUU)-3′] (bottom), the magnitudes of the chemical shift changes are also plotted as cyan bars, corresponding to the scale on the right vertical axis. The chemical shift perturbation values were obtained from the [^{1}H, ^{15}N]-HSQC spectrum. The absolute values of the chemical shift change $\Delta\delta$ were calculated as follows: $\Delta\delta = ((\Delta\delta_{15N}/6.5)^2+\Delta\delta_{1H}^2)^{1/2}$. (**B**) and (**C**) Chemical shift perturbations mapped on the tertiary structure of the CPEB3 RRM1 upon binding to [5′-(GAGGAU)-3′] (B) and [5′-(UUUUU)-3′] (C) (top and bottom panels show the structures with views of the β-sheet surface and the α-helical surface, respectively). With gradients corresponding to the amplitude of their chemical shift changes, the residues are colored magenta based on the black-bar data of the top histogram (B), and cyan based on the cyan-bar data of the bottom histogram (C). The side chains of Lys442, Phe444, Tyr480, Phe481, Phe488, Tyr512, and Trp530 are shown in green. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

several hydrophilic amino-acid residues around Trp530 in CPEB3 RRM1 (Arg529, Asn531, and Ser533) are replaced with hydrophobic residues in CPEB1 RRM1 (Ile, Val, and Ala, respectively). These replacements could modify the hydrophobic interactions of the corresponding Trp residues in the CPEB1 RRM1, and make the β-sheet surface of CPEB1 RRM1 accessible for RNA binding. The present structural information about the CPEB3 RRM1 is a step leading to the elucidation of the RNA binding specificities of the CPEB protein family members.

ACKNOWLEDGMENTS

We thank Dr. T. Someya and Ms. S. Suzuki for assistance with the NMR data analysis, structure calculations, and structure refinement. We also thank Dr. T. Matsuda, Ms. Y. Tomo, Dr. M. Aoki, Dr. T. Nagira, Dr. E. Seki, Mr. K. Hanada, Mr. M. Ikari, Ms. Y. Fujikura, Ms. Y. Tomabechi and Ms. Y. Kamewari-Hayami for sample preparation, and Ms. T. Imada, Mr. K. Ake and Ms. T. Nakayama for assistance with the manuscript preparation.

REFERENCES

1. Fernandez-Miranda G, Mendez R. The CPEB-family of proteins, translational control in senescence and cancer. Ageing Res Rev 2012; 11:460–472.
2. Kim JH, Richter JD. Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. Mol Cell 2006;24:173–183.
3. Huang YS, Kan MC, Lin CL, Richter JD. CPEB3 and CPEB4 in neurons: analysis of RNA-binding specificity and translational control of AMPA receptor GluR2 mRNA. EMBO J 2006;25:4865–4876.
4. Hosoda N, Funakoshi Y, Hirasawa M, Yamagishi R, Asano Y, Miyagawa R, Ogami K, Tsujimoto M, Hoshino S. Anti-proliferative protein Tob negatively regulates CPEB3 target by recruiting Caf1 deadenylase. EMBO J 2011;30:1311–1323.
5. Muto Y, Yokoyama S. Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. Wiley Interdiscip Rev RNA 2012;3:229–246.
6. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol 2008;18:290–298.
7. Chen Y, Varani G. Protein families and RNA recognition. FEBS J 2005;272:2088–2097.
8. Nagai K, Oubridge C, Ito N, Avis J, Evans P. The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. Trends Biochem Sci 1995;20:235–240.
9. Matsuda T, Koshiba S, Tochio N, Seki E, Iwasaki N, Yabuki T, Inoue M, Yokoyama S, Kigawa T. Improving cell-free protein synthesis for stable-isotope labeling. J Biomol NMR 2007;37:225–229.
10. Kay LE. NMR methods for the study of protein structure and dynamics. Biochem Cell Biol 1997;75:1–15.
11. Bax A. Multidimensional nuclear magnetic resonance methods for protein studies. Curr Opin Struct Biol 1994;4:738–744.
12. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE. Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry 1994;33: 5984–6003.
13. Vuister G, Bax A. Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNH.alpha.) coupling constants in 15N-enriched proteins. J Am Chem Soc 1993;115:7772–7777.
14. Powers R, Garrett DS, March CJ, Frieden EA, Gronenborn AM, Clore GM. The high-resolution, three-dimensional solution structure of human interleukin-4 determined by multidimensional heteronuclear magnetic resonance spectroscopy. Biochemistry 1993;32: 6744–6762.
15. 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, Muto Y. Structural basis for the dual RNA-recognition modes of human Tra2-beta RRM. Nucleic Acids Res 2011;39:1538–1553.
16. Nagata T, Suzuki S, Endo R, Shirouzu M, Terada T, Inoue M, Kigawa T, Kobayashi N, Guntert 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.
17. Wang S, Hu Y, Overgaard MT, Karginov FV, Uhlenbeck OC, McKay DB. The domain of the Bacillus subtilis DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold. RNA 2006;12:959–967.
18. Chen YC, Sargsyan K, Wright JD, Huang YS, Lim C. Identifying RNA-binding residues based on evolutionary conserved structural and energetic features. Nucleic Acids Res 2014;42:e15.