The FEBS Journal

FEBS PRESS science publishing by scientists

Structural investigation of human U6 snRNA recognition by spliceosomal recycling factor SART3 RNA recognition motifs

Iktae Kim[1,4], Kyeong-Mi Bang[2,3], So Young An[1,2], Changkon Park[1], Ji-Yeon Shin[2,3], Youngim Kim[1,*], Hyun Kyu Song[3], Jeong-Yong Suh[1,4] and Nak-Kyoon Kim[2]

1 Department of Agricultural Biotechnology, Seoul National University, Seoul, Korea
2 Advanced Analysis Data Center, Korea Institute of Science and Technology, Seoul, Korea
3 Department of Life Sciences, Korea University, Seoul, Korea
4 Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, Korea

Keywords
RNA recognition motif; SART3; solution structure; U6 snRNA

Correspondence
J. Y. Suh, Department of Agricultural Biotechnology, Seoul National University, Seoul 08826, Korea
Tel: +82-2-880-4879
E-mail: jysuh@snu.ac.kr
and
N. K. Kim, Advanced Analysis Data Center, Korea Institute of Science and Technology, Seoul 02792, Korea
Tel: +82-2-958-5996
E-mail: nkkim@kist.re.kr

Present address
*Dong-A ST R&D Center, Incheon, Korea

Iktae Kim, Kyeong-Mi Bang and So Young An contributed equally to this work.

(Received 19 March 2025, revised 12 July 2025, accepted 19 September 2025)

doi:10.1111/febs.70275

Human spliceosome-associated factor 3, SART3, is a key factor in spliceosome recycling and engages with U6 small nuclear RNA (snRNA) to promote the formation of the U4/U6 small nuclear ribonucleoprotein complex. Unlike its counterpart U4/U6 snRNA-associated-splicing factor PRP24 (Prp24) from *Saccharomyces cerevisiae*, which uses four RNA recognition motifs (RRMs) for the U6 snRNA interaction, SART3 has two RRMs at its C terminus. Here, we demonstrate that SART3 binds U6 snRNA as a dimer, and four RRM subunits recognize the asymmetric bulge of U6 snRNA. SART3 RRMs adopt a tandem βαββαβ motif of the canonical RRM fold to interact with the U6 bulge region via a conserved electropositive surface. We identified the cognate U6 elements that specifically bind SART3 RRM1, which is distinct from the Prp24–U6 interactions in yeast. Our findings suggest a divergent RRM binding mechanism for U6 snRNA recognition during spliceosome assembly and recycling.

Introduction

RNA splicing converts a precursor messenger RNA (pre-mRNA) into its mature form by removing introns and ligating adjacent exons in eukaryotes [1]. This process is facilitated by the spliceosome, a large complex comprising five small nuclear ribonucleoproteins (snRNPs; U1, U2, U4, U5, and U6) and their

Abbreviations
CSP, chemical shift perturbation; EMSA, electrophoretic mobility shift assay; HAT, half-a-tetracopeptide; HSQC, heteronuclear single quantum correlation; ISL, internal stem loop; MALS, multi-angle light scattering; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser enhancement spectroscopy; pre-mRNA, precursor messenger RNA; RMSD, root-mean-square deviation; RRM, RNA recognition motifs; SART3, squamous cell carcinoma antigen recognized by T cells 3; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein; TOCSY, total correlation spectroscopy.

associated proteins. The formation of the spliceosome involves the initial recognition of pre-mRNA introns by U1 and U2 snRNPs, followed by the recruitment of a preassembled U4, U5, and U6 snRNP complex for catalytic splicing. Dysregulation of this process can lead to the production of nonfunctional proteins implicated in various human diseases, emphasizing the importance of ordered spliceosome assembly and function [2–4]. A hallmark of spliceosome formation is the dynamic participation of individual snRNPs that transiently join or exit the spliceosomal assembly and undergo a series of large conformational rearrangements required for splicing. U6 snRNP is important in substrate binding and structural reorganization of the spliceosome throughout the splicing cycle [1,5].

Squamous cell carcinoma antigen recognized by T cells 3 (SART3), a mammalian homolog of yeast Prp24 from *Saccharomyces cerevisiae*, associates with U6 small nuclear RNA (snRNA) and promotes the formation of the U4/U6 complex [6]. Human SART3 has been detected in the binary U4/U6 snRNP complex in the nucleus, but not in the ternary U4/U6·U5 snRNP complex or the spliceosome assembly [6,7]. Therefore, it is supposed that SART3 participates in the early stage of snRNP assembly formation, mainly by recruiting U4 and U6 snRNPs. Human SART3 and yeast Prp24 share homologous RNA recognition motifs (RRMs), but they have different domain organizations. SART3 comprises 12 half-a-tetracopeptide (HAT) repeats and a nuclear localization sequence, followed by two RRMs and an LSm-binding domain [8,9]. By contrast, Prp24 contains four RRMs and a C-terminal LSm-binding domain for efficient snRNP complex formation [9,10].

In the Prp24–U6 snRNA complex, four RRMs of Prp24 associate tightly with the asymmetric bulge region of U6 snRNA [11]. Human SART3 could interact with U6 snRNA in a similar manner, but how two instead of four RRMs recognize U6 snRNA is unclear. Here we report that the RRM1 domains of two SART3 simultaneously bind U6 snRNA, so that dimeric SART3 can employ four RRMs for U6 snRNA binding. SART3 RRM1 specifically binds the U6 snRNA bulge region, and RRM2 enhances the binding affinity. Remarkably, the U6 sequence for SART3 binding is distinct from that of Prp24, suggesting divergent evolution of RNA recognition for splicing. The solution structures of SART3 RRM1 and RRM2 show that they both adopt a canonical βαββαβ fold and interact with the U6 bulge region via conserved electropositive surfaces. We propose a structural model of SART3 RRM1 bound to U6 snRNA based on nuclear magnetic resonance (NMR) titration, calorimetry, and fluorescence anisotropy combined with site-directed mutagenesis.

Results

Design of U6 snRNA and SART3 domain constructs

We explored the domain constructs of human U6 snRNA and SART3 (Fig. 1). We used a model of human U6 snRNA that incorporated the telestem, asymmetric bulge, and internal stem loop (ISL) regions, based on the crystal structure of yeast U6 snRNA (PDB code 4N0T; Fig. 1A and B). The crystal structure revealed that yeast U6 snRNA interacts with yeast Prp24 mainly via its asymmetric bulge and part of the telestem region (Fig. 1A, red shading) [11]. We hypothesized that human U6 snRNA would bind SART3 with similar interfaces. U6 snRNA with full stem and loop structures (U6$_{\mathrm{FL}}$; nucleotides 24–95) was prepared using *in vitro* transcription. Nucleotides in the telestem region were engineered to form Watson–Crick base pairs to enhance the stability of U6$_{\mathrm{FL}}$ (Fig. 1C). For the fluorescence measurement, we prepared FAM-labeled U6 snRNA with reduced stem regions that maintained the binding interface. The truncated U6 snRNA (U6$_{\mathrm{TR}}$; nucleotides 30–58 and 77–89) contained the asymmetric bulge together with nearby flanking telestem and ISL regions (Fig. 1D). The telestem of U6$_{\mathrm{TR}}$ was capped with GC base pairs at the termini, and the ISL loop was replaced with a stable UUCG tetraloop for stability (Fig. 1D). Lastly, we prepared U6 snRNA spanning nucleotides 33–54 (U6$_{33–54}$) to evaluate the function of the bulge region in SART3 binding (Fig. 1E).

Human SART3 (residues 1–963) comprises N-terminal HAT repeats, followed by two RRM domains at the C terminus. By contrast, the yeast homolog Prp24 contains four RRM domains without the HAT repeats (Fig. 1F). To investigate the functions of individual RRM domains in U6 snRNA binding, we prepared RRM domains of SART3 (RRM$_{1–2}$) and separate constructs for RRM1 and RRM2 (Fig. 1F). In addition, to investigate the effect of HAT-mediated dimerization of SART3 on the interaction with U6 snRNA, we prepared the N-terminal HAT repeat domain linked to RRM$_{1–2}$ (HAT-RRM).

Two SART3 RRMs bind U6 snRNA at the asymmetric bulge region

We measured the binding affinity of SART3 RRM to U6 snRNA using the fluorescence polarization assay (FPA). The equilibrium dissociation constant (K_{D}) of

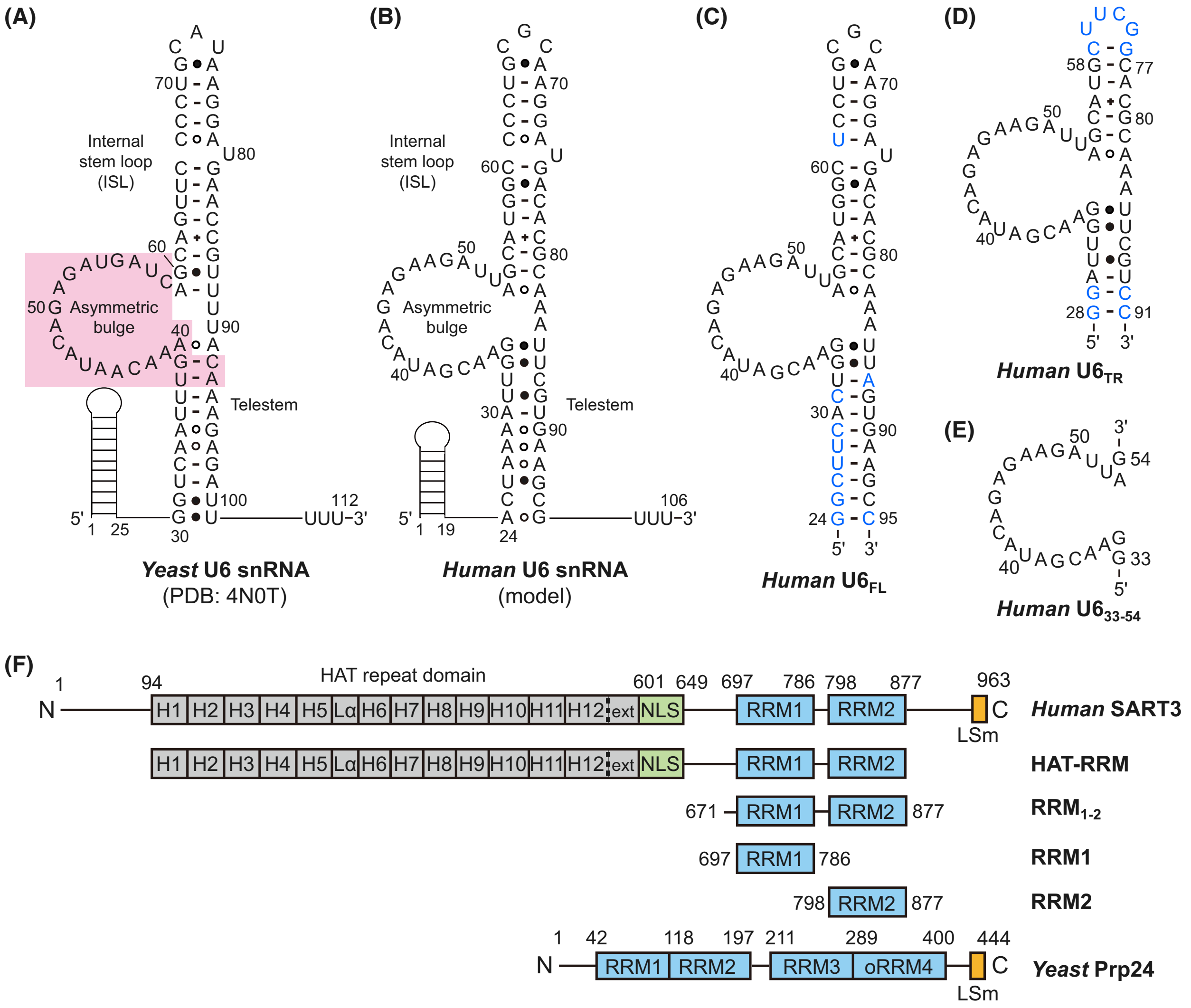

Fig. 1. Schematic of the U6 snRNA and SART3 domains, and the constructs used in this study. Secondary structures of (A) yeast U6 snRNA and (B) human U6 snRNA. The binding interfaces of yeast U6 snRNA for Prp24 are shaded in *red*. RNA constructs for (C) human U6$_{FL}$, (D) U6$_{TR}$ with truncated stems, and (E) U6$_{33-54}$ with an isolated bulge were used in this study. Mutations to stabilize the U6 secondary structures are in *blue*. (F) Domain organization of the human SART3 constructs used in this study and yeast Prp24. SART3 comprises 12 half-a-tetracopeptide (HAT) repeats, 2 RNA recognition motifs (RRMs), and 1 LSm-interaction motif. Other features include an extended helix (ext), linker α-helix (Lα), and nuclear localization sequence (NLS).

RRM$_{1-2}$ binding to fluorescein-labeled U6$_{TR}$ was determined to be 153 ± 10 nM (Fig. 2A). Slightly reduced binding affinity was observed for RRM$_{1-2}$ to U6$_{33-54}$, with a K_D of 188 ± 5 nM, demonstrating that the asymmetric bulge of U6 snRNA was the primary recognition site for SART3 RRM$_{1-2}$ (Fig. 2B). The small difference in affinity suggests that the U6 stem region flanking the bulge may also contribute to SART3 RRM binding. We next investigated the contribution of individual RRM1 and RRM2 to interactions with the U6 bulge and obtained a K_D of 1020 ± 160 nM for isolated RRM1 and U6$_{33-54}$, a fivefold weaker affinity compared to RRM$_{1-2}$ (Fig. 2C). The binding of RRM1 to U6$_{33-54}$ deviated slightly from the Langmuir adsorption isotherm, potentially indicating non-identical binding sites. RRM2 alone exhibited significantly reduced binding to U6$_{33-54}$, such that the lower bound of the K_D was estimated to be >50 μM (Fig. 2D). Our results demonstrate that SART3 RRM$_{1-2}$ recognizes the asymmetric bulge of U6 snRNA, with RRM1 playing a dominant role in the binding affinity.

Electrophoretic mobility shift assay (EMSA) showed that the U6$_{FL}$ band migrated in a stepwise manner

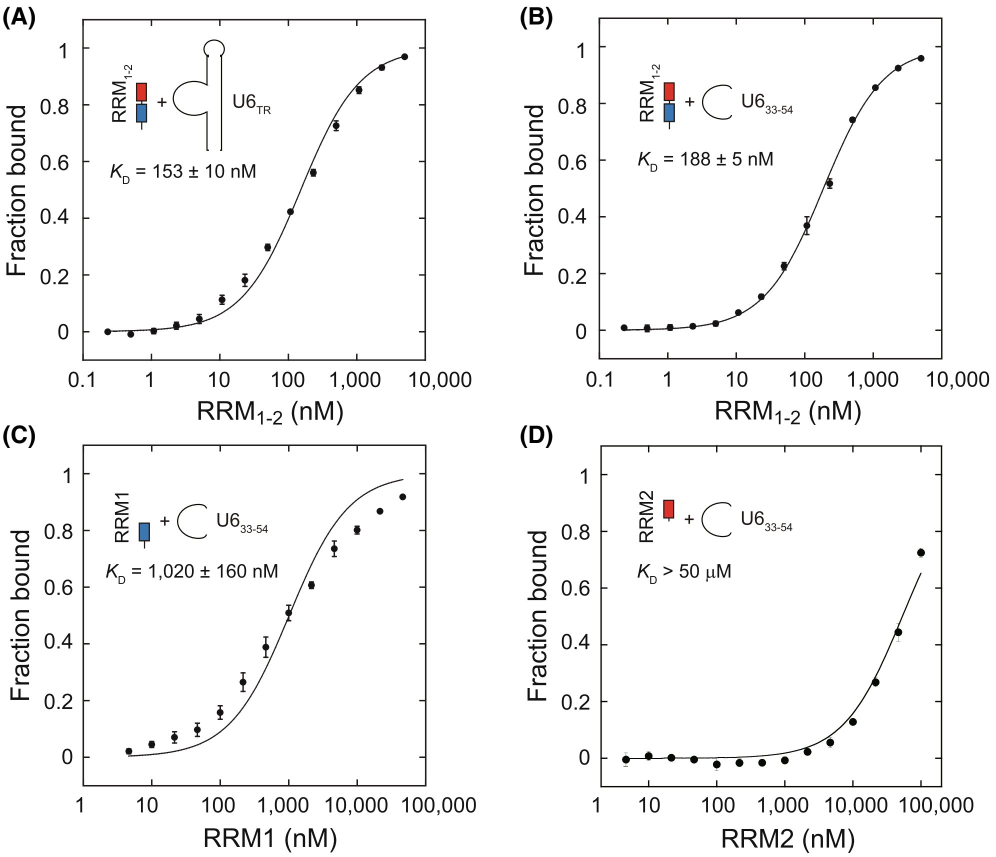

Fig. 2. Binding curves and K_D values from fluorescence polarization assays of the SART3 RNA recognition motifs (RRMs) and fluorescein-labeled U6 snRNA constructs. (A) $RRM_{1–2}$ with $U6_{TR}$, (B) $RRM_{1–2}$ with $U6_{33–54}$, (C) RRM1 with $U6_{33–54}$, and (D) RRM2 with $U6_{33–54}$. Constructs are shown as cartoon diagrams. Equilibrium dissociation constants (K_D) obtained by curve fitting are provided. The measurements were performed in triplicate, and average values and standard errors of the mean are reported.

upon forming a complex with $RRM_{1–2}$, suggesting that $RRM_{1–2}$ associates with $U6_{FL}$ at two binding sites (Fig. 3A). Similar band shifts were observed with $U6_{33–54}$, indicating that the $U6_{33–54}$ bulge region is capable of binding two $RRM_{1–2}$ molecules (Fig. 3B). The molecular mass of the $RRM_{1–2}$ (calculated mass 23.4 kDa) and $U6_{FL}$ (calculated mass 23.3 kDa) complex was 62 ± 7 kDa by multi-angle light scattering (MALS) coupled with size exclusion chromatography (SEC), which aligns with the 2:1 stoichiometry (Fig. 3C). Similarly, the molecular mass of the $RRM_{1–2}$ and $U6_{33–54}$ (calculated mass 7.2 kDa) complex was 52 ± 5 kDa based on the MALS data, further supporting the complexation of two $RRM_{1–2}$ and one $U6_{33–54}$ (Fig. 3D). We note that free $RRM_{1–2}$ eluted as a monomer from the size exclusion chromatography (Fig. S1). RRM1 alone also induced band shifts of $U6_{33–54}$, whereas RRM2 alone did not show noticeable band shifts due to its low affinity (Fig. S2A,B). In addition, the band shifts of RRM1 did not change in the presence of RRM2, and the band shifts of RRM2 did not change in the presence of RRM1, indicating that they do not compete for the same binding site on $U6_{33–54}$, nor do they cooperate with each other in the absence of the connecting linker (Fig. S2C,D). We carried out isothermal titration calorimetry (ITC) to analyze the binding of RRM1 and RRM2 to $U6_{33–54}$. RRM1 binding to $U6_{33–54}$ resulted in a 2:1 stoichiometry and K_D of 940 ± 160 nM, whereas RRM2 showed negligible binding heats, which were consistent with the FPA results (Fig. S2E,F). Taken together, our findings demonstrate that two SART3 $RRM_{1–2}$ bind the U6 bulge region and that RRM1 plays a primary role in the recognition of U6 snRNA.

Crystal structure analysis and cellular assays have shown that the isolated HAT domain of SART3 forms a functional dimer [12–14]. MALS analysis of HAT-RRM indicated a stable dimeric state in solution, where HAT-RRM (calculated mass 90.9 kDa) eluted as a single peak with an absolute molecular

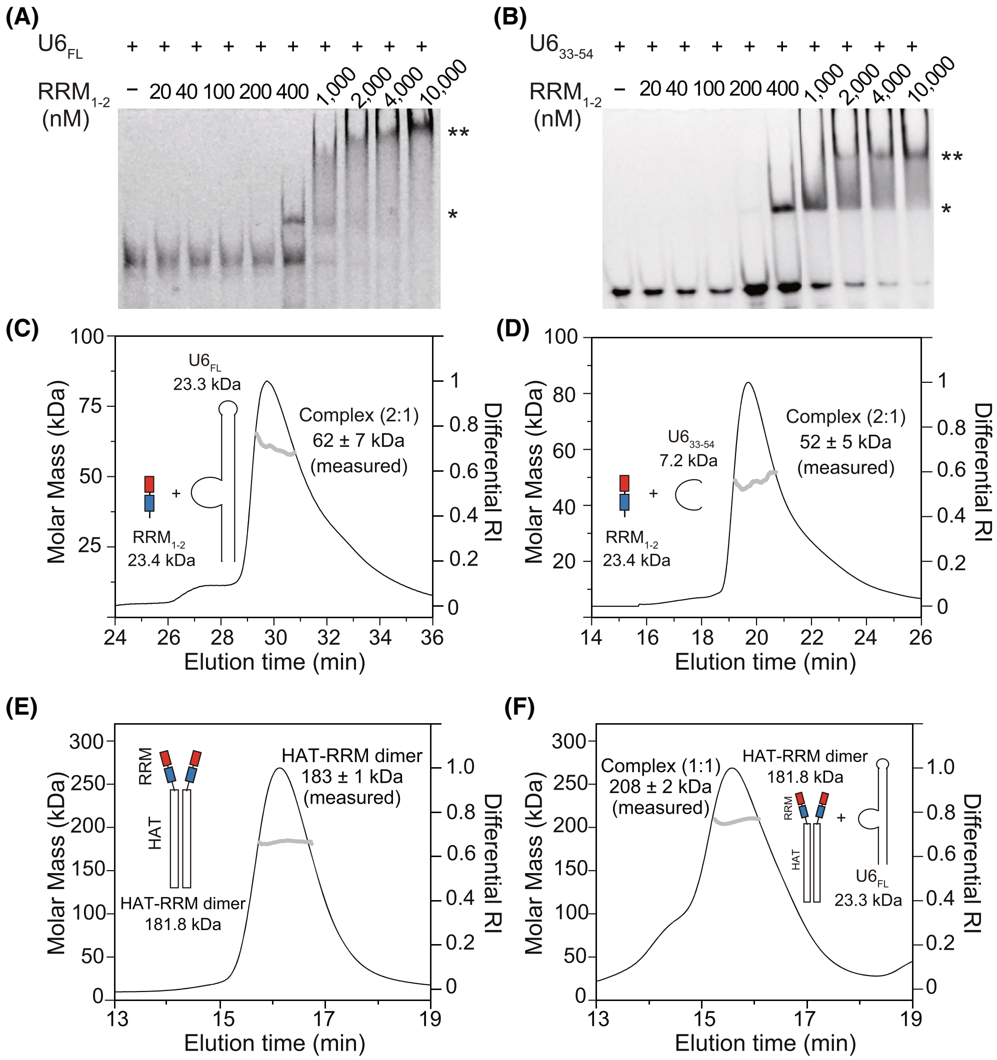

Fig. 3. Interaction between SART3 RNA recognition motifs (RRMs) and U6 snRNA. Gel shift profiles for the titration of (A) 20 nM FAM-U6$_{FL}$ and (B) 20 nM FAM-U6$_{33–54}$ with increasing RRM$_{1–2}$ (nM). Stepwise migration of U6 bands is marked by asterisks and double asterisks. Light scattering data of (C) the RRM$_{1–2}$ and U6$_{FL}$ complex, (D) the RRM$_{1–2}$ and U6$_{33–54}$ complex, (E) the HAT-RRM dimer, and (F) the HAT-RRM and U6$_{FL}$ complex. The differential refractive index is shown as a solid *black* line (y-axis on the *left*), and the molar mass is shown in *gray* (y-axis on the *right*). The gel shift assays were carried out in duplicate, and the light scattering data were collected in a single replicate.

mass of 183 ± 1 kDa (Fig. 3E). Remarkably, dimeric HAT-RRM and U6$_{FL}$ formed a 1:1 complex of 208 ± 2 kDa according to MALS, underpinning the binding of two RRM$_{1–2}$ to U6$_{FL}$ (Fig. 3F).

Identification of U6 snRNA elements for SART3 RRM binding

We analyzed the 2-D ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) spectra of ^{13}C-labeled U6$_{33–54}$ upon titration with SART3 RRM$_{1–2}$. The 22-nucleotide ^{13}C-U6$_{33–54}$ displayed 22 well-resolved resonances in the C$_6$H$_6$/C$_8$H$_8$ nucleobase region of the HSQC spectrum (Fig. 4A). Upon addition of RRM$_{1–2}$, U6$_{33–54}$ resonances gradually broadened and disappeared, indicating an intermediate exchange on the chemical shift time scale (Fig. S3A). Differential line broadening of NMR signals during titration provides information on the interaction surface of the complex [15]. We assigned the U6$_{33–54}$ resonances based on the sequential connectivity between nucleobase and ribose protons using 2-D ^{1}H-^{1}H nuclear Overhauser effect spectroscopy (NOESY) (Fig. 4B). The C$_6$H$_6$/C$_8$H$_8$ nucleobase resonances of U6$_{33–54}$ were completely assigned from their connectivity to ribose H1’ protons, and they are annotated in the HSQC spectrum in

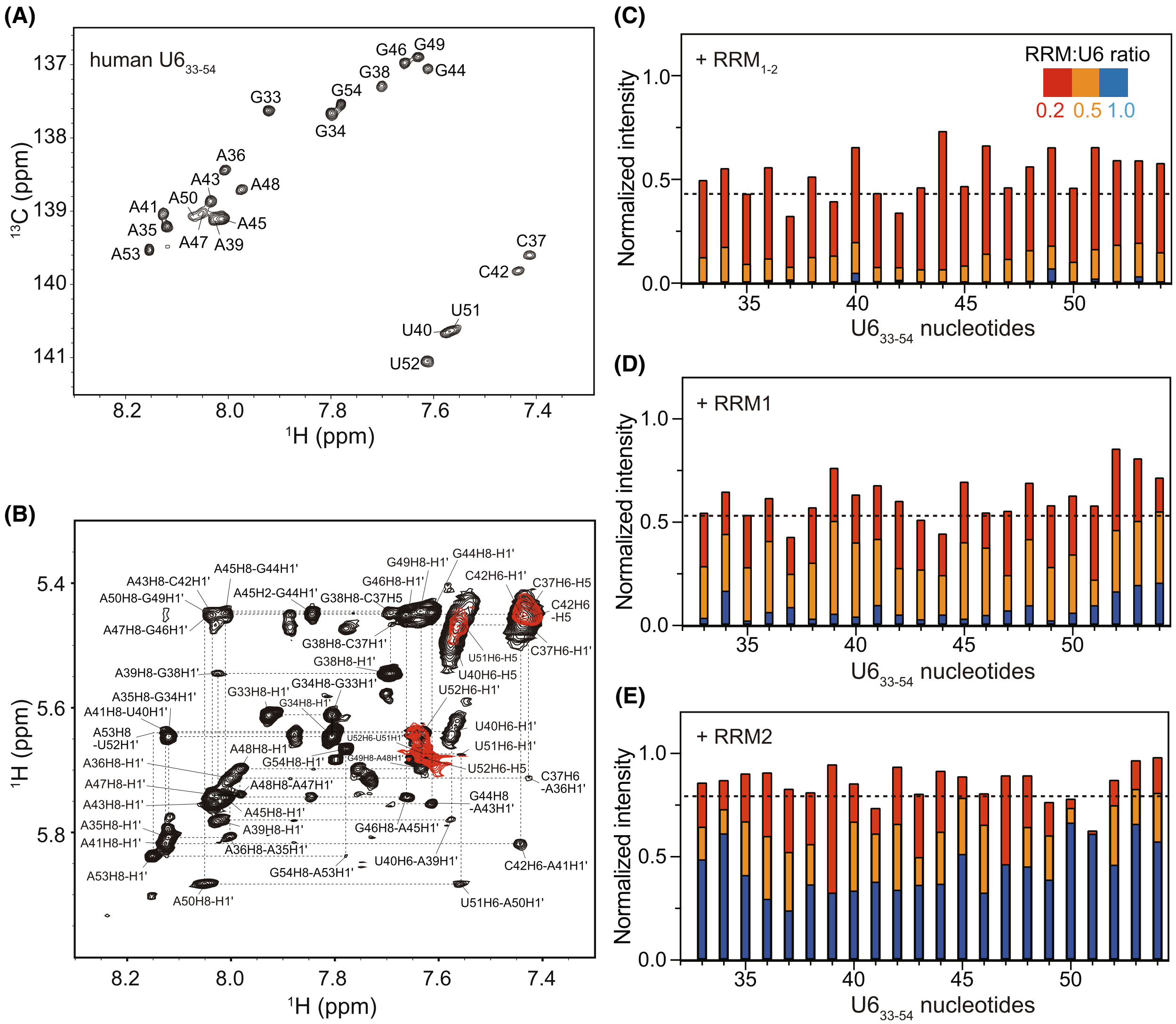

Fig. 4. Assignment of nucleobase protons of $U6_{33-54}$ and intensity ratios titrating with RNA recognition motifs (RRMs). (A) 2-D ^{1}H-^{13}C HSQC, and (B) superimposed 2-D ^{1}H-^{1}H NOESY (black) and 2-D ^{1}H-^{1}H TOCSY (red) spectra. Normalized intensity ratios of $U6_{33-54}$ resonances against nucleotide positions after titration with (C) SART3 RRM_{1-2}, (D) SART3 RRM1, and (E) SART3 RRM2. Dashed lines represent the average minus the standard deviation of the normalized intensity for each titration at an RRM:U6 ratio of 0.2. The intensity ratios were measured in a single replicate.

Fig. 4A. We next calculated the normalized intensity ratio between $U6_{33-54}$ resonances titrating with SART3 RRM_{1-2} and those in the free state (Fig. 4C). The dashed line shows the first quartile of the normalized intensity ratio, indicating regions with a large decrease in signal intensity. Titrating with RRM_{1-2} caused the greatest line broadening at Cyt37, Ade39, and Cyt42 (Fig. 4C). The intensities of Ade45, Ade47, and Ade50 also decreased markedly. By contrast, Ura40, Gua44, and Gua46, as well as terminal nucleotides (Gua33–Gua34 and Ura51–Gua54), showed the least line broadening. In sum, nucleotides with greater intensity decreases were located between Ade35 and Ade50, suggesting that the $U6_{35-50}$ region contains the RRM_{1-2} binding sites.

Titration with individual RRM1 and RRM2 also caused line broadening of $U6_{33-54}$ resonances, albeit less marked than the RRM_{1-2} titration (Fig. S3B,C). Titration of RRM1 resulted in larger intensity decreases at the $U6_{35-50}$ region, similar to RRM_{1-2}, whereas titration of RRM2 exhibited a weaker line-broadening profile (Fig. 4D and E). Overall, the extent of line broadening in each titration correlated with the binding affinity of $U6_{33-54}$ to RRM_{1-2},

Table 1. Binding stoichiometries (*N*), equilibrium dissociation constants (K_D), and sequences of the asymmetric bulge of U6 snRNA for the interaction with SART3 RRM1 as determined by ITC measurements. N.D., not detected.

U6	*N*	K_D (μM)	Sequence (5′–3′)[a]	Nt
33–54	2.3 ± 0.1	0.94 ± 0.16	GGAACGAUACAGAGAAGAUUAG	22
38–50	2.0 ± 0.1	0.85 ± 0.18	GAUACAGAGAAGA	13
38–49	1.1 ± 0.1	2.6 ± 0.4	GAUACAGAGAAG	12
38–46	1.4 ± 0.1	1.8 ± 0.4	GAUACAGAG	9
38–44	1.3 ± 0.1	1.9 ± 0.3	GAUACAG	7
38–43	1.1 ± 0.2	4.7 ± 1.1	GAUACA	6
38–42	N.D.	N.D.	GAUAC	5
39–44	N.D.	N.D.	AUACAG	6
40–46	N.D.	N.D.	UACAGAG	7
45–49	N.D.	N.D.	AGAAG	5
46–50	1.1 ± 0.1	2.6 ± 0.5	GAAGA	5
47–51	N.D.	N.D.	AAGAU	5
48–52	N.D.	N.D.	AGAUU	5

[a] U6 recognition sequences for specific RRM1 binding are underlined.

RRM1, and RRM2. However, the line-broadening profile was insufficient to identify the U6 elements for RRM binding. This led us to conduct a systematic scan of U6$_{33–54}$ to identify sequence elements for RRM recognition.

We searched for U6 elements essential for RRM binding by progressive truncations of U6$_{33–54}$ and measuring binding affinity using ITC. Since two RRM1 potently interacted with U6$_{33–54}$, we evaluated the U6 elements required for RRM1 binding. Trimming nucleotides from both the 5′- and 3′-ends of U6$_{33–54}$ revealed that U6$_{38–50}$ maintained the binding affinity and 2:1 stoichiometry observed for U6$_{33–54}$ (Table 1). Further truncations at either end of U6$_{38–50}$ altered the binding stoichiometry, such that U6$_{38–49}$ accommodated only a single RRM1, pinpointing the 13-nucleotide U6$_{38–50}$ construct as the minimal sequence to accommodate two RRM1s. This finding prompted us to delineate the binding regions for individual RRM1s. Truncations from the 3′-end of U6$_{38–50}$ showed that U6$_{38–44}$ maintained a comparable binding affinity ($K_D = 1.4 \pm 0.4$ μM) to a single RRM1 (Table S1; Fig. S4). In addition, U6$_{38–43}$ was capable of RRM1 binding with slightly reduced affinity ($K_D = 5.0 \pm 2.7$ μM), whereas U6$_{38–42}$ completely lost the affinity. Thus, we identified the six-nucleotide segment U6$_{38–43}$ as one of the RRM1 binding elements. Next, we scanned five-nucleotide windows along U6$_{38–50}$ to locate the RRM1 binding site and identified U6$_{46–50}$ as the other element for the RRM1 interaction ($K_D = 2.6 \pm 0.7$ μM). Neither U6$_{45–49}$ nor U6$_{47–51}$ was able to bind RRM1, indicating that RRM1 associates with U6$_{46–50}$ in a sequence-specific manner (Table S1).

SART3 RRM1 and RRM2 adopt a βαββαβ fold

For the structural analysis of SART3, we prepared separate RRM1 (residues 697–786) and RRM2 (residues 798–877) domains and examined their ^{1}H–^{15}N HSQC spectra. The spectra showed well-resolved amide resonances, indicating that the proteins adopt a compact folded conformation (Fig. S5A,B). The amide resonances of RRM1 and RRM2 closely matched those of the linked RRM$_{1–2}$ domain, except near the linker region (Fig. S5C–E). In addition, RRM2 showed a slight increase in chemical shift differences at the β1–α1 loop region. Overall, our results indicate that the two RRM domains are structurally independent and do not undergo significant interdomain interactions in the RRM$_{1–2}$ configuration. We determined the solution structures of the individual RRM1 and RRM2 domains using NMR spectroscopy.

We assigned the backbone and side chain ^{1}H, ^{15}N, and ^{13}C resonances for the RRM1 and RRM2 domains using a suite of triple-resonance heteronuclear correlation NMR experiments (Fig. S6). The RRM1 structure was determined using 1642 NMR restraints, comprising 1389 experimental NOE restraints, 144 dihedral angle restraints, 80 backbone $^{1}D_{NH}$ residual dipolar coupling (RDC) restraints, and 29 hydrogen bonding restraints (Table 2). Similarly, the RRM2 structure was solved with 1457 NMR restraints comprising 1204 experimental NOE restraints, 152 dihedral angle restraints, 74 backbone $^{1}D_{NH}$ RDC restraints, and 27 hydrogen bonding restraints (Table 2). RRM1 and RRM2 exhibited the characteristic βαββαβ motif of canonical RRM folds, with an additional β4’ strand (Fig. 5A,B). The antiparallel β-sheet composed of four

Table 2. Restraints and structural statistics for RRM1 and RRM2.

Experimental restraints	RRM1<SA>[a]	RRM2<SA>[a]
Nonredundant NOEs	1389	1204
Dihedral angles, φ/ψ	72/72	76/76
Hydrogen bonds	28	27
Residual dipolar coupling, $^{1}D_{NH}$	80	74
Total number of restraints	1642 (18.2 per residue)	1457 (17.3 per residue)
RMSD from experimental restraints		
Distances (Å)	0.028 ± 0.001	0.032 ± 0.001
Torsion angles (°)	0.41 ± 0.10	0.41 ± 0.01
RDC *R*-factor (%)[b]		
$^{1}D_{NH}$ (%) (64)	0.9 ± 0.08	0.5 ± 0.07
RMSD from idealized covalent geometry		
Bonds (Å)	0.002 ± 0	0.002 ± 0
Angles (°)	0.41 ± 0.01	0.41 ± 0.01
Impropers (°)	0.40 ± 0.02	0.55 ± 0.02
Coordinate precision (Å)[a,c]		
Backbone	0.38 ± 0.08	0.48 ± 0.08
Heavy atoms	1.23 ± 0.13	1.04 ± 0.07
Ramachandran statistics (%)		
Most favorable regions	98.0	98.0
Allowed regions	2.0	2.0

[a]For the ensemble of the final 20 lowest-energy simulated annealing (SA) structures; [b]The magnitudes of the axial and rhombic components of the alignment tensor were 7.9 Hz and 0.55 for RRM1, and 14.4 Hz and 0.37 for RRM2, respectively; [c]Regions with secondary structures (for RRM1, residues 705–709, 719–728, 732–738, 749–753, 756–767, 775–778; for RRM2, residues 801–807, 814–823, 826–834, 842–848, 852–860, 869–876).

β-strands (β4-β1-β3-β2) was juxtaposed with two α-helices on the same side, and a short β4′ strand is inserted between the α2 helix and the β4 strand. In addition, RRM1 contained a short α1′ helix at the N terminus preceding the β1 strand. Overall, the secondary structures and loops were well defined in both RRM1 and RRM2, except for the terminal tail regions (Fig. 5C,D). The amino acid sequence identity between SART3 RRM1 and RRM2 was 24%, and their backbone folds aligned with a Cα root-mean-square deviation (RMSD) of 2.3 Å over 72 atoms (Tables S2 and S3). The SART3 RRM domains share 12–30% sequence identity with those of Prp24, and the backbone of the SART3 and Prp24 RRM domains aligned with Cα RMSDs of 1.9–4.2 Å (Tables S2 and S3). Notably, the canonical βαββαβ motifs were highly conserved across RRM domains, and SART3 RRM1 and RRM2 closely resembled yeast Prp24 RRM2, showing C_{α} RMSDs of 1.9 and 2.7 Å for SART3 RRM1 (72 atoms) and RRM2 (72 atoms), respectively (Fig. 5E,F).

Interaction surface of SART3 RRM for U6 snRNA binding

To identify the binding interface of SART3 RRMs for U6 snRNA, we monitored the chemical shift perturbation (CSP) of ^{15}N-labeled RRM$_{1–2}$ in the HSQC spectra by titrating with U6 snRNA. Titrating RRM$_{1–2}$ with U6$_{FL}$ or U6$_{TR}$ caused severe line broadening in the HSQC spectra, and most amide resonances of RRM$_{1–2}$ disappeared at the beginning of the titration. We instead titrated RRM$_{1–2}$ with U6$_{33–54}$, since U6$_{33–54}$ had a similar binding affinity by FPA. The titration showed large CSPs in the RRM1 domain across the β1, β3, β4′, and β4 strands, and also in the β1–α1 and β2–β3 loops (Fig. 6A, Fig. S7A). Most CSPs in the RRM2 domain were smaller than those in RRM1, and similarly clustered in the β1, β3, and β4 strands as well as the β2–β3 loop (Fig. 6A). We note that separate ^{15}N-labeled RRM1 and ^{15}N-labeled RRM2 titrated with U6$_{33–54}$ yielded similar CSP profiles to RRM$_{1–2}$ (Fig. 6B,C, Fig. S7B,C). The third-quartile (Q3) values of CSPs are shown as dashed lines to indicate residues with large CSPs, where the Q3 values were 0.13, 0.18, and 0.08 for RRM$_{1–2}$, RRM1, and RRM2, respectively (Fig. 6A–C). The similar CSP profiles of the linked and separate RRM domains indicate that two RRM domains in RRM$_{1–2}$ make minimal interdomain contacts in the RRM$_{1–2}$–U6$_{33–54}$ complex.

When CSPs were mapped onto the structures of RRM1 and RRM2, they formed contiguous interaction surfaces for U6 snRNA binding (Fig. 6D). The interaction surfaces of SART3 RRM$_{1–2}$ and Prp24 RRMs for the U6 bulge region share similarities, in that the β1 and β3 strands, and the β2–β3 loop, mainly participate in U6 recognition. The binding interfaces derived from CSPs also featured highly electropositive surfaces clustered on each RRM, which facilitates U6 snRNA binding through electrostatic interactions (Fig. 6E, Fig. S8). We note that the β1 and β3 strands of RRM$_{1–2}$ contain aromatic residues that are conserved in canonical RRMs for RNA recognition [16,17]. Specifically, RRM1 contains Phe707 on the β1 strand, and Tyr748 and Tyr750 on the β3 strand, whereas RRM2 has Phe804 on the β1 strand and Tyr845 on the β3 strand. Sequence analysis of SART3 RRM1 and RRM2 homologs revealed that these aromatic residues were well conserved, underscoring their importance in U6 recognition (Fig. 7A,B). Mapping the sequence conservation scores onto the RRM$_{1–2}$ surface revealed a largely conserved U6 binding interface, in contrast to the variable surface on the other side (Fig. 7C, Fig. S8).

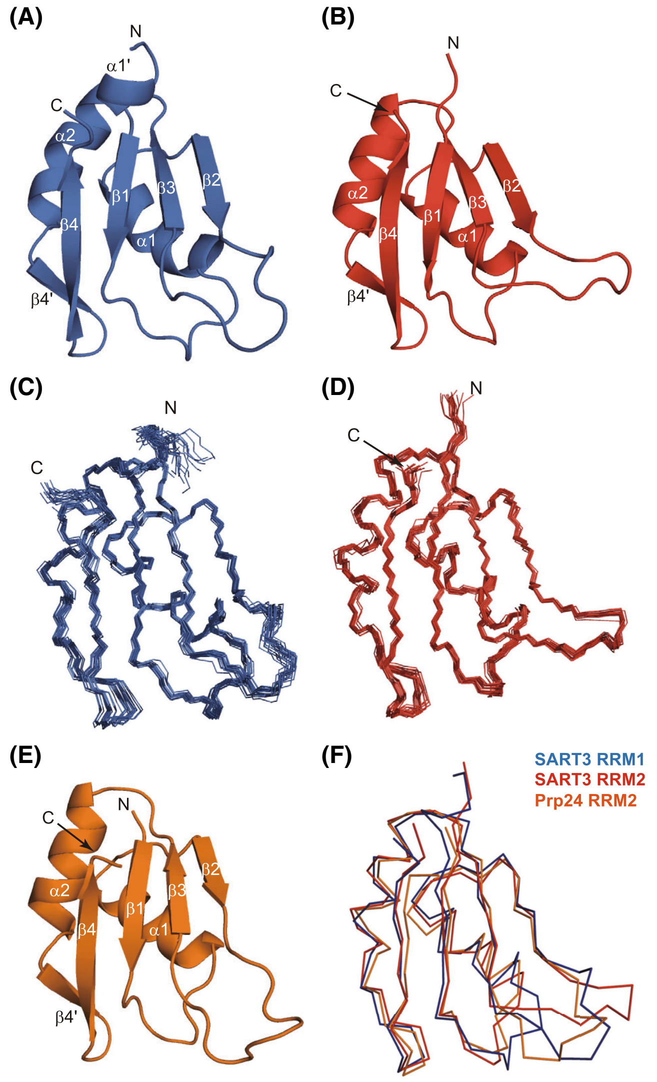

Fig. 5. Solution structures of SART3 RNA recognition motifs (RRMs). Cartoon diagram with secondary structure annotations of (A) RRM1 and (B) RRM2. Superposition of the backbone atoms of the 20 lowest-energy simulated annealing structures of (C) RRM1 and (D) RRM2. (E) The crystal structure of Prp24 RRM2 (PDB code 4N0T). (F) Superimposition of SART3 RRM1 (*blue*), SART3 RRM2 (*red*), and Prp24 RRM2 (*orange*) as a C_α trace representation. The structural figures were generated using the PyMOL software.

We next investigated the roles of polar and positively charged residues on the $RRM_{1–2}$ surface in the interaction with U6 snRNA. We examined the interfacial residues of Prp24 RRMs for U6 snRNA binding based on the crystal structure and selected conserved residues in the multiple sequence alignment of RRM1 and RRM2 (Fig. 7). Eight residues of $RRM_{1–2}$ were selected and replaced one at a time with alanine, and U6 binding affinity was measured using the FPA (Fig. 8). The mutations did not disrupt the structure of $RRM_{1–2}$ from the circular dichroism spectroscopy (Fig. S9). Mutations at the RRM1 surface resulted in 1.7- to 3.7-fold decreases in affinity for $U6_{33–54}$, whereas mutations at the RRM2 surface had negligible effects, except for the $RRM_{1–2}$ (K841A) mutant (2.2-fold decrease in affinity) (Fig. 8, Fig. S10). Thus, the $RRM_{1–2}$–U6 binding was more susceptible to changes in RRM1 than RRM2, supporting the importance of RRM1 in U6 recognition. We further examined how changes in ionic strength would affect the interaction between $RRM_{1–2}$ and U6 snRNA using the FPA. Increasing ionic strength significantly reduced $RRM_{1–2}$–U6 binding at 500 mM NaCl, and completely abrogated the interaction at 2 M NaCl (Fig. S11). We used the RRM1 residues that affected the U6 binding affinity

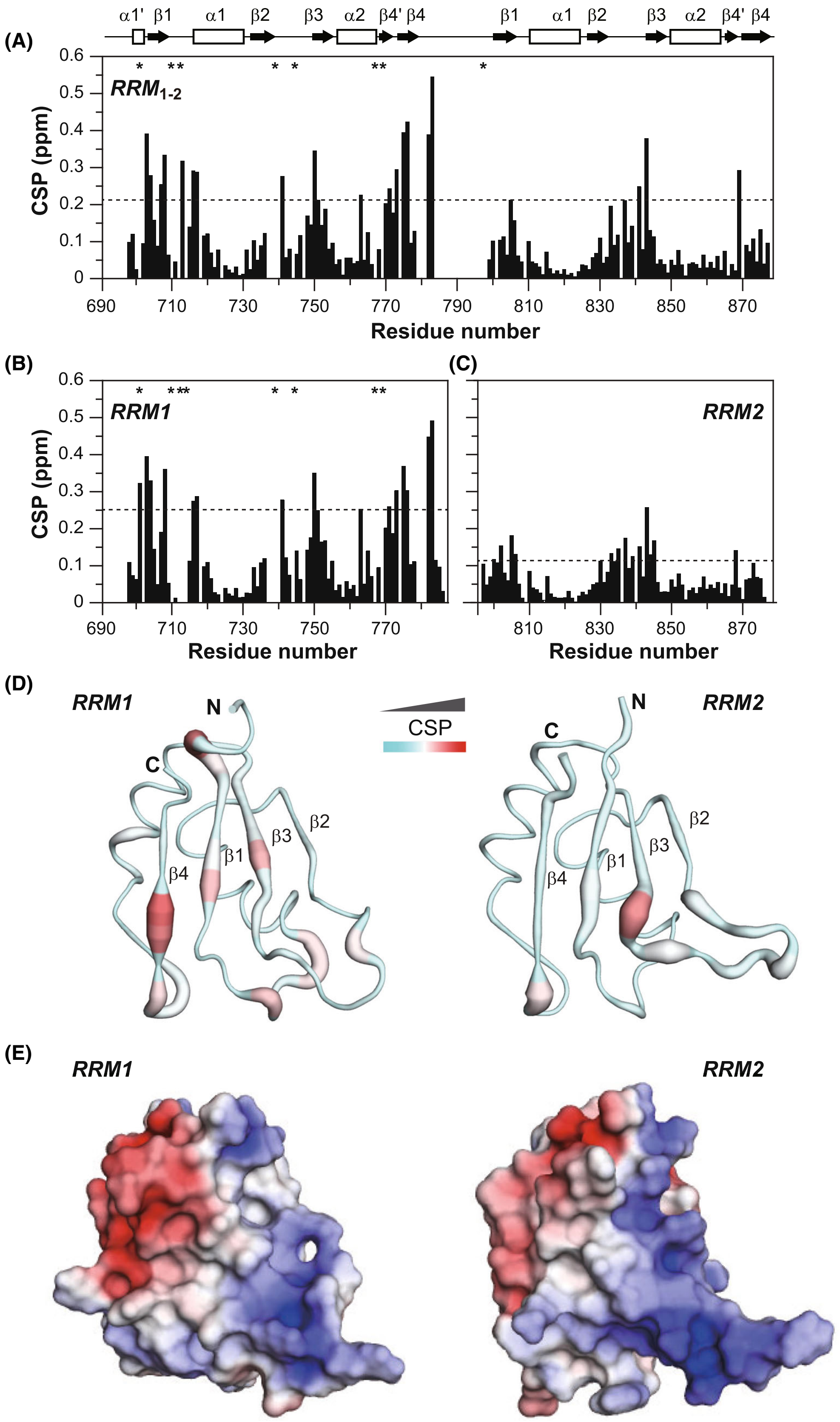
(A)
α1' β1 α1 β2 β3 α2 β4' β4 β1 α1 β2 β3 α2 β4' β4
RRM1-2
CSP (ppm)
0 0.1 0.2 0.3 0.4 0.5 0.6
690 710 730 750 770 790 810 830 850 870
Residue number
(B)
RRM1
CSP (ppm)
0 0.1 0.2 0.3 0.4 0.5 0.6
690 710 730 750 770
Residue number
(C)
RRM2
810 830 850 870
Residue number
(D)
RRM1
N
C
β1 β2 β3 β4
CSP
RRM2
N
C
β1 β2 β3 β4
(E)
RRM1
RRM2

Fig. 6. Chemical shift perturbations (CSPs) and surface charges of SART3 RNA recognition motif (RRM) domains. CSPs of (A) SART3 $RRM_{1–2}$, (B) SART3 RRM1, and (C) SART3 RRM2 upon titration with $U6_{33–54}$. Residues with line broadening upon titration are marked by asterisks, and the average plus the standard deviation of CSP is shown as a dashed line in each plot. (D) CSPs mapped onto structures of RRM1 (left) and RRM2 (right), with CSP magnitude shown as a spectrum from *cyan* (minimum) to *red* (maximum) with varying chain thickness. (E) Electrostatic surface potential maps of RRM1 (left) and RRM2 (right). RRM1 and RRM2 structures are shown in the same perspective as those in Fig. 5. The CSPs were measured in a single replicate, and the structural figures were generated using the PyMOL software.

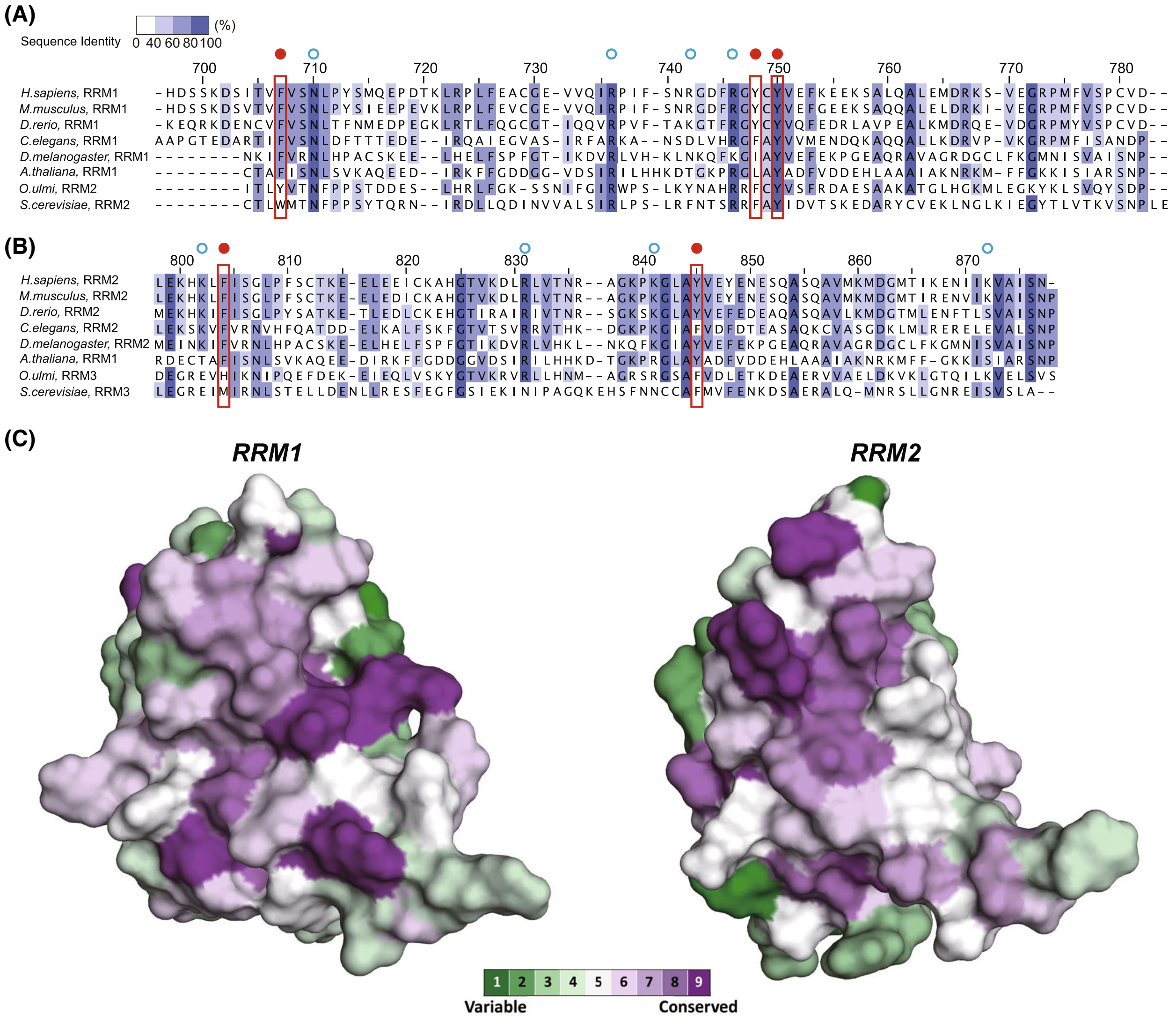

Fig. 7. Multiple sequence alignment and conserved surface representation of SART3 RRM1 and RRM2. Aligned sequences of (A) RRM1 and (B) RRM2 are shaded in *blue* according to the percentage identity score, with conserved aromatic residues highlighted in *red* boxes under filled *red* circles. Polar and charged residues selected for mutagenesis are marked by open *blue* circles. Primary UniProt accession numbers are as follows: *Homo sapiens*, Q15020; *Mus musculus*, Q9JLI8; *Danio rerio*, B3DJT0; *Caenorhabditis elegans*, Q17430; *Drosophila melanogaster*, Q9W4D2; *Arabidopsis thaliana*, F4JQ75; *Ophiostoma ulmi*, Q01491; *Saccharomyces cerevisiae*, P49960. *A. thaliana* has a single RRM domain. For *O. ulmi* and *S. cerevisiae*, which have four RRM domains, RRM2 and RRM3 are aligned. (C) Sequence conservation scores calculated and mapped on the surface models of RRM1 and RRM2 using ConSurf software [35]. Structures are shown from the same perspectives as in Fig. 5, and the structural figures were generated using the PyMOL software.

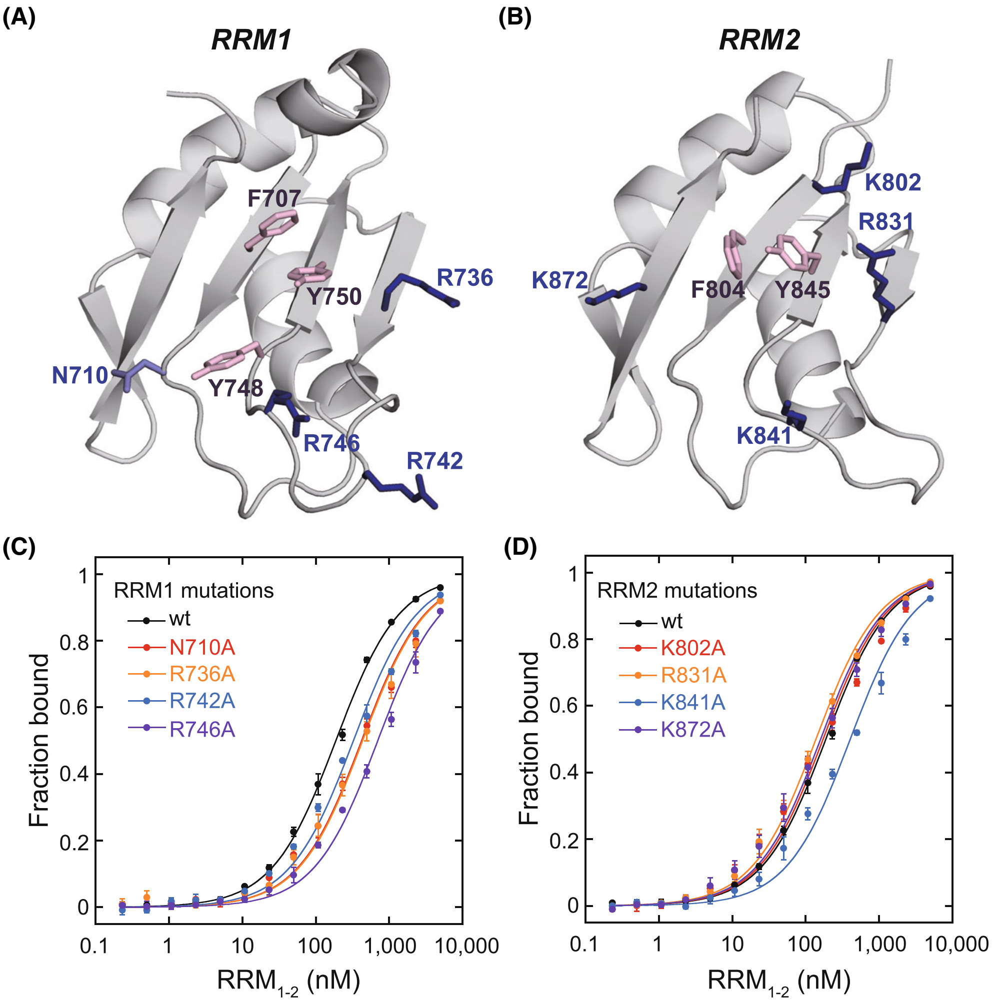

Fig. 8. Mutation positions of $RRM_{1–2}$ mutants and $U6_{33–54}$ binding analysis according to fluorescence polarization assay (FPA). Side chains of eight mutation sites are shown in *blue* on the structures of (A) RRM1 and (B) RRM2. Conserved aromatic residues at the interface are shown in *pink*. Binding curves from FPAs between fluorescein-labeled $U6_{33–54}$ and $RRM_{1–2}$ with mutations in the (C) RRM1 domain and (D) RRM2 domain. Error bars are standard deviations of triplicate measurements. Individual binding curves and K_D values are shown in Fig. S10, and the structural figures were generated using the PyMOL software.

as binding interfaces to build a structural model of the RRM1–$U6_{FL}$ complex.

Complex model for RRM1 binding to $U6_{FL}$

We constructed a structural model of the SART3 RRM1–$U6_{FL}$ complex using HADDOCK software, integrating a $U6_{FL}$ model, solution structures of RRM, and CSP-derived binding interfaces as ambiguous restraints. The human $U6_{FL}$ snRNA model was generated using ModeRNA software, based on the crystal structure of yeast U6 snRNA [18]. We docked two $RRM_{1–2}$ onto $U6_{38–43}$ and $U6_{46–50}$ in a sequential manner, and the two-step HADDOCK runs produced a ternary complex model. The structural model showed that two RRM1 domains were positioned in proximity to snugly bind the $U6_{38–43}$ and $U6_{46–50}$ regions (Fig. 9A). On the other hand, RRM2 domains did not converge into a reliable complex possibly due to insufficient restraints and weak affinity, and we showed RRM2 as dashed ovals next to linked RRM1 at hypothetical binding sites. RRM1 subunits were annotated as RRM1 and RRM1′ to distinguish between the two protomers of dimeric SART3. Binding interfaces in the complex were consistent with the experimental findings, showing that conserved aromatic and positively charged residues of RRM1 formed the interface for the cognate U6 snRNA sequence. RRM2 subunits are shown as dashed ovals next to the RRM1 subunits at hypothetical interfaces.

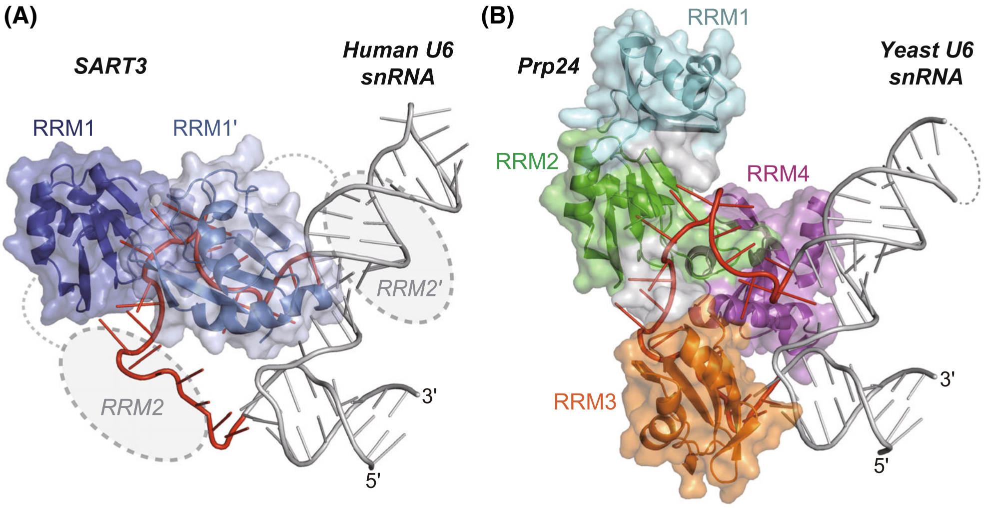

Fig. 9. Complex structures of RNA recognition motif (RRM) and U6 snRNA. (A) Model structure of SART3 RRM1 bound to human U6 snRNA. Two RRM1 subunits bound to the U6 bulge are annotated as RRM1 and RRM1′. Putative positions of RRM2 subunits are shown as dashed ovals. (B) Crystal structure of Prp24 bound to yeast U6 snRNA (PDB code 4N0T). Missing coordinates of yeast U6 snRNA in the crystal structure are shown by dashed lines. The asymmetric bulge region of U6 snRNA is colored in *red*; otherwise, it is in *gray*. The structural figures were generated using the PyMOL software.

Given that the U6 sequences for RRM1 and RRM1' binding are close to each other, we suppose that RRM2 and RRM2′ would bind the periphery of the $U6_{33–54}$ bulge region. We note that RRM1 and RRM1' do not make a steric clash at their binding sites (Fig. S12).

A comparison of our SART3–U6 model with the existing Prp24–U6 structure demonstrated that the SART3 RRM1 interaction with $U6_{38–43}$ is analogous to that of Prp24 RRM2 (Fig. 9B). SART3 RRM1 binds the GAUACA sequence of human $U6_{38–43}$, and Prp24 RRM2 binds the UACAGA sequence of yeast $U6_{46–51}$. However, the interaction of SART3 RRM1 with $U6_{46–50}$ lacks a counterpart in the Prp24 structure, and the corresponding U6 sequence in yeast does not tightly engage with Prp24 RRM domains (Fig. 9B). Our model structure highlights the differences in U6 snRNA binding between SART3 and Prp24, suggesting a novel binding mode of SART3 RRM for U6 recognition.

Discussion

Human SART3 was initially identified as a distant homolog of yeast Prp24 by searching a database using subregions of Prp24 as query sequences [6]. Both proteins interact with U6 snRNA via their RRMs, but SART3 is characterized by two RRMs, whereas Prp24 comprises four. Our results demonstrate that SART3 dimerization via the HAT repeat domain allows four RRM domains to participate in U6 snRNA recognition. SART3 specifically binds the asymmetric bulge region of U6 snRNA via four RRM domains, which is similar to Prp24. The overall architecture of the SART3–U6 complex, however, is distinct from that of the Prp24–U6 complex. Whereas RRM2, RRM3, and RRM4 of Prp24 tightly engage with yeast U6 snRNA, dimeric SART3 predominantly uses its RRM1 subunits to interact with human U6 snRNA. We identified the $U6_{38–43}$ and $U6_{46–50}$ as key recognition motifs for SART3 RRM1. The interaction between RRM1 and $U6_{38–43}$ is analogous to that of Prp24 RRM2, but RRM1 binding to $U6_{46–50}$ presents a unique mode not seen in the interaction of Prp24 with U6. Thus, U6 snRNA binding of SART3 and Prp24 highlights both conserved and divergent features of RNA recognition.

We predicted the structure of the complex of SART3 $RRM_{1–2}$ and human U6 snRNA using the AlphaFold3 server [19]. As a proof of concept, we first tested the AlphaFold3 prediction of the known complex between yeast Prp24 and U6 snRNA. AlphaFold3 accurately predicted the complex structure of Prp24 and yeast U6 snRNA, aligning closely with the crystal structure. However, Prp24 without U6 snRNA was predicted to form the experimental structure, whereas yeast U6 snRNA without Prp24 formed an incorrect fold with very low confidence scores (<50) based on the predicted local distance difference test (pLDDT). Similarly, when we predicted the complex structure of the SART3 RRM and human U6 snRNA, the SART3

RRM structures were correctly predicted with high pLDDT scores (>70), but the U6 snRNA failed to adopt the expected stem and bulge structures, displaying poor pLDDT scores (<50). We also predicted the complex of SART3 $\text{RRM}_{1\text{–}2}$ and $\text{U6}_{33\text{–}54}$, but the complex structure exhibited high expected position errors and low confidence scores for U6 snRNA, indicating limitations in RNA modeling accuracy. Given these challenges, we instead employed HADDOCK software to integrate NMR CSPs and mutagenesis data, generating a refined model of the SART3 RRM1–U6 snRNA complex. However, our model is still constrained by the accuracy of the U6 snRNA structure derived from template-based modeling. In addition, the U6 elements involved in RRM2 binding were not identified in this study. Despite these limitations, our data-driven complex model of two RRM1 domains bound to $\text{U6}_{33\text{–}54}$ highlights a noticeable difference in the binding mode between SART3 and Prp24.

RRM is the most abundant and well-characterized RNA-binding module in eukaryotes [20]. While most RRMs share a conserved backbone fold of 80–90 residues, their binding interfaces vary significantly depending on their interacting partners [21]. Our CSP analysis showed that SART3 RRMs employed canonical RNA-binding interfaces, specifically targeting the U6 bulge region. Given that RRM1 recognizes two distinct U6 sequences, GAUACA ($\text{U6}_{38\text{–}43}$) and GAAGA ($\text{U6}_{46\text{–}50}$), the shared GA motif may serve as a key element for sequence-specific interaction with SART3 RRM1. In addition, RRM1 binds GAUACA ($\text{U6}_{38\text{–}43}$), but not GAUAC ($\text{U6}_{38\text{–}42}$) or AUACAG ($\text{U6}_{39\text{–}44}$), suggesting that the 3' adenine could be important to RRM1 binding.

Mutagenic studies have demonstrated that Prp24 is involved in the unwinding of yeast U6 snRNA, thereby promoting U6/U4 annealing. Notably, Prp24 RRM1 does not directly contact U6 snRNA; instead, it has been proposed that RRM1 extends an electropositive groove formed by RRM2 and RRM4 to stabilize the U6 ISL, thereby aiding its annealing with U4 snRNA [11,22]. By contrast, our findings indicate that SART3 RRM2 alone exhibited a very low affinity for U6 snRNA, suggesting that it may transiently associate and dissociate within the functional complex. We speculate that RRM2 contributes to U4/U6 annealing by reinforcing the electropositive surfaces adjacent to RRM1.

In summary, we investigated the structural basis of SART3 RRM interaction with U6 snRNA. SART3 dimerizes via its HAT repeat domain, enabling two RRM1 domains in the dimer to specifically recognize the U6 snRNA bulge region. Our structural model suggests that one RRM1 binds U6 similar to Prp24 RRM2, while the other engages with U6 in a distinct mode, diverging from the Prp24 interaction. These findings provide structural insights into how Prp24 homologs with a varying number of RRM repeats associate with U6 snRNA, highlighting both common and unique aspects of RNA recognition.

Materials and methods

Cloning, expression, and purification of SART3 domains

The SART3 RRM1 (residues 697–786), RRM2 (residues 798–877), $\text{RRM}_{1\text{–}2}$ (residues 671–877), RRM mutants, and HAT-RRM (residues 94–877) genes were cloned into a pET28a vector (Novagen, Madison, WI, USA) with a His_6-tag and a thrombin cleavage site at the N terminus, and the cloned vectors were transformed into *Escherichia coli* BL21(DE3) cells. Cells were grown in lysogeny broth (LB) or M9 minimal medium, with $^{15}\text{NH}_4\text{Cl}$ and $^{13}\text{C}_6$-D-glucose as the sole nitrogen and carbon sources, respectively, at 37 °C until the optical density at 600 nm reached 0.6. Next, the culture was induced by 0.5 mM isopropyl β-D-thiogalactoside at 18 °C for 16 h. Cells were harvested by centrifugation, resuspended in 50 mM Tris/HCl, pH 8.0, and 500 mM NaCl, and lysed by sonication after adding protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO, USA). Cell debris was removed by centrifugation at 4 °C, and the supernatant was loaded onto a HisTrap column (GE Healthcare, Chicago, IL, USA). Bound proteins were washed with 20 mM imidazole and eluted with a linear gradient of 20–500 mM imidazole. SDS/PAGE fractions containing the protein were collected and treated with thrombin to cleave the N-terminal His_6-tag. The reaction was further purified by SEC using a HiLoad 26/600 Superdex 75 or 200 pg columns (GE Healthcare) pre-equilibrated with 20 mM HEPES, pH 7.0, 150 mM NaCl, and 1 mM dithiothreitol (DTT) or 0.5 mM Tris(2-carboxyethyl) phosphine (TCEP). Fractions containing target proteins were collected and stored at −80 °C until use.

For the purification of SART3 $\text{RRM}_{1\text{–}2}$ mutants, harvested cells were resuspended in 20 mM Tris/HCl pH 7.4, 500 mM NaCl, 10% glycerol, 5 mM 2-mercaptoethanol (BME), and 1 mM phenylmethylsulfonyl fluoride. The cells were lysed by sonication and centrifuged at 40 000 ***g*** for 30 min. Supernatants were filtered, loaded on a HisTrap column (GE Healthcare, Chicago, IL, USA), and eluted with a gradient of 0–500 mM imidazole. Fractions containing protein were treated with a TEV protease in 20 mM Tris/HCl, pH 7.4, 100 mM NaCl, 10% glycerol, 0.5 mM ethylenediaminetetraacetic acid, and 5 mM BME. The reaction was purified by SEC using a HiLoad 26/600 Superdex

75 pg column (GE Healthcare) equilibrated with 20 mM HEPES, pH 7.0, 100 mM NaCl, and 5 mM BME. Fractions containing protein were loaded on a monoS column (GE Healthcare) equilibrated with 20 mM HEPES, pH 7.0, 100 mM NaCl, and 5 mM BME and eluted with a gradient of 0–1 M NaCl. Protein purity was assessed by SDS/PAGE.

Preparation of U6 snRNAs

Human U6 snRNA constructs were prepared by *in vitro* transcription using synthetic DNA templates (Integrated DNA Technologies, Coralville, IA, USA). U6 snRNA was prepared by *in vitro* transcription using a purified P266L T7 RNA polymerase mutant [23]. 4 mM of each ribonucleotide triphosphate (rNTP; rATP, rGTP, rCTP, and rUTP), 25 mM $MgCl_2$, 2.5 mM DTT, 0.5 unit·mL^{-1} inorganic pyrophosphatase (Sigma-Aldrich, Cat No. I1643), and 40 units·μL^{-1} RNase inhibitor (Roche, Basel, Switzerland, Protector RNase Inhibitor) were incubated with the DNA template and polymerase in the transcription buffer (40 mM Tris/HCl, pH 8.0, 1 mM spermidine, and 0.01% (v/v) Triton X-100) at 37 °C for 4–6 h, and the transcribed RNA was precipitated with ethanol at −20 °C. After centrifugation, the RNA pellet was dissolved in water, resolved by 12–20% denaturing PAGE (19:1 cross-linking ratio), and extracted using an electroelution system (Elutrap; Whatman). RNA was further purified on an anion exchange column (HiTrap Q; GE Healthcare). Purified RNA was heated to 95 °C, cooled on ice, and concentrated to 0.2–1 mM using the Amicon Ultra centrifugal filter (Millipore, Burlington, MA, USA). For ^{13}C-U6 snRNA preparation, ^{13}C-labeled rNTPs (Biolog) were used for *in vitro* transcription. Truncated constructs of $U6_{33–54}$ snRNA were synthesized by Bioneer.

Fluorescence polarization assay

The FPA was performed in 20 mM sodium phosphate, pH 6.5, 150 mM NaCl, and 0.5 mM TCEP. 5'-FAM-labeled U6 snRNA constructs (5 nM) were incubated with 0–5 μM SART3 RRM1-2 (and its mutants) or 0–100 μM RRM1 and RRM2 for 30 min at 30 °C. Fluorescence polarization was measured using the Synergy H1 microplate reader (BioTek, Winooski, VT, USA). The polarization signal was corrected by subtracting the signal from buffer controls. The fluorescence polarization data were normalized to produce bound fraction values, and the K_D value was obtained from the Langmuir isotherm model using the KaleidaGraph software. The measurements were performed in triplicate, and the average values and standard errors of the mean are reported.

Electrophoretic mobility shift assay

The RNA–protein interaction was analyzed by EMSA. The FAM-labeled U6 snRNAs (20 nM) were incubated with increasing concentrations of SART3 RRM constructs (20–100 000 nM) for 60 min at room temperature in 20 mM sodium phosphate, pH 6.5, and 1 mM DTT. The mixtures were loaded on an 8% (29:1) native polyacrylamide gel and resolved at 100 V for 80–100 min at 4 °C in 0.5× Tris–-borate–EDTA buffer. The gel was visualized using an Amersham ImageQuant IQ800 imager (Cytiva).

Size exclusion chromatography with multi-angle light scattering

RNA, protein, and RNA–protein complex samples were separated by SEC using Superdex 75 Increase 10/300 GL or Superdex 200 Increase 10/300 GL (Cytiva) columns, equilibrated with 20 mM sodium phosphate, pH 6.5, 150 mM NaCl, and 0.5 mM TCEP. The SEC columns were coupled to a three-angle light scattering detector (miniDAWN) and a refractive index detector (Optilab) (Wyatt Technology). Data collection and analysis were performed using ASTRA 8 software (Wyatt Technology, Santa Barbara, CA, USA). Bovine serum albumin (Thermo Fisher Scientific, Waltham, MA, USA) was used for calibration before each measurement. For SEC-MALS measurements, 100 μL of injection samples were prepared as follows: $RRM_{1–2}$ (200 μM) with $U6_{FL}$ (100 μM), $RRM_{1–2}$ (400 μM) with $U6_{33–54}$ (200 μM), HAT-RRM (80 μM) alone, and HAT-RRM (80 μM) with $U6_{FL}$ (40 μM).

Isothermal titration calorimetry

Dissociation constants (K_D) between U6 snRNA and SART3 RRM constructs were measured at 30 °C in 20 mM sodium phosphate, pH 6.5, 150 mM NaCl, and 0.5 mM TCEP using an ITC_{200} calorimeter (Malvern). 5 μM U6 snRNA constructs were placed in the cell and titrated with 100 μM SART3 RRM constructs in the syringe. Nineteen 2-μL aliquots of proteins were titrated into the cell. Data were analyzed using Origin software provided by the manufacturer.

NMR spectroscopy

SART3 NMR samples were prepared as 0.8 mM ^{13}C,^{15}N-SART3 RRM1 or RRM2 in 20 mM sodium phosphate, pH 6.5, 1 mM DTT, and 10% D_2O. NMR spectra were collected at 25 °C on a Bruker Avance III HD 800 MHz NMR spectrometer equipped with a *z*-shielded gradient triple-resonance cryoprobe. Sequential assignments were performed by triple resonance through-bond scalar correlation experiments including HNCO, HN(CA)CO, HNCACB, and CBCA(CO)NH. Side chain assignment was performed by conducting HBHA(CO)NH and HCCH–total correlation spectroscopy (TOCSY) experiments. ^{13}C-seperated NOESY and ^{15}N-seperated NOESY experiments were performed using a mixing time of 120 ms. Residual $^{1}D_{NH}$ dipolar couplings were obtained by taking the difference between the $^{1}J_{NH}$ splitting values measured in aligned (10 mg·mL^{-1} *pf1* phage; ASLA Biotech) and isotropic media using 2-D in-phase/antiphase ^{1}H-^{15}N HSQC spectra.

For NMR CSP measurements, 2-D ^{1}H-^{15}N HSQC spectra were recorded for 0.1 mM ^{15}N-labeled RRM1, RRM2, and RRM$_{1–2}$ titrated with U6$_{33–54}$ at 25 °C. CSPs were calculated using the equation $\Delta\delta = \sqrt{\Delta\delta_{HN}^{2} + (\Delta\delta_{N}/5)^{2}}$, where $\Delta\delta_{HN}$ and $\Delta\delta_{N}$ are the chemical shift changes for amide proton and nitrogen resonances, respectively. The NMR spectra were processed using NMRPipe [24] and NMRView [25] software and analyzed using NMRFAM-SPARKY software [26].

The U6 snRNA NMR sample was prepared as 0.9 mM ^{13}C-U6$_{33–54}$ in 20 mM sodium phosphate, pH 6.5, 100 mM NaCl, and 100% D_2O. Then, 2-D ^{1}H-^{1}H NOESY (300 ms mixing time) and 2-D TOCSY were obtained at 25 °C for the chemical shift assignment of the nucleobase resonances of U6$_{33–54}$. NMR spectra were processed using Topspin 3.5pl7 software (Bruker) and analyzed using POKY software [27]. A series of 2-D ^{1}H-^{13}C HSQC spectra was collected for 0.1 mM ^{13}C-labeled U6$_{33–54}$, titrating with unlabeled RRM1, RRM2, and RRM$_{1–2}$ at 25 °C. The intensity of the U6 resonance was plotted as I/I_0, where I is the intensity along the titration with RRM constructs, and I_0 is the intensity of free U6 snRNA.

CD spectroscopy

The circular dichroism was measured using a Jasco J-815 CD spectropolarimeter (Jasco). Protein samples (10 μM) were prepared in 20 mM sodium phosphate, pH 6.5, and 0.5 mM tris(2-carboxyethyl) phosphine (TCEP) at 25 °C.

Structure calculation

The distance restraints from NOESY experiments were classified into short, medium, and long distance ranges according to peak intensity. The backbone chemical shifts of N, Cα, Cβ, C′, H_N, and H_α atoms were used to derive φ/ψ dihedral angle restraints using TALOS-N software [28]. Structures were calculated with simulated annealing in torsion angle space using the Xplor-NIH software [29]. Twenty lowest-energy structures that do not violate the experimental distance and dihedral angle restraints were selected for both RRM1 and RRM2. The target function for simulated annealing included covalent geometry, a quadratic van der Waals repulsion potential, square-well potentials for interproton distance and torsion angle restraints, hydrogen bonding, harmonic potentials for $13C_\alpha/13C_\beta$ chemical shift restraints [30], and a multidimensional torsion angle database potential of mean force [31]. The structure calculation protocol has been described in our previous report [32]. Structures were displayed using PyMOL software (The PyMOL Molecular Graphics System, ver. 3.1.6.1; Schrödinger, LLC).

Molecular docking

For modeling of the human U6 snRNA structure, the crystal structure of yeast U6 snRNA in complex with Prp24 (PDB code 4N0T) was used as the template. The alignments of the human and yeast U6 snRNA sequences and the template structure were used to build the model structure using ModeRNA software [18]. Subsequently, QRNAS software was used to refine the model structure, which improved on the covalent geometry and local clashes in the nucleic acid model to increase accuracy [33]. Finally, the complex structure of SART3 RRM1 and human U6 snRNA was built using HADDOCK software based on the solution structure of RRM1 and the U6 snRNA model combined with the binding interfaces as distance restraints [34]. RRM interfaces were selected using two key criteria: (1) residues with CSPs larger than the average plus one standard deviation in the NMR titration (Fig. 6A), and (2) mutations that reduced U6 snRNA binding affinity. The U6 snRNA interfaces (U6$_{38–43}$ and U6$_{46–50}$ sequences) for RRM1 binding were derived from ITC measurements of truncated U6 constructs. These selected interfaces from experiments were employed as ambiguous interaction restraints for the HADDOCK run.

Acknowledgements

This work was supported by the Cooperative Research Program for Agricultural Science and Technology Development funded by the Rural Development Administration (PJ01495901), National Research Foundation of Korea grants funded by the Korean government (MSIT; 2021R1A2C1003917, RS-2023-00207820), and the Institutional Research Program of Korea Institute of Science and Technology (KIST; 2 V10491). We thank Prof. Janusz Bujnicki at the International Institute of Molecular and Cell Biology in Warsaw, Poland, for his help with building the RNA model structure using ModeRNA software.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

IK, KMB, SYA, HKS, N-KK, and Je-YS conceived the study. KMB, SYA, and Ji-YS prepared the proteins and RNA, and N-KK and Je-YS collected and analyzed the NMR data. IK and KMB determined the structure. SYA, CP, and YK performed the binding measurements. IK, KMB, SYA, N-KK, and Je-YS wrote the manuscript.

Peer review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/febs.70275.

Data availability statement

The atomic coordinates of the solution structures of RRM1 and RRM2 have been deposited in the Protein Data Bank with accession codes 7XX8 and 7XX9, respectively. The chemical shifts and NMR restraints of RRM1 and RRM2 can be retrieved from the Biological Magnetic Resonance Bank with accession codes 36490 and 36491, respectively. Chemical shifts of U6$_{33–54}$ have been deposited in the Biological Magnetic Resonance Bank with accession code 53 223.

References

1 Will CL & Luhrmann R (2011) Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3**, a003707.

2 Wang E & Aifantis I (2020) RNA splicing and cancer. *Trends Cancer* **6**, 631–644.

3 Montes M, Sanford BL, Comiskey DF & Chandler DS (2019) RNA splicing and disease: animal models to therapies. *Trends Genet* **35**, 68–87.

4 Scotti MM & Swanson MS (2016) RNA mis-splicing in disease. *Nat Rev Genet* **17**, 19–32.

5 Didychuk AL, Butcher SE & Brow DA (2018) The life of U6 small nuclear RNA, from cradle to grave. *RNA* **24**, 437–460.

6 Bell M, Schreiner S, Damianov A, Reddy R & Bindereif A (2002) p110, a novel human U6 snRNP protein and U4/U6 snRNP recycling factor. *EMBO J* **21**, 2724–2735.

7 Agafonov DE, Kastner B, Dybkov O, Hofele RV, Liu WT, Urlaub H, Luhrmann R & Stark H (2016) Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* **351**, 1416–1420.

8 Medenbach J, Schreiner S, Liu S, Luhrmann R & Bindereif A (2004) Human U4/U6 snRNP recycling factor p110: mutational analysis reveals the function of the tetratricopeptide repeat domain in recycling. *Mol Cell Biol* **24**, 7392–7401.

9 Rader SD & Guthrie C (2002) A conserved Lsm-interaction motif in Prp24 required for efficient U4/U6 di-snRNP formation. *RNA* **8**, 1378–1392.

10 Licht K, Medenbach J, Luhrmann R, Kambach C & Bindereif A (2008) 3′-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. *RNA* **14**, 1532–1538.

11 Montemayor EJ, Curran EC, Liao HH, Andrews KL, Treba CN, Butcher SE & Brow DA (2014) Core structure of the U6 small nuclear ribonucleoprotein at 1.7-a resolution. *Nat Struct Mol Biol* **21**, 544–551.

12 Park JK, Das T, Song EJ & Kim EE (2016) Structural basis for recruiting and shuttling of the spliceosomal deubiquitinase USP4 by SART3. *Nucleic Acids Res* **44**, 5424–5437.

13 Huang M, Zhou B, Gong J, Xing L, Ma X, Wang F, Wu W, Shen H, Sun C, Zhu X *et al.* (2018) RNA-splicing factor SART3 regulates translesion DNA synthesis. *Nucleic Acids Res* **46**, 4560–4574.

14 Klimesova K, Petrzilkova H, Barinka C & Stanek D (2023) SART3 associates with a post-splicing complex. *J Cell Sci* **136**, jcs.260380.

15 Matsuo H, Walters KJ, Teruya K, Tanaka T, Gassner GT, Lippard SJ, Kyogoku Y & Wagner G (1999) Identification by NMR spectroscopy of residues at contact surfaces in large, slowly exchanging macromolecular complexes. *J Am Chem Soc* **121**, 9903–9904.

16 Maris C, Dominguez C & Allain FH (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**, 2118–2131.

17 Daubner GM, Clery A & Allain FH (2013) RRM-RNA recognition: NMR or crystallography…And new findings. *Curr Opin Struct Biol* **23**, 100–108.

18 Rother M, Rother K, Puton T & Bujnicki JM (2011) ModeRNA: a tool for comparative modeling of RNA 3D structure. *Nucleic Acids Res* **39**, 4007–4022.

19 Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J *et al.* (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500.

20 Lunde BM, Moore C & Varani G (2007) RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8**, 479–490.

21 Afroz T, Cienikova Z, Clery A & Allain FHT (2015) One, two, three, four! How multiple RRMs read the genome sequence. *Methods Enzymol* **558**, 235–278.

22 Martin-Tumasz S, Richie AC, Clos LJ 2nd, Brow DA & Butcher SE (2011) A novel occluded RNA recognition motif in Prp24 unwinds the U6 RNA internal stem loop. *Nucleic Acids Res* **39**, 7837–7847.

23 Guillerez J, Lopez PJ, Proux F, Launay H & Dreyfus M (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc Natl Acad Sci USA* **102**, 5958–5963.

24 Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J & Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277–293.

25 Johnson BA & Blevins RA (1994) NMR view: a computer program for the visualization and analysis of NMR data. *J Biomol NMR* **4**, 603–614.

26 Lee W, Tonelli M & Markley JL (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327.

27 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.

28 Shen Y & Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56**, 227–241.

29 Schwieters CD, Kuszewski JJ, Tjandra N & Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson* **160**, 65–73.

30 Kuszewski J, Qin J, Gronenborn AM & Clore GM (1995) The impact of direct refinement against 13C alpha and 13C beta chemical shifts on protein structure determination by NMR. *J Magn Reson B* **106**, 92–96.

31 Clore GM & Kuszewski J (2002) Chi(1) rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force. *J Am Chem Soc* **124**, 2866–2867.

32 Kim I, Koo J, An SY, Hong S, Ka D, Kim EH, Bae E & Suh JY (2020) Structural and mechanistic insights into the CRISPR inhibition of AcrIF7. *Nucleic Acids Res* **48**, 9959–9968.

33 Stasiewicz J, Mukherjee S, Nithin C & Bujnicki JM (2019) QRNAS: software tool for refinement of nucleic acid structures. *BMC Struct Biol* **19**, 5.

34 van Zundert GCP, Rodrigues J, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ & Bonvin A (2016) The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J Mol Biol* **428**, 720–725.

35 Ashkenazy H, Erez E, Martz E, Pupko T & Ben-Tal N (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **38**, W529–W533.

Supporting information

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

Fig. S1. Size exclusion chromatogram of SART3 $\text{RRM}_{1\text{–}2}$.

Fig. S2. EMSA profiles of $\text{U6}_{33\text{–}54}$ titrating with RRM.

Fig. S3. 2-D ^{1}H-^{13}C HSQC spectrum of $\text{U6}_{33\text{–}54}$.

Fig. S4. Integrated heats of injection from ITC analyses for the titration between RRM1 and various U6 snRNA constructs in Table 1.

Fig. S5. 2-D ^{1}H-^{15}N HSQC spectra and chemical shift perturbations of RRM domains.

Fig. S6. 2-D ^{1}H-^{15}N HSQC spectra with annotated assignment.

Fig. S7. Superimposed 2-D ^{1}H-^{15}N HSQC spectra of RRM titrating with $\text{U6}_{33\text{–}54}$.

Fig. S8. Electrostatic surface potential and sequence conservation profiles of RRM.

Fig. S9. Circular dichroism spectra of $\text{RRM}_{1\text{–}2}$ and its mutants.

Fig. S10. Binding curves and K_D values as determined by fluorescence polarization assays of fluorescein-labeled $\text{U6}_{33\text{–}54}$ and $\text{RRM}_{1\text{–}2}$ mutants.

Fig. S11. Anisotropy values of SART3 $\text{RRM}_{1\text{–}2}$ and fluorescein-labeled $\text{U6}_{33\text{–}54}$ from fluorescence polarization assays at different ionic strengths.

Fig. S12. Complex model structure of SART3 RRM1 and human U6 snRNA.

Table S1. Thermodynamic parameters from ITC analyses of binding between SART3 RRM domains and human U6 snRNA truncation constructs.

Table S2. Percentage identity matrix analysis of the RRM domains in SART3 and Prp24 obtained from multiple sequence alignment using Clustal Omega software.

Table S3. Root-mean-square deviation (RMSD) of Cα positions from the structural alignment of the RRM domains of SART3 and Prp24.