



Oligomeric transition and dynamics of RNA binding by the HuR RRM1 domain in solution

Carolina Lixa¹ · Amanda Mujo¹ · Mariana T. Q. de Magalhães² · Fabio C. L. Almeida³ · Luis Mauricio T. R. Lima⁴ · Anderson S. Pinheiro¹

Received: 17 July 2018 / Accepted: 4 December 2018 / Published online: 8 December 2018
© Springer Nature B.V. 2018

Abstract

Human antigen R (HuR) functions as a major post-transcriptional regulator of gene expression through its RNA-binding activity. HuR is composed by three RNA recognition motifs, namely RRM1, RRM2, and RRM3. The two N-terminal RRM domains are disposed in tandem and contribute mostly to HuR interaction with adenine and uracil-rich elements (ARE) in mRNA. Here, we used a combination of NMR and electrospray ionization–ion mobility spectrometry–mass spectrometry (ESI–IMS–MS) to characterize the structure, dynamics, RNA recognition, and dimerization of HuR RRM1. Our solution structure reveals a canonical RRM fold containing a 19-residue, intrinsically disordered N-terminal extension, which is not involved in RNA binding. NMR titration results confirm the primary RNA-binding site to the two central β -strands, $\beta 1$ and $\beta 3$, for a cyclooxygenase 2 (Cox2) ARE I-derived, 7-nucleotide RNA ligand. We show by ^{15}N relaxation that, in addition to the N- and C-termini, the $\beta 2$ – $\beta 3$ loop undergoes fast backbone dynamics (ps–ns) both in the free and RNA-bound state, indicating that no structural ordering happens upon RNA interaction. ESI–IMS–MS reveals that HuR RRM1 dimerizes, however dimer population represents a minority. Dimerization occurs via the α -helical surface, which is oppositely oriented to the RNA-binding β -sheet. By using a DNA analog of the Cox2 ARE I, we show that DNA binding stabilizes HuR RRM1 monomer and shifts the monomer–dimer equilibrium toward the monomeric species. Altogether, our results deepen the current understanding of the mechanism of RNA recognition employed by HuR.

Keywords HuR · RRM · RNA · Structure · Dynamics · Ion-mobility · NMR

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10858-018-0217-y>) contains supplementary material, which is available to authorized users.

✉ Anderson S. Pinheiro
pinheiro@iq.ufrj.br

¹ Department of Biochemistry, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro 21941-909, Brazil

² Department of Biochemistry and Immunology, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil

³ National Center for Nuclear Magnetic Resonance Jiri Jonas, Institute of Medical Biochemistry, Federal University of Rio de Janeiro, Rio de Janeiro 21941-902, Brazil

⁴ Faculty of Pharmacy, Federal University of Rio de Janeiro, Rio de Janeiro 21941-590, Brazil

Introduction

Eukaryotes employ a variety of mechanisms to fine tune gene expression. Many of those act post-transcriptionally by controlling the levels of mRNA. Post-transcriptional regulation of gene expression starts in the nucleus with the addition of a 7-methylguanosine to the 5' end, a poly-adenosine tail to the 3' end, as well as editing and splicing of pre-mRNA (Moore 2005). The concentration of mature mRNA in the cytoplasm, however, is dependent on several post-transcriptional mechanisms that regulate mRNA transport, localization, stability, and degradation (Moore 2005). Such mechanisms are determined by a dynamic interplay between *cis*-elements, specific mRNA sequences and/or secondary structures, and *trans* factors, which by interacting with *cis*-elements exert control over the mRNA level. A well-characterized *cis*-regulatory element is the adenine and uracil-rich element (ARE) located primarily in the 3' untranslated regions of mRNA (Migone et al. 2002; Barreau

et al. 2006). AREs are bound by a plethora of *trans*-acting factors, including non-coding RNAs (ncRNAs) and RNA-binding proteins (RBPs), which regulate mRNA stability and translation (Bevilacqua et al. 2003; Glisovic et al. 2008).

To date, several ARE-specific RBPs (ARE-BPs) have been characterized. Depending on the nature of the ARE-BP and the sequence of the ARE-containing mRNA, the interaction between them may lead to different functional outcomes. Many ARE-BPs, such as TTP, AUF1 and KSRP, exert their functions by increasing the rate of mRNA decay (Wilusz and Wilusz 2004). Others, like TIA1 and TIAR, silence translation by sequestering mRNA to stress granules (Buchan and Parker 2009). In contrast, few ARE-BPs act by stabilizing mRNA transcripts, including the embryonic lethal abnormal vision/human antigen (ELAV/Hu) family members, specifically human antigen R (HuR) (Myer et al. 1997; Fan and Steitz 1999; Brennan and Steitz 2001).

HuR is a 32-kDa, ubiquitous protein that predominantly localizes in the nucleus of resting cells. However, in response to proliferative and stress stimuli, HuR undergoes cytoplasmic translocation (Doller et al. 2008). In the nucleus, HuR is engaged in the regulation of pre-mRNA splicing and polyadenylation (Doller et al. 2008; Zhu et al. 2007; Izquierdo 2008). In the cytoplasm, HuR binds and stabilizes a wide range of mRNA transcripts, modulating translation (Hinman and Lou 2008).

Many HuR targets encode proteins that play key roles in the establishment and development of tumors, including cyclins, cytokines, growth factors, proto-oncogenes, tumor suppressors, apoptosis regulators, among others (de LópezSilanes et al. 2005). High expression levels and abnormal nuclear-cytoplasmic localization of HuR have been found in various types of cancers, including lung, colon, breast, ovarian (Blaxall et al. 2000; López de Silanes et al. 2003; Denkert et al. 2004a, b). Clinical evidences suggest a direct correlation between cytoplasmic HuR overexpression and specific clinicopathological features of tumors, such as size and drug resistance, as well as low patient survival (Zhu et al. 2013; Giaginis et al. 2015). Thus, HuR emerges as a potential therapeutic target and prognostic marker of cancer. Inhibition of HuR through compound MS-444 in colorectal carcinoma cells decreased growth and increased apoptosis, pointing HuR as an effective therapeutic strategy against cancer (Blanco et al. 2016).

HuR has a tripartite architecture; it is composed by three functionally distinct RNA recognition motif (RRM) domains, which contribute individually for mRNA binding. The N-terminal RRM1 and RRM2 domains are disposed in tandem, while the C-terminal RRM3 domain is connected to RRM2 by a basic hinge region (Brennan and Steitz 2001). This RRM2–RRM3 linker region contains an HNS sequence and thus regulates HuR nuclear-cytoplasmic shuttling. ARE recognition is primarily mediated by the N-terminal RRM1

and RRM2 domains (Doller et al. 2008); whereas RRM3 interacts with the polyA tail and displays a 3' adenosyl transferase activity capable of elongating mRNA (Meisner et al. 2009). In addition, RRM3 binds uracil-rich RNA sequences and it has been implicated in the stabilization of HuR-ARE complexes (Sheiba et al. 2014).

The crystal structure of the isolated RRM1 domain of HuR reveals a typical RRM fold, composed of a four-stranded β -sheet packed against two α -helices (Benoit et al. 2010). RNA binding is accomplished by two canonical RNP1 and RNP2 sequences located at the central β -strands, β 1 and β 3 (Benoit et al. 2010). Crystal structures of the two tandem RRM1-2 domains both free and in complex with an 11-mer RNA oligonucleotide reveal that ARE binding is accompanied by a conformational change in HuR RRM1-2 (Wang et al. 2013). In both cases, the protein constructs used in the crystallization experiments lacked the first 19 residues of the RRM1 sequence. Moreover, the role played by HuR RRM1 dynamics in RNA interaction is unclear, preventing a deeper understanding of the mechanism of ARE recognition. Previous studies suggest that full-length HuR has the ability to dimerize in the absence of RNA (Meisner et al. 2007). Nevertheless, the contribution of the RRM1 domain to HuR dimerization has not been investigated. Here, we report a detailed solution characterization of the dynamics and homotropic contribution of dimerization to ARE binding by HuR RRM1 (residues 1–99, herein referred as RRM1_{1–99}).

Materials and methods

RRM1_{1–99} expression and purification

The pET-RP1B plasmid carrying the DNA sequence encoding RRM1_{1–99} was purchased from GenScript USA (Piscataway, USA). RRM1_{1–99} was expressed as a fusion protein to an N-terminal Thio₆His₆ tag followed by a TEV protease cleavage site and a C-terminal, non-cleavable His₆ tag. *Escherichia coli* BL21 (DE3) cells were transformed with pET-RP1B-RRM1_{1–99}. Cells were grown at 37 °C until mid-exponential phase ($A_{600} \sim 0.6$). Protein expression was induced with 1 mM IPTG and cells were allowed to grow for additional 16 h at 37 °C. For the production of ¹³C/¹⁵N- and ¹⁵N-labeled samples, cells were grown in M9 minimal medium containing 3 g/L ¹³C-D-glucose and/or 1 g/L ¹⁵NH₄Cl g/L as the sole sources of carbon and nitrogen, respectively. Cells were harvested by centrifugation (8000×g, 20 min, 4 °C), resuspended in lysis buffer [50 mM Tris–HCl (pH 8.0), 500 mM NaCl, 5 mM imidazol, 0.1% Triton-X 100, 250 μ M PMSF, 5 mM DTT], and lysed by sonication (20 cycles 30 s on and 60 s off, 100 W). The cell lysate was centrifuged (8000×g, 40 min, 4 °C) and the

clarified supernatant was loaded onto a His-Trap HP column (GE Healthcare, USA). The protein of interest was eluted with an imidazole gradient ranging from 5 to 500 mM and dialyzed against fresh buffer [50 mM Tris–HCl (pH 7.2), 5 mM DTT] in the presence of TEV protease to remove the N-terminal Thio₆His₆ tag. After complete cleavage, the protein sample was loaded onto an anionic exchange Q FF 16/10 column (GE Healthcare, USA), from which it was eluted with a NaCl gradient varying from 0 to 1 M. Finally, RRM1_{1–99} was purified by size-exclusion chromatography using a Superdex 75 16/60 column (GE Healthcare, USA) equilibrated with 20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT. Fractions containing the protein of interest were analyzed by SDS PAGE, pooled and concentrated. Purified, recombinant RRM1_{1–99} contained two N-terminal residues (GH) and a C-terminal His₆ tag as cloning artifacts.

NMR spectroscopy

NMR experiments were performed in the analytical platform of the National Center for NMR of the Federal University of Rio de Janeiro (CNMRN-UFRJ). Measurements were acquired at 25 °C on Bruker Avance III 800 and 700 MHz spectrometers equipped with inverse-detection 5 mm TXI HCN z-gradient probes, and a Bruker DRX 600 MHz spectrometer equipped with an inverse-detection 5 mm TCI HCN z-gradient cryogenic probe. The NMR samples contained 0.2–1.5 mM RRM1_{1–99} in 20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT, supplemented with 10% (v/v) ²H₂O. The sequence-specific backbone and side chain assignments of RRM1_{1–99} were reported previously (Mujo et al. 2015). NMR spectra were recorded with Topspin 3.2 (Bruker Biospin), processed with either Topspin 3.2 or NMRPipe (Delaglio et al. 1995), and analyzed with CcpNmr Analysis (Vranken et al. 2005).

Structure determination from NMR data

NMR data were collected on a 1.5 mM ¹³C/¹⁵N-labeled RRM1_{1–99} sample in 20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT, 10% (v/v) ²H₂O at 600 and 700 MHz. The three-dimensional structure of RRM1_{1–99} was determined using the programs Aria 2.3 (Linger et al. 2001; Rieping et al. 2007) and CNS 1.2 (Brunger 2007). NOE-based distance restraints were derived from 3D ¹⁵N-edited (80 ms) and ¹³C-edited (80 ms, aliphatic and aromatic) [¹H, ¹H] NOESY-HSQC spectra. NOESY cross peaks were picked manually and assigned in a semi-automatic fashion using Aria 2.3 (Linger et al. 2001; Rieping et al. 2007). Dihedral angle restraints were derived from ¹H^N, ¹⁵N, ¹³C^α, ¹H^α, ¹³C^β, and C' chemical shifts using TALOS-N (Shen and Bax 2013). The amino acid sequence, the chemical

shift lists, the dihedral angle values and the three NOESY data sets were used as input files for structural determination. Structures were calculated using a standard torsion and simulated annealing protocol as implemented in CNS 1.2 (Brunger 2007). In the final calculation, 400 structures were generated and the 20 lowest-energy structures, derived from the water refinement step (Linge et al. 2003), were selected as representative of the ensemble of RRM1_{1–99} conformations in solution. Structures were visualized with MOLMOL (Koradi et al. 1996) and PyMOL (DeLano 2002). The geometric and stereo chemical quality of the structural ensemble was analyzed with the *Protein Structure Validation Software Suite*. The atomic coordinates of RRM1_{1–99} were deposited in the *Protein Data Bank* under accession code 5SZW.

NMR relaxation measurements

¹⁵N relaxation parameters (R_1 , R_2 , and ¹⁵N–{¹H} NOE) were measured on ¹⁵N-labeled RRM1_{1–99} samples (1.5 and 0.25 mM) in 20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT, 10% (v/v) ²H₂O at 800 MHz using standard pulse sequences (hsqct1etf3gpsi3d, hsqct2etf3gpsi3d, and hsqcnoef3gpsi, respectively, as implemented in Topspin 3.2). NMR spectra were collected with 1024 × 256 total points in the F_2 (¹H) and F_1 (¹⁵N) dimensions, respectively. The longitudinal (R_1) and transverse (R_2) relaxation rates were measured at eight different relaxation times: 100, 200, 300, 400, 500, 700, 900, and 1100 ms for R_1 ; 16.96, 33.92, 50.88, 84.80, 118.72, 135.65, 169.90, and 203.52 ms for R_2 . All R_2 measurements used a CPMG π pulse of 160 μ s at 63.5 W. Relaxation delays of 1.2 and 3 s were used for the R_1 and R_2 experiments, respectively. The steady-state heteronuclear ¹⁵N–{¹H} NOE were measured in spectra acquired with and without hydrogen presaturation, in an interleaved manner. To ensure that maximum ¹⁵N–{¹H} NOE were built up, a relaxation delay of 5 s was used. The proton saturation scheme for ¹⁵N–{¹H} NOE measurements used a train of high-power (12.3 W) 120° pulses for a total duration of 20.3 ms. All relaxation experiments were acquired as pseudo-3D spectra and converted to 2D data sets. NMR spectra were processed with NMRPipe (Delaglio et al. 1995) and analyzed with CcpNmr Analysis (Vranken et al. 2005). R_1 and R_2 values were determined from the fit of the peak intensities using a two-parameter monoexponential function as described in Eq. (1) (Barbato et al. 1992; Stone et al. 1992; Farrow et al. 1994):

$$I(t) = I_0 \exp(-R_{1,2}t) \quad (1)$$

where I_0 and $I(t)$ are the peak intensities at times zero and t , respectively. The ¹⁵N–{¹H} NOE values were determined from the ratio of the peak intensities measured in the reference (I_0) and the presaturated (I_S) spectra according to Eq. (2) (Barbato et al. 1992; Stone et al. 1992):

$$\text{NOE} = I_S/I_0 \quad (2)$$

The experimental errors were estimated from the standard deviation of the background noise in relation to that of the peak intensities (Farrow et al. 1994). For the $^{15}\text{N}\{-^1\text{H}\}$ NOE, the background noise and peak intensities standard deviation, measured for the reference and the presaturated spectra, were used in the calculation of the errors (Farrow et al. 1994).

The overall rotational correlation times (τ_c) of RRM1_{1–99} were determined from the mean values of T_1 and T_2 measured at 1.5 and 0.25 mM concentration, considering only the residues involved in secondary structure, according to Eq. (3):

$$\tau_c \approx \frac{1}{4\pi\nu_N} \sqrt{6 \frac{T_1}{T_2}} - 7 \quad (3)$$

where ν_N is the ^{15}N resonance frequency (Hz), and T_1 and T_2 are the mean values of the ^{15}N relaxation times (Kay et al. 1989).

NMR chemical shift perturbation experiments

NMR titration experiments were performed on ^{15}N -labeled RRM1_{1–99} samples (0.25 mM for RNA binding and various concentrations for self-association) in 20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT, 10% (v/v) $^2\text{H}_2\text{O}$ at 800 and 700 MHz. A 7-mer RNA oligonucleotide corresponding to the Cox-2 mRNA ARE I (5'-UUUAAU U-3'), a previously known HuR binding site, was used as RNA ligand. Synthetic oligonucleotides were purchased from Integrated DNA Technologies (Coraville, USA) and dissolved in the same buffer used in the assay. To identify the RRM1_{1–99}:RNA binding interface, [^1H , ^{15}N] HSQC spectra were collected at increasing RNA–protein molar ratios: 0.0, 0.25, 0.5, 1.0, and 1.5. To characterize RRM1_{1–99} self-association in solution, [^1H , ^{15}N] HSQC spectra were recorded at increasing protein concentrations: 0.25, 0.5, 0.75, 1.0, 1.25, and 1.5 mM. Chemical shift changes were monitored as a function of RNA or protein concentration and quantified as chemical shift perturbation values. Apparent dissociation constants (K_d) were estimated using CcpNMR Analysis (Vranken et al. 2005).

Electrospray ionization–ion mobility spectrometry–mass spectrometry (ESI–IMS–MS)

All experiments were performed on a MALDI-Synapt G1 HDMS mass spectrometer (Waters Brazil; from the UEMP-IBqM-UFRJ analytical facility) with direct injection system, with Waters MassLynx 4.1 software for instrument control and data processing. RRM1_{1–99}, both free and in complex

(1:1 molar ratio) with a single-stranded DNA oligonucleotide of sequence 5'-TTTAATT-3' (Integrated DNA Technologies, Coraville, USA), were prepared at 50 μM in 50 mM ammonium acetate (pH 7.4), 5 mM DTT. Samples were manually injected at 20 $\mu\text{L}/\text{min}$, ionized by electrospray and entered the travelling-wave IMS–MS instrument, which was operated in the following conditions: capillary 2.8 kV, sampling cone 30 V, extraction cone 5.0 V, source temperature 70 $^\circ\text{C}$, and nanoflow gas pressure of 0.4 bar. Collision energies in the trap and transfer cells were set to 6.0 and 4.0 V, respectively. The travelling-wave ion-mobility spectrometry cell was operated with a wave velocity of 248 m/s and a wave high of 3.0 V. All spectra were collected at a scan rate of 2.0/s. Data were processed and analyzed with DriftScope 2.1 (Williams et al. 2009). The charge state plotted in each part of the figure was chosen based on sufficient signal intensity across all dataset.

Results

Solution structure of RRM1_{1–99} and comparison with the crystallographic structure

We first solved the solution structure of the RRM1 domain of HuR, comprising residues 1–99, using heteronuclear multidimensional NMR. Figure 1a shows the superposition of the backbone traces of the 20 lowest-energy structures calculated for RRM1_{1–99}. Structures were determined on the basis of 1688 NOE distance restraints, of which 680 were intra-residue, 336 were sequential, 155 were short-range, 57 were medium-range, and 460 were long-range NOEs, in addition to 56 dihedral angle restraints. Table 1 summarizes the experimental restraints used in the last calculation step, as well as the statistics of the structural models generated for RRM1_{1–99}.

The calculated structures met the experimental restraints with only one NOE distance violation > 0.5 Å and no dihedral angle violations $> 5^\circ$. In addition, only 7 out of the 20 RRM1_{1–99} lowest-energy structures showed a single NOE distance violation greater than 0.5 Å, which was not consistent among the 7 models. The root-mean square deviation (RMSD) of the 20 lowest-energy structures was 0.30 Å for the backbone atoms and 0.70 Å for the heavy atoms, considering only residues in secondary structure (Fig. 1a). The calculated structures displayed a high stereochemical quality with 99.5% of the amino acids in the most favorable and additionally allowed regions of the Ramachandran plot (Table 1).

The solution structure of RRM1_{1–99} is formed by a central β -sheet composed of four β -strands surrounded by two α -helices, arranged in a $\beta 1\text{--}\alpha 1\text{--}\beta 2\text{--}\beta 3\text{--}\alpha 2\text{--}\beta 4$ topology, typical of RRM domains (Fig. 1). Helices $\alpha 1$ and $\alpha 2$ are

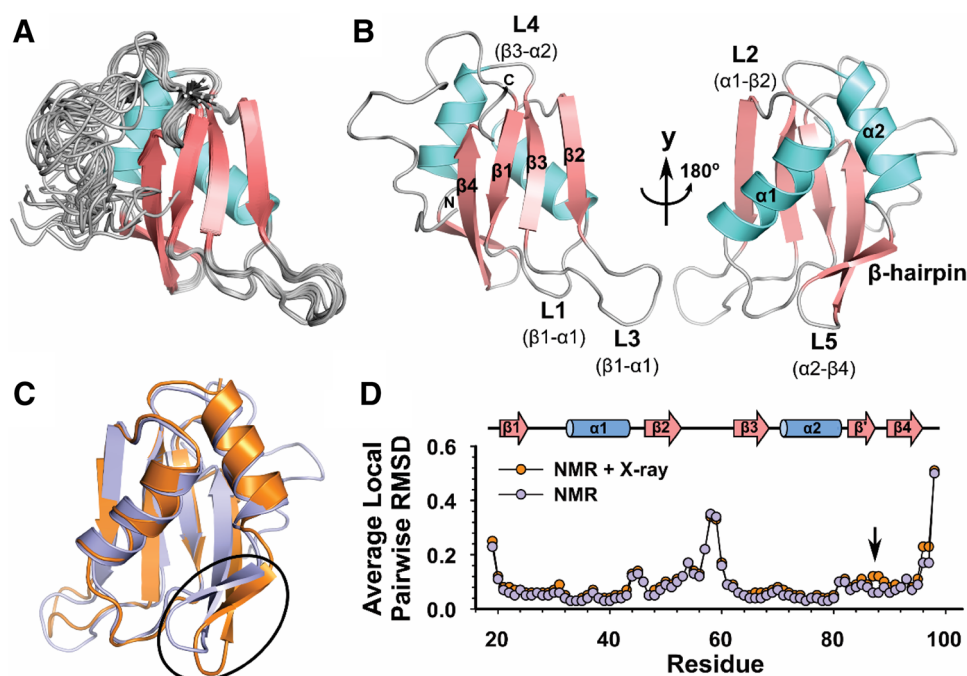


Fig. 1 NMR structure of RRM1_{1–99}. **a** Ensemble of the 20 lowest-energy conformers of RRM1_{1–99} superimposed on the backbone atoms of residues 20–99. α -Helices are colored blue, β -sheet is colored red, while loops are colored gray. **b** Ribbon representation of the lowest-energy structure calculated for RRM1_{1–99}, highlighting the secondary structure elements (helices $\alpha 1$ and $\alpha 2$, and strands $\beta 1$ – $\beta 4$). Loops are named L1–L4. The N- and C-termini and the β -hairpin are indicated.

c Superposition of the lowest-energy structure of RRM1_{1–99} (purple; PDB 5SZW) with the crystal structure of HuR RRM1 (residues 18–96) (orange; PDB 3HI9), showing the difference in orientation of the β -hairpin. **d** Average local pairwise RMSD calculated for the NMR ensemble and the NMR ensemble plus the crystal structure of HuR RRM1 (PDB 3HI9)

positioned on one side of the central β -sheet and oriented almost perpendicularly to each other (Fig. 1b). Secondary structure elements are connected by five loop regions. Except for the $\beta 2$ – $\beta 3$ loop, all loops are well-defined and overlap well within the ensemble of calculated structures (Fig. 1a). In addition, residues in the $\alpha 2$ – $\beta 4$ loop adopt a β -hairpin structure, often observed in other RRM domains (Fig. 1b) (Cléry et al. 2008).

The first 19 amino acids form an N-terminal intrinsically disordered region that extends from the canonical RRM fold. The disordered nature of the N-terminal segment is supported by the absence of NOE-derived distances within this region. In addition, chemical shift-based secondary structure propensity analysis revealed no preferential secondary structure propensity (Supplemental Figure S1A). The function of the N-terminal disordered region remains unclear. Primary sequence analysis indicated that the N-terminal extension of HuR is unique and shares no sequence conservation with other members of the ELAV/Hu family (Supplemental Figure S1B). In addition, this segment is enriched in acidic amino acids and thus it may not engage in RNA recognition. In contrast, the C-terminal region, comprising residues Ala96–Ser99 and corresponding to the RRM1-RRM2 linker is well-defined in solution (Fig. 1b).

The solution structure of RRM1_{1–99} was compared to the crystal structure of the RRM1 domain of HuR (residues 18–96) (Benoit et al. 2010). Overall, the structures superimpose well with a RMSD of 1.49 and 2.07 Å for backbone and heavy atoms, respectively. However, average local pairwise RMSD between the NMR ensemble and the NMR ensemble plus the crystal structure show that the β -hairpin adopts a different orientation in solution (Fig. 1c, d). We used the proteins, interfaces, structures and assemblies (PISA) server to determine the RRM1 residues involved in the interaction interface between monomers in the crystal. According to PISA, residues Leu84, Arg85, Leu86 and Gln87, located in the β -hairpin, are involved in the interface between chains C and D of the crystal structure and thus constitute crystal contacts. It is possible that the observed structural differences between the solution and crystal structures of HuR RRM1 result from crystal packing.

NMR analysis of RRM1_{1–99} binding to RNA

To characterize the interaction of RRM1_{1–99} with RNA ligands in solution, a 7-mer RNA oligonucleotide derived from the ARE I of the cyclooxygenase 2 mRNA was titrated into ¹⁵N-labelled RRM1_{1–99} and amide chemical shifts were

Table 1 Statistics of the 20 lowest-energy structures calculated for RRM1_{1–99}

Number of experimental restraints	
Total NOE distance restraints	1688
Intraresidue	680
Sequential	336
Short range ($2 \leq i - j \leq 3$)	155
Medium range ($4 \leq i - j \leq 5$)	57
Long range ($ i - j > 5$)	460
Torsion angle (phi/psi)	56
H-bond restraints	–
RMSD from average structure (Å)	
Backbone (21–25, 33–43, 46–53, 60–68, 72–80, 85–86, 89–90, 92–95)	0.3
Backbone, all residues	1.6
Heavy atoms (21–25, 33–43, 46–53, 60–68, 72–80, 85–86, 89–90, 92–95)	0.7
Heavy atoms, all residues	1.8
Restraint violations	
NOE violations (> 0.5 Å)	1
Dihedral angle ($> 5^\circ$)	0
Deviations from ideal geometry ^a	
RMS for bond lengths (Å)	0.007 ± 0.0003
RMS for bond angles ($^\circ$)	0.9 ± 0.02
Ramachandran plot of ordered residues (%) ^b	
Most favored regions	85
Allowed regions	14.5
Generously allowed	0.5
Disallowed	0
Energy (kcal mol ^{−1})	
E _{Total}	-3303.95 ± 119.826
E _{Bond}	22.1722 ± 3.04527
E _{Angle}	136.118 ± 10.092
E _{Vdm}	-996.631 ± 10.2688
E _{Elec}	-3073.94 ± 117.226
E _{Impar}	31.186 ± 3.34158

^aDeviation from ideal geometry calculated using CNS Solve protocols^bThe RMSD calculation and PROCHECK analyses were performed for the ensemble of the 20 lowest-energy structures using the PSVS suite

monitored. Figure 2a shows the overlap of subsets of 2D [¹H, ¹⁵N] HSQC spectra acquired for RRM1_{1–99} in the presence of varying concentrations of RNA. Increasing RNA–protein molar ratios resulted in a continuous shift of certain ¹H, ¹⁵N resonance signals, indicating that the RRM1_{1–99}:RNA interaction occurs in the fast exchange regime on the NMR time scale.

Figure 2b displays the chemical shift perturbation (CSP) values as a function of the RRM1_{1–99} residue number. Residues Leu22, Val24, Asn25, Gly64, and Phe65, located at the two central β-strands, β1 and β3, exhibited

the largest CSP values (greater than two standard deviations of the mean) and thus constitute the primary RNA binding site, which contains the consensus RNP1 and RNP2 binding sequences. In addition, residues located in helix α2 (Ala71), β2–β3 loop (Ala57, His59), α2–β4 loop (Ser88), β4 (Lys89, Thr90, Lys92), and the RRM1–RRM2 linker (Ser99) displayed CSP values greater than one standard deviation of the mean, indicating that these residues are either engaged in direct contacts or are indirectly affected by conformational changes induced by RNA binding. Moreover, RNA titration led to the disappearance of Arg97 ¹H, ¹⁵N resonance signals, suggesting an interaction that occurs in the intermediate exchange regime.

Figure 2c shows that residues displaying the highest CSP values map to a surface of the RRM1 domain comprising the central β-sheet, which is implicated in RNA binding. Furthermore, most of these residues have solvent-exposed side chains on the surface of the RRM1 structure, suggesting a direct interaction with RNA (Fig. 2d). The primary binding site is mostly hydrophobic but surrounded by positively charged residues that may contact the RNA phosphate backbone (Fig. 2e). Taken into account the five residues displaying the largest CSP values, an apparent dissociation constant (K_d) of 15.7 ± 0.98 μM was estimated for the RRM1_{1–99}:RNA interaction, adopting a one-site saturation binding mode (Supplemental Figure S2). The K_d measured from NMR titration experiments is in great agreement with previous results (Wang et al. 2013).

We compared the NMR titration results with the crystallographic structure of HuR RRM1-2 domain complexed with the RNA ligand 5'-AUUUUUAUUUU-3' derived from the *c-fos* mRNA (Wang et al. 2013) (Supplemental Figure S3). Most residues displaying significantly higher CSP values make direct contacts with the RNA ligand in the crystal structure, namely Asn25 (β1; position 2 of RNP2), Phe65 (β3; position 5 of RNP1), Lys89, Thr90, Lys92 (β4), Arg97, and Ser99 (RRM1–RRM2 linker), suggesting that the binding interface probed in solution is the same as in the crystal (Supplemental Figure S3A). However, certain residues that exhibited large CSP values do not engage in direct RNA contacts in the HuR RRM1-2:*c-fos* mRNA complex structure. Some of these residues, such as Leu22, Val24 (β1), and Leu27 (β1–α1 loop) have side chains pointing to the internal hydrophobic core of RRM1_{1–99}, suggesting that the large CSPs are a result of conformational changes induced by RNA interaction (Supplemental Figure S3B). Moreover, other solvent-accessible residues that bind RNA directly in the complex structure did not show significant CSP values, including Asn21, Tyr26 (β1), Gln29 (β1–α1 loop), and Leu61 (β3), despite the fact that their side chain conformations are highly similar to those of the crystal structure (Supplemental Figures S3C and D).

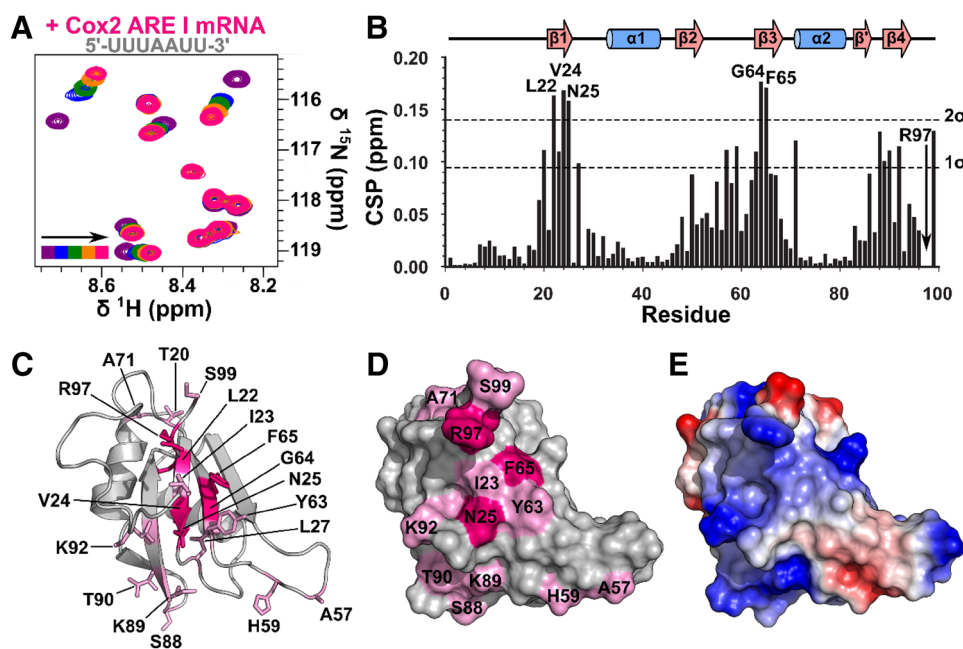


Fig. 2 NMR analysis of RRM1_{1–99} binding to the Cox2 ARE I-derived RNA ligand. **a** Overlay of subsets of 2D [¹H, ¹⁵N] HSQC spectra collected for ¹⁵N-RRM1_{1–99} in [20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT] at increasing RNA–protein molar ratios (0.0, purple; 0.25, blue; 0.5, green; 1.0, orange; 1.5, pink). **b** CSP values calculated for RRM1_{1–99} upon titration of 1.5 molar excess of 5′-UUUAAUU-3′ RNA oligonucleotide. CSPs are shown as a function of RRM1_{1–99} residue number. Secondary structure elements are depicted on top of the figure. The dashed lines represent 1 and 2 SDs of the mean CSP values. Residues display-

ing CSPs larger than 2 SDs are labeled. The arrow indicates Arg97 ¹H, ¹⁵N resonances that disappear upon RNA titration. **c** Mapping of residues exhibiting the highest CSPs (more than 2 SDs, dark pink; between 1 and 2 SDs, light pink) on top of the lowest-energy structure of RRM1_{1–99}. Side chains are represented in sticks and labeled. **d** Surface representation of RRM1_{1–99} structure in the same orientation as **c**, showing the residues that display the largest CSP values. **e** Electrostatic surface potential of RRM1_{1–99} structure. Negative potential is colored red, positive potential is colored blue, and neutral potential, white

Dynamics of free and RNA-bound RRM1_{1–99}

Dynamics of the RRM1_{1–99} backbone was monitored by the ¹⁵N relaxation parameters R_1 , R_2 , and $\{^1\text{H}\}-^{15}\text{N}$ NOE. The large chemical shift dispersion of the [¹H, ¹⁵N] HSQC spectrum of RRM1_{1–99} enabled the dynamics characterization of most amide resonances, except for residues Asp7, Asp12, Tyr63, Ala71, Asn82, and Val93, in which signal overlapping prevented the accurate measurement of intensities. Figure 3a shows the relaxation parameters R_1 , R_2 , $\{^1\text{H}\}-^{15}\text{N}$ NOE, and their associated errors, as a function of RRM1_{1–99} residue number. The mean values of the relaxation parameters calculated for the folded portion of RRM1_{1–99} (residues 20–99) are R_1 ($1.15 \pm 0.05 \text{ s}^{-1}$), R_2 ($13.52 \pm 0.53 \text{ s}^{-1}$), and $\{^1\text{H}\}-^{15}\text{N}$ NOE (0.78 ± 0.02).

Overall, RRM1_{1–99} exhibited uniform dynamics throughout most of its backbone. However, three regions showed an increase in fast backbone dynamics, motions that happen in the subnanosecond time scale, namely (1) the N-terminal extension, (2) the $\beta 2$ – $\beta 3$ loop, (3) the C-terminal RRM1–RRM2 linker. The first 19 amino acids of RRM1_{1–99} showed significantly higher than average values of R_1 accompanied by lower values of R_2 and $\{^1\text{H}\}-^{15}\text{N}$ NOE, suggesting that

they are highly flexible in solution. This correlates well with the results obtained in the RRM1_{1–99} structure calculation, where these residues constitute an intrinsically disordered segment of the protein. In addition, $\{^1\text{H}\}-^{15}\text{N}$ NOE values below 0.65 were observed for residues located in the $\beta 2$ – $\beta 3$ loop and the RRM1–RRM2 linker, suggesting that they experience increased flexibility.

From the mean values of the transverse and longitudinal relaxation rates, we estimated a global rotational correlation time (τ_c) for RRM1_{1–99} of 8 ns, taking into consideration residues in secondary structure only (Supplemental Figure S4A). This τ_c value is higher than what is expected for a globular protein of ~11 kDa, which is typically around 6 ns. The increase in τ_c is mostly due to the additive effect of the disordered region on the tumbling of the folded part of the protein, but also suggests a tendency of RRM1_{1–99} to self-associate in solution. To test this hypothesis, we collected a set of ¹⁵N relaxation measurements on RRM1_{1–99} at lower protein concentration (0.25 mM) (Fig. 3a). An overall decrease in R_2/R_1 was observed at 0.25 mM, leading to a decrease in τ_c from 8 to 7.1 ns (Supplemental Figure S4A). The fact that τ_c varies with protein concentration suggests that at least a subpopulation of RRM1_{1–99} exists in solution

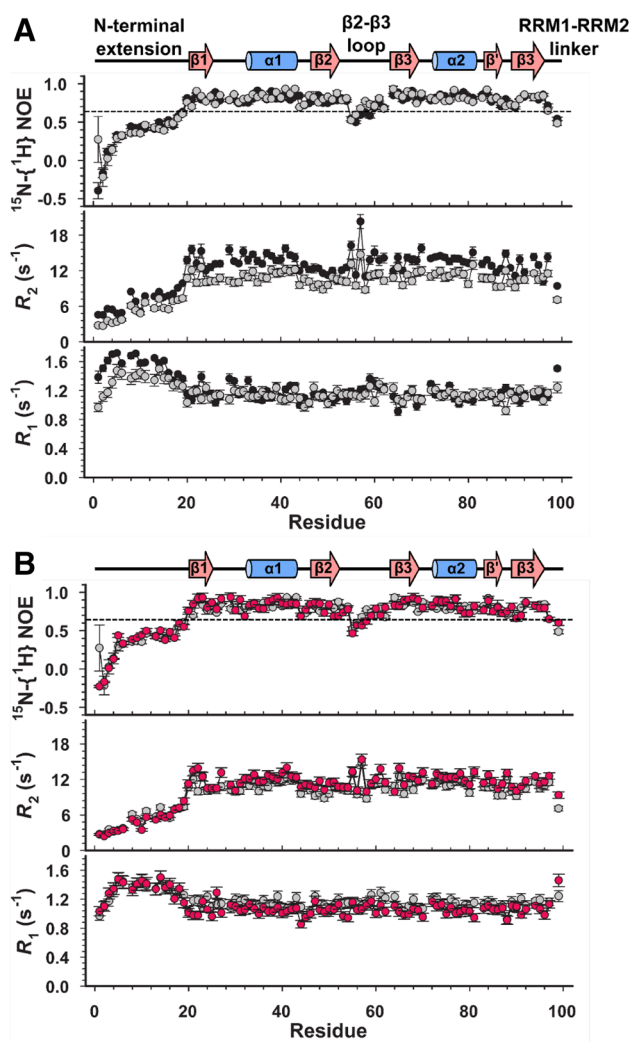


Fig. 3 Backbone dynamics of RRM1₁₋₉₉ free and in complex with RNA. **a** The $\{^1\text{H}\}\text{-}^{15}\text{N}$ NOEs, longitudinal R_1 , and transverse R_2 measured for RRM1₁₋₉₉ in the free state are shown as a function of residue number. Secondary structure elements are depicted on top of the figure. RRM1₁₋₉₉ regions undergoing fast backbone dynamics are labeled. Relaxation parameters were measured at 1.5 mM (black) and 0.25 mM (gray) ^{15}N -RRM1₁₋₉₉ in [20 mM sodium phosphate (pH 6.0), 100 mM NaCl, 5 mM DTT]. The dashed line represents $\{^1\text{H}\}\text{-}^{15}\text{N}$ NOE values of 0.65. **b** ^{15}N relaxation parameters ($\{^1\text{H}\}\text{-}^{15}\text{N}$ NOE, R_1 , and R_2) measured for RRM1₁₋₉₉ (0.25 mM) in the presence of 1.5 molar excess of Cox2 ARE I-derived RNA oligonucleotide (5'-UUUAAUU-3') shown as a function of residue number (pink). For clarity, values are superimposed with those measured for free RRM1₁₋₉₉ at 0.25 mM protein concentration (gray). Experimental conditions are the same as in **a**

in an oligomeric state larger than monomer. It is noteworthy that while R_2 values increased with the increase in protein concentration, R_1 values remained the same, suggesting that there is monomer-oligomer exchange contribution to the observed increase in R_2 .

To investigate the dynamics of the RRM1₁₋₉₉:RNA complex, we measured ^{15}N relaxation data on RRM1₁₋₉₉ in the

presence of 1.5 molar excess of the RNA ligand 5'-AAA UUAA-3' (Fig. 3b). The data indicate that the internal motions of the RRM1₁₋₉₉:RNA complex are similar to those of free RRM1₁₋₉₉. The high values of R_1 combined with the low values of R_2 and $\{^1\text{H}\}\text{-}^{15}\text{N}$ NOE displayed by the N-terminal extension suggests that it remains flexible in the RNA-bound state, consistent with the observation that this region does not engage in RNA interaction. In addition, $\{^1\text{H}\}\text{-}^{15}\text{N}$ NOE values demonstrate that, in spite of the observed CSPs, residues in the $\beta 2\text{-}\beta 3$ loop and the RRM1-RRM2 linker do not become ordered upon binding of RNA. Based on the mean R_2/R_1 ratio, a τ_c of 7.8 ns was estimated for the RRM1₁₋₉₉:RNA complex, which is slightly higher than that observed for RRM1₁₋₉₉ in the free state (Supplemental Figure S4B).

RRM1₁₋₉₉ dimerizes through the α -helical surface

Our relaxation results suggested an ability of RRM1₁₋₉₉ to self-associate in solution. The presence of highly populated oligomers could not be detected by size exclusion chromatography, where only a single peak corresponding to a molecular mass compatible with the monomer (~11 kDa) was detected. This could be due either to the technique, including in-run dilution effects, and/or to the concentration used in measurements (low μM range). Thus, at this experimental condition, no other oligomeric form besides monomer is observed. However, the decrease in the overall correlation time as a function of protein concentration suggested the existence of a small population of RRM1₁₋₉₉ oligomers.

To further investigate the presence and characterize the nature of these RRM1₁₋₉₉ oligomers, we employed electrospray ionization-ion mobility spectrometry-mass spectrometry (ESI-IMS-MS). In short, following transfer to the gas-phase, proteins are ionized and separated in one dimension according to their mass-to-charge ratio (m/z) and, in a second dimension, according to their drift times, which depend on the size and shape of the ions. Thus, ESI-IMS-MS provides valuable insights about the structural properties of gas-phase proteins, including the distribution of their individual oligomeric species. Using ESI-IMS-MS, we unambiguously identified the presence of RRM1₁₋₉₉ dimers in the absence of a nucleic acid ligand (Fig. 4a). In addition, some ions compatible with trimers and tetramers were also detected but they represent a minor total count in the ESI-IMS-MS spectrum, indicating that free RRM1₁₋₉₉ ranged predominantly from monomers to dimers (Supplemental Figure S5).

To identify the amino acid residues mediating RRM1₁₋₉₉ dimerization, we collected a series of [^1H , ^{15}N] HSQC spectra at increasing protein concentrations and amide chemical shifts were measured (Fig. 5a). Figure 5b shows the CSPs, caused by the increase in protein concentration, as a function

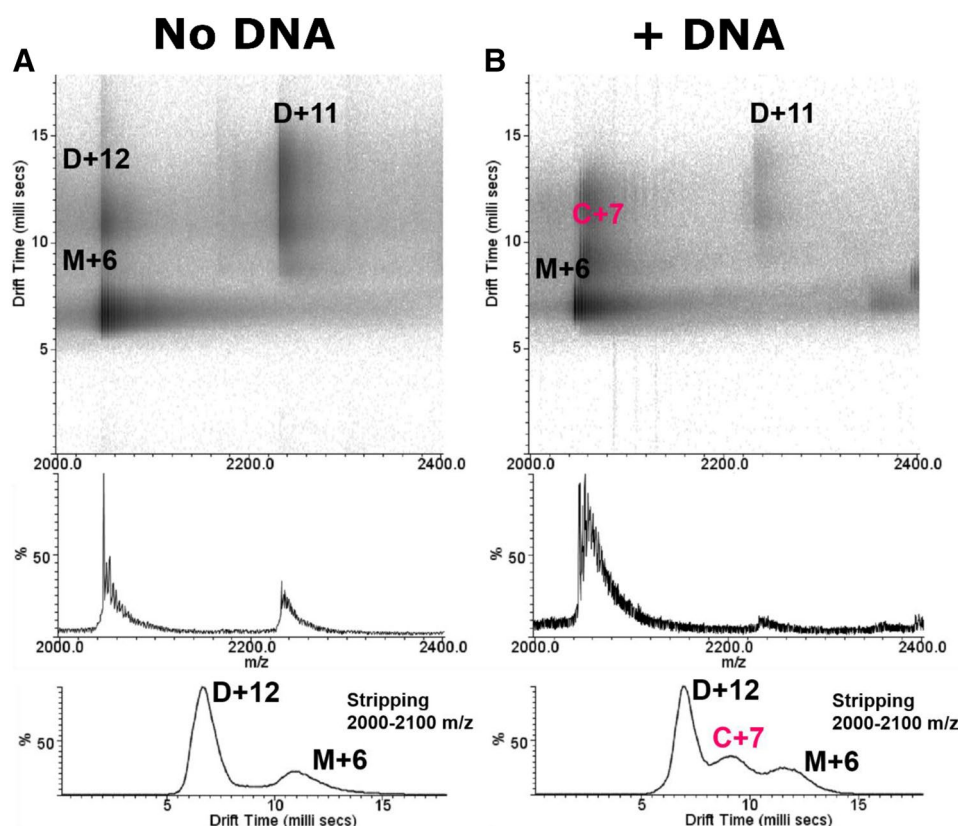


Fig. 4 Oligonucleotide binding shifts the RRM1₁₋₉₉ monomer–dimer equilibrium. A region of the ion electrospray ionization–ion mobility spectrometry–mass spectrometry (ESI–IMS–MS) data obtained for **a** apo-RRM1₁₋₉₉ and **b** RRM1₁₋₉₉:DNA complex. 50 μ M of free and DNA-bound RRM1₁₋₉₉ in [50 mM ammonium acetate (pH 7.4), 5 mM DTT] was subjected to ESI–IMS–MS. A DNA analog (5′-TTT AATT-3′) of the 7-mer Cox 2 ARE I RNA ligand was used. Ions were assigned to a given protein species, represented by a letter (*D* dimer;

M monomer; *C* complex), according to their mass-to-charge ratio and drift times. Numbers followed by the “+” sign represent the charge state of the respective ion. RRM1₁₋₉₉ dimers were identified in the free state. Addition of a DNA ligand makes the D+11 signal much weaker in the ESI–IMS–MS spectrum. In addition, only monomers in complex with one ligand (C+7; drift time ~9 ms) were detected. Ion count in log scale

of RRM1₁₋₉₉ residue number. Residues showing the highest CSP values (> 1 and 2 standard deviations of the mean) are located in the β 1– α 1 loop (Thr32), helix α 1 (Leu39, Ser41, Ile43), β 2– β 3 loop (Lys55, Val56, Ala57, Gly58, His59, Ser60), helix α 2 (Thr80, Leu81), and the β -hairpin (Leu84, Arg85, Leu86, Gln87, Ser88) (Fig. 5b). Interestingly, these residues cluster on a surface of RRM1₁₋₉₉ oriented opposite to the central β -sheet, which is responsible for RNA binding (Fig. 5c). Therefore, our results suggest that RRM1₁₋₉₉ uses one interface for protein–RNA (β -sheet) and an opposite surface for protein–protein interactions (α -helices/ β -hairpin).

To investigate the effect of nucleic acid binding on RRM1₁₋₉₉ self-association, we conducted ESI–IMS–MS measurements in the presence of a single-stranded DNA oligonucleotide of sequence 5′-TTTAATT-3′. The presence of a DNA ligand leads to the disappearance of the dimer D+11 peak, suggesting that DNA binding occurs at the expense of RRM1₁₋₉₉ dimerization. In addition, only 1:1 protein:DNA complexes were observed, indicating this

heterodimer as the most plausible consolidated of the monomeric RRM1₁₋₉₉:DNA complex (Fig. 4b). Furthermore, low intensity charge states for trimers and tetramers disappeared when in the presence of DNA, suggesting that, although their existence seems likely at those experimental conditions, such oligomers are promptly reversible toward the formation of a heterodimeric complex of a monomeric RRM1₁₋₉₉:DNA complex (Supplemental Figure S6). Therefore, our results show that nucleic acid binding shifts the monomer–dimer equilibrium of RRM1₁₋₉₉ toward the monomeric species.

Discussion

Despite small, local differences in orientation of the β -hairpin, the NMR structure is quite similar to the previously reported crystallographic structure of HuR RRM1 (Benoit et al. 2010). However, our solution structure reveals

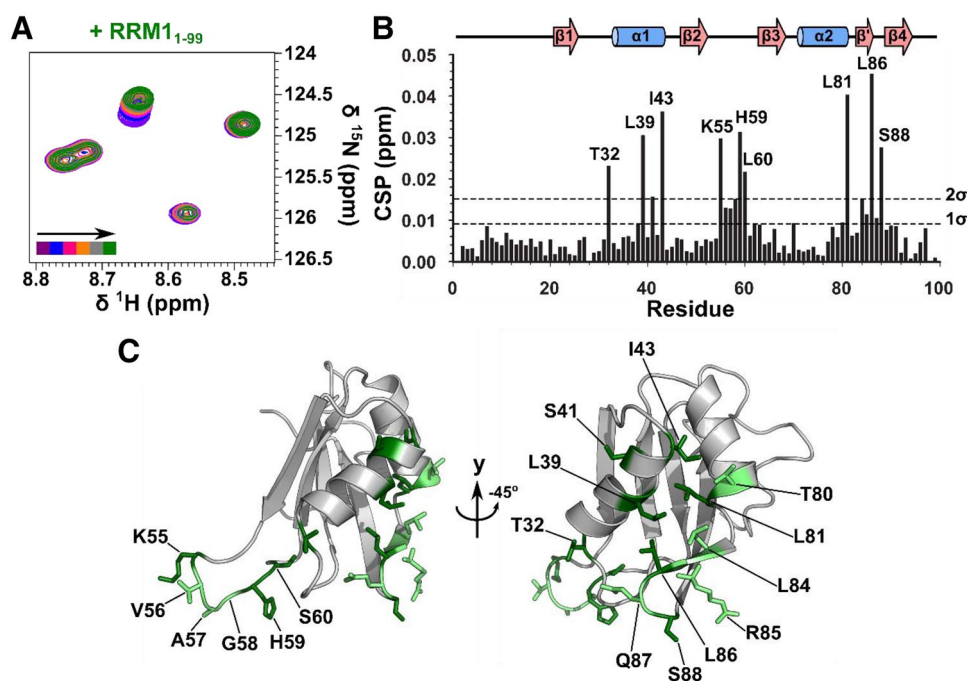


Fig. 5 NMR analysis of RRM1₁₋₉₉ dimerization interface. **a** Overlay of subsets of 2D [¹H, ¹⁵N] HSQC spectra collected for ¹⁵N-RRM1₁₋₉₉ at increasing protein concentrations (0.25, purple; 0.5, blue; 0.75, pink; 1.0, orange; 1.25 mM, gray; 1.5 mM, green). **b** CSP values calculated for RRM1₁₋₉₉ at 1.5 and 0.25 mM protein concentration are shown as a function residue number. Secondary structure elements

are depicted on top of the figure. The dashed lines represent 1 and 2 SDs. Residues displaying CSPs larger than 2 SDs are labeled. **c** Mapping of residues exhibiting the highest CSPs (more than 2 SDs, dark green; between 1 and 2 SDs, light green) on top of the lowest-energy structure of RRM1₁₋₉₉. Side chains are represented in sticks and labeled

an extended RRM fold, preceded by a 19-residue, intrinsically disordered segment, which is absent in the crystal structure. This N-terminal extension is unique of HuR and it does not show any sequence similarity with other members of the ELAV/Hu family. In addition, this feature is conserved among HuR protein sequences from human to other species. Structural variations of the canonical RRM fold have been reported to other RRM-containing proteins (Afroz et al. 2015). Among them, N- and C-terminal expansions are the most prevalent. In some cases, such as in TIAR RRM2 (Kim et al. 2013), TIA1 RRM3 (Wang et al. 2014), and SR proteins pseudo-RRMs (Cléry et al. 2011, 2013), the N- and C-termini contribute to additional RNA contacts. In others, including yeast Nop15 (Zhang and Hall 2017) and the human La family proteins (Singh et al. 2012, 2013; Eichhorn et al. 2016), a dynamic, α -helical C-terminal extension occludes the RNA binding site. Our NMR titration results indicate that the N-terminal segment of HuR does not mediate RNA binding, as no CSPs were observed for residues in this region.

Although several RNA binding mechanisms have been described to RRM domains (Cléry et al. 2008; Samatanga et al. 2017), they are known to bind RNA via base stacking interactions with aromatic side chains within the β -sheet, namely β 1 and β 3. NMR titration experiments confirm the

primary RNA binding site of RRM1₁₋₉₉ to these two central β -strands for the Cox2 ARE I-derived 5'-UUUAAUU-3' oligonucleotide. In addition, residues around the central β -strands, as those in the RRM1-RRM2 linker, contribute to additional RNA contacts. Comparison of the solution analysis of the RRM1₁₋₉₉:Cox2 AREI RNA interaction and the crystal structure of the HuR RRM1-2:*c-fos* RNA complex uncovers distinct residues that bind RNA directly. Our results suggest that the mode of binding is dependent on the nucleotide sequence and that each RNA ligand may achieve sequence specificity slightly differently.

NMR relaxation data indicates that the N-terminal extension and the β 2- β 3 loop experience dynamics on the ps-ns timescale. Structural extensions of the RRM fold are often highly dynamic (Zhang and Hall 2017). In addition, the increased flexibility shown by the β 2- β 3 loop seems to be a general feature of RRM domains in the free state (Tsuda et al. 2014; Duviganud et al. 2016; Soufari and Mackereth 2017). Interestingly, RNA binding has no effect in RRM1₁₋₉₉ overall ps-ns dynamics, suggesting that complex formation does not lead to any significant structural ordering. Similar results were found for other RRM-oligonucleotide complexes, including heterogeneous nuclear ribonucleoprotein hnRNP D0 RNA-binding domain 2 (RBD2) interaction with 5'-TTAGGG-3' DNA (Katahira et al. 2001) and

RBM5 RRM2 interaction with 5′-CUCUUC-3′ RNA (Song et al. 2012). Furthermore, binding of the PHD3 finger of MLL to Cyp33 RRM results in RRM stabilization, with no significant changes in the β 2– β 3 loop mobility (Hom et al. 2010). Although it is remarkable that RRM1_{1–99} ^{15}N backbone relaxation parameters showed little change in the RNA-bound state, we cannot rule out that dynamics of other bond vectors, such as those on side chains, might be sensitive to RNA binding.

Our NMR relaxation data indicates that RRM1_{1–99} is mostly monomeric in solution. However, we observed a smaller R_2/R_1 ratio when protein concentration was decreased from 1.5 to 0.25 mM, suggesting that RRM1_{1–99} experiences an equilibrium with a protein species exhibiting either exchange contribution to R_2 due to monomer–dimer equilibrium and/or a larger rotational correlation time. Most frequently, this is caused by nonspecific protein aggregation due to concentration effects, as seen for other RRM domains (Wang et al. 2014). However, by taking advantage of the high sensitivity of ESI–ESI–IMS–MS, we detected a gas-phase RRM1_{1–99} dimer, in the absence of RNA, at protein concentration as low as 50 μM . Moreover, NMR titration experiments identified a protein interface formed by residues located in the two α -helices, β -hairpin and the β 2– β 3 loop as responsible for RRM1_{1–99} dimerization. If dimerization were a nonspecific process, one would expect the CSPs to be distributed throughout the entire protein surface instead of being clustered in one defined region. Therefore, we suggest that RRM1_{1–99} dimerization is indeed specific. Further studies are necessary to evaluate if the observed dimerization is a biologically relevant effect.

Previous studies have indicated the capability of full-length Hu/ELAV proteins to dimerize or oligomerize using electrophoretic-mobility shift assays (Dean et al. 2001) or yeast two-hybrid screening (Kasashima et al. 2002). In addition, isolated RRM3, in its free state, undergoes monomer–dimer equilibrium, as probed by NMR and molecular dynamics simulations. RRM3 exists as a 25%-populated dimer, which is mediated by a specific residue, Trp261, lying on helix α 1 (Sheiba et al. 2014; Dean et al. 2001). Furthermore, other RRMs have been shown to experience a monomer–dimer equilibrium in the absence of RNA, including *Citrus sinensis* PABPN1 RRM (Díaz-Quintana and García-Mauriño 2015) and CPEB4 RRM1–2 tandem domains (Domingues et al. 2015). This highlights a general ability of RRM domains to self-associate through homotypic interactions.

As the central β -sheet has been implicated in RNA recognition, the opposite, α -helical surface has been described as a putative protein–protein interaction site. For example, residues in helix α 1 and the α 2– β 4 loop mediate interactions between the C-terminal RRM3 and RRM4 domains of hnRNPL (Schelhorn et al. 2014). In addition, helix α 2

of RRM4 interacts with a hydrophobic pocket formed by helices α 1 and α 2 of RRM3 in polypyrimidine tract binding protein (PTB) (Blatter et al. 2015). Mapping of the CSPs observed upon increase in protein concentration on the RRM1_{1–99} structure reveals a cleft formed by hydrophobic side chains located in helices α 1 and α 2, and the β -hairpin, including Leu39, Ile 43, Leu81, and Leu86, which are conserved among RRM domains. Interestingly, this hydrophobic cleft has been shown to participate in the interaction of RRM domains with a variety of peptide ligands, including the interaction between SF3B49 RRM and a SF3b145-derived peptide (Joshi et al. 2014). At present, our data cannot precise the orientation of the two domains in the dimer structure. The crystal structure of HuR RRM1 reveals a tetramer in the asymmetric unit, with two possible dimer conformations (Benoit et al. 2010). Both dimeric interfaces are formed by the α -helices; however, none of them engages the β 2– β 3 loop. Therefore, the CSPs observed for the β 2– β 3 loop upon protein concentration suggest a dimerization mode in which the loop of one monomer make contacts with the hydrophobic cleft of another monomer, in a manner similar to RRM-peptide interactions, including PTPB RRM2–Raver 1 (Kuwasako et al. 2017) and RNSP1 RRM–Acinus (Joshi et al. 2011). It is worth mentioning that dimerization of the RRM domain of the splicing factor Pub60 is mediated by the flexible β 2– β 3 loop (Murachelli et al. 2012).

Curiously, RRM1_{1–99} dimer does not bind a single-stranded DNA oligonucleotide of sequence 5′-TTTAATT-3′, related to Cox2 ARE I. A dimer–monomer transition upon RNA binding has been described for other RRM-containing proteins. Poly(A) interaction induces dimer dissociation of *Xenopus laevis* XlePABP2 (Corsini et al. 2009) as well as *Citrus sinensis* PABPN1 (Song et al. 2008). Based on these results, we suggest that oligonucleotide binding stabilizes RRM1_{1–99} monomer, leading to dissociation of the low-populated dimeric species. It is noteworthy that the RRM1_{1–99} dimer interface partially overlaps with the RNA-binding interface at the N-terminal region of strand β 4, suggesting a possible mechanism for RNA-induced dimer dissociation in which RNA interaction competes for a partially overlapping binding site. Further investigations using full-length HuR will be necessary to confirm this hypothesis.

Conclusion

In conclusion, we used NMR to solve the structure, characterize the dynamics and the binding properties of RRM1_{1–99} in solution. Our data show that RRM1_{1–99} contains an N-terminal, highly flexible, intrinsically disordered extension, which is not involved in RNA binding. In addition, the β 2– β 3 loop undergoes fast backbone dynamics, as evidenced by ^{15}N relaxation measurements. By employing a combination of

NMR and ESI–IMS–MS, we show that RRM1_{1–99} dimerizes; however, dimer population represents a minority. We identified helices $\alpha 1$ and $\alpha 2$, the β -hairpin and the $\beta 2$ – $\beta 3$ loop as critical regions for RRM1_{1–99} dimerization. Remarkably, ligand binding shifts RRM1_{1–99} oligomeric equilibrium toward the monomeric species, which is the one capable of binding oligonucleotide.

Acknowledgements We would like to thank the staff of the UEMP-IBqM-UFRJ (Proteomic/Mass Spectrometry Platform) and the CNRMN-UFRJ (NMR Platform) for providing access to their facilities and excellent technical support. This work was supported by grants from *Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro* (FAPERJ), *Conselho Nacional Científico e Tecnológico* (CNPq), and by a Brazil Initiative Collaboration Grant from Brown University to ASP. CL is recipient of a *Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro* (FAPERJ) “outstanding student” graduate fellowship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Afroz T, Cienikova Z, Cléry A, Allain FHT (2015) One, two, three, four! How multiple RRM domains read the genome sequence. *Methods Enzymol* 558:235–278
- Barbato G, Ikura M, Kay LE, Pastor RW, Bax A (1992) Backbone dynamics of calmodulin studied by ¹⁵N relaxation using inverse detected two-dimensional NMR spectroscopy: the central helix is flexible. *Biochemistry* 31:5269–5278
- Barreau C, Paillard L, Osborne HB (2006) AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Res* 33:7138–7150
- Benoit RM, Meisner NC, Kallen J, Graff P, Hemmig R, Cèbe R, Ostermeier C, Widmer H, Auer M (2010) The X-ray crystal structure of the first RNA recognition motif and site-directed mutagenesis suggest a possible HuR redox sensing mechanism. *J Mol Biol* 397:1231–1244
- Bevilacqua A, Ceriani MC, Capaccioli S, Nicolini A (2003) Post-transcriptional regulation of gene expression by degradation of messenger RNAs. *J Cell Physiol* 195:356–372
- Blanco FF, Preet R, Aguado A, Vishwakarma V, Stevens LE, Vyas A, Padhye S, Xu L, Weir SJ, Anant S, Meisner-Kober N, Brody JR, Dizon DA (2016) Impact of HuR inhibition by the small molecule MS-444 on colorectal cancer cell tumorigenesis. *Oncotarget* 7:74043–74058
- Blatter M, Dunin-Horkawicz S, Grishina I, Maris C, Thore S, Maier T, Bindereif A, Bujnicki JM, Allain FHT (2015) The signature of the five-stranded vRRM fold defines functional, structural and computational analysis of the hnRNP L protein. *J Mol Biol* 427:3001–3022
- Blaxall BC, Dwyer-Nield LD, Bauer AK, Bohlmeier TJ, Malkinson AM, Port JD (2000) Differential expression and localization of the mRNA binding proteins, AU-rich element mRNA binding protein (AUF1) and Hu antigen R (HuR), in neoplastic lung tissue. *Mol Carcinog* 28:76–83
- Brennan CM, Steitz JA (2001) HuR and mRNA stability. *Cell Mol Life Sci* 58:266–277 (2001)
- Brunger AT (2007) Version 1.2 of the crystallographic and NMR system. *Nat Protoc* 2:2728–2733
- Buchan JR, Parker R (2009) Eukaryotic stress granules: the ins and outs of translation. *Mol Cell* 36:932–941
- Cléry A, Blatter M, Allain FHT (2008) RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* 18:290–298
- Cléry A, Jayne S, Benderska N, Dominguez C, Stamm S, Allain FHT (2011) Molecular basis of purine-rich RNA recognition by the human SR-like protein Tra2-beta 1. *Nat Struct Mol Biol* 18:443–450
- Cléry A, Sinha R, Anczuków O, Corriero A, Moursy A, Daubner GM, Valcárcel J, Krainer AR, Allain FHT (2013) Isolated pseudo-RRM-recognition motifs of SR proteins can regulate splicing using a noncanonical mode of RNA recognition. *Proc Natl Acad Sci USA* 110:E2802–E2811
- Corsini L, Hothorn M, Stier G, Rybin V, Scheffzek K, Gibson TJ, Sattler M (2009) Dimerization and protein binding specificity of the U2AF homology motif of the splicing factor Pub60. *J Biol Chem* 284:630–639
- Dean JLE, Wait R, Mahtani KR, Sully G, Clark AR, Sklatvala J (2001) The 3'-untranslated region of tumor necrosis factor alpha mRNA is a target of the mRNA-stabilizing factor HuR. *Mol Cell Biol* 21:721–730
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293
- DeLano WL (2002) The PyMol molecular graphics system. <http://www.pymol.org>
- Denkert C, Weichert W, Winzer KJ, Muller BM, Noske A, Niesporek A, Kristiansen G, Guski H, Dietel M, Hauptmann S (2004a) Expression of the ELAV-like protein HuR is associated with tumor grade and increased cyclooxygenase-2 expression in human breast carcinoma. *Clin Cancer Res* 64:5580–5586
- Denkert C, Weichert W, Pest S, Koch I, Licht D, Köbel M, Reles A, Schouli J, Dietel M, Hauptmann S (2004b) Overexpression of the embryonic-lethal abnormal vision-like protein HuR in ovarian carcinoma is a prognostic factor and is associated with increased cyclooxygenase 2 expression. *Cancer Res* 64:189–195
- Díaz-Quintana A, García-Mauriño SM, Díaz-Moreno I (2015) Dimerization model of the C-terminal RNA recognition motif of HuR. *FEBS Lett* 589:1059–1066
- Doller A, Pfeilschifter J, Eberhardt W (2008) Signalling pathways regulating nucleocytoplasmic shuttling of the mRNA-binding protein HuR. *Cell Signal* 20:2165–2173
- Domingues MN, Sforça ML, Soprano AS, Lee J, de Souza TACB, Cassago A, Portugal RV, de Mattos Zeri AC, Murakami MT, Sadanandom A, de Oliveira PSL, Benedetti CE (2015) Structure and mechanism of dimer-monomer transition of a plant poly(A)-binding protein upon RNA interaction: insights into its poly(A) tail assembly. *J Mol Biol* 427:2491–2506
- Duviganud JB, Bédard M, Nagata T, Muto Y, Yokoyama S, Gagné SM, Vincent M (2016) Structure, dynamics, and interaction of p54(nrb)/NonO RRM1 with r' splice site RNA sequence. *Biochemistry* 10:2553–2566
- Eichhorn CD, Chug R, Feigon J (2016) hLARP7 C-terminal domain contains an xRRM that binds the 3' hairpin of 7SK RNA. *Nucleic Acids Res* 44:9977–9989
- Fan XC, Steitz JA (1999) Overexpression of HuR, a nuclear-cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs. *EMBO J* 17:3448–3460
- Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, Gish G, Shoelson SE, Pawson T, Forman-Kay JD, Kay LE (1994) Backbone dynamics of a free and a phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* 33:5984–6003

- Giaginis C, Alexandrou P, Tsoukalas N, Sfiniadakis I, Kavantzis N, Agapitos E, Patsouris E, Theocharis S (2015) Hu-antigen receptor (HuR) and cyclooxygenase-2 (COX-2) expression in human non-small-cell lung carcinoma: associations with clinicopathological parameters, tumor proliferative capacity and patient's survival. *Tumour Biol* 36:315–327
- Glisovic T, Bachorik JL, Yong J, Dreyfuss G (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 582:1977–1986
- Hinman MN, Lou H (2008) Diverse molecular functions of Hu proteins. *Cell Mol Life Sci* 65:3168–3181
- Hom RA, Chang PY, Roy S, Musselman CA, Glass KC, Selezneva AI, Gozani O, Ismagilov RF, Cleary ML, Kutateladze TG (2010) Molecular mechanism of MLL PHD3 and RNA recognition by the Cyp33 RRM domain. *J Mol Biol* 400:145–154
- Izquierdo JM (2008) Hu antigen R (HuR) functions as an alternative pre-mRNA splicing regulator of Fas apoptosis-promoting receptor on exon definition. *J Biol Chem* 283:19077–19084
- Joshi A, Coelho MB, Kotik-kohan O, Simpson PJ, Matthews SJ, Smith CW, Curry S (2011) Crystallographic analysis of polypyrimidine tract-binding protein-Raver 1 interactions involved in regulation of alternative splicing. *Structure* 19:1816–1825
- Joshi A, Esteve V, Buckroyd AN, Blatter M, Allain FHT, Curry S (2014) Solution and crystal structures of a C-terminal fragment of the neuronal isoform of the polypyrimidine tract binding protein (nPTB). *Peer J* 2:e305
- Kasashima K, Sakashita E, Saito K, Sakamoto H (2002) Complex formation of the neuron-specific ELAV-like Hu RNA binding proteins. *Nucleic Acids Res* 30:4519–4526
- Katahira M, Miyanoiri Y, Enokizono Y, Matsuda G, Nagata T, Ishikawa F, Usegi S (2001) Structure of the C-terminal RNA-binding domain of hnRNP D0 (AUF1), its interactions with RNA and DNA, and change in backbone dynamics upon complex formation with DNA. *J Mol Biol* 311:973–988
- Kay LE, Torchia DA, Bax A (1989) Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28:8972–8979
- Kim HS, Headey SJ, Yoga YM, Scanlon MJ, Gorospe M, Wilce MC, Wilce JA (2013) Distinct binding properties of TIAR RRMs and linker region. *RNA Biol* 10:579–589
- Koradi R, Billeter M, Wüthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14:51–55
- Kuwasako K, Nameki N, Tsuda K, Takahashi M, Sato A, Tochio N, Inoue M, Terada T, Kigawa T, Kobayashi N, Shirouzu M, Ito T, Sakamoto T, Wakamatsu K, Güntert P, Takahashi S, Yokoyama S, Muto Y (2017) Solution structure of the first RNA recognition motif domain of human spliceosomal protein SF3b49 and its mode of interaction with SF3b145 fragment. *Protein Sci* 26:280–291
- Linge JP, Williams MA, Spronk CA, Bonvin AM, Nilges M (2003) Refinement of protein structures in explicit solvent. *Proteins* 50:496–506
- Linger JP, O'Donoghue SI, Nilges M (2001) Automated assignment of ambiguous overhauser effects with ARIA. *Methods Enzymol* 339:71–90
- López de Silanes I, Lal A, Gorospe M (2005) HuR: post-transcriptional paths to malignancy. *RNA Biol* 2:11–13
- López de Silanes I, Fan J, Yang X, Zonderman AB, Potapova O, Pizer ES, Gorospe M (2003) Role of the RNA-binding platform HuR in colon carcinogenesis. *Oncogene* 22:7146–7154
- Meisner NC, Hintersteiner M, Mueller K, Bauer R, Seifert JM, Naegele HU, Ottl J, Oberer L, Guenat C, Moss S, Harrer N, Woitschlaeger M, Buehler C, Uhl V, Auer M (2007) Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nat Chem Biol* 3:508–515
- Meisner NC, Hintersteiner M, Seifert JM, Bauer R, Benoit RM, Widmer A, Schindler T, Uhl V, Lang M, Gstach H, Auer M (2009) Terminal adenosyl transferase activity of posttranscriptional regulator HuR revealed by confocal on-bead screening. *J Mol Biol* 386:435–450
- Migone F, Gissi C, Liuni S, Pesole G (2002) Untranslated regions of mRNAs. *Genome Biol* 3:reviews0004.1-0004.10
- Moore M (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science* 309:1514–1518
- Mujo A, Lixa C, Carneiro LA, Anobom CD, Almeida FC, Pinheiro, A (2015) S. ^1H , ^{15}N and ^{13}C resonance assignments of the RRM1 domain of the key post-transcriptional regulator HuR. *Biomol NMR Assign* 9:281–284
- Murachelli AG, Ebert J, Basquin C, Le Hir H, Conti E (2012) The structure of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex. *Nat Struct Mol Biol* 19:378–386 (2012)
- Myer VE, Fan XC, Steitz JA (1997) Identification of HuR as a protein implicated in AUUUA-mediated mRNA decay. *EMBO J* 16:2130–2139
- Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE, Nilges M (2007) ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* 23:381–382
- Samatanga B, Cléry A, Barraud P, Allain FHT, Jelesarov I (2017) Comparative analysis of the thermodynamics RNA binding signatures of different types of RNA recognition motifs. *Nucleic Acids Res* 45:6037–6050
- Schelhorn C, Gordon JM, Ruiz L, Alguacil J, Pedrosa E, Macias MJ (2014) RNA recognition and self-association of CPEB4 is mediated by its tandem RRM domains. *Nucleic Acids Res* 42:10185–10195
- Sheiba RM, de Opakua AI, Díaz-Quintana A, Cruz-Gallardo I, Martínez-Cruz LA, Martínez-Chantar ML, Blanco FJ, Díaz-Moreno I (2014) The C-terminal binding motif of HuR is a multi-functional domain leading to HuR oligomerization and binding to U-rich RNA targets. *RNA Biol* 11:1250–1261
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241
- Singh M, Wang Z, Koo BK, Patel A, Cascio D, Collins K, Feigon J (2012) Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* 47:16–26
- Singh M, Choi CP, Feigon J (2013) xRRM: a new class of RRM found in the telomerase La family protein p65. *RNA Biol* 10:353–359
- Song J, McGivern JV, Nichols KW, Markley JL, Sheets MD (2008) Structural basis for RNA recognition by a type II poly(A)-binding protein. *Proc Natl Acad Sci USA* 105:15317–15322
- Song Z, Wu P, Ji P, Zhang J, Gong Q, Wu J, Shi Y (2012) Solution structure of the second RRM domain of RBM5 and its unusual binding characters for different RNA targets. *Biochemistry* 51:6667–6678
- Soufari H, Mackereth CD (2017) Conserved binding of GCAC motifs by MEC-8, couch potato, and the RBPMs protein family. *RNA* 23:308–316
- Stone MJ, Fairbrother WJ, Palmer III, Reizer AG, Saier J Jr, M. H. & Wright PE (1992) Backbone dynamics of the Bacillus subtilis glucose permease IIA domain determined from ^{15}N NMR relaxation measurements. *Biochemistry* 31:4394–4406
- Tsuda K, Kuwasako K, Nagata T, Takahashi M, Kigawa T, Kobayashi N, Güntert P, Shirouzu M, Yokoyama S, Muto Y (2014) Novel RNA recognition motif domain in the cytoplasmic polyadenylation element binding protein 3. *Proteins* 82:2879–2886
- Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue E (2005) D. The CCPN

- data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59:687–696
- Wang H, Zeng F, Liu Q, Liu H, Liu Z, Niu L, Teng M, Li X (2013) The structure of the ARE-binding domains of Hu antigen R (HuR) undergoes conformational changes during RNA binding. *Acta Crystallogr D* 69:373–380
- Wang I, Hennig J, Jagtap PK, Sonntag M, Valcárcel J, Sattler M (2014) Structure, dynamics, and RNA binding of the multi-domain splicing factor TIA-1. *Nucleic Acids Res* 42:5949–5966
- Williams JP, Lough JA, Campuzano I, Richardson K, Sadler PJ (2009) Use of ion mobility mass spectrometry and a collision cross-section algorithm to study organometallic ruthenium anticancer complex and its adducts with a DNA oligonucleotide. *Rapid Commun Mass Spectrom* 23:3563–3569
- Wilusz CJ, Wilusz J (2004) Bringing the role of mRNA decay in the control of gene expression into focus. *Trends Genet* 20:491–497
- Zhang J, Gonzalez, Hall TMT (2017) Structural analysis reveals the flexible C-terminus of Nop15 undergoes rearrangement to recognize a pre-ribosomal RNA folding intermediate. *Nucleic Acids Res* 45:2829–2837
- Zhu H, Zhou HL, Hasman RA, Lou H (2007) Hu proteins regulate polyadenylation by blocking sites containing U-rich sequences. *J Biol Chem* 282:2203–2210
- Zhu Z, Wang B, Bi J, Zhang C, Guo Y, Chu H, Liang X, Zhong C, Wang J (2013) Cytoplasmic HuR expression correlates with P-gp, HER-2 positivity, and poor outcome in breast cancer. *Tumour Biol* 34:2299–2308