Structural analysis reveals the flexible C-terminus of Nop15 undergoes rearrangement to recognize a pre-ribosomal RNA folding intermediate

Jun Zhang, Lauren E. Gonzalez and Traci M. Tanaka Hall[*]

Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, National Institutes of Health, P.O. Box 12233, MD F3-05, Research Triangle Park, NC 27709, USA

ABSTRACT

The RNA recognition motif (RRM) is the most abundant RNA-binding domain in eukaryotes, and it plays versatile roles in RNA metabolism. Despite its abundance, diversity of RRM structure and function is generated by variations on a conserved core. Yeast Nop15 is an RRM protein that is essential for large ribosomal subunit biogenesis. We determined a 2.0 Å crystal structure of Nop15 that reveals a C-terminal α-helical region obscures its canonical RNA-binding surface. Small-angle X-ray scattering, NMR and RNA-binding analyses further reveal that the C-terminal residues of Nop15 are highly flexible, but essential for tight RNA binding. Moreover, comparison with a recently reported cryo-electron microscopy structure indicates that dramatic rearrangement of the C-terminal region of Nop15 in the pre-ribosome exposes the RNA-binding surface to recognize the base of its stem-loop target RNA and extends a newly-formed α helix to the distal loop where it forms protein interactions.

INTRODUCTION

The RNA recognition motif (RRM) is the most abundant RNA-binding domain, accounting for 50% of identified RNA-binding domains and 2% of human protein-coding genes (1). RRMs have been found to function in a wide range of cellular RNA regulatory pathways, such as RNA processing, RNA editing, translation and RNA degradation (2–4). Numerous structural studies have revealed that the conserved core RRM is a domain of ∼90 amino acids folded into a β1-α1-β2-β3-α2-β4 fold (1,5,6). β1 and β3 contain characteristic RNA-binding motifs, RNP2 and RNP1, respectively, which typically bear aromatic residues that form stacking interactions with cognate RNA bases. Most RRMs recognize single-stranded RNA, with or with-

out sequence specificity (1,5,6). Despite the structural conservation of RRMs, variance to the prototype is frequently documented. The core RRM may be flanked by additional α-helices or β-strands that consequently modify the RNA-binding properties of RRMs (1,7). These modifications make it difficult, if not impossible, to predict RRM/RNA interaction. Therefore, a case-by-case approach is required to understand RRM recognition properties, and structural analysis has been central to illuminating these properties.

Nop15 consists of a single RRM domain embedded between N- and C-terminal flexible regions (Figure 1A). In yeast, Nop15 is essential for large (60S) ribosomal subunit biogenesis, as depletion of Nop15 abolishes synthesis of the 5.8S and 25S ribosomal RNAs (rRNAs) (8,9). Ribosome biogenesis is a highly complex process requiring hundreds of factors to correctly assemble and arrange the translational machine. Nop15 has been suggested to play a role in a major event in large subunit biogenesis: the removal of Internal Transcribed Spacer 2 (ITS2) RNA, which lies between the 5.8S and 25S rRNA sequences (10–13). During pre-rRNA processing, ITS2 must transition between two specific structures, an energetically unfavorable 'ring' structure and the more energetically favorable 'hairpin' structure. Nop15 has been shown to stabilize ITS2 in the unfavorable ring structure before the pre-rRNA refolds into the favorable hairpin structure (14–16).

We took a structural and biochemical approach to understand Nop15 interaction with ITS2 RNA and its function in ribosome biogenesis. We determined a crystal structure of the RRM of yeast Nop15 that revealed a core RRM with an accessory C-terminal α-helical region that sequesters the aromatic residues on the classical RNA-binding surface. The C-terminus, including the accessory α-helical region, was required for tight ITS2 RNA binding. Small-angle X-ray scattering (SAXS) and NMR experiments demonstrated that the C-terminal residues are highly flexible in solution. Comparing our crystal structure with Nop15 in a cryo-EM model of the pre-60S ribosome (17) indicates that the α-helical region is dramatically refolded, which re-

*To whom correspondence should be addressed. Tel: +1 919 541 1017; Fax: +1 301 480 3055; Email: hall4@niehs.nih.gov
Present address: Jun Zhang, Department of Chemistry, University of Alabama at Birmingham, Birmingham, AL 35294, USA.
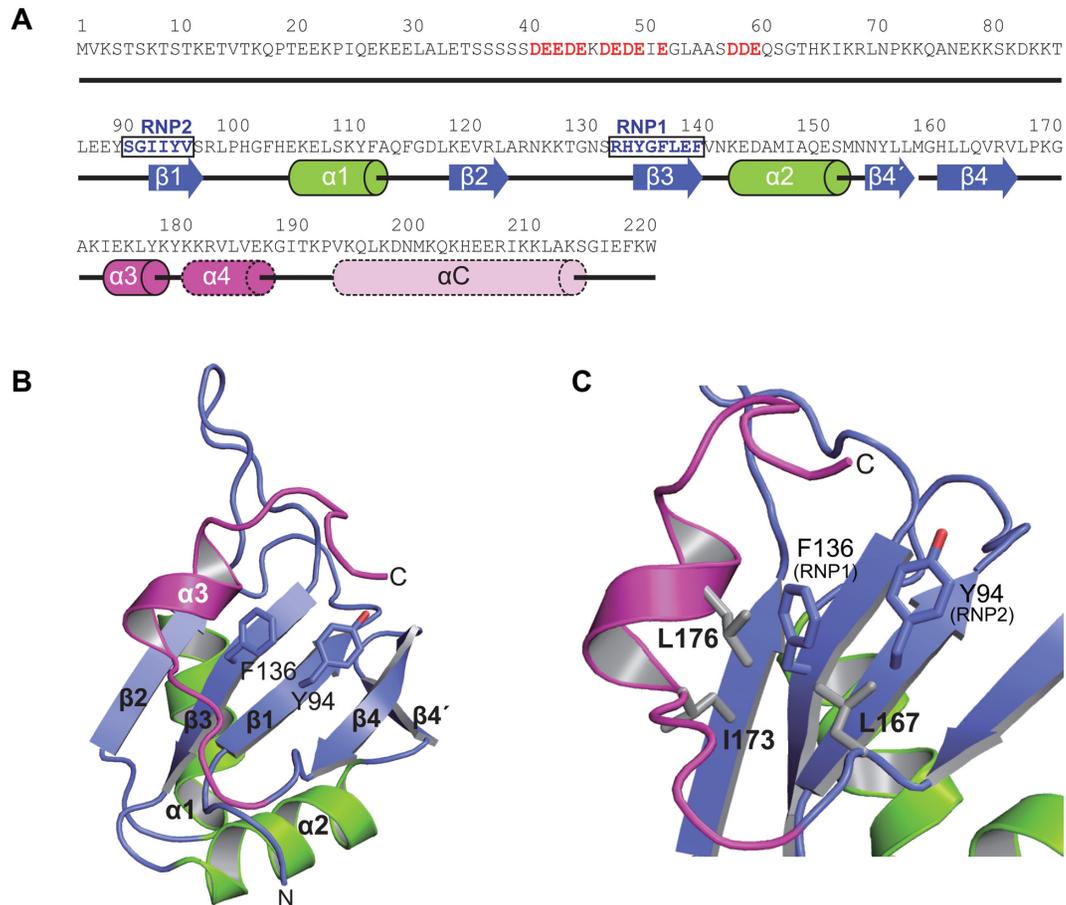
Figure 1. Crystal structure of yeast Nop15 reveals masking of canonical RNA-binding residues by its C-terminal region. (**A**) Amino acid sequence and secondary structural elements of Nop15. Nop15 comprises a single RRM domain (middle row), flanked by N-terminal (top row) and C-terminal (bottom row) regions. α-helices, β-strands and coil regions are shown as cylinders, arrows and black lines, respectively. The electronegative cluster of the N-terminal region is colored red. RNA-binding motifs, RNP1 and RNP2, are denoted. Helices α4 and αC are not visible in the Nop15 crystal structure (chain A) and are drawn with dotted lines. (**B**) Ribbon diagram of the crystal structure of Nop15$^{81-191}$. The core RRM domain is shown with blue β-strands and green α-helices, and the C-terminal region, including helix α3, is colored magenta. Y94 in RNP2 and F136 in RNP1 are shown as stick models. (**C**) Close-up view of the interaction of the C-terminal region with the RNA-binding surface. Hydrophobic residues involved in the interaction are shown in grey. This figure, and Figures 4 and 5 and Supplementary Figure S1 were prepared using PyMol (Schroedinger).

leases the classical RNA-binding residues for interaction at the base of an ITS2 stem-loop and allows interaction of the C-terminal residues near the distal loop of the stem-loop.

MATERIALS AND METHODS

Nop15 expression and purification

A DNA fragment encoding full-length Nop15 (residues 1–220) was amplified from *S. cerevisiae* genomic DNA and cloned into pSMT3 (kindly provided by Christopher Lima, Memorial Sloan Kettering Cancer Center, New York, NY) using *BamH* I and *Hind* III restriction sites (18). Expression constructs for truncated proteins (81–220, 81–180 and 81–191) were cloned similarly. Nop15 proteins were expressed at 22°C overnight in *E. coli* strain BL21-CodonPlus (DE3) in the presence of 0.4 mM IPTG, which was added when the OD$_{600}$ reached 0.6. Cell pellets were resuspended in sonication buffer (25 mM HEPES, pH 7.5, 1 M NaCl, 1 mM TCEP [tris(2-carboxyethyl)phosphine], 25 mM imidazole) plus 1 mg/ml lysozyme and lysed by sonication, followed by

centrifugation to remove cell debris. The supernatant was applied to 5 ml of HisPur Ni-NTA resin (Thermo Scientific), washed with 200 ml of sonication buffer and eluted with 25 mM HEPES, pH 7.5, 500 mM NaCl, 1 mM TCEP, 500 mM imidazole. The N-terminal SUMO-tagged protein was cleaved overnight with 2 μg/ml of Ulp1 at 4°C. The cleaved sample was diluted 5-fold with 20 mM HEPES, pH 7.5, 20 mM NaCl, 1 mM TCEP and loaded onto a 5-ml HiTrap Heparin column (GE Healthcare). The sample was eluted with a linear gradient from 0 to 2 M NaCl in 20 mM HEPES, 1 mM TCEP. Nop15 was eluted from the heparin column when the salt concentration reached ∼1.4 M NaCl and was further purified using a HiLoad 16/60 Superdex75 column (GE Healthcare) equilibrated with 25 mM HEPES, pH 7.5, 500 mM NaCl, 1 mM TCEP. Truncated proteins were expressed and purified with the same protocol. The identities of the proteins were confirmed by mass spectrometry, and the purities were >95% based on SDS-PAGE.

Crystallization and structure determination

The Nop15[81–191] construct was used for crystallization. Initial crystals were obtained by hanging drop vapor diffusion at 22°C, mixing 2 μl of 6 mg/ml Nop15 in 25 mM HEPES, pH 7.5, 500 mM NaCl, 1 mM TCEP with 2 μl of a crystallization solution containing 18% (w/v) PEG 4000, 0.2 M ammonium sulfate, pH 5.25 and 9% (v/v) Jeffamine M-600 (Hampton Research). Crystals were improved by iterative microseeding. Crystallization solution was supplemented with 30% (v/v) ethylene glycol as a cryo-protectant.

X-ray diffraction data were collected at beamline 22-ID of the Advanced Photon Source at 100 K with a wavelength of 1.000 Å for native crystals. The data were processed using HKL2000 (19). Phases were determined by molecular replacement using Phenix Phaser-MR (20,21) with a crystal structure of Y14 as a search model (PDB ID 1HL6). Iterative refinement and model building were carried out using Phenix.Refine (20) and WinCoot (22). There are four molecules in the asymmetric unit, and 45 N- and C-terminal residues were not modeled (A81-A87, A185-A191, B81-B86, C81-C85, D81-D87 and D179-D191). Over 99% of dihedral angles are in favored regions of the Ramachandran plot with no outliers.

RNA preparation

5′ Fluorescein-labeled RNAs corresponding to ITS2 nucleotides 26–60 (UGAGUGAUACUCUUUGGAGU-UAACUUGAAAUUGCU) and its variants, were purchased from Dharmacon (Thermo Scientific), de-protected using 400 μl 100 mM acetic acid (pH adjusted to 3.8 by TEMED) at 60°C for 30 min, and used without further purification. RNA secondary structures were predicted using Mfold (23) and the graphics in Figures 2 and 3 and Supplementary Figure S2 were prepared using VARNA (24).

Fluorescence polarization assays

For protein/RNA interaction assays, 9 nM fluorescein-labeled RNAs were mixed with protein samples prepared at concentrations ranging from 10 000 to 0.6 nM by 2-fold dilution. The mixtures (50 μl) were incubated at room temperature for 40 min in 50 mM HEPES, pH 7.5 and 150 mM NaCl in black flat bottom 96-well plates (Costar).

The fluorescence polarization measurements were collected at room temperature with a POLARstar Omega microplate reader (BMG Labtech) using excitation and emission wavelengths of 485 and 520 nM, respectively. The binding affinities (K_d) were determined by non-linear regression analysis for one-site interaction with GraphPad Prism 6. The fluorescence polarization anisotropy F_P was fit using the following quadratic equation, where fitting parameters F_{min}, F_{max} and K_d are fluorescence polarization anisotropy baseline, plateau and dissociation constant, respectively; $[P_T]$ and $[L_T]$ are experimental total protein and total RNA probe concentration, respectively. $[L_T]$ was kept to 9 nM in our experiments.

$$F_p = F_{min} + (F_{max} - F_{min})\{[([P_T] + [L_T] + K_d) -$$
$$\{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]\}^{1/2}]/2[L_T]\}$$

We conducted triplicate technical replicates, a customary sample size that provides the power to detect statistically significant differences, if present. *P*-values were calculated using an unpaired, two-sided t-test without Welch's correction.

Small angle X-ray scattering

SAXS data (0.013 Å^{-1} < q < 0.328 Å^{-1}) were collected at the SIBYLS beamline (12.3.1) of the Advanced Light Source at room temperature. Nop15 (81–220) was exchanged into 25 mM HEPES, pH 7.5, 500 mM NaCl and 2 mM DTT by size exclusion chromatography. The sample was prepared at three concentrations (Supplementary Table S2), and the corresponding chromatography buffer or dialysate was used as SAXS reference.

SAXS data were analyzed with the ATSAS package (2.5.2) (25). Guinier analysis was carried out using PRIMUS to determine radius of gyration (R_g). The Ensemble optimization method (EOM) was used to simulate the conformation ensemble of Nop15[81–220]. The crystal structure of Nop15 (residue 81–180) was input as a rigid domain and the C-terminal region (residues 181–220) was simulated by the EOM. Five thousand conformers were generated for the C-terminal region, and an ensemble of three conformers was selected by genetic algorithm.

NMR experiments

^{15}N-HSQC were collected at 298 K on a 14 T MHz magnet equipped with a cryo-probe for Nop15[81–180] and Nop15[81–220]. The protein samples (∼ 200 μM) were dissolved into 20 mM HEPES, pH 7.5, 100 mM NaCl and 10% D$_2$O. Sixteen transients were accumulated for each fid signal. The NMR data were processed by NMRPipe (26), and peak intensities were measured using NMRViewJ (27).

RESULTS

The Nop15 RNA-binding surface is masked by C-terminal residues

We determined a 2.0 Å crystal structure of yeast Nop15's RRM domain encompassing residues 81–191 (Figure 1, Supplementary Table S1). Our crystal structure confirms that Nop15 contains a core RRM fold similar to the canonical RRM with a four-stranded β sheet and two α helices (β1-α1-β2-β3-α2-β4) (Figure 1B). This core RRM fold is modified with a short additional β strand (β4′), inserted between α2 and β4, that pairs with and extends the β4 strand. Moreover, one or two short α-helices (α3/α4) are appended to the core RRM domain in the four copies of Nop15 in the asymmetric unit (Supplementary Figure S1A). The root mean square deviation ranges from 0.25 to 0.56 Å over all atoms between the four different copies of Nop15, but the C-terminal regions are more variable (residues 171 – C-term) than the core RRM domains (residues 88–170) (Supplementary Figure S1B, Supplementary Table S2). Although all conformations in the crystals may exist in solution, hereafter we present and analyze chain A, because its C-terminal residues are not involved in crystal packing interactions (Supplementary Figure S1C).

The C-terminal residues of Nop15 mask the conserved RNP1/RNP2 RNA-binding motifs in the central β strands (Figure 1A and B). A closer inspection reveals that aliphatic residues (L167, I173 and L176) in the C-terminal region interact with Y94 of RNP1 and F136 of RNP2 (Figure 1C), aromatic residues in canonical RRM domains that typically form stacking interactions with RNA bases. Truncating the C-terminus up to K180 produced soluble Nop15 fragments. However, removing C-terminal residues to the end of the core RRM (Nop15$^{81-169}$), including residues that form hydrophobic interactions with RNP1/RNP2 residues, resulted in aggregated protein, including a portion that eluted in the void volume of a size exclusion column. This suggests that sequestration of the aromatic residues on the RNA-binding surface by C-terminal residues promotes Nop15 solubility.

Regions of Nop15 flanking the core RRM are important for recognition of the ITS2 III.A stem-loop structure

Nop15 has been shown to crosslink *in vivo* to the ITS2 III.A region (nt 26–60), comprising a stem-loop and 3′ single-stranded region (Figure 2A) (8). We developed a fluorescence polarization binding assay to determine the affinity of Nop15 for its pre-rRNA target and to analyze the importance of regions outside the core RRM for RNA recognition (Figure 2B and C). We found that full-length Nop15 bound with an apparent K_d of 293 nM to ITS2 III.A RNA (Table 1). Deletion of the N-terminal region that was not included in our crystal structure (residues 1–80) increased RNA-binding affinity 8 fold (K_d = 35.4 nM for Nop15$^{81-220}$) (Figure 2C and D, Table 1). The N-terminal region contains an electronegative cluster of 19 acidic residues between residues 20–60 that may compete with and therefore weaken RNA binding (Figure 1A). The C-terminal truncation used for crystallization of Nop15 (Nop15$^{81-191}$) bound 6-fold weaker to ITS2 III.A RNA than Nop15$^{81-220}$ (Figure 2D, Table 1). Moreover, removal of an additional 12 residues from the C-terminus (Nop15$^{81-180}$) further weakened RNA-binding affinity to 30-fold weaker than Nop15$^{81-220}$. Together these results indicate that the C-terminal residues are paradoxically important for RNA recognition, despite sequestering the classical RNA-binding residues.

We next examined ITS2 III.A RNA elements that are essential for Nop15 recognition by measuring the binding of Nop15$^{81-220}$ to RNA variants that modified the structure of the ITS2 III.A RNA (Figure 3A and B, Table 1). We found that the 3′ single-stranded region contributes to binding, as its deletion (ITS2 26–51) reduced binding affinity 3-fold (K_d = 105 nM, P = 0.0001 versus ITS2 26–60 WT, $\Delta\Delta G$ = 0.65 kcal/mol). This is consistent with *in vivo* crosslinking of Nop15 to the single-stranded region of ITS2 III.A and increased methylation when Nop15 expression is depleted (8). The integrity of the structure at the stem-loop base is also important for Nop15 recognition, as disrupting two base pairs in the lower stem weakened Nop15 binding 4 fold (K_d = 144 nM, P = 0.0003 versus ITS2 26–60 WT, $\Delta\Delta G$ = 0.84 kcal/mol). In contrast, disruption of the upper stem had no effect on binding affinity (K_d = 35.5 nM, p = 0.96 versus ITS2 26–60 WT). Altering the structure of

the stem by replacing the central bulge with a paired duplex (ITS2 paired) had a small, but significant effect, weakening binding affinity 2.5-fold (K_d = 89 nM, P = 0.0006 versus ITS2 26–60 WT, $\Delta\Delta G$ = 0.56 kcal/mol). Mutations that disrupted the structure of these regions resulted in no mature 25S rRNA accumulation *in vivo* (16), but it is not clear whether the defect can be attributed to disruption of Nop15 binding.

We also tested the effects of sequence mutations that preserved the secondary structure of the ITS2 III.A RNA (Figure 3C, Table 1). Sequence changes to the lower stem, upper stem or bulge had only small effects, if any (up to 2-fold weaker affinity), on Nop15 binding. Equivalent mutations to the upper stem and bulge sequence did not affect mature 25S rRNA accumulation *in vivo* (16), substantiating the relative importance of the RNA structure versus sequence. Consistent with the importance of the structure for recognition, Nop15/ITS2 interaction is salt-sensitive. Increasing the concentration of NaCl to 250 mM decreased RNA binding 10 fold (K_d = 336.9 nM, P = 0.0001 versus 150 mM NaCl, $\Delta\Delta G$ = 0.84 kcal/mol) (Table 1). Together these results suggest that the overall structure of the RNA is important for Nop15 recognition.

The C-terminal residues of Nop15 (181–220) are flexible and refold upon RNA binding

Our binding assays indicate that the residues of Nop15 C-terminal to the α3/α4 helices (181–220) are important for RNA binding, but they were not present in our crystal structure. To characterize the structural behavior of the C-terminus, we used SAXS to study Nop15$^{81-220}$ (Figure 4A, Supplementary Table S3). We evaluated the SAXS data using the EOM, which allows fitting of an ensemble of structures for flexible protein regions. The selected set of three conformers that optimally fit the SAXS data indicated that the C-terminal residues (181–220) are highly flexible, widely sampling around the core RRM domain (Figure 4A and B).

We also analyzed the flexibility of this region by NMR, and the results reinforced the SAXS model. When evaluated by NMR, flexible N- or C-terminal residues display faster tumbling rates than folded domains. These faster tumbling rates correlate with longer transverse relaxation times (T_2) and consequently yield higher peak intensities or narrower resonance linewidths. Therefore, the relationship between molecular tumbling and NMR peak intensity can be used to identify highly flexible regions in proteins (28,29). We collected and measured peak intensities of ^{15}N-HSQC for Nop15$^{81-180}$ and Nop15$^{81-220}$. Although the chemical shifts for residues are unassigned, most resonances of Nop15$^{81-220}$ could be superimposed with those of Nop15$^{81-180}$, the core RRM domain. We assigned the superimposed resonances to the core RRM domain. We attributed the resonances unique to Nop15$^{81-220}$ to the C-terminal residues (Figure 4C). We plotted the intensities of the two groups from the Nop15$^{81-220}$ data set, core RRM (residues 81–180) and C-terminal region (residues 181–220), and calculated the mean intensity for each group, normalizing the mean for N-terminal resonances to 1. The average intensity of the C-terminal resonances was 4 times higher than those of the central RRM domain (P < 0.0001). Therefore, our NMR
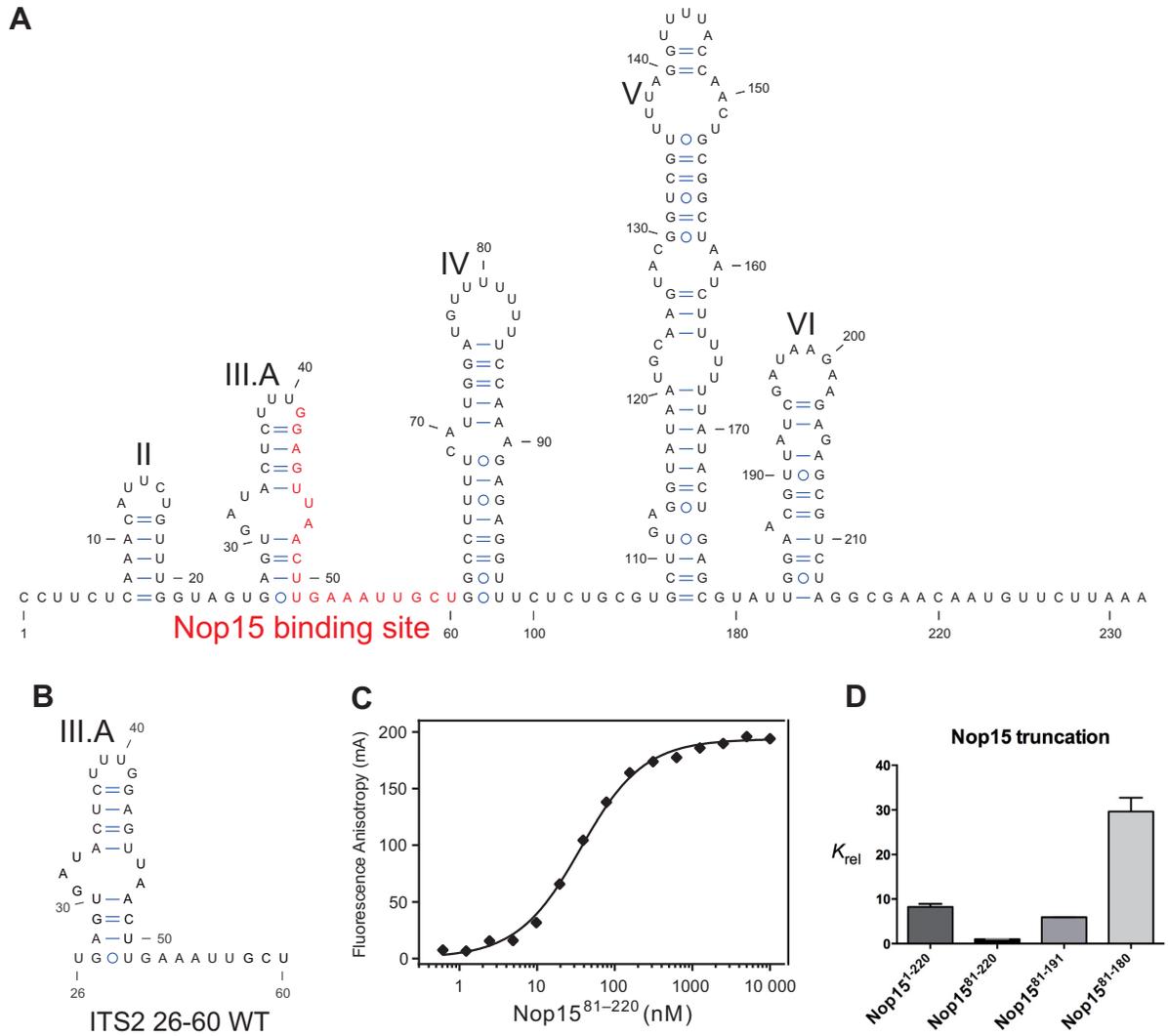
Figure 2. The C-terminal region of Nop15 enhances ITS2 RNA-binding affinity. (**A**) Secondary structure of ITS2 in the ring conformation. A:U, G:C and G:U pairs are denoted by single lines, double lines and circles, respectively. The stem-loop regions are labeled II to VI. Nucleotides that are part of the Nop15 binding region are colored red (8). (**B**) Secondary structure of ITS2 26–60 WT RNA. (**C**) Representative fluorescence polarization binding curve for Nop15$^{81-220}$ binding to ITS2 26–60 WT RNA. (**D**) Relative binding affinities of N- and C-terminally truncated Nop15. K_{rel} for Nop15$^{81-220}$ binding to ITS2 26–60 WT RNA was set equal to 1. Binding assays were conducted in triplicate, and K_{rel} values are shown as mean ± s.e.m..

Table 1. Nop15 RNA-binding analyses

Protein	RNA	K_d (nM)	NaCl (mM)	K_{rel}	*P*-value
Nop15 1–220	ITS2 26–60 WT	293 ± 23	150	8.3	0.0003
Nop15 81–220	ITS2 26–60 WT	35.4 ± 1.3	150	1	-
Nop15 81–191	ITS2 26–60 WT	209 ± 15	150	5.9	0.0003
Nop15 81–180	ITS2 26–60 WT	1048 ± 111	150	30	0.0008
Nop15 81–220	ITS2 26–51	104.5 ± 3.6	150	3	0.0001
Nop15 81–220	ITS2 26–60 disrupted lower	144 ± 16	150	4.1	0.0003
Nop15 81–220	ITS2 26–60 disrupted upper	35.5 ± 3.2	150	1	0.96
Nop15 81–220	ITS2 26–60 paired	89 ± 8.9	150	2.5	0.0006
Nop15 81–220	ITS2 26–60 lower1	52.1 ± 3.6	150	1.5	0.0016
Nop15 81–220	ITS2 26–60 lower2	39.3 ± 2.7	150	1.1	0.087
Nop15 81–220	ITS2 26–60 upper	44.9 ± 4.3	150	1.3	0.022
Nop15 81–220	ITS2 26–60 bulge	73.6 ± 5.6	150	2.1	0.0003
Nop15 81–220	ITS2 26–60 WT	337 ± 19	250	9.5	0.0001

Protein and RNA constructs are indicated. Mean K_d ± standard error of the mean (s.e.m.) from three technical replicates are shown. K_{rel} values were calculated relative to the K_d of Nop15$^{81-220}$ binding to ITS2 26–60 WT RNA. *P*-values were calculated using an unpaired, two-sided t-test without Welch's correction.
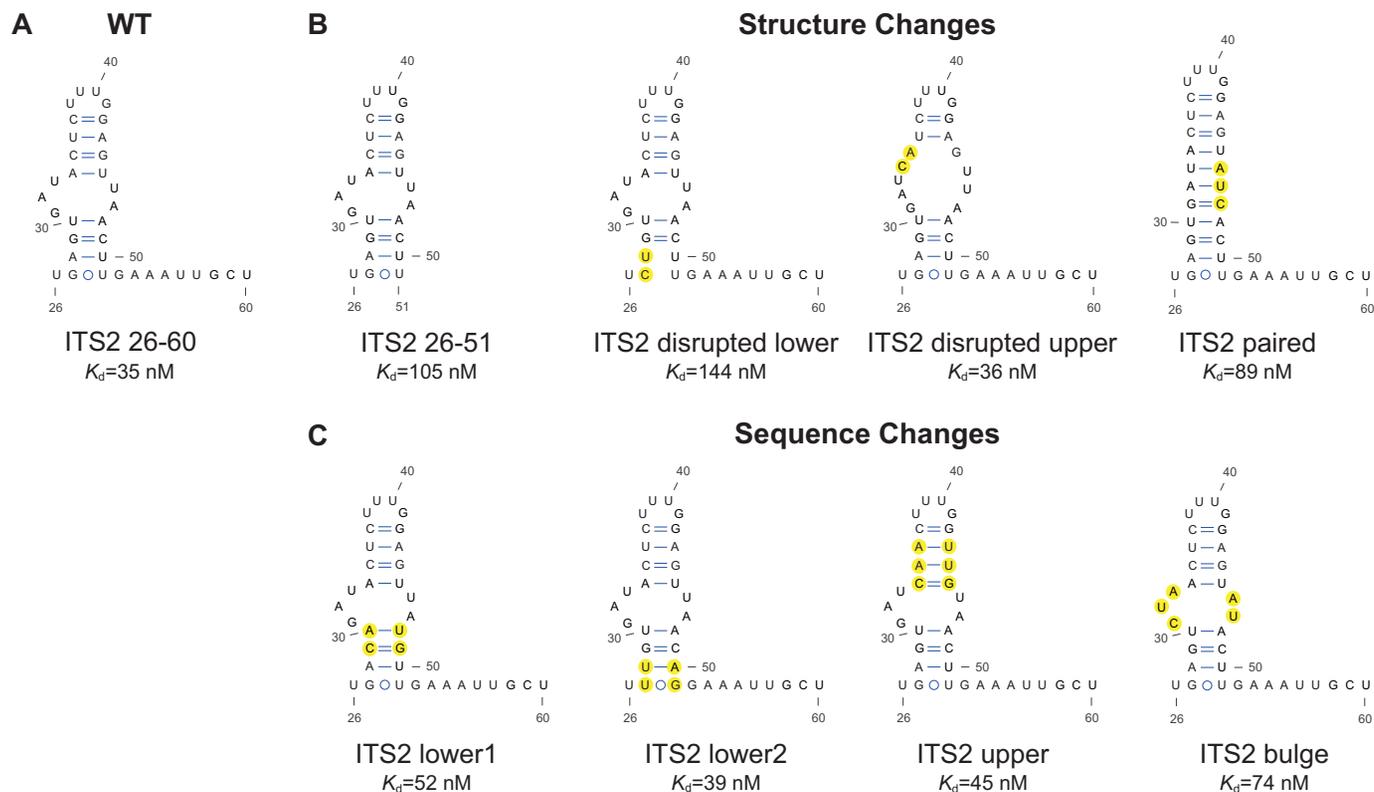
Figure 3. Nop15 recognizes the overall structure of ITS2 RNA. Secondary structure of ITS2 variants tested and K_d for Nop15[81–220] binding are shown for (**A**) ITS2 26–60 WT, (**B**) variations that change the structure of the ITS2 III.A region and (**C**) variations that change the sequence of the ITS2 III.A region. Variations are highlighted yellow. Mutant ITS2 disrupted lower is similar to III.2, mutant ITS2 disrupted upper is similar to F&G, mutant ITS2 paired is the same as 3A-close, mutant ITS2 upper is the same as 3A-3 and mutant ITS2 bulge is the same as 3A-change in Cote *et al.* (16).

results indicate that the C-terminal region tumbles as a flexible moiety at a rate faster than the core RRM domain (Figure 4D).

A recent cryo-EM study of a pre-60S ribosomal complex includes Nop15 and its ITS2 binding site (Figure 5) (17). In contrast to our crystal structure of Nop15 (Figure 5A), the α3/α4 helices are unfolded in the pre-rRNA-bound Nop15 and the structure of this region is different from the conformations represented by the four molecules in our crystal structure. The conformational change when bound to pre-rRNA exposes the RNP1 and RNP2 motifs to allow interaction with nucleotides U56 and U57, respectively (Figure 5B). In addition, residues R132 and H133 in the loop between β2 and β3 form stacking interactions with nucleotides G58 and A55, respectively. The interactions with the 3′ ITS2 nucleotides explain the importance of this 3′ flanking region for Nop15 binding affinity (Figure 3B). Residues (H100, N125, N130) appear positioned to interact with the RNA structure at the stem-loop base, whose integrity we showed to be important for interaction (Figure 3B).

DISCUSSION

The core βαββαβ RRM fold is conserved across different classes of RRMs (1,6,30). However, flanking helices can modify this prototype. These flanking helices may block the canonical RNP1/RNP2 RNA-binding surface, as exemplified by the N-terminal α-helix of ABH8 (31), the C-terminal α-helices of La (32), U1A (33) and hnRNP F (34,35), and the N- and C-terminal α-helices of Prp24 RRM4 (36) and CstF-64 (37). In Prp24 RRM4 and hnRNP F, binding via the RNP1/RNP2 sites is further dampened by substitutions for the aromatic residues that would typically form stacking interactions. Instead, new RNA-binding interfaces are formed, such as the surface of the additional α-helices in RRM4 of Prp24 (36) or loop residues in human hnRNP F (35). In this study, we found that the core RRM domain of Nop15 is appended by one or two short α-helices that obstruct the classical RNA-binding surface. However, in contrast to hnRNP F (34,35), consensus RNP1 and RNP2 motifs of Nop15 are intact and they are used to recognize ITS2 RNA in a cryo-EM model of the pre-60S ribosome (17). Moreover, the C-terminal region that blocks access to the RNP1/2 motifs is rearranged to form protein/RNA and protein/protein interactions in the pre-60S ribosome.

The cryo-EM model suggests that unwinding the α3/α4 helices allows extension of the C-terminal region of Nop15 away from the core RRM and formation of a new terminal α-helix (αC, residues 193–214) to interact with the distal end of the ITS2 stem-loop (Figure 5B). These observed differences suggest that the α3/α4 helices in Nop15 may shield the hydrophobic RNP1/RNP2 residues in the absence of RNA. During RNA binding, the α3/α4 helices appear to unfold to expose the RNP1 and RNP2 motifs. The flexibil-
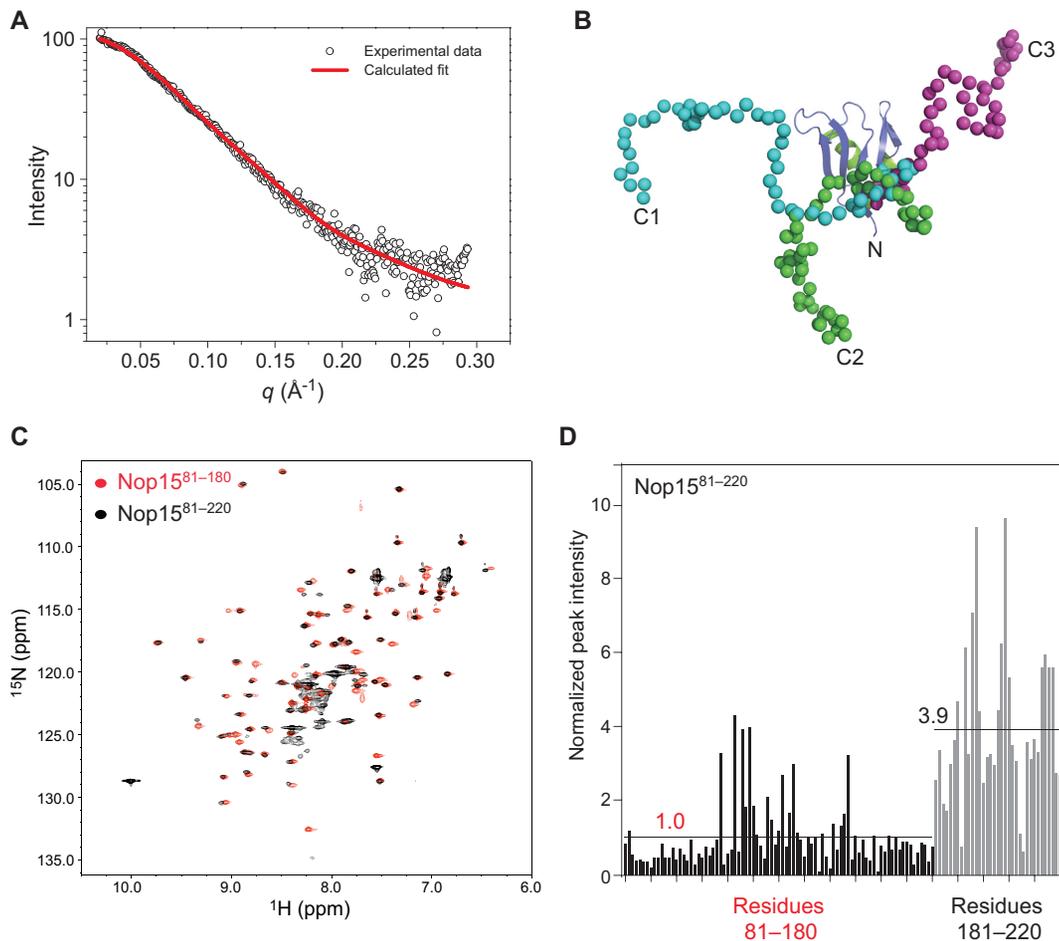
Figure 4. The C-terminal region of Nop15 is flexible in solution. (**A**) Superposition of the SAXS curve of Nop15[81–220] (black) and the calculated fit of the EOM ensemble (red). The χ^2 for the calculated fit of the EOM ensemble and the corresponding experimental data is 0.98. The χ^2 for the fitting of a single conformer and the experimental data is 1.52. In contrast, a model lacking the C-terminal region (Nop15[81–191]) results in a χ^2 value of 7.24. (**B**) SAXS model of the EOM ensemble comprising three representative conformers. The core RRM domain is shown as a ribbon diagram and the three C-terminal regions are shown as blue, green or magenta spheres. (**C**) ^{15}N-HSQC of Nop15[81–180] (red) and Nop15[81–220] (black). (**D**) Peak intensities measured from ^{15}N-HSQC of Nop15[81–220]. Peaks that were observed for both Nop15[81–180] and Nop15[81–220] were assigned to residues 81–180 and are shown as black bars. The remaining non-overlapping intensities were assigned to the C-terminal region (residues 181–220) and are shown as grey bars. The peak intensities for each region are ordered along the x-axis in the graph according to the automated peak picking using NMRViewJ (27). The mean peak intensities for each region were calculated. The mean peak intensity of residues 81–180 was normalized to 1, and the mean peak intensity of residues 181–220 relative to residues 81–180 is indicated. The black bars indicate the mean values for each group.

ity of the C-terminal region that we observed may lower the energy barrier to extend away from the core RRM and bind the distal RNA loop. Extension of the C-terminal residues may create an arm corresponding to the length of the ITS2 III.A stem-loop. The Nop15 αC helix also interacts with Cic1, which binds to stem-loop II of ITS2 (Figure 5C). This protein–protein interaction may also contribute to stabilizing the ring conformation. Although an αC helix is predicted based on sequence analysis, our NMR results indicated that the C-terminus is flexible in the absence of RNA, and therefore, the αC helix is induced by RNA or protein binding.

During pre-rRNA processing, ITS2 switches from the ring conformation to the hairpin conformation (14–16). The ring conformation is characterized by multiple local stem-loops that are required for early-stage ITS2 assembly (Figure 6 and Supplementary Figure S2). In contrast, the

hairpin conformation contains more extensive and longer-distance base pairing. Therefore, the ring conformation is favorable kinetically, but unfavorable energetically. A major rearrangement involves nucleotides 26–60, the Nop15 binding site, which form a local stem loop (III.A) in the ring conformation, but must unfold and become paired with nucleotides 214–231 in the hairpin conformation (Supplementary Figure S2). Nop15 appears to hold ITS2 in the ring conformation by recognizing and stabilizing the III.A stem-loop, which would also permit other ribosome biogenesis factors to assemble.

Bioinformatics analysis (23) predicts that ITS2 could be trapped in incorrect conformations of similar stabilities (Figure 6). Our data suggest that Nop15 binding may also prevent III.A mispairing and reduce ITS2 misfolding by recognizing the base of the III.A stem and single-stranded nucleotides 52–57 in the ring conformation. As ribosome bio-
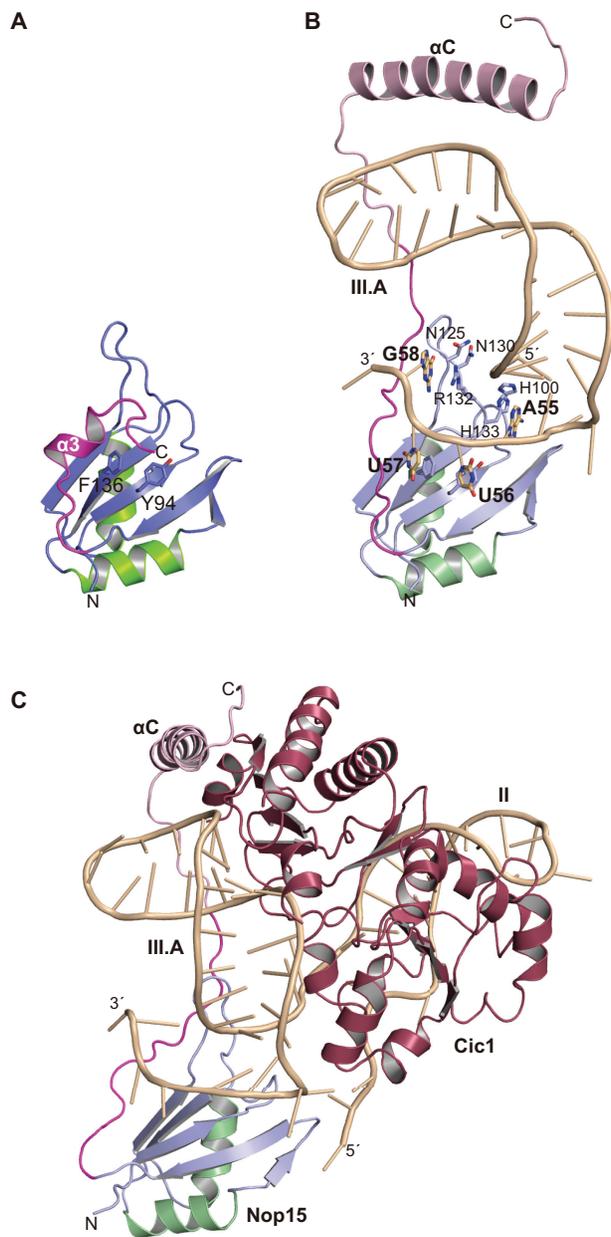
Figure 5. The C-terminal region of Nop15 refolds to interact with ITS2 RNA. (**A**) Ribbon diagram of a crystal structure of Nop15. (**B**) Cryo-EM model of Nop15/ITS2 from the structure of a pre-60S ribosome (PDB ID: 3JCT) [17]. ITS2 nucleotides 26–59 that were visible in the cryo-EM model are shown. The α3/α4 region is colored magenta in both structures. Residues in Nop15 that contact the nucleotides are shown. The surface area buried at the Nop15/ITS2 RNA interface is 3234 Å^2, calculated as the sum of the solvent accessible surface area of the Nop15 crystal structure and ITS2 RNA from the pre-60S ribosome cryo-EM model minus the solvent accessible surface area of the complex of Nop15/ITS2 from the cryo-EM model using PyMol (Schroedinger) with water represented as a sphere of radius 1.4 Å. This results in a surface energy density of 3.2 cal/mol Å^2, which is consistent with the expected range of 3–4 cal/mol Å^2 for protein–protein interfaces burying the equivalent surface area [38]. (**C**) Cryo-EM model of Nop15/ITS2/Cic1 from the structure of a pre-60S ribosome (PDB ID: 3JCT) [17]. ITS2 nucleotides 1–59 that were visible in the cryo-EM model are shown. Cic1 is shown in red.
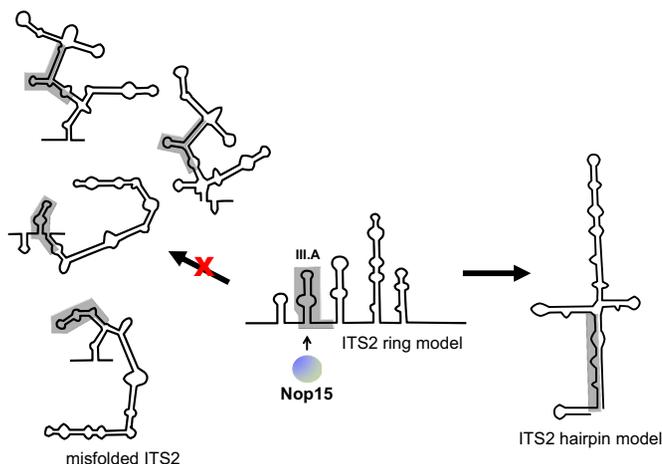
Figure 6. Nop15 recognizes the structure of ITS2 III.A, promoting the ring conformation of ITS2 over the hairpin conformation and preventing misfolding into other energetically equivalent forms. Nucleotides 26–60 in each RNA conformation are indicated by grey shading. The cartoons of misfolded ITS2 RNAs were based on four predicted energy equivalent conformations [23].

genesis proceeds, Nop15 is lost from the pre-ribosome before export to the cytoplasm [17], and this release allows continued maturation of the pre-rRNA folding through the hairpin conformation as other ribosome biogenesis factors and ribosomal proteins join the pre-ribosome.

DATA AVAILABILITY/ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank with accession ID 5T9P.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Maris,C., Dominguez,C. and Allain,F.H. (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J.*, **272**, 2118–2131.
2. Kuhn,U. and Wahle,E. (2004) Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta*, **1678**, 67–84.
3. Birney,E., Kumar,S. and Krainer,A.R. (1993) Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.*, **21**, 5803–5816.
4. Teplova,M., Farazi,T.A., Tuschl,T. and Patel,D.J. (2015) Structural basis underlying CAC RNA recognition by the RRM domain of dimeric RNA-binding protein RBPMS. *Q. Rev. Biophys.*, 1–11.
5. Clery,A., Blatter,M. and Allain,F.H. (2008) RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.*, **18**, 290–298.
6. Daubner,G.M., Clery,A. and Allain,F.H. (2013) RRM-RNA recognition: NMR or crystallography...and new findings. *Curr. Opin. Struct. Biol.*, **23**, 100–108.
7. Conte,M.R., Grune,T., Ghuman,J., Kelly,G., Ladas,A., Matthews,S. and Curry,S. (2000) Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. *EMBO J.*, **19**, 3132–3141.
8. Granneman,S., Petfalski,E. and Tollervey,D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J.*, **30**, 4006–4019.
9. Oeffinger,M. and Tollervey,D. (2003) Yeast Nop15p is an RNA-binding protein required for pre-rRNA processing and cytokinesis. *EMBO J.*, **22**, 6573–6583.
10. Woolford,J.L. Jr and Baserga,S.J. (2013) Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics*, **195**, 643–681.
11. Lafontaine,D.L. (2015) Noncoding RNAs in eukaryotic ribosome biogenesis and function. *Nat. Struct. Mol. Biol.*, **22**, 11–19.
12. Henras,A.K., Plisson-Chastang,C., O'Donohue,M.F., Chakraborty,A. and Gleizes,P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA*, **6**, 225–242.
13. de la Cruz,J., Karbstein,K. and Woolford,J.L. Jr2015) Functions of ribosomal proteins in assembly of eukaryotic ribosomes *in vivo*. *Annu. Rev. Biochem.*, **84**, 93–129.
14. Yeh,L.C. and Lee,J.C. (1990) Structural analysis of the internal transcribed spacer 2 of the precursor ribosomal RNA from Saccharomyces cerevisiae. *J. Mol. Biol.*, **211**, 699–712.
15. Joseph,N., Krauskopf,E., Vera,M.I. and Michot,B. (1999) Ribosomal internal transcribed spacer 2 (ITS2) exhibits a common core of secondary structure in vertebrates and yeast. *Nucleic Acids Res.*, **27**, 4533–4540.
16. Cote,C.A., Greer,C.L. and Peculis,B.A. (2002) Dynamic conformational model for the role of ITS2 in pre-rRNA processing in yeast. *RNA*, **8**, 786–797.
17. Wu,S., Tutuncuoglu,B., Yan,K., Brown,H., Zhang,Y., Tan,D., Gamalinda,M., Yuan,Y., Li,Z., Jakovljevic,J. *et al.* (2016) Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature*, **534**, 133–137.
18. Mossessova,E. and Lima,C.D. (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell*, **5**, 865–876.
19. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Method Enzymol.*, **276**, 307–326.
20. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
21. McCoy,A.J., Grosse-Kunstleve,R.W., Adams,P.D., Winn,M.D., Storoni,L.C. and Read,R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.
22. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
23. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.
24. Darty,K., Denise,A. and Ponty,Y. (2009) VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
25. Petoukhov,M.V., Konarev,P.V., Kikhney,A.G. and Svergun,D.I. (2007) ATSAS 2.1 – towards automated and websupported small-angle scattering data analysis. *J. Appl. Cryst.*, **40**, s223–s228.
26. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
27. Johnson,B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.*, **278**, 313–352.
28. Nematollahi,L.A., Garza-Garcia,A., Bechara,C., Esposito,D., Morgner,N., Robinson,C.V. and Driscoll,P.C. (2015) Flexible stoichiometry and asymmetry of the PIDDosome core complex by heteronuclear NMR spectroscopy and mass spectrometry. *J. Mol. Biol.*, **427**, 737–752.
29. Zhang,J., Lewis,S.M., Kuhlman,B. and Lee,A.L. (2013) Supertertiary structure of the MAGUK core from PSD-95. *Structure*, **21**, 402–413.
30. Muto,Y. and Yokoyama,S. (2012) Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. *Wiley Interdiscip. Rev. RNA*, **3**, 229–246.
31. Pastore,C., Topalidou,I., Forouhar,F., Yan,A.C., Levy,M. and Hunt,J.F. (2012) Crystal structure and RNA binding properties of the RNA recognition motif (RRM) and AlkB domains in human AlkB homolog 8 (ABH8), an enzyme catalyzing tRNA hypermodification. *J. Biol. Chem.*, **287**, 2130–2143.
32. Jacks,A., Babon,J., Kelly,G., Manolaridis,I., Cary,P.D., Curry,S. and Conte,M.R. (2003) Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure*, **11**, 833–843.
33. Nagai,K., Oubridge,C., Jessen,T.H., Li,J. and Evans,P.R. (1990) Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature*, **348**, 515–520.
34. Dominguez,C. and Allain,F.H. (2006) NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Res.*, **34**, 3634–3645.
35. Dominguez,C., Fisette,J.F., Chabot,B. and Allain,F.H. (2010) Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nat. Struct. Mol. Biol.*, **17**, 853–861.
36. Martin-Tumasz,S., Richie,A.C., Clos,L.J. 2nd, Brow,D.A. and Butcher,S.E. (2011) A novel occluded RNA recognition motif in Prp24 unwinds the U6 RNA internal stem loop. *Nucleic Acids Res.*, **39**, 7837–7847.
37. Perez Canadillas,J.M. and Varani,G. (2003) Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *EMBO J.*, **22**, 2821–2830.
38. Chen,J., Sawyer,N. and Regan,L. (2013) Protein-protein interactions: general trends in the relationship between binding affinity and interfacial buried surface area. *Protein Sci.*, **22**, 510–515.