





Structural and dynamic studies of the human RNA binding protein RBM3 reveals the molecular basis of its oligomerization and RNA recognition

Sayantani Roy¹ , Soumendu Boral¹ , Snigdha Maiti¹, Tushar Kushwaha², Aditya J. Basak¹, Woonghee Lee³ , Amit Basak^{1,4}, Shivajirao L. Gholap⁵, Krishna K. Inampudi² and Soumya De¹ 

¹ School of Bioscience, Indian Institute of Technology Kharagpur, India

² Department of Biophysics, All India Institute of Medical Sciences, New Delhi, India

³ Department of Chemistry, University of Colorado Denver, CO, USA

⁴ Department of Chemistry, Indian Institute of Technology Kharagpur, India

⁵ Department of Chemistry, Indian Institute of Technology Delhi, New Delhi, India

Keywords

cold-shock protein; molecular dynamics simulation; NMR relaxation; protein–RNA interaction; RRM domain

Correspondence

S. Roy and S. De, School of Bioscience, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India
 Tel: +91-8116158624 (SR); +91 (03222) 284552 (SD)

E-mail: sayantani23@gmail.com (SR);
 somde@iitkgp.ac.in (SD)

Website: <http://www.iitkgp.ac.in/departments/BS/faculty/bs-somde> (SD)

Sayantani Roy and Soumendu Boral contributed equally to this work.

(Received 30 March 2021, revised 30 September 2021, accepted 25 November 2021)

doi:10.1111/febs.16301

Human RNA-binding motif 3 protein (RBM3) is a cold-shock protein which functions in various aspects of global protein synthesis, cell proliferation and apoptosis by interacting with the components of basal translational machinery. RBM3 plays important roles in tumour progression and cancer metastasis, and also has been shown to be involved in neuroprotection and endoplasmic reticulum stress response. Here, we have solved the solution NMR structure of the N-terminal 84 residue RNA recognition motif (RRM) of RBM3. The remaining residues are rich in RGG and YGG motifs and are disordered. The RRM domain adopts a $\beta\alpha\beta\alpha\beta$ topology, which is found in many RNA-binding proteins. NMR-monitored titration experiments and molecular dynamic simulations show that the beta-sheet and two loops form the RNA-binding interface. Hydrogen bond, pi–pi and pi–cation are the key interactions between the RNA and the RRM domain. NMR, size exclusion chromatography and chemical cross-linking experiments show that RBM3 forms oligomers in solution, which is favoured by decrease in temperature, thus, potentially linking it to its function as a cold-shock protein. Temperature-dependent NMR studies revealed that oligomerization of the RRM domain occurs via nonspecific interactions. Overall, this study provides the detailed structural analysis of RRM domain of RBM3, its interaction with RNA and the molecular basis of its temperature-dependent oligomerization.

Introduction

RNA-binding proteins (RBPs), as the name suggests, are a class of proteins which form complexes with nucleic acids, such as microRNAs, messenger RNAs, small inducible RNAs, small nuclear RNAs, small nucleolar RNAs, transfer RNAs and noncoding RNAs [1]. They can also, on occasion, bind with single

stranded DNA [2]. Approximately 7.5% of the human protein-coding genome consists of RBPs [1]. RBPs play key roles in pre-mRNA splicing, editing, polyadenylation, localization, turnover, translation and degradation [3,4]. As a result of their involvement in such critical cellular functions, RBPs are tightly

Abbreviations

HSQC, heteronuclear single quantum correlation; RBP, RNA-binding protein; RRM, RNA recognition motif; SEC, size exclusion chromatography.

regulated and their dysregulation, mutations or mislocalization are implicated in several diseases. RBPs, such as TDP-43, Matrin-3 and FUS, are speculated to be involved in amyotrophic lateral sclerosis (Lou Gehrig's disease), a fatal neurodegenerative disease [5,6]. RBPs have also been linked to Alzheimer's disease [7], Parkinson's disease [8], cancers [9,10], and cardiovascular diseases [11]. In recent years, RBPs have emerged as important therapeutic targets [12]; thus, renewing efforts in elucidating the molecular mechanisms of their functions.

RNA-binding proteins interact with nucleic acids via one or more domains [1]. There are several types of domains that recognize RNA, such as RNA-binding domain (RBD) also known as ribonucleoprotein (RNP) domain or RNA recognition motif (RRM), double-stranded RNA-binding domain (dsRBD), K-homology (KH) domain (type I and type II), zinc finger (ZnF, mostly C-x8-X-x5-X-x3-H), Sm domain, DEAD/DEAH box, cold-shock domain, Pumilio/FBF (PUF or Pum-HD) domain and the Piwi/Argonaute/Zwille (PAZ) domain [1]. One such domain can bind 2–6 nucleotides. Higher copies of these domains facilitate binding of more complex nucleotide sequences and add to sequence specificity. The RRM domain is the most widespread in higher organisms [13]. Apart from these folded domains, glycine–arginine-rich (GAR/RGG/RG/GRG) motifs are often found as auxiliary regions in several human RBPs [14]. These motifs are intrinsically disordered, bind RNA and have been implicated in several diseases, including neurodegenerative and neuromuscular diseases, and cancer [15,16].

The RNA-binding motif protein 3 (RBM3) belongs to the glycine–arginine (GR)-rich RBP family protein with one N-terminal RNA-binding domain (RRM) and highly flexible C-terminal GR-rich region [17]. Because of one RRM domain, RBM3 belongs to the class IVa-GRP subfamily, members of which contain up to three RRM domains [18]. Human RBM3 is expressed in the X chromosome [19], and is reported to have increased expression in conditions of cold stress [20]. Originally discovered as a cold-shock protein, recent discoveries indicate that it also has cytoprotective functions under diverse stress conditions, such as hypoxia [21] and infection-induced fever [22]. RBM3 plays important roles in the basal translational machinery, influencing global protein synthesis, cell proliferation and acts as a modulator of apoptosis. RBM3 is also implicated in tumour progression, cancer metastasis as well as in neuroprotection [18].

In this study, we have characterized the structure and nucleic acid binding function of two truncated RBM3 constructs RBM3^{1–84} and RBM3^{1–110}, which

contain the minimal RRM domain and sequences containing several arginine–glycine (RG) motifs appended to the RRM domain respectively. We report the atomic structure of the 84-residue RRM domain of human RBM3, as determined by solution NMR spectroscopy. The RRM domain adopts the typical RRM fold consisting of a beta-sheet packed against two alpha helices. By NMR-based titration experiments, we have identified the RNA-binding surface of RBM3 and presented an RNA-docked model of the protein. A 500-ns molecular dynamics simulation of the protein–RNA complex reveals the important features of RNA recognition by the RRM domain. Furthermore, we show that both RBM3^{1–84} and RBM3^{1–110} form dimer, trimer and higher oligomers in solution. RBM3 oligomerization is favoured with decrease in temperature. Overall, this study presents the structural basis of RBM3 function and should provide the necessary molecular framework for investigation of its function at the cellular and organismal levels.

Results

Identification of the RRM domain in RBM3

Based on sequence analysis using InterPro [23] and sequence alignment using MultAlin [24] of several RNA-binding domains (Fig. 1A), the 1–84 N-terminal residues of RBM3 were found to contain an RRM domain. The remaining C-terminal residues were predicted to be disordered by Rosetta Residue Disordered Prediction [25] and DisProt [26]. The disordered part of RBM3 is rich in arginine, glycine and tyrosine residues and contains several RGG/RG motifs, which are known to facilitate binding to nucleic acids [14]. The complete RBM3 protein (157 residues) is highly aggregation prone and forms inclusion bodies after overexpression in bacteria. The inclusion bodies were purified using established protocol [27] and solubilized in guanidinium solution but did not refold. Hence, two truncated constructs, RBM3^{1–84} and RBM3^{1–110}, were designed where RBM3^{1–84} should contain the RRM domain and RBM3^{1–110} consists of additional 26 amino acid residues from the disordered region containing several RGG/RG motifs. RBM3^{1–84} is highly soluble and conducive to structural studies while RBM3^{1–110} is less stable in solution.

In order to characterize the RBM3 structure at atomic resolution, we used solution NMR spectroscopy. Two-dimensional ¹H–¹⁵N heteronuclear single-quantum coherence (HSQC) spectrum of ¹⁵N-labeled RBM3^{1–84} has well-dispersed cross-peaks, indicating a properly folded domain (Fig. 1B). The

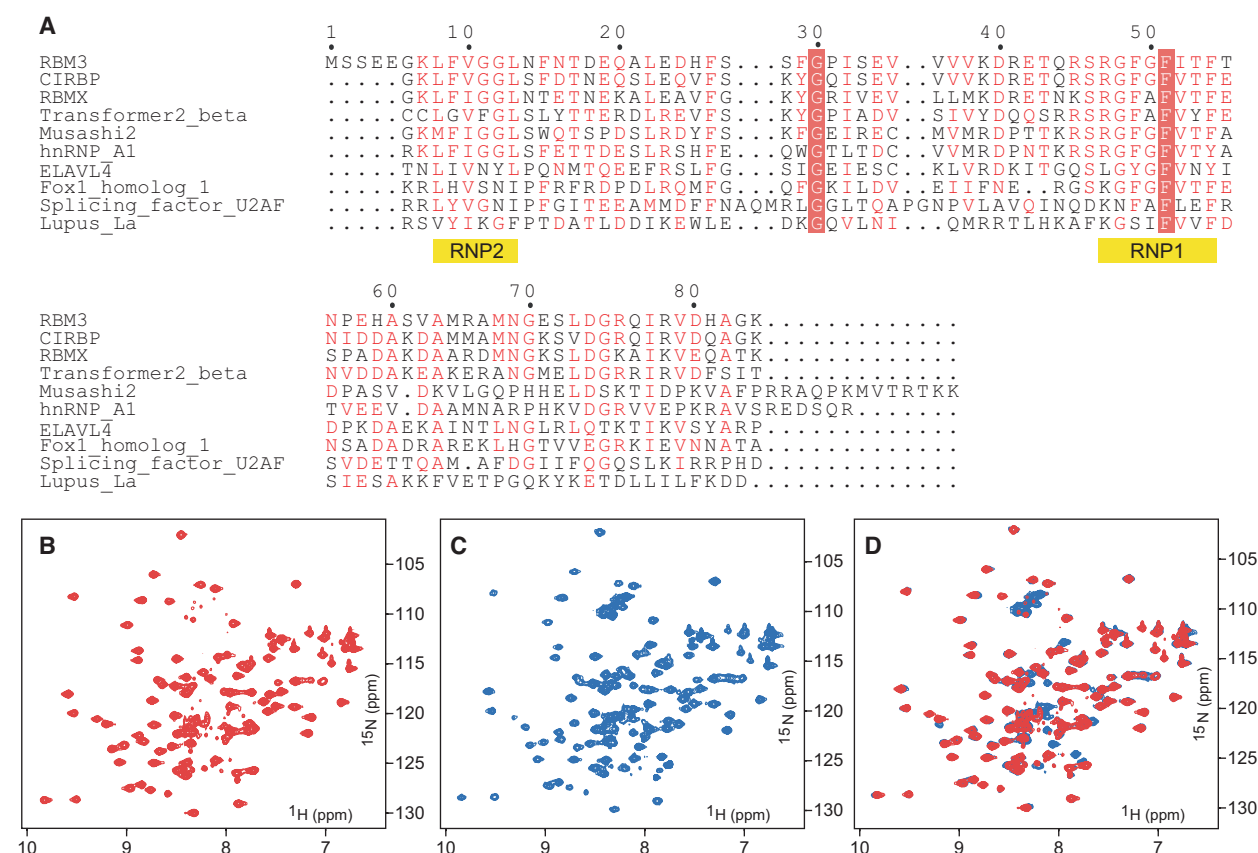


Fig. 1. Determination of the folded domain in RBM3. (A) The sequence of RBM3 is aligned with other RRM domains. The semi-conserved residues are in red and conserved are highlighted in red in this alignment. The RNP1 and RNP2 motifs are also shown. ^1H - ^{15}N -HSQC spectra of (B) RBM3¹⁻⁸⁴, the minimal RRM domain and (C) RBM3¹⁻¹¹⁰, the RRM domain with RG/RGG motif sequences. (D) Overlay of the two spectra shows perfect overlap of the well-dispersed peaks. There are extra peaks at the centre for RBM3¹⁻¹¹⁰ (blue).

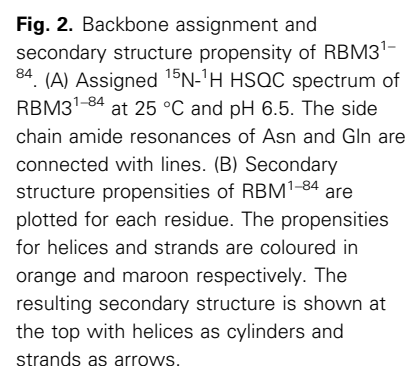
^1H - ^{15}N HSQC spectra of RBM3¹⁻¹¹⁰ had a few strong signals in the centre of the spectra, which are characteristic of an unfolded protein sequence (Fig. 1C). The additional peaks from the glycine residues in the RG/RGG motifs can be clearly identified. Overlay of the two ^1H - ^{15}N HSQC spectra shows near-perfect overlap of the well-dispersed peaks from the structured regions (Fig. 1D). This indicates that RBM3 consists of a folded domain in its N terminus and an unstructured region in C terminus, and the truncation of RBM3 did not alter the structure of its N-terminal folded domain.

Solution structure of the RRM domain of RBM3

A ^{13}C and ^{15}N -labeled RBM3¹⁻⁸⁴ sample was prepared to assign its backbone and sidechain atoms. Of 82 nonproline residues, 80 backbone amides were unambiguously assigned (Fig. 2A). The diastereotopic methyl groups present in valines and leucines were

assigned using a fractional (10%) ^{13}C -labeled sample [28]. The chemical shifts of backbone $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, ^{13}CO , $^1\text{H}^\alpha$, $^1\text{H}^\text{N}$ and ^{15}N were used to calculate the secondary structure propensity with MICS [29]. Residues K7 to H81 are folded and contain secondary structures consisting of two α -helices and four β -strands (Fig. 2B), which is consistent with the RRM fold [13]. The ^1H -, ^{15}N - and ^{13}C -assigned chemical shifts of RBM3¹⁻⁸⁴ have been deposited in the BMRB (<http://www.bmr.io/>) under the accession number 50804 [30].

The structural ensemble of RBM3¹⁻⁸⁴ was determined with PONDEROSA-C/S coupled with Xplor-NIH using established protocols [27,31]. The distance and dihedral angle restraints were used for structure calculation. In the final step, 200 structures were calculated, from which the statistics of the 20 lowest energy structures are summarized in Table 1. These structures align with a RMSD of 0.22 Å for backbone atoms and 0.96 Å for all heavy atoms across all ordered residues (Fig. 3A). On average, 14.5 distance restraints



Such N-capping boxes have been observed in several RBPs [35–38]. The N-capping boxes are formed by the residues D18 and N56 for the $\alpha 1$ and $\alpha 2$ helices respectively. RBM3^{1–84} also has three long loops, a 6-residue loop between $\beta 1$ and $\alpha 1$, a 7-residue loop between $\beta 2$ and $\beta 3$ and another 7-residue loop between $\alpha 2$ and $\beta 4$. Interestingly, the $\beta 1$ - $\alpha 1$ and $\alpha 2$ - $\beta 4$ loops, when seen in the ensemble structure, show tight conformation, indicating low flexibility of these long loops (Fig. 3).

Backbone ^{15}N relaxation measurement of RBM3

To characterize the global hydrodynamic and local dynamic properties of RBM3, amide ^{15}N R_1 , R_2 and heteronuclear NOE relaxation data were collected for RBM3 $^{1-84}$ and RBM3 $^{1-110}$ at 25 and 35 °C. Excellent fits were obtained for all sets of data except for RBM3 $^{1-110}$ at 25 °C. Heteronuclear $\{^1\text{H}\}\text{-}^{15}\text{N}$ NOE values are a sensitive measure of the fast-internal backbone motions of proteins in the picosecond to nanosecond timescale [39]. The heteronuclear NOE values of residues in helices and sheets are uniformly high with an average of 0.74 ± 0.05 at 25 °C and

Table 1. Restraints and statistics for 20 best solution NMR structures of RBM3^{1–84}. Structure quality was evaluated using wwPDB validation pipeline (wwPDB-VP: 2.17.1) and PSVS v1.5 [62].

Parameters	Value
a) Distance and dihedral restraints	
Distance restraints	
Total number of NOE	1214
Short range ($ i-j \leq 1$)	605
Medium range ($1 < i-j < 5$)	156
Long range ($ i-j \geq 5$)	453
Hydrogen bond restraints	50
Dihedral angle restraints	
Total	142
Phi (ϕ)	71
Psi (ψ)	71
Total number of restricting constraints	1406
(b) Average RMSD (Å) against the lowest energy model for ordered residues ^a	
Backbone atoms (N, C $^{\alpha}$, C, O)	0.224 \pm 0.05
All heavy atoms	0.965 \pm 0.13
(c) RMSD from ideal geometry	
For bond lengths (Å)	0.012
For bond angles ($^{\circ}$)	1.4
(d) Consistent violations	
Distance constraints (> 0.5 Å)	0
Dihedral angle constraints ($> 5^{\circ}$)	0
Van der Waals constraints (> 0.2 Å)	0
(e) Ramachandran plot summary (%) from PROCHECK for selected residues ^b	
Residues in most favoured regions	91.5
Residues in additionally allowed regions	8.5
Residues in generously allowed regions	0
Residues in disallowed regions	0

^aOrdered CYRANGE [74] residues: 6–35, 46–84.; ^bSelected residues: 5–83.

0.68 \pm 0.06 at 35 $^{\circ}$ C for RBM3^{1–84} and 0.71 \pm 0.07 at 35 $^{\circ}$ C for RBM3^{1–110}, which demonstrates a well-ordered domain (Fig. 4A). The loops β 1– α 1 and α 2– β 4 have similarly high NOE values indicating rigid conformation, which is consistent with the structural ensemble (Fig. 3A). The β 2– β 3 loop is relatively flexible, again consistent with the determined structural ensemble.

The transverse relaxation rate constants (R_2) for the folded domain in RBM3^{1–84} have average values of 12.3 \pm 1.8 s^{–1} and 14.1 \pm 2.5 s^{–1} at 25 and 35 $^{\circ}$ C respectively. The folded domain in RBM3^{1–110} has an average R_2 of 17.1 \pm 3.0 s^{–1} at 35 $^{\circ}$ C (Fig. 4B). The longitudinal relaxation rate constants (R_1) show much less variation (Fig. 4C). The average R_1 value for RBM3^{1–84} of 1.7 \pm 0.1 s^{–1} at 25 $^{\circ}$ C increased to 2.5 \pm 0.2 s^{–1} at 35 $^{\circ}$ C. RBM3^{1–110} has an average R_1 of 2.0 \pm 0.2 s^{–1} at 35 $^{\circ}$ C. Collectively, these data

(Tables S1–S3) indicate that the RRM domain is properly folded and has a stable three-dimensional structure.

RBM3 forms oligomers in solution, favoured by lowering of temperature

During sample preparation, it was observed that both RBM3 constructs behave well at room temperature but tend to aggregate at lower temperatures, RBM3^{1–110} more drastically than RBM3^{1–84} (Fig. S1). In order to get a better insight into this temperature-dependent behaviour, the ¹⁵N relaxation data collected at 25 and 35 $^{\circ}$ C were used to determine the isotropic rotational correlation time (τ_c), which is proportional to the mass of a protein and, hence, reports on its oligomeric state in solution. Using the R_2/R_1 ratios for well-ordered amides in the RRM domain, the correlation times (τ_c) were determined to be 7.84 \pm 0.04 ns at 25 $^{\circ}$ C and 6.52 \pm 0.06 ns at 35 $^{\circ}$ C for RBM3^{1–84} and 8.77 \pm 0.06 ns at 35 $^{\circ}$ C for RBM3^{1–110}. The expected correlation times (τ_c) for the monomeric proteins are 5.41 ns at 25 $^{\circ}$ C and 5.23 ns at 35 $^{\circ}$ C for RBM3^{1–84} and 6.67 ns at 35 $^{\circ}$ C for RBM3^{1–110} [40]. Interestingly, NMR dynamic studies on other RRM domains also report higher than expected correlation times [35,41].

The determined correlation times are in between those expected for monomeric and dimeric proteins. Hence, it can be postulated that both RBM3^{1–84} and RBM3^{1–110} exist in an equilibrium of monomeric and dimeric forms resulting in an intermediate correlation time. However, higher order oligomeric forms, at lower populations, could not be ruled out. Also, a 17% decrease in correlation time for RBM3^{1–84} as the temperature is increased as opposed to an expected 3% decrease indicates that increase in temperature disfavors oligomerization.

Since the monomer and the oligomers are in dynamic equilibrium, the surface residues of the RRM domain involved in oligomerization should experience chemical exchange resulting in their enhanced R_2 relaxation and chemical shift perturbation. The product of the relaxation rate constants R_1R_2 recognizes residues with chemical exchange without the influence of anisotropic motion [42]. The surface residues 15, 16, 28, 41, 67, 69, 73 and 74 in the RRM domain show significantly enhanced R_1R_2 compared to the average of the domain (Fig. 4D). Furthermore, ¹H–¹⁵N HSQC spectra were collected for both RBM3^{1–84} and RBM3^{1–110} at 35, 30, 25 and 20 $^{\circ}$ C. Greater than average chemical shift perturbations between the temperatures 20 and 35 $^{\circ}$ C were observed for the surface residues 15, 16, 18, 20, 28, 35, 40, 41, 55 and 72 (Fig. 4E). The chemical shift perturbation of amide protons

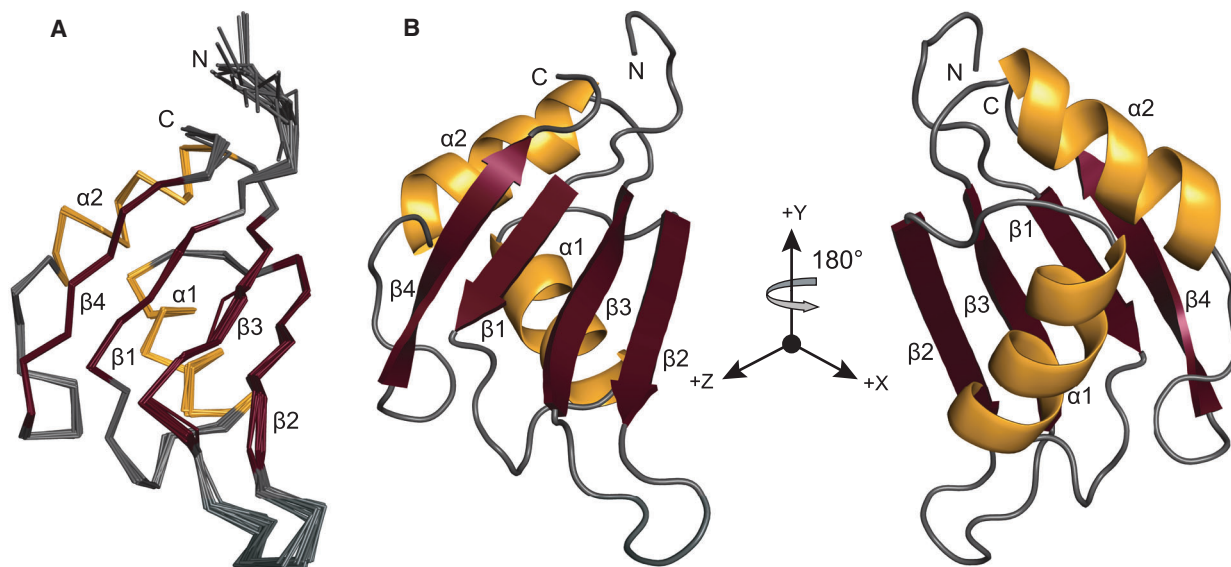


Fig. 3. Solution NMR structure of RBM3^{1–84}. (A) The C_α trace of 20 superimposed energy-minimized structures represent the three-dimensional structure of RBM3^{1–84}, with a backbone RMSD of 0.22 Å. (B) The lowest energy conformer of RBM3^{1–84} is shown in ribbon representation. RBM3^{1–84} adopts four-stranded antiparallel β -sheet (coloured in maroon) packed against two α -helices (coloured in orange). Loops are coloured in grey. Structures were generated using PYMOL [75].

with more than -4.6 ppb·K⁻¹ also indicate their involvement in hydrogen bond formation [43]. Almost all such residues are found in the α -helices and β -strands of the RRM domain, thus, indicating consistency between the temperature-dependent studies and the determined structure (Fig. 4E).

To further investigate the oligomeric states of RBM3^{1–84} and RBM3^{1–110}, chemical cross-linking and size exclusion chromatography (SEC) experiments were performed. Chemical cross-linking with glutaraldehyde followed by SDS/PAGE clearly showed protein bands with molecular masses corresponding to monomeric, dimeric, trimeric and higher order oligomers for both RBM3^{1–84} and RBM3^{1–110} (Fig. 5A). In the SEC, RBM3^{1–84} elutes at 91.02 mL and RBM3^{1–110} elutes at 86 and 93.4 mL (Fig. 5B). These correspond to molecular masses of 19.1 kDa for RBM3^{1–84} and 28.5 and 15.8 kDa for RBM3^{1–110} respectively (Fig. 5C). The calculated molecular mass of monomeric RBM3^{1–84} and RBM3^{1–110} are 9.7 and 12.2 kDa respectively. This is again consistent with the presence of higher oligomeric states of both RBM3^{1–84} and RBM3^{1–110} in solution. The temperature dependence of oligomerization was further confirmed by performing cross-linking experiments at 5–40 °C, which clearly shows increase in oligomeric states with decrease in temperature for both proteins (Fig. 5D and Fig. S2).

Identification of the nucleic acid binding residues of RBM3 by NMR titration and molecular dynamics simulation

In order to determine the nucleic acid binding surface of RBM3, two titration experiments were performed. A seven base RNA, 5'-GGAGGUG-3', was titrated into ¹⁵N-labeled RBM3^{1–84} and RBM3^{1–110}, and ¹H-¹⁵N HSQC spectra were collected for each titration point. For both titrations, continuous changes in chemical shifts of several residues were observed (Fig. 6A,B) indicating exchange in intermediate to fast timescale [44]. In order to get structural insight of the binding interface, a model of the RNA was docked onto the RBM3^{1–84} structure by HADDOCK [45] using the perturbed residues as restraints and the docked RNA-RBM3^{1–84} complex was subjected to molecular dynamics (MD) simulation (Fig. 6C,D).

For comparison, both the free RBM3^{1–84} and the RNA-RBM3^{1–84} structures were subjected to all atom MD simulation for 500 ns under aqueous conditions. After the first few nanoseconds, the RMSD remained constant for the entire period of simulation indicating no major conformational changes in both free and RNA-bound RBM3^{1–84} (Fig. S3A). The β 2- β 3 loop showed the most flexibility in the MD simulations as seen from the RMSF plots (Fig. S3B), which is consistent with the NMR-derived dynamics (Fig. 4A–C) and

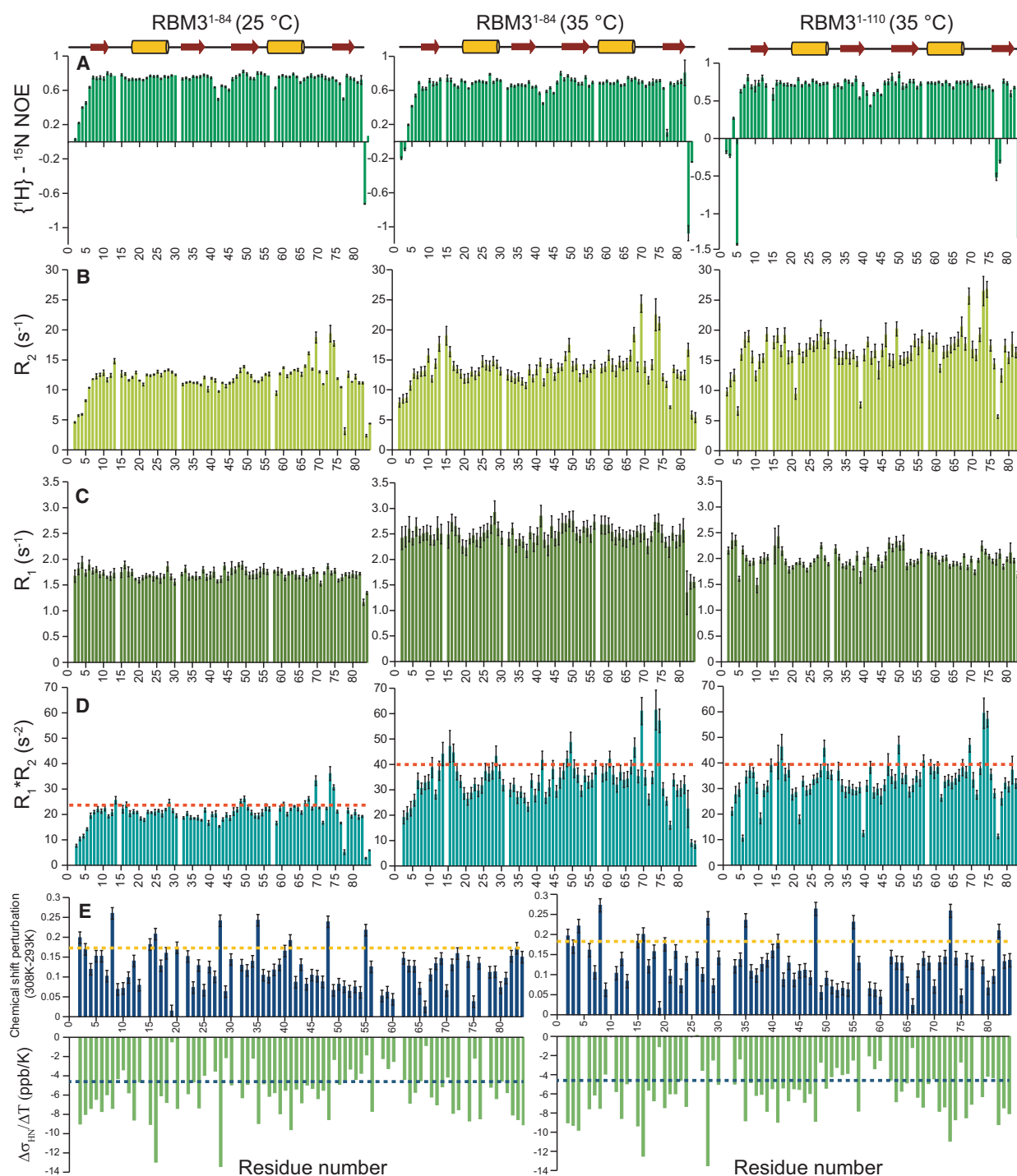


Fig. 4. Fast ps-ns timescale dynamics of RBM3¹⁻⁸⁴ and RBM3¹⁻¹¹⁰. The (A) $\{^1\text{H}\}-^{15}\text{N}$ -NOE, (B) R_2 , (C) R_1 and (D) $R_1 R_2$ values of each residue for RBM3¹⁻⁸⁴ at 25 and 35 °C and RBM3¹⁻¹¹⁰ at 35 °C are shown. For $R_1 R_2$, the dashed red line represents the sum of average and one standard deviation values. (E) The chemical shift perturbation between 35 and 20 °C for both RBM3¹⁻⁸⁴ and RBM3¹⁻¹¹⁰ is shown. The slope of amide proton chemical shift change with temperature is plotted. The dashed horizontal lines represent the sum of average and one standard deviation values.

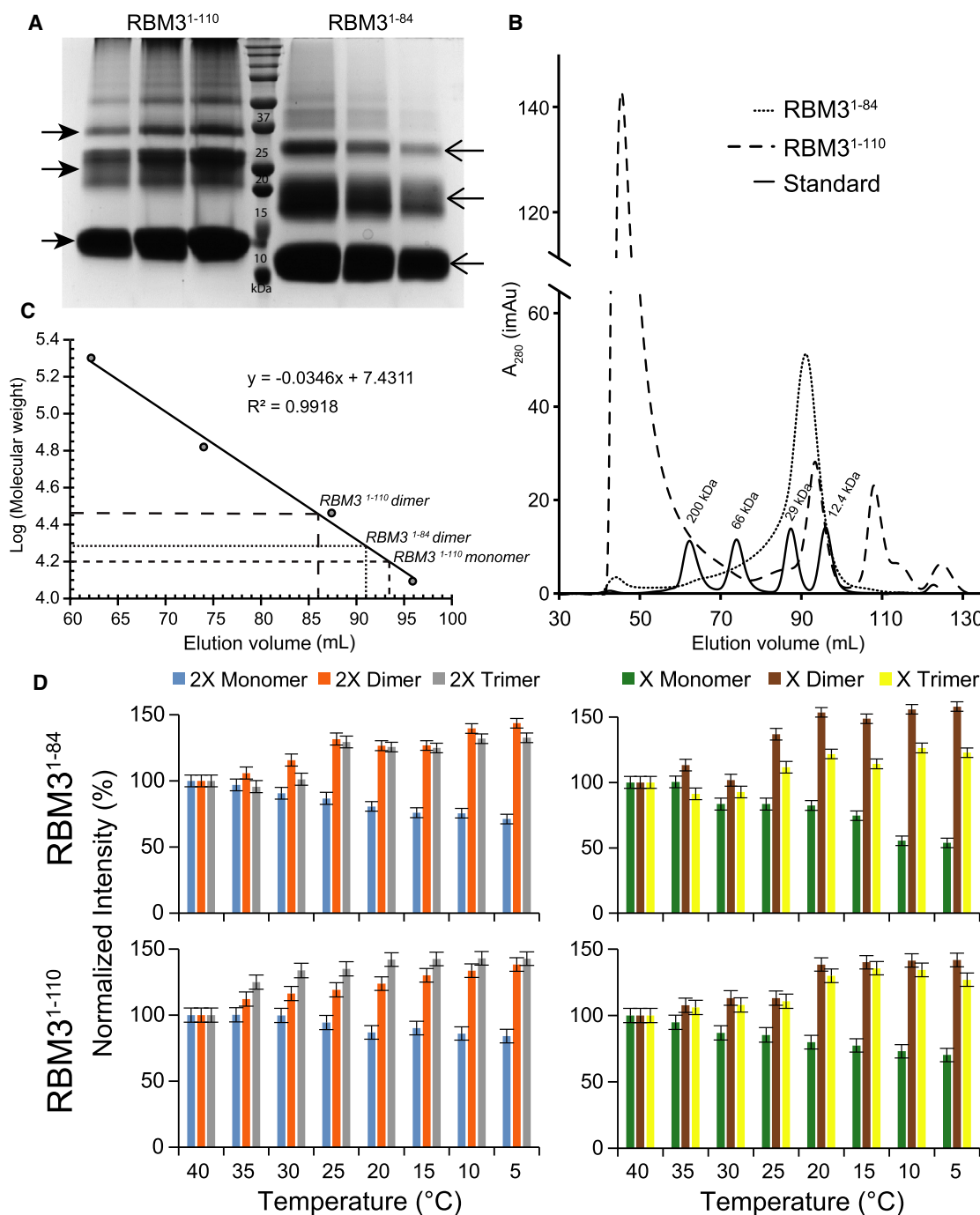


Fig. 5. RBM3¹⁻⁸⁴ and RBM3¹⁻¹¹⁰ form higher oligomeric states. (A) Chemical cross-linking of RBM3¹⁻¹¹⁰ (lanes 1–3) and RBM3¹⁻⁸⁴ (lanes 5–7) using glutaraldehyde show protein bands with molecular masses corresponding to monomeric, dimeric, trimeric (indicated by arrows) and higher order oligomeric states. Lane 4 is the molecular weight marker. (B) In SEC, the elution of the protein standards with known molecular masses is shown in solid line, RBM3¹⁻⁸⁴ in dotted line and RBM3¹⁻¹¹⁰ in broken line. (C) The Log of molecular weight is plotted against the corresponding elution volume for the protein standards. The standard curve is a solid straight line, based on which the molecular masses of the eluted proteins are determined. (D) Intensity of the protein bands corresponding to monomer, dimer and trimer, obtained from chemical cross-linking experiments, are plotted as function of temperature ranging from 5 to 40 °C for both RBM3¹⁻⁸⁴ (top panels) and RBM3¹⁻¹¹⁰ (bottom panels). The left and right panels correspond to two different concentrations of each protein. Bands in each lane are normalized to the corresponding band at 40 °C.

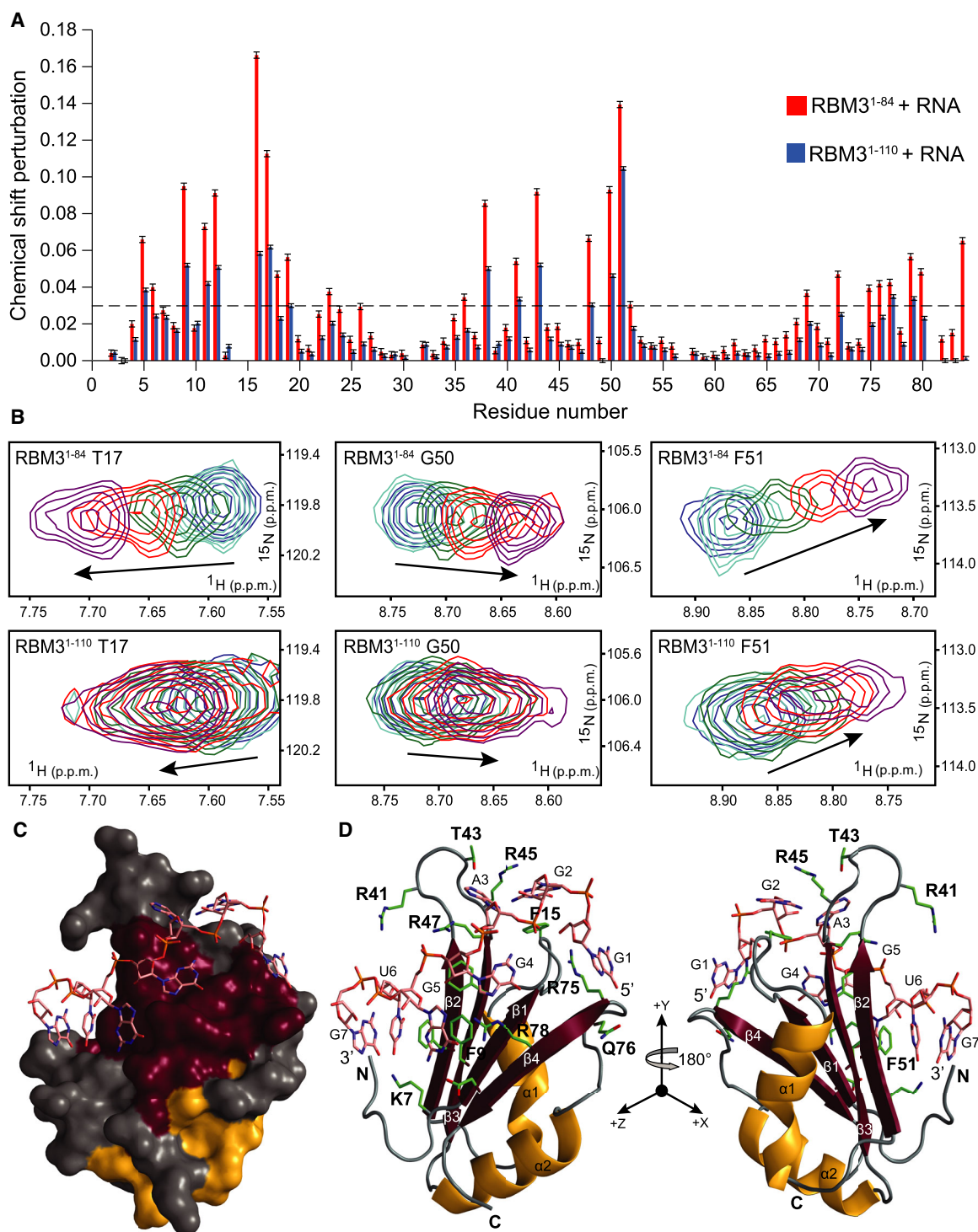


Fig. 6. Titration of ^{15}N -labeled RBM3¹⁻⁸⁴ and RBM3¹⁻¹¹⁰ with RNA. (A) Chemical shift perturbation of each residue upon titration with RNA is plotted for RBM3¹⁻⁸⁴ (red) and RBM3¹⁻¹¹⁰ (blue). The broken line represents the sum of average and one standard deviation values for RBM3¹⁻⁸⁴. Residues with greater values are considered to be significantly perturbed. (B) Changes in chemical shifts of a few residues in RBM3¹⁻⁸⁴ (top) and RBM3¹⁻¹¹⁰ (below) are shown. (C) The molecular model of RNA (5'-GGAGGUG-3'), shown in sticks, bound to RBM3¹⁻⁸⁴, shown as surface, obtained after 500 ns of MD simulation is shown. The RNA and the protein residues in the binding site are labelled (D). Structures were generated using PYMOL [75].

structure (Fig. 3A). Radius of gyration plots show that the overall dimensions of both structures remained constant (Fig. S3C).

During MD simulation of the RNA-RBM3^{1–84} complex, in the first 10 ns, the RNA reoriented from its compact conformation around $\beta 1-\alpha 1$, $\beta 2-\beta 3$ loops to an extended conformation that was tightly packed against the β -sheets and remained stable throughout the simulation (Fig. S4). The stability of the complex can be attributed to multiple interactions, such as hydrogen bonds, hydrophobic and ionic interactions (Fig. S5), between the protein and the RNA (Fig. 7). All the interactions were determined based on geometric constraints and visual inspection in VMD. Most of these interactions revealed by MD simulation are supported by the chemical shift perturbations from the NMR titration experiments (Table S4).

Discussion

The RBM3 protein has three distinct regions: a folded RNA-binding RRM domain, a stretch of RGG/RG repeats and a low sequence complexity YGG-rich region. The RRM domain is properly folded while the C-terminal region of the protein is structurally disordered. Each of these regions have distinct functions. The RRM domain binds RNA through a canonical binding interface. The disordered RGG/RG repeats facilitate high affinity nonspecific interactions with RNA for RBPs [14]. The flexible YGG sequences are also reported to modulate protein–protein and protein–RNA interactions [46]. The flexible RGG/RG and the YGG motifs are also reported to promote self-aggregation resulting in liquid–liquid phase separation in several proteins [46,47].

In this study, solution NMR spectroscopy was used to determine the structure of the RRM domain of RBM3. NMR dynamics also show that the appended RGG/RG region is disordered. The structured domain has a typical RRM fold with four β -strands forming a β -sheet and two α -helices. Using a combination of NMR experiments and molecular dynamics simulation, the binding of the RRM domain to a seven residue RNA strand (5' G1-G2-A3-G4-G5-U6-G7 3') has been characterized with atomic detail. The bound RNA is in an extended conformation and forms extensive contacts with the β -sheet, $\beta 1-\alpha 1$ and $\beta 2-\beta 3$ loops of the RRM domain (Fig. 7 and Table S4). The conserved residues F9 and F51 form π -stacking interactions with G5 and U6 bases respectively. These interactions are present for almost the entire simulation indicating their high stability. Another π -stacking interaction, with less duration, is observed between a nonconserved residue F15 and

G2 base. This interaction is unique to RBM3 and most likely contributes to the extended binding conformation of the RNA. Several arginine sidechains form ionic and hydrogen bonding interactions with the RNA, resulting in a stable complex.

This study also revealed that both RBM3^{1–84} and RBM3^{1–110} form asymmetric dimer, trimer and higher order oligomers in solution as evident from NMR-based dynamics, cross-linking and SEC experiments. Also, this oligomerization is temperature dependent and is favoured with decrease in temperature. Thus, in solution the RBM3 protein can be envisioned to be present in a dynamic equilibrium of monomer and higher order oligomeric states (Fig. 8A). Upon RNA titration, both RBM3^{1–84} and RBM3^{1–110} bind RNA at the same interface (Fig. 6B). However, the ¹H–¹⁵N HSQC peaks of the interface residues keep on broadening with addition of RNA (Fig. 6B). This behaviour is different from a simple two-site exchange between free and bound states [44]. Indeed, the chemical shift perturbation did not fit to a 1 : 1 binding model. The broadened peaks in the ¹H–¹⁵N HSQC spectra (Fig. 6B) most likely result from exchange between the oligomeric states of the RBM3–RNA complex. Thus, it can be postulated that the RBM3–RNA complex also exists in a dynamic equilibrium of monomeric and oligomeric forms (Fig. 8A).

In order to identify the residues involved in oligomerization, temperature-dependent chemical shift perturbation and backbone relaxation were measured. The surface residues F15, N16, S28 and R41 were identified by both measurements. Interestingly, these residues are scattered on the surface of the protein. Thus, the RBM3^{1–84} domain has multiple 'sticky' patches [48] on its surface, which drive the formation of the asymmetric oligomers (Fig. 8B).

RBM3 is a cold-shock protein with increased expression in cold stress [20]. It is predominantly found in the nucleus but can undergo nucleocytoplasmic shuttling. For example, RBM3 shuttles to endoplasmic reticulum (ER) during ER stress [18]. RBM3 has also been reported to increase the global protein synthesis under mild hypothermia [49]. It interacts with 60S ribosomes, helps in formation of active polysomes, dephosphorylates eukaryotic initiation factor 2 alpha (eIF2 α); facilitates the phosphorylation of eukaryotic initiation factor 4E (eIF4E) and regulates microRNA level in cells [50]. However, the molecular details of how RBM3 modulates global protein expression levels in stressed conditions is not yet known. It has been proposed that during ER stress response RBM3 participates in stress granules in the cytosol and chaperones mRNAs, thus, preventing apoptosis [51]. It is also reported that RBM3 interacts with GTPase-activating

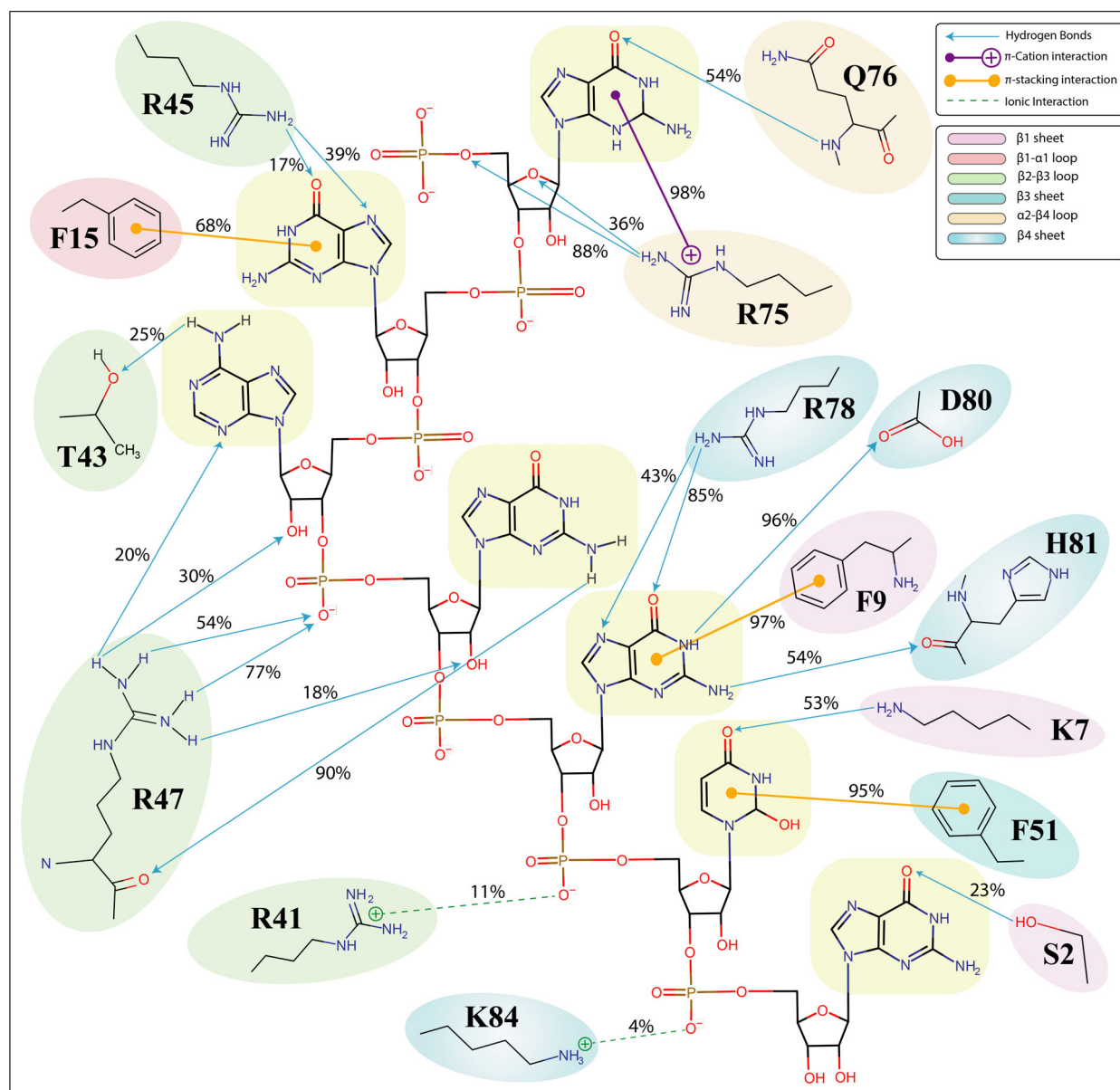


Fig. 7. Important interactions at the RNA-RBM3¹⁻⁸⁴ interface. Chemical structures of the seven residue RNA and the RBM3¹⁻⁸⁴ interface residues are shown. Top right panel indicates the type of interaction. The occurrence of each interaction between RNA and RBM3¹⁻⁸⁴, as observed during the 500-ns simulation, is shown as percentage of the simulation time.

protein-binding protein 1 (G3BP1) and promotes stress granules formation in ischemic stress to protect the cells from apoptosis [52]. Here, we have solved the RRM domain structure of RBM3 and shown that it undergoes oligomerization, which is favoured by decrease in temperature. We postulate that the RBM3 oligomerization upon decrease in temperature can result in liquid-liquid phase separation and might trigger the formation of stress granules under cold-shock conditions. Overall, our structural studies should

provide the necessary foundation for the detailed investigation of RBM3 function at the cellular and organismal levels.

Materials and methods

Cloning, protein expression and purification

DNA fragments of 252 bp and 330 bp, encoding the RRM domain (residues 1-84) and extended RRM-RGG domain

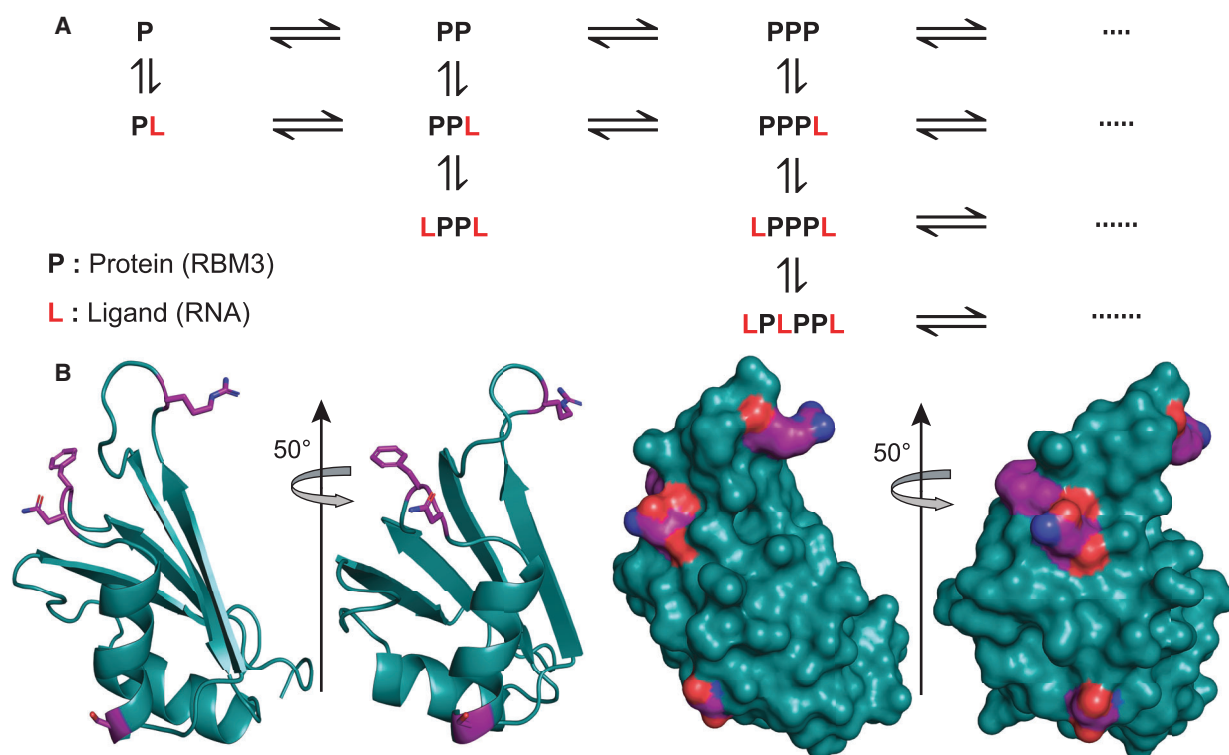


Fig. 8. Model of the oligomerization of RBM3 and RBM3–RNA complex. (A) RBM3 (P) exists in a dynamic equilibrium among monomer (P), dimer (PP), trimer (PPP) and higher order oligomers (top line). RNA is represented as L (ligand). Monomeric RBM3 can bind one RNA molecule (PL); dimeric RBM3 can bind one (PPL) or two (LPPL) RNA molecules; and trimeric RBM3 can bind one (PPPL), two (LPPPL) or three (LPLPPL) RNA molecules. Each of these multimeric forms (free and RNA-bound RBM3) also exchange among themselves. (B) Several residues on the surface of the RRM domain of RBM3 were identified. These residues, forming ‘sticky’ patches, are coloured in purple. Structures were generated using *PYMOLE* [75].

(residues 1–110), respectively, of human RNA-Binding Motif 3 protein (RBM3, UniProt entry P98179/RBM3_HUMAN), were amplified from human cDNA. These were cloned in pET28a vector (Novagen, Merck, Darmstadt, Germany). The resulting construct contained the gene sequence for the target protein with an N-terminal His₆-affinity tag followed by a thrombin cleavage site. The constructs were confirmed by DNA sequencing.

The constructs were transformed into *Escherichia coli* BL21(ΔDE3) cells. For NMR studies, uniformly ¹⁵N-labeled RBM3^{1–84} (residues Met1 to Lys84) and RBM3^{1–110} (residues Met1 to Gln110) were obtained by growing transformed bacteria containing these constructs in M9 minimal media supplemented with 1 g·L^{−1} ¹⁵NH₄Cl as the only source of nitrogen. Similarly, in order to obtain uniformly ¹³C- and ¹⁵N-labeled RBM3^{1–84}, transformed bacteria were cultured in M9 minimal media containing 3 g·L^{−1} ¹³C₆H₁₂O₆ and 1 g·L^{−1} ¹⁵NH₄Cl as the sole carbon and nitrogen sources respectively. For cross-linking and SEC studies, bacteria were grown in LB media in order to produce unlabelled proteins. All cultures were grown at 37 °C until cell density OD₆₀₀ reached ~ 0.6, induced with 0.4 mM

IPTG at 22 °C and then grown overnight. Cells were harvested by centrifugation, lysed by sonication in a lysis buffer [100 mM Tris-Cl, 200 mM NaCl, 10 mM Imidazole (pH 8.2)] and further centrifuged to remove cell debris. The protein was purified from the resulting supernatant using a Ni²⁺ affinity column, following which, the His₆-tag was cleaved using a Thrombin CleanCleave Kit (Sigma-Aldrich, St. Louis, MO, USA). The final proteins RBM3^{1–84} and RBM3^{1–110} have four remnant N-terminal residues (Gly-Ser-His-Met) and have molecular mass of 9.7 and 12.2 kDa respectively. The proteins were exchanged into the final buffer [10 mM sodium phosphate (pH 6.5) and 200 mM NaCl]. For the long-term stability of the proteins, 0.8 mM PMSF, 2 μL of protease inhibitor cocktail and 0.04% sodium azide were also added to the final samples for NMR spectroscopy.

Due to the absence of tryptophan residues, the concentrations of RBM3^{1–84} and RBM3^{1–110} were determined from a standard BSA plot. Different concentrations (0.2–1.0 μg·μL^{−1}) of BSA were run in a 12% SDS/PAGE and the peak intensities (obtained from IMAGEJ [53]) were plotted against the known BSA concentrations. Further, various amounts of purified protein (10, 7, 3 and 1 μg) were

run on a 12% SDS/PAGE and the concentrations were calculated from the peak intensities, using the slope and y-intercept generated from the BSA standard plot.

NMR experiments and structure calculations

NMR experiments were performed on TCI cryoprobe-equipped Bruker 600 MHz NMR spectrometer at 25 and 35 °C. Sample concentrations were 0.5–1 mM with 7% D₂O for spin lock. The backbone resonances of the ¹³C- and ¹⁵N-labeled RBM3^{1–84} were manually assigned using two-dimensional ¹⁵N-¹H HSQC and three-dimensional NMR experiments, such as CBCA(CO)NH, HNCACB, HNCO and HN(CA)CO experiments [54]. The sidechain assignments were obtained using three-dimensional (H)CC(CO)NH, HBHA(CO)NH, HCCH-TOCSY, H(CC)(CO)NH and ¹⁵N-HSQC-TOCSY experiments. These spectra were processed with NMRPipe [55] and analysed using NMRFAM-Sparky [56] with the help of I-PINE automated assignments [57]. The secondary structure propensity was calculated from the backbone chemical shifts (¹H^N, ¹⁵N, ¹³C_α, ¹³C_β and ¹³CO) using MICS [29].

For the structure calculation, 3D ¹H-¹H-¹⁵N and ¹H-¹H-¹³C NOESY-HSQC experiments were recorded with a mixing time of 110 ms. Nuclear overhauser effect (NOE)-derived distance restraints and chemical shift-based torsion angle restraints were used to calculate the three-dimensional structure of RBM3^{1–84}. Well-resolved cross-peaks were manually assigned, followed by automated assignment using AUDANA [58]. The solution structure was determined using PONDEROSA-C/S [59] in conjunction with Xplor-NIH [60] as described previously [27,31]. We also used the POKY suite, a successor of NMRFAM-SPARKY and PONDEROSA-C/S, for the restraint refinement for convenience [61]. Structure calculations were performed in iterative cycles. In the initial stages, 40 structures were calculated and of these, the 20 lowest energy structures were used for further analysis and structural calculations. In the final cycle, 200 structures were calculated using explicit water refinement. Out of these, the 20 lowest energy structures were selected for submission. The final structures were validated using the online servers, such as wwPDB validation service (<https://validate-rcsb-2.wwpdb.org/>) and PSVS 1.5 [62]. We have developed and used an in-house POKY Notepad script, *ponderosa_rmsd_script.py*, for the superimposition of 20 structures and RMSD calculations of backbone and heavy atoms from ordered residues (https://github.com/pokynmr/POKY/blob/master/User_Modules/ponderosa_rmsd_script.py). Secondary structural boundaries were determined using DSSP [63], and structures were visualized using PyMOL (PyMOL molecular graphics system, version 2.4.1, Schrodinger, LLC, New York, NY).

Temperature-dependent ¹H-¹⁵N HSQC experiments were obtained for RBM3^{1–84} and RBM3^{1–110} at 20, 25, 30 and

35 °C. The actual chemical shift changes by correcting for the change in D₂O lock was obtained using the webserver Shift-T [64]. It was also used to determine the slope of the amide proton chemical shifts with temperature in ppb/K.

Backbone amide ¹⁵N relaxation

Backbone amide ¹⁵N R₁, R₂ and steady-state heteronuclear NOE experiments were collected at 25 and 35 °C. The heteronuclear {¹H}-¹⁵N-NOE spectra were acquired with and without 3s of ¹H saturation and a recycle delay of 5s. Heteronuclear NOE for each residue was determined from the ratios of the peak heights in the two spectra and uncertainty in NOE was estimated by propagation of error using the spectral noise. The R₁ and R₂ relaxation rate constants were determined by fitting the peak intensities to single exponential decay {*I*_t = *I*₀*exp(−*t**R_i)}, where '*I*_t' is the peak intensity, '*t*' is the relaxation delay, *I*₀ is the Initial intensity and R_i is either R₁ or R₂ [65]. Uncertainties in the rate constants were estimated by Monte Carlo simulation. The relaxation data for the folded domain were used to determine the global correlation times τ_c using Tensor2 [66].

The product of R₁R₂ reports residues undergoing chemical exchange. Residues involved in chemical exchange (R_{ex}) have slower motion at μs-ms timescale and observe higher R₁R₂ values [42]. In order to determine the residues undergoing significant chemical exchange, average and standard deviation of R₁R₂ for RBM3^{1–84} and RBM3^{1–110} were calculated at 25 and 35 °C (Fig. 4). Residues with greater than the average + 2*standard deviation values were discarded and the average was recalculated with the rest of the amino acids. Residues with R₁R₂ values greater than the recalculated average plus one standard deviation value are considered for further analysis. Errors in R₁R₂ were calculated by propagation of error.

RNA titration into RBM3

HPLC purified RNA (5' GGAGGUG 3') was purchased from Sigma. A 1.107 mM stock solution of RNA (72 nmoles) was titrated into ¹⁵N-labeled RBM3^{1–84} or RBM3^{1–110}, and ¹⁵N-HSQC spectra were collected at each titration point. The initial concentrations of RBM3^{1–84} and RBM3^{1–110} were 80 and 100 μM respectively. The molar ratios of RNA to RBM3^{1–84} in the titration sets were 0, 0.33, 0.67, 1.0 and 1.8, whereas for RBM3^{1–110}, the molar ratios were 0, 0.33, 0.67, 1.0 and 1.4. The chemical shifts in the ¹⁵N-HSQC spectra at each titration point were assigned by tracking the shifts relative to the initial free protein. Combined amide chemical shift changes for each residue were obtained as

$$\Delta\delta_{\text{obs}} = \left\{ (\Delta\delta_{\text{H}})^2 + (0.154 * \Delta\delta_{\text{N}})^2 \right\}^{1/2},$$

where $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{N}}$ are the observed change of chemical shifts from the free state to the bound state, in the proton and nitrogen dimension respectively. Errors in the chemical shift changes were calculated from the changes of unperturbed peaks.

In order to determine the residues undergoing significant perturbation after titration with RNA, average and standard deviation of chemical shift perturbation ($\Delta\delta_{\text{obs}}$) for RBM3^{1–84} were calculated. Residues with greater than the average + 2*standard deviation values were discarded and the average was recalculated with the rest of the amino acids. Residues with $\Delta\delta_{\text{obs}}$ greater than the recalculated average plus one standard deviation value are considered for further analysis.

Chemical cross-linking experiments

Several protein concentrations of purified RBM3^{1–84} and RBM3^{1–110} (0.5, 0.25 and 0.125 mg·mL^{–1}) were mixed with 0.05% of glutaraldehyde solution in 20- μ L reaction volume at room temperature for 30 min. Reactions were quenched by adding SDS and heating at 95 °C for 10 min. These samples were run on 12% SDS/PAGE and stained with Coomassie blue.

For temperature-dependent experiments, protein samples at two concentrations were equilibrated in a PCR thermocycler for 10 min at 5, 10, 15, 20, 25, 30, 35 and 40 °C. Glutaraldehyde was added to these samples and equilibrated for another 20 min at their respective temperatures, quenched by SDS and resolved on a 12% SDS/PAGE. Bands corresponding to monomer, dimer and trimer were quantified by IMAGEJ [53]. Errors in the band intensity were estimated from background noise.

Size exclusion chromatography

RBM3^{1–84} and RBM3^{1–110} were run on a HiLoad 16/600 Superdex PG200 column (GE Healthcare, Uppsala, Sweden) at ambient temperature. A set of standard proteins with molecular weights of 200, 66, 29 and 12.4 kDa were run under the same conditions to serve as molecular weight markers. The molecular mass of RBM3^{1–84} and RBM3^{1–110} was determined from the calibration plot of standard molecular weight markers.

Structure calculations of the RBM3^{1–84}-RNA complex using HADDOCK

The structural model of the complex between RBM3^{1–84} and RNA is generated using HADDOCK [45]. For RBM3^{1–84}, the active residues (directly involved in the interactions) used in the docking were 9, 11, 12, 15–17, 38, 43, 48, 50, 51. The active nucleotides in the RNA strand were defined as nucleotides 1–7. The active residues were

selected on the basis of the chemical shift perturbation data. For the RNA, we used the X-Ray crystal structure of the Sxl (RBD1–RBD2)-RNA complex (PDB entry 1B7F) as model [67]. We mutated the Sxl RNA using PYMOL to obtain the structure of our 7-nucleotide (5'-GGAGGUG-3') RNA fragment. Docking was performed using the ensemble of 20 lowest energy RBM3^{1–84} free structures (PDB entry 7EB1). During docking, the protein was kept rigid, while the RNA was kept flexible. The HADDOCK score is derived by combining electrostatic, van der Waals energies along with additional empirical terms (desolvation, buried surface area) and restraints violation energies [45]. The structure with the best HADDOCK score and lowest Z-score was selected for further analysis.

Molecular dynamics simulations of free and RNA-bound RBM3^{1–84}

Both free RBM3^{1–84} and RNA-bound RBM3^{1–84} (obtained from HADDOCK docking) were subjected to molecular dynamics simulations using GROMACS software suite (version 2018.7) [68]. The topologies were built employing AMBER99SB force field and solvated in a cubic box using TIP3P water model [69]. Solvation was followed by addition of 4 and 10 sodium ions for neutralization of the free RBM3^{1–84} and RNA-bound RBM3^{1–84} systems respectively. This was followed by energy minimization using a force convergence criteria of 1000 kJ·mol^{–1}·nm^{–1} for elimination of steric clashes. The energy-minimized systems were then equilibrated under NVT and NPT ensemble sequentially for 500 ps each. Unrestrained production simulations for both free RBM3^{1–84} and RNA-bound RBM3^{1–84} were then performed for 500 ns using leapfrog dynamics integrator with an integration step size of 2 fs. Periodic boundary conditions (PBC) were considered in all three dimensions. A cut-off of 1.2 nm was used for calculating Coulombic and van der Waals interactions; long-ranged electrostatic interactions were computed using Particle-mesh Ewald algorithm and P-LINC algorithm was used for constraining the bond lengths [70,71]. Berendsen's modified thermostat was used for simulating constant temperature at 300 K and Parrinello–Rahman pressure coupling was used to maintain a constant pressure of 1 bar [72,73]. The PBC corrected trajectory was used for analysis after completion of the production simulation using modules available in GROMACS and in-house python scripts.

Acknowledgements

The authors thank the Central Research Facility (CRF) at IIT Kharagpur for the use of NMR, and SEC facilities and IIT-Delhi HPC facility for the molecular dynamics simulations. This work has been supported by funding from the Department of

Biotechnology (DBT), India (BT/PR34319/MED/29/1488/2019) to SD and from Department of Science & Technology (DST), India (SR/WOS-A/LS-659/2016) to SR. SB acknowledges a NET-JRF fellowship [918/(CSIR-UGC NET DEC. 2016)] provided by the University Grants Commission (UGC), Government of India. SM and AJB acknowledge IIT Kharagpur for their fellowships.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

SR, AB and SD conceived the project. SR, SB, AJB and SM performed NMR assignments, cross-linking and SEC experiments. SB, AJB, WL and SD solved the solution NMR structure. SM and SD performed and analysed NMR dynamics experiments. SB and SD performed and analysed NMR titration experiments. TK, SLG and KKI performed and analysed MD experiments. All authors wrote and edited the manuscript.

Peer Review

The peer review history for this article is available at <https://publons.com/publon/10.1111/febs.16301>.

Data Availability Statement

The atomic coordinates and structure factors for the solution NMR structure of RBM3^{1–84} have been deposited in the RCSB Protein Data Bank with accession number [7EB1](#). The chemical shifts are deposited in the BioMagResBank under the accession number 50804.

References

- Gerstberger S, Hafner M & Tuschl T (2014) A census of human RNA-binding proteins. *Nat Rev Genet* **15**, 829–845.
- Masuzawa T & Oyoshi T (2020) Roles of the RGG domain and RNA recognition motif of nucleolin in G-quadruplex stabilization. *ACS Omega* **5**, 5202–5208.
- Glisovic T, Bachorik JL, Yong J & Dreyfuss G (2008) RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* **582**, 1977–1986.
- Khan F, Daniëls MA, Folkers GE, Boelens R, Saqlan Naqvi SM & van Ingen H (2014) Structural basis of nucleic acid binding by *Nicotiana tabacum* glycine-rich RNA-binding protein: implications for its RNA chaperone function. *Nucleic Acids Res* **42**, 8705–8718.
- Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, Ackerley S, Durnall JC, Williams KL, Buratti E *et al.* (2008) TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* (80-) **319**, 1668–1672.
- Tyzack GE, Luisier R, Taha DM, Neeves J, Modic M, Mitchell JS, Meyer I, Greensmith L, Newcombe J, Ule J *et al.* (2019) Widespread FUS mislocalization is a molecular hallmark of amyotrophic lateral sclerosis. *Brain* **142**, 2572–2580.
- Vanderweyde T, Apicco DJ, Youmans-Kidder K, Ash PEA, Cook C, Lummertz da Rocha E, Jansen-West K, Frame AA, Citro A, Leszyk JD *et al.* (2016) Interaction of tau with the RNA-binding protein TIA1 regulates tau pathophysiology and toxicity. *Cell Rep* **15**, 1455–1466.
- Repici M, Hassanjani M, Maddison DC, Garção P, Cimini S, Patel B, Szegő ÉM, Straatman KR, Lilley KS, Borsello T *et al.* (2019) The Parkinson's disease-linked protein DJ-1 associates with cytoplasmic mRNP granules during stress and neurodegeneration. *Mol Neurobiol* **56**, 61–77.
- Wang E, Lu SX, Pastore A, Chen X, Imig J, Chun-Wei Lee S, Hockemeyer K, Ghebrehrestos YE, Yoshimi A, Inoue D *et al.* (2019) Targeting an RNA-binding protein network in acute myeloid leukemia. *Cancer Cell* **35**, 369–384.e7.
- Qian J, Hassanein M, Hoeksema MD, Harris BK, Zou Y, Chen H, Lu P, Eisenberg R, Wang J, Espinosa A *et al.* (2015) The RNA binding protein FXR1 is a new driver in the 3q26-29 amplicon and predicts poor prognosis in human cancers. *Proc Natl Acad Sci USA* **112**, 3469–3474.
- Chenette DM, Cadwallader AB, Antwine TL, Larkin LC, Wang J, Olwin BB & Schneider RJ (2016) Targeted mRNA decay by RNA binding protein AUF1 regulates adult muscle stem cell fate, promoting skeletal muscle integrity. *Cell Rep* **16**, 1379–1390.
- Julio AR & Backus KM (2021) New approaches to target RNA binding proteins. *Curr Opin Chem Biol* **62**, 13–23.
- Cléry A, Blatter M & Allain FHT (2008) RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290–298.
- Thandapani P, O'Connor TR, Bailey TL & Richard S (2013) Defining the RGG/RG motif. *Mol Cell* **50**, 613–623.
- Tradewell ML, Yu Z, Tibshirani M, Boulanger M-C, Durham HD & Richard S (2012) Arginine methylation by PRMT1 regulates nuclear-cytoplasmic localization and toxicity of FUS/TLS harbouring ALS-linked mutations. *Hum Mol Genet* **21**, 136–149.

- 16 Côté J & Richard S (2005) Tudor domains bind symmetrical dimethylated arginines. *J Biol Chem* **280**, 28476–28483.
- 17 Ciuzan O, Hancock J, Pamfil D, Wilson I & Ladomery M (2015) The evolutionarily conserved multifunctional glycine-rich RNA-binding proteins play key roles in development and stress adaptation. *Physiol Plant* **153**, 1–11.
- 18 Zhu X, Bühner C & Wellmann S (2016) Cold-inducible proteins CIRP and RBM3, a unique couple with activities far beyond the cold. *Cell Mol Life Sci* **73**, 3839–3859.
- 19 Derry MJ, Kerns JA & Francke U (1995) RBM3, a novel human gene in Xp11.23 with a putative RNA-binding domain. *Hum Mol Genet* **4**, 2307–2311.
- 20 Danno S, Nishiyama H, Higashitsuji H, Yokoi H, Xue J-H, Itoh K, Matsuda T & Fujita J (1997) Increased transcript level of RBM3, a member of the glycine-rich RNA-binding protein family, in human cells in response to cold stress. *Biochem Biophys Res Commun* **236**, 804–807.
- 21 Wellmann S, Bühner C, Moderegger E, Zelmer A, Kirschner R, Koehne P, Fujita J & Seeger K (2004) Oxygen-regulated expression of the RNA-binding proteins RBM3 and CIRP by a HIF-1-independent mechanism. *J Cell Sci* **117**, 1785–1794.
- 22 Wong JLL, Au AYM, Gao D, Pinello N, Kwok C-T, Thoeng A, Lau KA, Gordon JEA, Schmitz U, Feng Y *et al.* (2016) RBM3 regulates temperature sensitive miR-142-5p and miR-143 (thermomiRs), which target immune genes and control fever. *Nucleic Acids Res* **44**, 2888–2897.
- 23 Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasamy S, Mitchell A, Nuka G, Paysan-Lafosse T, Qureshi M, Raj S *et al.* (2021) The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* **49**, D344–D354.
- 24 Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* **16**, 10881–10890.
- 25 Kim SS, Seffernick JT & Lindert S (2018) Accurately predicting disordered regions of proteins using rosetta ResidueDisorder application. *J Phys Chem B* **122**, 3920–3930.
- 26 Hatos A, Hajdu-Soltész B, Monzon AM, Palopoli N, Álvarez L, Aykac-Fas B, Bassot C, Benítez GI, Bevilacqua M, Chasapi A *et al.* (2019) DisProt: intrinsic protein disorder annotation in 2020. *Nucleic Acids Res* **48**, D269–D276.
- 27 Basak AJ, Maiti S, Hansda A, Mahata D, Duraivelan K, Kundapura SV, Lee W, Mukherjee G, De S & Samanta D (2020) Structural insights into N-terminal IgV domain of BTNL2, a T cell inhibitory molecule, suggests a non-canonical binding interface for its putative receptors. *J Mol Biol* **432**, 5938–5950.
- 28 Johnson PE, Tomme P, Joshi MD & McIntosh LP (1996) Interaction of soluble cellooligosaccharides with the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC. 2. NMR and ultraviolet absorption spectroscopy †. *Biochemistry* **35**, 13895–13906.
- 29 Shen Y & Bax A (2012) Identification of helix capping and β -turn motifs from NMR chemical shifts. *J Biomol NMR* **52**, 211–232.
- 30 Ulrich EL, Akutsu H, Doreleijers JF, Harano Y, Ioannidis YE, Lin J, Livny M, Mading S, Maziuk D, Miller Z *et al.* (2007) BioMagResBank. *Nucleic Acids Res* **36**, D402–D408.
- 31 Boral S, Maiti S, Basak AJ, Lee W & De S (2020) Structural, dynamic, and functional characterization of a DnaX Mini-intein derived from spirulina platensis provides important insights into intein-mediated catalysis of protein splicing. *Biochemistry* **59**, 4711–4724.
- 32 Schubert M, Labudde D, Oschkinat H & Schmieder P (2002) A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on ¹³C chemical shift statistics. *J Biomol NMR* **24**, 149–154.
- 33 Berman H, Henrick K, Nakamura H & Markley JL (2007) The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. *Nucleic Acids Res* **35**, D301–D303.
- 34 Nagai K, Oubridge C, Ito N, Avis J & Evans P (1995) The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. *Trends Biochem Sci* **20**, 235–240.
- 35 Lu J & Hall KB (1997) Tertiary structure of RBD2 and backbone dynamics of RBD1 and RBD2 of the human U1A protein determined by NMR spectroscopy. *Biochemistry* **36**, 10393–10405.
- 36 Garrett DS, Lodi PJ, Shamoo Y, Williams KR, Clore GM & Gronenborn AM (1994) Determination of the secondary structure and folding topology of an rna binding domain of mammalian hnRNP A1 protein using three-dimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry* **33**, 2852–2858.
- 37 Wittekind M, Grollach M, Friedrichs M, Dreyfuss G & Mueller L (1992) ¹H, ¹³C, and ¹⁵N NMR assignments and global folding pattern of the RNA-binding domain of the human hnRNP C proteins. *Biochemistry* **31**, 6254–6265.
- 38 Lee AL, Kanaar R, Rio DC & Wemmer DE (1994) Resonance assignments and solution structure of the second RNA-binding domain of sex-lethal determined by multidimensional heteronuclear magnetic resonance. *Biochemistry* **33**, 13775–13786.
- 39 Kay LE, Torchia DA & Bax A (1989) Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* **28**, 8972–8979.

- 40 Halle B & Davidovic M (2003) Biomolecular hydration: from water dynamics to hydrodynamics. *Proc Natl Acad Sci USA* **100**, 12135–12140.
- 41 Nagata T, Kanno R, Kurihara Y, Uesugi S, Imai T, Sakakibara S, Okano H & Katahira M (1999) Structure, backbone dynamics and interactions with RNA of the C-terminal RNA-binding domain of a mouse neural RNA-binding protein, Musashi1. *J Mol Biol* **287**, 315–330.
- 42 Kneller JM, Lu M & Bracken C (2002) An effective method for the discrimination of motional anisotropy and chemical exchange. *J Am Chem Soc* **124**, 1852–1853.
- 43 Cierpicki T & Otlewski J (2001) Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J Biomol NMR* **21**, 249–261.
- 44 De S, Greenwood AI, Rogals MJ, Kovrigin EL, Lu KP & Nicholson LK (2012) Complete thermodynamic and kinetic characterization of the isomer-specific interaction between Pin1-WW domain and the amyloid precursor protein cytoplasmic tail phosphorylated at Thr668. *Biochemistry* **51**, 8583–8596.
- 45 van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastiris PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ & Bonvin AMJJ (2016) The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J Mol Biol* **428**, 720–725.
- 46 Castello A, Fischer B, Frese CK, Horos R, Alleaume A-M, Foehr S, Curk T, Krijgsveld J & Hentze MW (2016) Comprehensive identification of RNA-binding domains in human cells. *Mol Cell* **63**, 696–710.
- 47 Chong PA, Vernon RM & Forman-Kay JD (2018) RGG/RG motif regions in RNA binding and phase separation. *J Mol Biol* **430**, 4650–4665.
- 48 Choi J-M, Holehouse AS & Pappu RV (2020) Physical principles underlying the complex biology of intracellular phase transitions. *Annu Rev Biophys* **49**, 107–133.
- 49 Chappell SA, Owens GC & Mauro VP (2001) A 5' leader of Rbm3, a cold stress-induced mRNA, mediates internal initiation of translation with increased efficiency under conditions of mild hypothermia. *J Biol Chem* **276**, 36917–36922.
- 50 Dresios J, Aschrafi A, Owens GC, Vanderklish PW, Edelman GM & Mauro VP (2005) Cold stress-induced protein Rbm3 binds 60S ribosomal subunits, alters microRNA levels, and enhances global protein synthesis. *Proc Natl Acad Sci USA* **102**, 1865–1870.
- 51 Zhu X, Zelmer A, Kapfhammer JP & Wellmann S (2016) Cold-inducible RBM3 inhibits PERK phosphorylation through cooperation with NF90 to protect cells from endoplasmic reticulum stress. *FASEB J* **30**, 624–634.
- 52 Si W, Li Z, Huang Z, Ye S, Li X, Li Y, Kuang W, Chen D & Zhu M (2020) RNA binding protein motif 3 inhibits oxygen-glucose deprivation/reoxygenation-induced apoptosis through promoting stress granules formation in PC12 cells and rat primary cortical neurons. *Front Cell Neurosci* **14**, 1–14.
- 53 Schneider CA, Rasband WS & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675.
- 54 Sattler M (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* **34**, 93–158.
- 55 Delaglio F, Grzesiek S, Vuister G, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.
- 56 Lee W, Tonelli M & Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327.
- 57 Lee W, Bahrami A, Dashti HT, Eghbalian HR, Tonelli M, Westler WM & Markley JL (2019) I-PINE web server: an integrative probabilistic NMR assignment system for proteins. *J Biomol NMR* **73**, 213–222.
- 58 Lee W, Petit CM, Cornilescu G, Stark JL & Markley JL (2016) The AUDANA algorithm for automated protein 3D structure determination from NMR NOE data. *J Biomol NMR* **65**, 51–57.
- 59 Lee W, Stark JL & Markley JL (2014) PONDEROSA-C/S: client-server based software package for automated protein 3D structure determination. *J Biomol NMR* **60**, 73–75.
- 60 Schwieters CD, Kuszewski JJ, Tjandra N & Marius Clore G (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65–73.
- 61 Lee W, Rahimi M, Lee Y & Chiu A (2021) POKY: a software suite for multidimensional NMR and 3D structure calculation of biomolecules. *Bioinformatics* **37**, 3041–3042.
- 62 Bhattacharya A, Tejero R & Montelione GT (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778–795.
- 63 Kabsch W & Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
- 64 Trainor K, Palumbo JA, MacKenzie DWS & Meiering EM (2020) Temperature dependence of NMR chemical shifts: tracking and statistical analysis. *Protein Sci* **29**, 306–314.
- 65 Maiti S, Acharya B, Boorla VS, Manna B, Ghosh A & De S (2019) Dynamic studies on intrinsically disordered regions of two paralogous transcription factors reveal rigid segments with important biological functions. *J Mol Biol* **431**, 1353–1369.
- 66 Dosset P, Hus JC, Blackledge M & Marion D (2000) Efficient analysis of macromolecular rotational

- diffusion from heteronuclear relaxation data. *J Biomol NMR* **16**, 23–28.
- 67 Handa N, Nureki O, Kurimoto K, Kim I, Sakamoto H, Shimura Y, Muto Y & Yokoyama S (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398**, 579–585.
 - 68 Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE & Berendsen HJC (2005) GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701–1718.
 - 69 Hornak V, Abel R, Okur A, Strockbine B, Roitberg A & Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725.
 - 70 Darden T, York D & Pedersen L (1993) Particle mesh Ewald: an $N \cdot \log(N)$ method for Ewald sums in large systems. *J Chem Phys* **98**, 10089–10092.
 - 71 Hess B (2008) P-LINCS: a parallel linear constraint solver for molecular simulation. *J Chem Theory Comput* **4**, 116–122.
 - 72 Bussi G, Donadio D & Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* **126**, 014101.
 - 73 Parrinello M & Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. *J Appl Phys* **52**, 7182–7190.
 - 74 Kirchner DK & Güntert P (2011) Objective identification of residue ranges for the superposition of protein structures. *BMC Bioinformatics* **12**, 170.
 - 75 The PyMOL molecular graphics system, Version 2.4.1. Schrödinger, LLC, New York, NY.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Observation of turbidity in RBM3^{1–84} and RBM3^{1–110} at lower temperature.

Fig. S2. Temperature-dependent chemical cross-linking experiments.

Fig. S3. Analysis of the MD trajectories.

Fig. S4. RNA–RBM3^{1–84} complex has a stable conformation.

Fig. S5. Interaction fractions of each interaction in the RNA–RBM3^{1–84} complex.

Table S1. NMR relaxation parameters of RBM3^{1–84} at 25 °C.

Table S2. NMR relaxation parameters of RBM3^{1–84} at 35 °C.

Table S3. NMR relaxation parameters of RBM3^{1–110} at 35 °C.

Table S4. Comparison of interactions in RNA–RBM3^{1–84} complex from MD simulation with chemical shift perturbation (CSP) from NMR titration data.