Published online 31 July 2014 *Nucleic Acids Research, 2014, Vol. 42, No. 15* ***10185–10195***
doi: 10.1093/nar/gku700

RNA recognition and self-association of CPEB4 is mediated by its tandem RRM domains

Constanze Schelhorn[1], James M.B. Gordon[1], Lidia Ruiz[1], Javier Alguacil[2], Enrique Pedroso[2] and Maria J. Macias[1,3,*]

[1]Institute for Research in Biomedicine (IRB Barcelona), Baldiri Reixac 10, Barcelona 08028, Spain, [2]Departament de Química Orgànica and IBUB, Universitat de Barcelona, Martí i Franquès 1-11, 08028 Barcelona, Spain and [3]Catalan Institution for Research and Advanced Studies (ICREA), Passeig Lluis Companys 23, Barcelona 08010, Spain

Received June 3, 2014; Revised July 11, 2014; Accepted July 19, 2014

ABSTRACT

Cytoplasmic polyadenylation is regulated by the interaction of the cytoplasmic polyadenylation element binding proteins (CPEB) with cytoplasmic polyadenylation element (CPE) containing mRNAs. The CPEB family comprises four paralogs, CPEB1–4, each composed of a variable N-terminal region, two RNA recognition motif (RRM) and a C-terminal ZZ-domain. We have characterized the RRM domains of CPEB4 and their binding properties using a combination of biochemical, biophysical and NMR techniques. Isothermal titration calorimetry, NMR and electrophoretic mobility shift assay experiments demonstrate that both the RRM domains are required for an optimal CPE interaction and the presence of either one or two adenosines in the two most commonly used consensus CPE motifs has little effect on the affinity of the interaction. Both the single RRM1 and the tandem RRM1–RRM2 have the ability to dimerize, although representing a minor population. Self-association does not affect the proteins' ability to interact with RNA as demonstrated by ion mobility–mass spectrometry. Chemical shift effects measured by NMR of the apo forms of the RRM1–RRM2 samples indicate that the two domains are orientated toward each other. NMR titration experiments show that residues on the β-sheet surface on RRM1 and at the C-terminus of RRM2 are affected upon RNA binding. We propose a model of the CPEB4 RRM1–RRM2–CPE complex that illustrates the experimental data.

INTRODUCTION

The cytoplasmic polyadenylation element binding protein (CPEB) is involved in translational regulation of cytoplasmic polyadenylation element (CPE) containing mRNAs. The mechanism by which cytoplasmic polyadenylation controls the translation of many key mRNAs has been elucidated in meiotic maturation of *Xenopus* oocytes. Many maternal mRNAs are translationally repressed in oocytes arrested at the end of meiosis prophase I. This repression is achieved via the interaction of CPEB to the CPE U-rich sequence (U4-6A1-2U1-2) in the 3′UTR of the specific mRNAs (1). Moreover, CPEB recruits protein factors that facilitate translational regulation. The poly A-specific ribonuclease PARN has an antagonistic effect on the non-canonical poly A polymerase Gld2 resulting in mRNAs with a short poly A tail and therefore transitionally repressed (2). Other factors that interact with CPEB include members of the mRNA 3′-end processing machinery including symplekin and CPSF. An interaction with the protein Maskin and its interaction with the translational initiation factor eIF4E inhibits the assembly of the translation initiation complex (3,4). A progesterone initiated signaling cascade results in the activation of kinase Aurora A that in turn phosphorylates CPEB (5). The phosphorylation leads to the inhibition of the PARN–CPEB interaction resulting in the extension of the poly A tail of the mRNA by Gld2 allowing for the binding of poly A binding proteins and subsequent translational activation (6).

The CPEB family of proteins comprises of four paralogs (CPEB1–4). CPEBs 2–4 are closely related and CPEB1 is the most distant relative (7). The proteins are widely expressed in a variety of tissues, cell types and tumors with partially overlapping patterns (7). Additional functions to that of oocyte maturation have been identified including roles in synaptic plasticity and cellular senescence (8). Members of the family comprise of a largely conserved C-terminal region that comprises of two RNA recognition

*To whom correspondence should be addressed. Tel: +34 934037189; Fax: +34 934039976; Email: maria.macias@irbbarcelona.org
Present address:
Javier Alguacil, Institut des Biomolécules Max Mousseron (IBMM), Université Montpellier 2, Place E. Bataillon, Montpellier cedex 5, 34095, France.

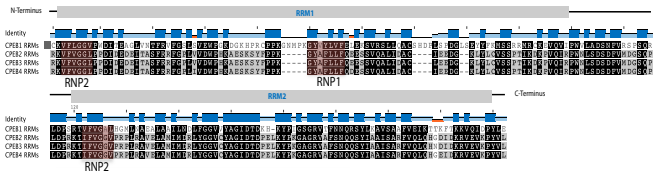

Figure 1. Alignment was prepared using the Geneious R6 software (biomatters). The position of the RRM domains is indicated and the RNP motifs highlighted. Black residues are conserved between all family members. Gray residues are partially conserved and blank residues are unconserved.

motifs (RRM1 and RRM2) that have been shown to be essential for the CPE interaction, and a ZZ-type domain that is postulated to be a protein–protein interaction module (9,10). The N-terminal half of the protein is more variable between the different members and contains a PEST-degradation motif, the site for Aurora A mediated phosphorylation and several uncharacterized phosphorylation sites. The two most commonly used CPE consensus motifs are UUUUAAU and UUUUAUU (11,12). Although there are other non-canonical motifs, studies have shown that CPEBs 2 and 4 recognize and regulate mRNA transcripts containing the same canonical CPE motifs that CPEB1 regulate (13,14–17).

The structure of the RRM1 domain of CPEB3, which is completely conserved when aligned to CPEB2 and 4, has been solved byNuclear Magnetic Resonance (NMR) (PDB code: 2DNL) and shows the canonical RRM domain fold of a αβ-sandwich with a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology with the addition of two β strands immediately before the β4 strand (β4′ and β4″). Two conserved motifs, which are generally involved in RRM–RNA interactions, RNP1 and RNP2 lie in the central strands of the RRM β-sheet (18). The consensus sequence of RNP1 is (RK)-G-(FY)-(GA)-(FY)-(ILV)-X-(FY), while that of RNP2 is (ILV)-(FY)-(ILV)-X-N-L. The first RRM domain (RRM1) contains, for the most part, both RNP motifs. However, the second (RRM2), more C-terminal domain lacks the consensus sequence of an RNP1 motif. This characteristic is shared among all CPEB family members (Figure 1).

Previously, Lin *et al.* (19) showed that the full-length CPEB1 protein has the ability to dimerize. Deletions of the RRM motifs from the full-length protein showed that this oligomerization was mediated by the pair of RRM domains. Moreover, the dimers of CPEB1 seem to lose the ability to interact with RNA (19), suggesting that the formation of dimers prevents RNA recognition, maintaining the protein in an inactive state.

Prompted by these observations in CPEB1 and in order to characterize the binding properties of the RRM domains present in CPEB4, we have examined the RNA interactions of the CPEB4 RRM tandem and also of the individual RRM1 domain, using a combination of biochemical, biophysical and NMR-based techniques. The chemical shift effects seen by NMR of the apo form proteins, when comparing RRM1 to the tandem construct RRM1–RRM2, indicates that the domains are in close proximity to each other. Furthermore, isothermal titration calorimetry (ITC) and electrophoretic mobility shift assay (EMSA) experiments have shown that both domains are necessary for optimal binding to the CPE RNA motifs and that the presence of one or two adenosines in the motif has little or no effect on the binding affinity. The binding site for this RNA interaction has been mapped using NMR titrations of the CPEs with the RRM1–RRM2 tandem construct and the observations made indicate that the binding site is localized largely on the β-sheet surface, including the conserved RNP motifs of RRM1, and the C-terminus of RRM2. We also report the ability of both the single RRM1 domain and the RRM1–RRM2 pair to form dimers, as previously reported for the full-length CPEB1 protein (19). In the case of the recombinant RRM1 and the RRM1–RRM2 pair, the dimer population represents a minority species in solution, below 10%. Remarkably, unlike CPEB1, the CPEB4–RRM dimers retain the ability to bind RNA as demonstrated by ion mobility–mass spectrometry (IM-MS). We have proposed a model of the RRM1–RRM2–CPE complex based on the data generated and reported in this manuscript. This model and the chemical shift assignments of the RRM1–RRM2 pair will provide a platform for further studies aimed at the design of protein–RNA inhibitors, target to the CPEB family of proteins.

MATERIALS AND METHODS

Cloning of human CPEB4 RRM domains

Complementary DNA (cDNA) from human CPEB4, isoform 1, was used as a template to clone into the pETM-11 vector (a gift from the EMBL-Heidelberg Protein Expression Facility). The constructs cloned correspond to the following boundaries: RRM1 (463–573) and RRM1–RRM2 (463–665). Residue numbering is according to isoform 1 of CPEB4 (Uniprot: Q17RY0). All clones were confirmed by DNA sequencing. *Escherichia coli* DH5α strain was used for cloning.

Protein expression and purification

Both constructs were expressed at 37°C in *E. coli* BL21 (DE3) following induction at an optical density of 0.6 (600 nm) with 0.5 mM IPTG overnight. Unlabeled and labeled samples were prepared using LB and minimal media (M9) cultures, respectively. D_2O (99.89%, CortecNet), $^{15}NH_4Cl$ and/or D-[^{13}C] glucose were used as sole hydrogen, nitrogen and carbon sources respectively to prepare the labeled samples as described (20). All proteins were expressed fused to an N-terminal His6-tag followed by a TEV-protease cleavage site. Cells were lysed using an EmulsiFlex-C5 (Avestin) in the presence of lysozyme and DNAseI. The soluble supernatant was purified by nickel-affinity chromatography (HiTrap Chelating HP column, GE Healthcare Life Science, Uppsala, Sweden) using 150 mM NaCl, 20 mM Tris-HCl, 10 mM imidazole pH 8.0 and 50 mM EDTA. Eluted His_6-RRM constructs were digested with TEV protease at room temperature overnight, and further purified by size-exclusion chromatography on a HiLoad^{TM} Superdex 75 16/60 prepgrade columns (GE Healthcare) equilibrated in 20 mM Tris pH 7.0 and 130 mM NaCl.

EMSA

Binding reactions were carried out for 20 min at 4°C in 10 μl of binding buffer [40 mM Tris pH 7.5, 150 mM NaCl, 0.1% (w/v) βME]. Increasing amounts of protein were incubated with a fixed concentration of ^{32}P labeled RNA (∼2 nM). Electrophoresis was performed in non-denaturing 8.0% (29:1) polyacrylamide gels. The gels were run for 4 h in 1× TBE buffer (90 mM Tris, 64.6 mM boric acid, 2.5 mM EDTA pH 8.4) at 100 V at 4°C. The gels were dried and exposed to a PhosphorImager screen overnight. The screen was scanned on a Molecular Dynamics Storm 840 PhosphorImager.

RNA synthesis and purification

The CPE-containing RNA fragments U5A2U1 and U5A1U2 were synthesized at the 4×1 mmol scale using the standard phosphite triester approach in a 3400 Applied Biosystems synthesizer (21). After cleavage and deprotection, the oligonucleotides were purified by reverse-phase HPLC using a Jupiter C18 semipreparative column (10 μm, 300 Å, 250 × 10 mm) from Phenomenex with a Waters 600 HPLC System (buffer A: 0.1 M triethylammonium acetate, buffer B: acetonitrile). The RNA triethylammonium counter-ion was exchanged by sodium using a cation exchange resin (Dowex 50Wx4, 200–400 mesh, Sigma-Aldrich). The final products were analyzed by MALDI-TOF MS, lyophilized and stored at −20°C until further use.

Isothermal titration calorimetry

ITC measurements were performed using a nano ITC calorimeter (TA Instruments) at 5°C, 12°C or 25°C. To be consistent with the condition used for the NMR studies, protein samples were prepared in the NMR buffer (20 mM Tris-HCl pH 7.0 and 130 mM NaCl). Buffer from concentrating the protein samples was used to resuspend the RNA. All samples were centrifuged and degassed prior to the experiments. Protein concentration was measured in a NanoDrop™ 2000 measuring the UV absorption at 280 nm. RNA concentration was determined spectrophotometrically. Depending on the expected affinity, sigmoidal curves were optimized by injecting 5- to 10-fold concentrated RNA in 16 × 3 μl steps in a cell containing 190 μl of protein at 20–50 μM adjusted concentration. The delay between injections was 3 min and data was collected while stirring at 200 revolutions per minute (rpm). The NanoAnalyze software (TA Instruments) was used to analyze the binding isotherms assuming a single binding site in each molecule. Baseline controls were acquired with buffer and pure RNA solutions. Measurements have been repeated at least twice.

Ion mobility–mass spectrometry

An ion mobility mass spectrometer features an ion mobility cell before the mass analyzer. Gaseous ions are injected into the cell and accelerated by a weak electric field. Due to the presence of buffer gas in this cell, low-energy collisions with the buffer gas occur. The higher the collision cross sections (CCS) of the ion, the greater the number of collisions with the buffer gas. As collisions increase, an energy loss of the ions occur and accordingly ions take a longer time to cross the IM cell (the 'drift time'). Consequently, ions are injected into the mass analyzer, achieving a simultaneous separation on the basis of the CCS to charge ratio and the m/z ratio. The three-dimensional spectrum obtained consists of mass, drift time and relative intensity. Traveling wave IM-MS experiments were performed on a Synapt G1 HDMS mass spectrometer (Waters, Manchester, UK). Samples were placed on a 384-well plate refrigerated at 15°C and sprayed using a Triversa NanoMate® (Advion BioSciences) automated Chip-Base nano-electrospray working in the positive ion mode. The instrument was calibrated over the 500–8000 Da m/z range using a cesium iodide solution. The software MassLynx 4.1 SCN 704 and Driftscope 2.1 were used for data processing. Samples containing the RNA complexes with either CPEB4–RRM1 or CPEB4 RRM1–RRM2 tandem or with the independent RRM domains (final concentrations of 30–50 μM) were prepared in 20–50 mM NH_4OAc pH 7.2. Prior to analysis, a 1D-^{1}H NMR spectrum was acquired on all samples in order to check the stability of the samples and to compare it with spectra obtained in NMR buffer conditions. Spray voltage was set to 1.75 kV and delivery pressure to 0.5 psi. A reduction of the source pumping speed in the backing region (5.85 mbar) was set for optimal transmission of high mass non-covalent ions. Cone voltage, extraction cone and source temperature were set to 20 V, 6 V and 20°C respectively. Ions passed through a quadrupol mass filter to the IM-MS section of the instrument.

NMR chemical shift perturbation experiments

Experiments were recorded at 303 K using a Bruker AVIII 600 MHz spectrometer equipped with a 5 mm TXI cryprobe, z-gradient. Protein samples of the CPEB4 RRM1 and the CPEB4 RRM1–RRM2 were equilibrated in a buffer containing 20 mM Tris-d_{11}, 130 mM NaCl and 5% DMSO-d_6. All samples were supplemented with 10% D_2O and pH adjusted to value 7. Spectra were acquired using 200 μM ^{15}N-labeled protein samples equilibrated together with progressively increasing amounts of the unlabeled RNA fragments until saturation was achieved. Chemical shift perturbation analyses were performed on CcpNmr Analysis (22) with a 0.15 weighting of ^{15}N with respect to ^{1}H.

NMR backbone assignment

Spectra were acquired at 303 K using a Bruker AVIII 600 MHz spectrometer, equipped with cryogenic or room temperature triple resonance gradient probes. Backbone ^{1}H, ^{13}C and ^{15}N resonance assignments were obtained by analyzing the 3D HNCACB and HN(CO)CACB experiment pair in the case of the CPEB4 RRM1 (fully protonated sample) or the CBCANH and CBCA(CO)NH pair in the case of the CPEB4 RRM1–RRM2 (^{2}H, ^{13}C, ^{15}N samples). Transverse relaxation-optimized spectroscopy (TROSY) versions of these experiments and/or the non-uniform sampling (NUS) acquisition strategy were used in selected cases to reduce experimental time and increase resolution. All buffer conditions were as mentioned. NMRPipe (23) was

used for spectra processing. CARA (24) and CcpNmr Analysis (22) were used for spectra analysis and assignment.

Relaxation measurements

Amide relaxation measurements were acquired on a 500 μM ^{15}N-labeled RRM1–RRM2 sample essentially as described (25). NMR experimental setup details were essentially as reported (26). Briefly, the T_1 and T_2 experiments were acquired using 135 (t_1) × 2048 (t_2) total real points. Twelve different relaxation time values (22, 54, 108, 162, 270, 432, 540, 702, 864,1080, 1404 and 1728 ms) were measured to determine T_1. To determine T_2 ten experiments were recorded with the following ^{15}N mixing times: 16.74, 33.48, 50.22, 66.96, 100.44, 117.18, 133.92, 167.4, 184.14, 200.88 ms. All relaxation experiments were acquired as pseudo-3D experiments and converted to 2D data sets during processing. Peak integration values were fitted to a two-parameter function as described in Equation (1):

$$I(t) = I_0 e^{\left(\frac{-t}{T_{1,2}}\right)} \quad (1)$$

where I_0 and $I(t)$ are the peak intensities at times 0 and t, respectively. The optimum value of the I_0 and the $T_{1,2}$ parameters was determined using the Levenberg–Marquardt optimisation algorithm for minimization of χ^2 goodness of fit parameter as reported.

The rotational correlation time of the RRM1–RRM2 pair was calculated with Equation (2), using the approximation of slow molecular motion τ_c larger than 0.5 ns and assuming that only $J(0)$ and $J(\omega N)$ spectral density terms contribute to the overall value. ν_N is the ^{15}N resonance frequency (60,08 ×10^6 Hz)

$$\tau \approx \frac{1}{4\pi\nu_N}\sqrt{\left(6\frac{T_1}{T_2} - 7\right)} \quad (2)$$

Regarding the heteronuclear Nuclear Overhauser Effect (NOE) experiment, the reference and the presaturated Heteronuclear single quantum correlation spectroscopy (HSQC) spectra were acquired in an interleaved manner. The values of the steady-state ^{1}H-^{15}N NOEs resulted from the ratios of the peak intensities measured in the reference (I_0) and the presaturated (I_S) spectra during the relaxation delay as described (27). Background noise levels σ_S and σ_0 were measured and used to determine the NOE standard deviation through the following relationship:

$$\frac{\sigma_{NOE}}{NOE} = \left(\left(\frac{\sigma_{I_S}}{I_S}\right)^2 + \left(\frac{\sigma_{I_0}}{I_0}\right)^2\right)^{\frac{1}{2}}.$$

RESULTS

Interactions of the RRM domains in the absence of RNA

To gain insight into which role the two RRM domains play in RNA binding we prepared several constructs of the isolated RRM1, RRM2 and the RRM1–RRM2 tandem domains using His-tagged fusion expression vectors. The isolated RRM1 domain and the tandem construct were in the soluble fraction of the expression cultures and they were purified by Ni-affinity chromatography. However, overexpression of the isolated RRM2 using various experimental conditions resulted in the formation of inclusion bodies containing the RRM2 protein. Protein purified under denaturing conditions, in the presence of guanidium chloride, could not be refolded by dialysis under the various buffer conditions we have assayed.

We therefore focused our work on the construct containing the pair of RRM domains and also on the isolated RRM1, which was used for comparison. In order to assign the amide and backbone resonances and to facilitate the investigation into whether the first RRM domain adopts a similar fold when being independently expressed or in the construct of both consecutive RRM domains ^{13}C/^{15}N and ^{2}H/^{13}C/^{15}N labeled samples were prepared respectively. Due to its size and compact fold, the acquisition of the triple resonance experiments of the tandem required the deuteration of the sample in order to minimize the loss of the signal caused by transverse T_2 relaxation. Applying TROSY versions of all NMR experiments was essential to improve the resolution and sensitivity of the backbone experiments.

A comparison of the secondary ^{13}C chemical shifts obtained for the samples to reference values indicate that both RRM1 and RRM2 in the tandem construct adopt the canonical αβ sandwich structure with a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology with an additional β-strand β_4 at the C-terminus of RRM1 (Figure 2A). Comparison of the ^{13}C chemical shifts of RRM1, when assigned independently or in the RRM1–RRM2 construct, indicates that the secondary structure is not altered due to the presence of RRM2 (Figure 2B). For RRM1, a homology model was built using SWISS-MODEL (Template: PDB entry 2DNL, RRM1 of CPEB3, sequence identity 97%, http://swissmodel.expasy.org/; a quality report is shown in Supplementary Figure S1). The topology of the model obtained is consistent with the elements of secondary structure indicated by the analysis of the ^{13}C chemical shifts. The spatial arrangement of the canonical four-stranded antiparallel β-sheet is $\beta_4\beta_1\beta_3\beta_2$. The homology model we obtained indicates that the additional β_4 strand is arranged antiparallel to β_4 resulting in the following order for the extended β-sheet: $\beta_4\beta_4\beta_1\beta_3\beta_2$.

HSQC-TROSY experiments for the single domain (Supplementary Figure S2) and for the pair (Supplementary Figure S3) respectively were compared to assess the influence that the domains may have on the chemical shifts of each other. The superimposition of both spectra (Figure 2C) acquired under identical conditions reveals that the comparison of the amide resonances only matches reasonably well in certain areas. Significant differences in the chemical shift of many residues are observed between the RRM1 and the RRM1–RRM2 construct. The linker connecting both domains is very short in length (six residues), limiting the freedom of the domains in the tandem. The comparison of the linewidths (Supplementary Figure S4) of the independent RRM1 domain to that of the construct containing both domains indicates a clear broadening of the amide signals characteristic of a 23 KDa sample and that the RRM tandem behaves as a unit. Moreover, this observation is supported by the homogeneous T_1, T_2 and het-

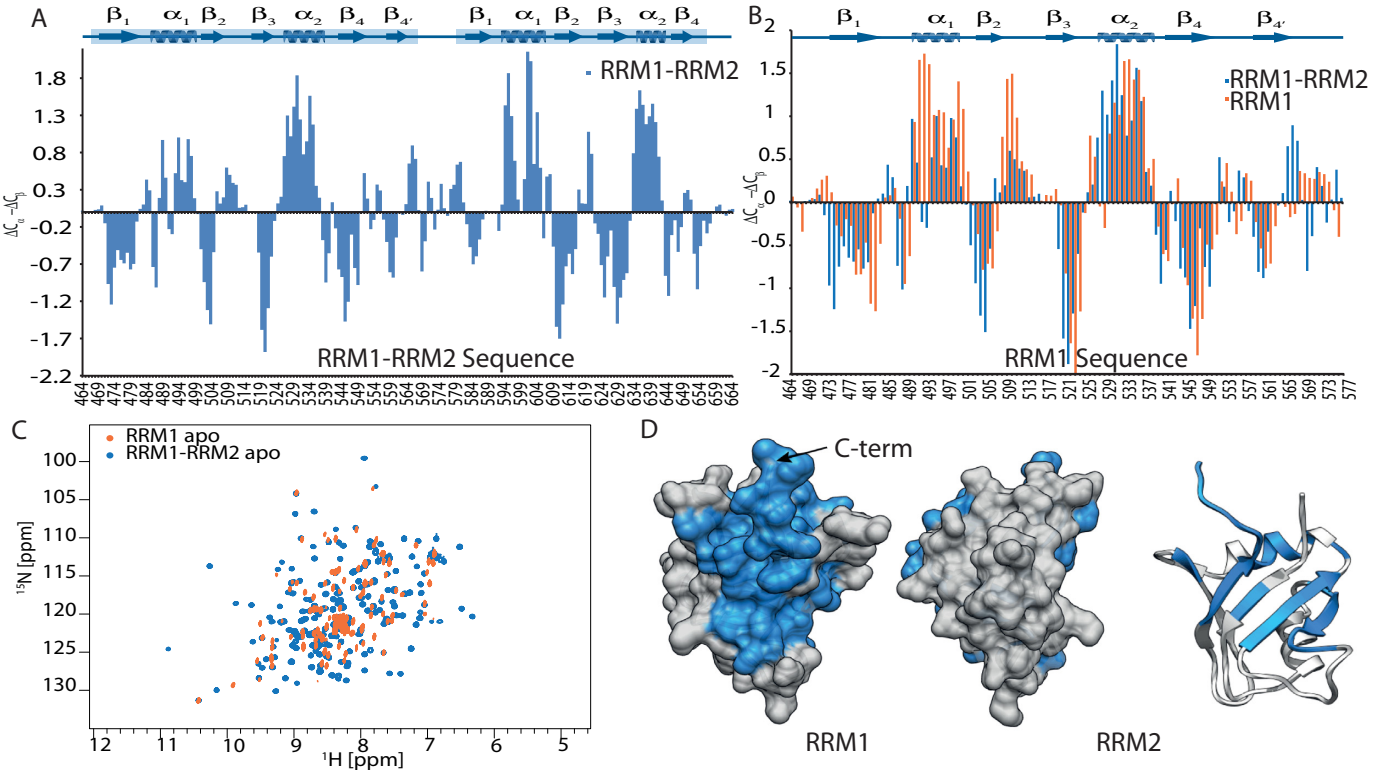

Figure 2. (**A**) Comparison of ^{1}H, ^{15}N-HSQC spectra of the apo RRM1 (orange) and apo RRM1–RRM2 (blue). Backbone assignment of both constructs allows us to identify the chemical shift of the residues in both samples. (**B**) Significant differences in chemical shifts are mapped onto the homology model for RRM1. As those differences are clustering on one surface of the model, we assume that this surface lies in proximity of the RRM2. Secondary ^{13}C chemical shifts $\Delta C_\alpha - \Delta C_\beta$ of CPEB4 RRM1–RRM2 (D) and comparison of the CSI of RRM1 single and in the RRM1–RRM2 tandem (**C**). Residues with negative CSI indicate β-strand conformation whereas positive values indicate turn or α-helix. (**D**) Secondary structure of RRM1 is not significantly altered due to the presence of RRM2.

eronuclear NOE values (Figure 3A, B and D respectively) measured for the tandem along its entire sequence. The calculated rotational correlation time ($\tau_c = 16$ ns) is in agreement with the average value obtained for a globular protein of this size (Figure 3C) (28). This further confirms that the CPEB4 RRM1–RRM2 domains do not tumble independently. Based on these evidence, we interpret that the chemical shift changes observed in the RRM1 domain indicate the presence of a conformation where both RRM domains are close to one another. Furthermore, when mapping the most affected residues onto the homology model of the RRM1 domain, they cluster on one surface (Figure 2D), indicating that this surface should be in proximity of the RRM2 domain.

CPEB4 uses both RRM domains to bind CPE motifs

In order to characterize the interaction of the RRM domains with short RNAs containing the CPE motifs, we have used several complementary techniques that allow us to detect and quantify the interactions. For this purpose, we performed EMSA as well as ITC binding assays using both RRM1 and RRM1–RRM2 and different RNA fragments. The EMSA experiments showed that RRM1–RRM2 tandem domain has a higher affinity for the CPE RNA when compared to the RRM1 alone (Figure 4A).

The binding ligands used for ITC experiments were two octamer RNA fragments containing the two consensus CPE motifs, U5A2U1 and U5A1U2. ITC experiments with RRM1 and both ligands showed low-affinity binding with dissociation constants in the high μM range. The experiments with RRM1–RRM2 as titrate yielded dissociation constants of 323 ± 34 nM for U5A2U1 and 299 ± 28 nM for U5A1U2 (Figure 4B). Both ligands, which differ only in one nucleotide, show very similar K_D values (both lying within the error limit of each other), thus not allowing us to significantly distinguish between the two binding affinities. The affinity increase due to the presence of the RRM2 domain is about 100-fold with respect to the values obtained for the single RRM1 domain. As we have not been able to obtain pure RRM2, the experiment could not be repeated for the single RRM2. However, from our NMR titration data, which allows the identification of the regions affected by RNA binding (Figure 5), it is clear that chemical shift perturbations are observed for residues in both domains. The titration of the RRM1 with U5A2U1 is shown in Supplementary Figure S5. Therefore, our data indicate that both RRM domains are important to maintain optimal RNA-binding activity and affinity. The improvement of the affinity by 100 times in the construct containing the RRM2 domain is quite remarkable since this domain contains a degenerate RNP motif and it was assumed to be a poor

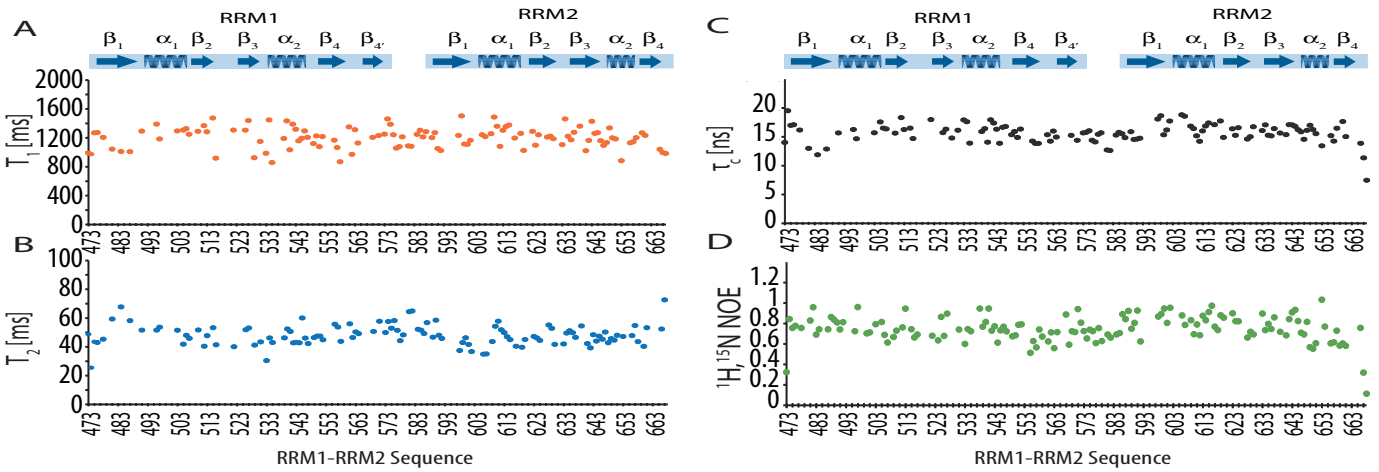

Figure 3. ^{15}N amide relaxation properties of CPEB4 RRM1–RRM2 in their free state. Measurements of longitudinal T_1 (**A**), transverse T_2 (**B**) relaxation were performed at 303 K and 600 MHz. Also shown is the rotational diffusion correlation time τ_c derived from the ratio of the ^{15}N T_1/T_2 relaxation times (**C**). The value $\tau_c \approx 16$ ns is consistent with a compact molecule of 23 kDa. Furthermore, ^{1}H-^{15}N heteronuclear NOE (**D**) experiments were acquired and the results are consistent with the T_1/T_2 relaxation data. Proline and overlapped residues were not included in the analysis.

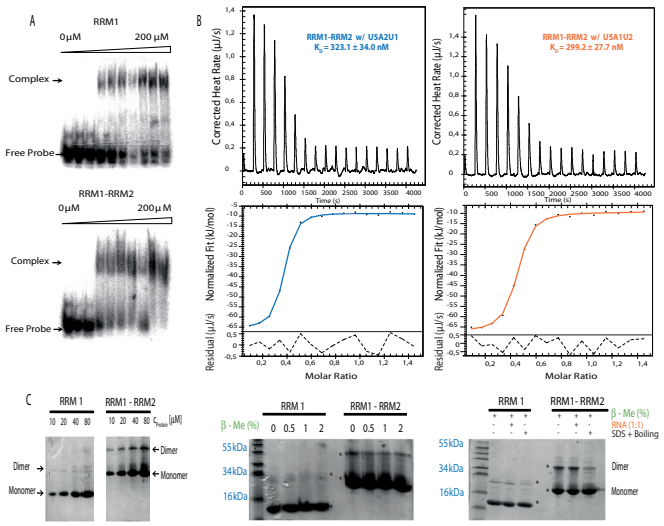

Figure 4. (**A**) EMSA conducted with increasing amounts of purified CPEB RRM1 or RRM1–RRM2 protein: 0, 1, 25, 50, 75, 100, 125, 150, 175 and 200 (mM) and radiolabelled U5A2U1 RNA (2 nM). Samples were fractionated in native gels and visualized by autoradiography. (**B**) ITC curves and affinity values obtained for the titration of CPEB4 RRM1–RRM2 with U5A2U1 (blue) and U5A1U2 (orange) RNAs. Data were acquired at 5°C in Tris buffer pH 7.0. The stoichiometry obtained was 0.6 and 0.7 respectively. The data were fitted using the independent model assuming a single binding site. (**C**) Protein samples were analyzed under semi-native conditions using SDS 12% PAGE. SDS was not present in the loading buffer and the samples were not boiled unless indicated.

RNA binder. This increment in the affinity together with the length of the RNA recognized by the RRM1–RRM2 pair suggests that both RRM domains act cooperatively resulting in high-affinity binding of the RNA. We suggest that the role of the RRM2 in the RNA interaction is dual: it increases the global surface of the pair to recognize the RNA with respect to the single RRM1 and helps adjusting the relative orientation of the two domains to best accommodate the RNA ligand.

Identifying the RNA binding site of RRM1-RRM2 via NMR

To characterize the RNA binding properties of the RRM1–RRM2 construct to the two CPE containing ligands, U5A1U2 and U5A2U1, and to investigate whether there are different binding modes between the consensus CPE motifs, amide chemical shift changes were monitored upon titration of the U5A1U2 and U5A2U1 to the ^{15}N-labeled RRM1–RRM2 domain (Figure 5A). The chemical shift perturbations regarding their range and also the residues affected showed no significant difference between U5A1U2 and U5A2U1 (Figure 5B). We therefore assume that the number of adenosines within the CPE motif does not alter the binding mode. Binding kinetics observed in the NMR titrations of RRM1–RRM2 are in fast to intermediate range on the NMR chemical shift time scale as some residues disappear and reappear with increasing U5A1U2 and U5A2U1 concentration (e.g. Arg^{559}; Figure 5A (lower left)).

For both RNA ligands we observe chemical shift perturbations in the canonical RNP motifs of the RRM1 domain, which, according to the secondary structure predictions and the homology model, lie on the β_1 and β_3 strand. In addition, large chemical shift perturbations are also observed for residues lying on β_2 (Trp^{502}, Lys^{505}) and a C-terminal region of RRM1. These C-terminal residues are predicted to form the β_4-strand (affected residues: Gln^{557}, Ile^{558}, Arg^{559}) and additional β-strand β_4 (affected residues: Ser^{547}, Thr^{550}). When mapping the most affected residues of RRM1 on the homology model, they cluster on a region on the β-sheet, which corresponds to the canonical RNA binding surface of RRM domains (Figure 5B).

For RRM2, which only features the RNP2 consensus sequence on β_1, the strongest chemical shift perturbations were detected in its C-terminal region containing residues with positively charged side groups (e.g. His^{648}, Lys^{653}, Arg^{654}), indicating that these regions are affected by binding to the RNA fragments either by engaging in direct con-

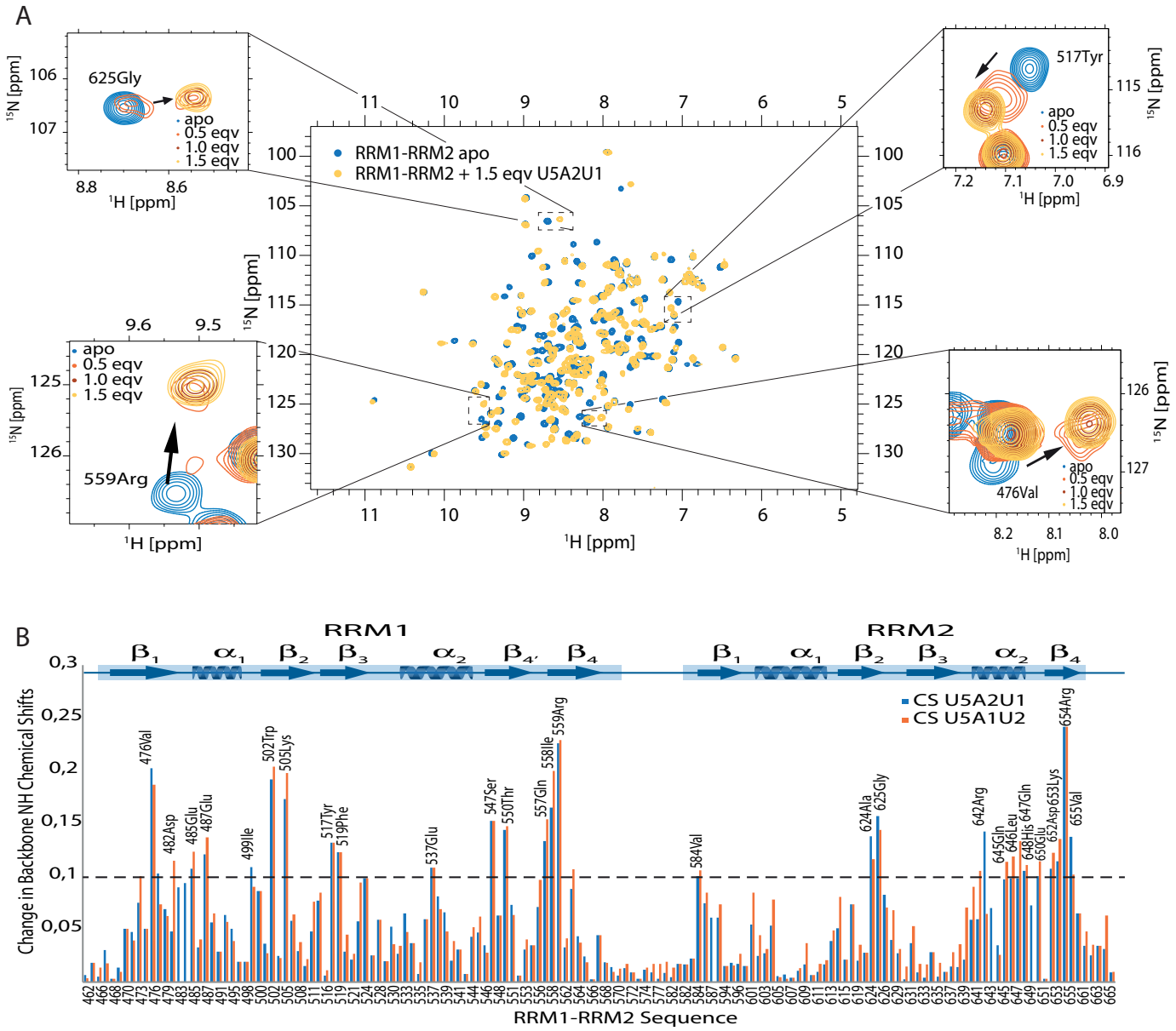

Figure 5. (**A**) Superimposition of ^{1}H, ^{15}N-HSQC spectra of RRM1–RRM2 free (blue) and in complex with U5A2U1 (yellow) 1.5 molar equivalents. Zooms in various regions of the spectra are displaying all titration points. Spectra are colored as blue (free); orange, 0.5; dark red, 1.0, and yellow, 1.5 molar equivalents. (**B**) Chemical shift perturbations of RRM1–RRM2 domains upon binding to U5A2U1 (blue) and U5A1U2 (orange). The Chemical Shift Perturbations (CSPs) are plotted versus residue numbers. The dashed line in gray shows a cut-off $\Delta\delta > 0.1$ ppm (average chemical shift perturbation over all resonances + standard deviation). The CSPs for both RNA ligands do not show any significant differences. Within RRM1, the largest shift perturbations are found in residues that are supposed to lie on the β-sheet; for RRM2, however, they clearly cluster in the C-terminal part. The RNP2 of RRM2 shows shifts just above the cut-off, which suggests that RNP2 is not the main binding site for RRM2, instead the C-terminal positively charged residues seem to play an essential role.

tacts to the RNA or by indirectly induced conformational changes.

RRM1 and RRM1-2 have the ability to dimerize

It has been previously reported by Lin *et al.* (19) that CPEB1 dimerization via its RNA binding regions functions as a regulatory mechanism during cell cycle. We therefore consider the possibility that protein dimerization could also occur in CPEB4. The presence of highly populated dimers cannot be detected at the protein concentrations used during gel filtration chromatography and also our NMR data were acquired at low protein concentrations (0.2–0.5 mM) to minimize protein precipitation. At this concentration, the main component of the protein solution corresponds to the monomeric protein, as revealed by the relaxation experiments and rotational correlation time measured for the tandem. However, when we performed the ITC experiments to determine the affinities of the different complexes, we observed that the stoichiometries for both the interactions of

RRM1 and RRM1–RRM2 proteins with the RNA ligands were close to but smaller than 1. Repetition of the experiments at several temperatures (5°C, 12°C and 25°C) and using different buffer solutions (tris, phosphate and bis-tris) yielded similar values and stoichiometries. We attributed these small discrepancies in the stoichiometries to the presence of a small population of oligomers that—together with the main complex formed by monomers—may also interact with the RNA.

In an attempt to detect the presence of monomers or other oligomers, and to characterize their effect in the interaction with the RNA, the proteins were analyzed by semi-native sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE). Figure 4C shows that when SDS and reducing agent was absent from the loading buffer and the samples were not denatured by high-temperature incubation, both RRM1 and RRM1-RRM2 migrate as two species on the gel representing both monomeric and dimeric forms of the proteins. The presence of the dimer is maintained through a series of protein concentrations (10–80 μM).

Due to the presence of cysteins in the protein, we were interested in analyzing whether the presence of disulfide interactions could play a role in the formation of the dimeric species. Therefore, we have tested the effect of adding increasing amounts of a reducing agent (0, 0.5, 1 and 2% β-mercaptoethanol) to the loading buffer (Figure 4D). Again, little or no effect on the formation of the dimeric species was observed. Furthermore, the addition of CPE RNA (1:1) to both the protein samples has no effect on the formation of the dimeric species (Figure 4E). However, when SDS was added to both samples and the samples were denatured with temperature, a significant reduction in the dimeric species was observed (Figure 4E). This result indicates that the oligomerization of the proteins is not maintained through disulfide linkages and is dependent on the proteins preserving a proper fold, therefore being mediated by surface electrostatic and/or hydrophobic interactions.

The ratio of dimer to monomer as estimated from the semi-native gels is shifted toward the monomeric population with a maximum dimeric population of 10% of molecules in solution. Given this low percentage of dimer population, its large molecular weight (46 kDa), and the concentration used for the NMR experiments, its contribution to our NMR data is imperceptible.

Dimerization of RRM domains does not disrupt RNA binding

To further study the potential presence of oligomers at the protein concentration used in ITC experiments (20–50 μM) and to investigate the RNA-binding activity of the dimeric species, we used IM-MS on both constructs in their apo form as well as in complex with RNA.

Ion mobility coupled to mass spectrometry (IM-MS) is a technique that simultaneously separates gaseous ions on the basis of their mass, shape and size. The continuous advances in native mass spectrometry applied to IM-MS have prompted its application in the structural study of biomolecular complexes (29). Hence, challenging systems in terms of molecular size, complexity and heterogeneity can be transferred to the gas phase and their structural properties can be analyzed. IM-MS can provide valuable information about the occurrence and population of the species present in a sample. This technique has been efficiently applied in the characterization of protein–peptide interactions, in particular in the determination of the presence of several species and their role in the calculation of stoichiometries of complexes (30).

Using this technique we could identify in the apo RRM1 sample (Figure 6A) and in the apo RRM1–RRM2 (Figure 6B) the presence of monomeric as well as dimeric species, consistent with the ITC results and the semi-native SDS PAGE. For the complex of RRM1 with RNA, a weak binding affinity was expected. Accordingly, we detected only a minor part of the species in a monomeric complex, the majority remained in its free form (Figure 6C). For the complex with the RRM-tandem domains, the presence of monomeric complexes as well as protein dimers bound to two ligands was detected (Figure 6D). Moreover, no unbound species were detected. The different species (see schematic representation in Figure 6E) were unambiguously identified based on their specific mass-to-charge ratio (MS) and/or their characteristic drift-times that correlate their size-to-charge ratio (IM). Therefore, IM-MS confirms the hypothesis that indeed dimeric species are present. Perhaps, the presence of monomeric and dimeric forms renders the determination of the protein concentration in native conditions inaccurate, explaining the stoichiometry values inferior to one. The results also clearly indicate that only 1:1 protein–RNA complexes are present in both the monomeric and dimeric forms (ML, D2L). The presence of the M1L and D2L species probably contributes to the global affinity measured by the ITC experiments, which is mainly governed by the contribution of the M1L form. The results of the IM-MS experiments with RRM1–RRM2 in complex with RNA reveal that the presence of the ligand does not prevent the formation of dimeric species and does not render the RNA binding site inaccessible. We therefore conclude that the dimerization surface of the RRM1–RRM2 is separate from its RNA-binding interface.

DISCUSSION

In this study we provide evidence that CPEB4 uses both its RRM domains to maintain optimal RNA binding. RRM1, whose sequence is the more conserved of the two tandem domains, relative to other RRM motifs, is only able to bind CPE-containing RNAs in the high μM range. ITC experiments performed with the RRM1–RRM2 pair yielded affinity values around 300 nM, 100 times higher than the isolated RRM1 domain.

An overview of a repertoire of nucleic acid binding modes by proteins containing two RRM domains and the principles of multi-domain protein–RNA recognition has been recently reviewed by Mackereth and Sattler (31) and reported by Barraud and Allain (32). Based on the structures available, they classified different domain arrangements in the free as well as in the bound state (31,32). In the free state, the domains can either be independent, with pre-formed domain contacts, or adopt a closed, autoinhibited form. Upon ligand binding, RRMs can, for example, independently bind separate RNA motifs or form a con-

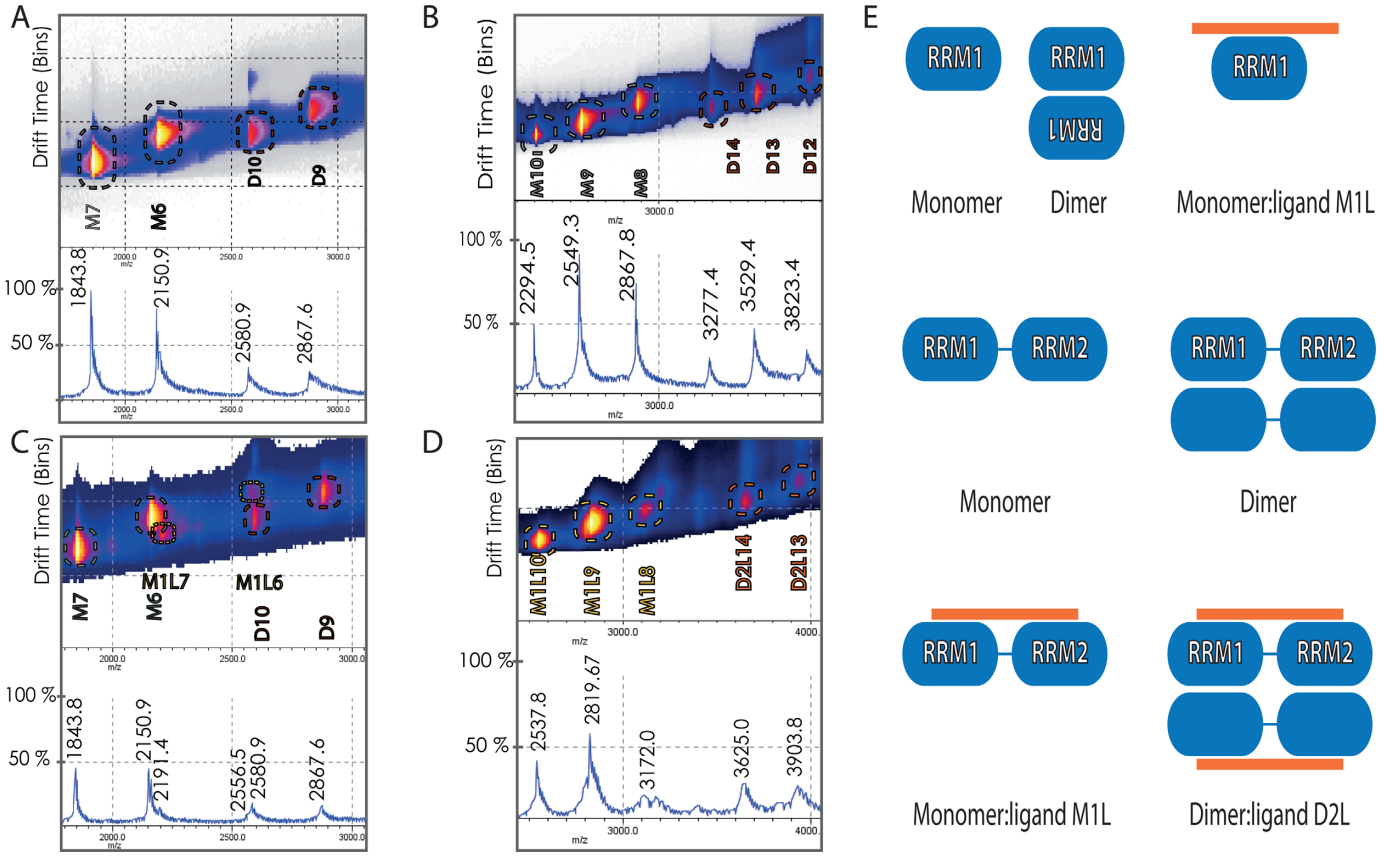

Figure 6. A region of the ion mobility–mass spectrometry data obtained for apo RRM1 (**A**), apo RRM1–RRM2 (**B**), RRM1–U5A2U1 complex (**C**) and RRM1–RRM2–U5A2U1 complex (**D**). Each ion was assigned to a given species based on its characteristic mobility and mass-to-charge ratio. Abbreviations used are M (monomer, apo state), D (dimer, apo state), M1L (monomer with one ligand) and D2L (dimer with two ligands). Numbers following the species' name indicate the ionization state. For exact ion masses, both observed and theoretical values, please refer to Supplementary Figure S6. We have unambiguously detected the presence of dimers for both RRM1 (A) and RRM1–RRM2 (B) in their apo state. Consistent with ITC results, for the RRM1–U5A2U1 complex (C) only a minor part of monomeric complex was detected (M1L7 and M1L6); the majority of the species represent the ones already detected for the apo state (A). (D) In the RRM1–RRM2–U5A2U1 sample, monomers in complex with one ligand as well as dimers with two ligands were identified. (**E**) A schematic representation of the species identified by IM-MS. The RRM domains are represented in blue ellipses and the RNA octamer as an orange line.

tinuous RNA-binding platform, stabilized by inter-domain contacts or increased linker rigidity. Another mechanism to stabilize the RNA binding is mediated by protein–protein interactions with an auxiliary protein.

In the case of CPEB4 RRM domains, we observed that both domains contribute to the interaction with the consensus eight nucleotide CPE motifs, presumably increasing not only the affinity but also the specificity of the interaction. Given the length of the RNA motif, we assume that the RRM domains form a continuous binding platform. Based on the ^{15}N amide relaxation properties and the residues affected upon RNA binding the CPEB4 RRM1–RRM2 pair behaves as a compact molecule already in the unbound state, suggesting that the pre-formed domain contacts provide an arrangement that might be finely tuned in the presence of RNA.

Our NMR titration data for CPEB4's tandem RRMs show that in RRM1 the most affected residues upon binding lie on its β-sheet, the canonical binding surface of RRMs, containing the conserved RNPs on the β_1 and β_3 strands. RRM2 is the less canonical domain of the pair containing three of six conserved residues of RNP2 motif (I-F-V) and completely lacking the RNP1 motif on the β_3 strand. Moreover, it seems to interact with the CPE in a non-canonical manner, since we detect chemical shift changes also at positively charged residues located in its C-terminus. As reported previously, the β-sheet surface can bind up to four nucleotides. Therefore, the recognition of a longer RNA requires more than one RRM to create a sufficiently large binding platform. The two consensus CPE sequences contain each seven nucleotides (UUU UAA U and UUU UAU U), thus it seems plausible that a motif of this length binds more than one RRM. Moreover, the use of regions other than the β-sheet binding platform has been demonstrated in other tandem RRM structures. For example, the RRMs in the protein nucleolin use the linker and a loop between the β-sheets to interact with a stem loop RNA structure (33). The seemingly different modes of RNA interaction of CPEB's RRMs, with RRM2 using C-terminal residues, suggest an asymmetrical orientation of the domains relative to each other.

In an attempt to visualize all the results obtained from this study, we built a homology model containing the two RRM domains and an octameric polyU stretch to illustrate

the relative size of a single-stranded octameric RNA compared to the RRM domains (Supplementary Figure S7). We believe that this preliminary model based on the CS changes obtained for the amide resonances of the RRM1–RRM2 sample in the presence of RNA provides an idea about the relative orientation of the two domains. Further investigation of the RRM's CPE interaction and their relative orientation to each other is required using structural techniques such as NMR or crystallography.

Our results demonstrate that CPEB4 RRM1 and RRM1–RRM2 can sample monomer and dimer conformations under the experimental conditions we have assayed. However, the inability to produce a soluble RRM2 fragment did not allow us to identify whether both domains are involved in the self-association. The dimeric species are present in low abundance, precluding its characterization using NMR spectroscopy. However, the potential to form dimers has been observed for other RRM domains. For example, the U1A protein forms a homodimer via its RRM domain. The dimer interface is separate from the polyadenylation inhibition element (PIE) RNA interaction interface (34). IM-MS assays have unambiguously detected both monomeric as well as dimeric 1:1 protein–RNA complexes, indicating that both the monomeric and dimeric states of RRM1–RRM2 possess RNA-binding activity. This suggests that dimerization is not required for RNA recognition. Previously, dimerization has been reported for another member of the CPEB-family CPEB1 (19). Lin *et al.* (2012) postulated that the dimerization functions to inactivate spare proteins, preventing them from inducing the polyadenylation of RNAs with low affinity binding sites and also serving as molecular hubs that release polyadenylation factors upon dimer destruction. The same study identified both the RRMs and the ZZ-domain to be essential for dimerization. Interestingly, and in contrast to our experiments, the results suggest that CPEB1 dimerization occurs at the expense of RNA binding. However, whereas in our study, isolated RRM domains were characterized, Lin *et al.* (2012) assessed full-length CPEB1 and two dimeric constructs, two full-length CPEB1 proteins separated by a linker sequence of 40 amino acids and full-length CPEB1 with an N-terminal coiled-coil dimerization domain. The sequence identity of CPEB1 and CPEB4 RRM domains is 45% and the overall identity is 22%, although it is possible that the sequence disparity between the two protein RRM domains may account for differences in the mode of the RNA interaction and self-association. We suggest, it is more likely, that the reason for the loss of binding activity in the full-length CPEB1 dimer is due to the spatial arrangement of the N-terminal unstructured regions and the C-terminal ZZ-domains, which renders the RNA-binding interface inaccessible. Our IM-MS data clearly show that the dimeric form of the RRM pair maintains their RNA-binding activity. We therefore conclude that the dimerization surface is separate from the identified RNA-binding interface.

Experiments with the full-length CPEB4, as investigated with CPEB1, would ascertain if the inhibition of RNA binding upon dimerization is exclusive of CPEB1 or is present in other family members. Further investigation into the function of the potential formation of CPEB oligomers is required since it may reveal the critical differences between the four functional CPEB paralogs of vertebrates. It would be interesting to see if the postulated molecular hubs are detectable using the available cell biology techniques or could indeed be purified for *in vitro* analysis.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

We thank the Mass Spectrometry Core Facility at IRB Barcelona (Dr M. Vilaseca) for support with the IM-MS experiments as well as the NMR Facility of FMP Berlin (Dr H. Oschkinat and Dr P. Schmieder) for NMR measurement time and support during the acquisition of the experiments.

FUNDING

Ministry of Science and Innovation [Programme Consolider RNAREG CSD2009-00080 to M.J.M. and E.P., CTQ2010-21567-C02-01 to E.P., SAF2011-25119 to M.J.M.]. C.S. is a recipient of a 'laCaixa'/IRB Barcelona International Ph.D. Programme fellowship. M.J.M. is an ICREA Programme Investigator. Source of open access funding: Publication charges are funded by the Programme Consolider RNAREG CSD2009-00080 (to M.J.M.) grant.
Conflict of interest statement. None declared.

REFERENCES

1. Richter,J.D. (2007) CPEB: a life in translation. *Trends Biochem. Sci.*, **32**, 279–285.
2. Kim,J.H. and Richter,J.D. (2006) Opposing polymerase-deadenylase activities regulate cytoplasmic polyadenylation. *Mol. Cell*, **24**, 173–183.
3. Stebbins-Boaz,B., Cao,Q., De Moor,C.H., Mendez,R. and Richter,J.D. (1999) Maskin is a CPEB-associated factor that transiently interacts with eIF-4E. *Mol. Cell*, **4**, 1017–1027.
4. Cao,Q., Kim,J.H. and Richter,J.D. (2006) CDK1 and calcineurin regulate Maskin association with eIF4E and translational control of cell cycle progression. *Nat. Struct. Mol. Biol.*, **13**, 1128–1134.
5. Mendez,R., Hake,L.E., Andresson,T., Littlepage,L.E., Ruderman,J.V. and Richter,J.D. (2000) Phosphorylation of CPE binding factor by Eg2 regulates translation of c-mos mRNA. *Nature*, **404**, 302–307.
6. Barnard,D.C., Ryan,K., Manley,J.L. and Richter,J.D. (2004) Symplekin and xGLD-2 are required for CPEB-mediated cytoplasmic polyadenylation. *Cell*, **119**, 641–651.
7. Wang,X.-P. and Cooper,N.G.F. (2010) Comparative in silico analyses of cpeb1–4 with functional predictions. *Bioinform. Biol. Insights*, **4**, 61–83.
8. Villalba,A., Coll,O. and Gebauer,F. (2011) Cytoplasmic polyadenylation and translational control. *Curr. Opin. Genet. Dev.*, **21**, 452–457.
9. Hake,L.E., Mendez,R. and Richter,J.D. (1998) Specificity of RNA binding by CPEB: requirement for RNA recognition motifs and a novel zinc finger. *Mol. Cell. Biol.*, **18**, 685–693.
10. Merkel,D.J., Wells,S.B., Hilburn,B.C., Elazzouzi,F., Pérez-Alvarado,G.C. and Lee,B.M. (2013) The C-terminal region of cytoplasmic polyadenylation element binding protein is a ZZ domain with potential for protein-protein interactions. *J. Mol. Biol.*, **425**, 2015–2026.
11. Fox,C.A., Sheets,M.D. and Wickens,M.P. (1989) Poly(A) addition during maturation of frog oocytes: distinct nuclear and cytoplasmic activities and regulation by the sequence UUUUUAU. *Genes Dev.*, **3**, 2151–2162.

12. McGrew,L.L., Dworkin-Rastl,E., Dworkin,M.B. and Richter,J.D. (1989) Poly(A) elongation during Xenopus oocyte maturation is required for translational recruitment and is mediated by a short sequence element. *Genes Dev.*, **3**, 803–815.
13. Novoa,I., Gallego,J., Ferreira,P.G. and Mendez,R. (2010) Mitotic cell-cycle progression is regulated by CPEB1 and CPEB4-dependent translational control. *Nat. Cell Biol.*, **12**, 447–456.
14. Igea,A. and Méndez,R. (2010) Meiosis requires a translational positive loop where CPEB1 ensues its replacement by CPEB4. *EMBO J.*, **29**, 2182–2193.
15. Huang,Y.-S., Kan,M.-C., Lin,C.-L. and Richter,J.D. (2006) CPEB3 and CPEB4 in neurons: analysis of RNA-binding specificity and translational control of AMPA receptor GluR2 mRNA. *EMBO J.*, **25**, 4865–4876.
16. Ortiz-Zapater,E., Pineda,D., Martínez-Bosch,N., Fernández-Miranda,G., Iglesias,M., Alameda,F., Moreno,M., Eliscovich,C., Eyras,E., Real,F.X. *et al.* (2012) Key contribution of CPEB4-mediated translational control to cancer progression. *Nat. Med.*, **18**, 83–90.
17. Pavlopoulos,E., Trifilieff,P., Chevaleyre,V., Fioriti,L., Zairis,S., Pagano,A., Malleret,G. and Kandel,E.R. (2011) Neuralized1 activates CPEB3: a function for nonproteolytic ubiquitin in synaptic plasticity and memory storage. *Cell*, **147**, 1369–1383.
18. Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.
19. Lin,C.-L., Huang,Y.-T. and Richter,J.D. (2012) Transient CPEB dimerization and translational control. *RNA*, **18**, 1050–1061.
20. Marley,J., Lu,M. and Bracken,C. (2001) A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR*, **20**, 71–75.
21. Bellon,L. (2001) Oligoribonucleotides with 2′-O-(tert-Butyldimethylsilyl) groups. In: *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons, Inc., HOBOKEN, NJ.
22. Vranken,W.F., Boucher,W., Stevens,T.J., Fogh,R.H., Pajon,A., Llinas,M., Ulrich,E.L., Markley,J.L., Ionides,J. and Laue,E.D. (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*, **59**, 687–696.
23. 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.
24. Bartels,C., Xia,T.H., Billeter,M., Güntert,P. and Wüthrich,K. (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. *J. Biomol. NMR*, **6**, 1–10.
25. Barbato,G., Ikura,M., Kay,L.E., Pastor,R.W. and Bax,A. (1992) Backbone dynamics of calmodulin studied by 15N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry (Mosc.)*, **31**, 5269–5278.
26. Sánchez-Hernández,N., Ruiz,L., Sánchez-Álvarez,M., Montes,M., Macias,M.J., Hernández-Munain,C. and Suñé,C. (2012) The FF4 and FF5 domains of transcription elongation regulator 1 (TCERG1) target proteins to the periphery of speckles. *J. Biol. Chem.*, **287**, 17789–17800.
27. Farrow,N.A., Muhandiram,R., Singer,A.U., Pascal,S.M., Kay,C.M., Gish,G., Shoelson,S.E., Pawson,T., Forman-Kay,J.D. and Kay,L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. *Biochemistry (Mosc.)*, **33**, 5984–6003.
28. Daragan,V.A. and Mayo,K.H. (1997) Motional model analyses of protein and peptide dynamics using 13C and 15N NMR relaxation. *Prog. Nucl. Magn. Reson. Spectrosc.*, **31**, 63–105.
29. Hoaglund,C.S., Valentine,S.J., Sporleder,C.R., Reilly,J.P. and Clemmer,D.E. (1998) Three-dimensional ion mobility/TOFMS analysis of electrosprayed biomolecules. *Anal. Chem.*, **70**, 2236–2242.
30. Aragón,E., Goerner,N., Xi,Q., Gomes,T., Gao,S., Massagué,J. and Macias,M.J. (2012) Structural basis for the versatile interactions of smad7 with regulator WW domains in TGF-β pathways. *Structure*, **20**, 1726–1736.
31. Mackereth,C.D., Madl,T., Bonnal,S., Simon,B., Zanier,K., Gasch,A., Rybin,V., Valcárcel,J. and Sattler,M. (2011) Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature*, **475**, 408–411.
32. Barraud,P. and Allain,F.H.-T. (2013) Solution structure of the two RNA recognition motifs of hnRNP A1 using segmental isotope labeling: how the relative orientation between RRMs influences the nucleic acid binding topology. *J. Biomol. NMR*, **55**, 119–138.
33. Allain,F.H., Bouvet,P., Dieckmann,T. and Feigon,J. (2000) Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin. *EMBO J.*, **19**, 6870–6881.
34. Varani,L., Gunderson,S.I., Mattaj,I.W., Kay,L.E., Neuhaus,D. and Varani,G. (2000) The NMR structure of the 38 kDa U1A protein-PIE RNA complex reveals the basis of cooperativity in regulation of polyadenylation by human U1A protein. *Nat. Struct. Biol.*, **7**, 329–335.