

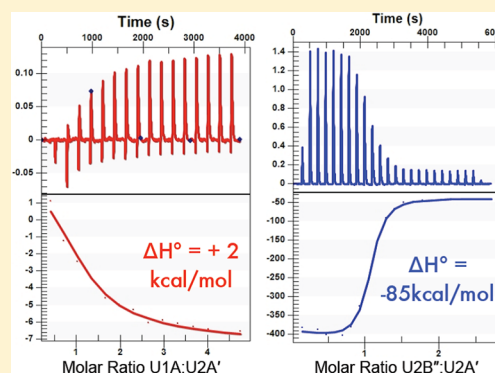
Binding Affinity and Cooperativity Control U2B''/snRNA/U2A' RNP Formation

Sandra G. Williams and Kathleen B. Hall*

Department of Biochemistry and Molecular Biophysics, Washington University Medical School, St. Louis, Missouri 63110, United States

Supporting Information

ABSTRACT: The U1A and U2B'' proteins are components of the U1 and U2 snRNPs, respectively, where they bind to snRNA stemloops. While localization of U1A and U2B'' to their respective snRNP is a well-known phenomenon, binding of U2B'' to U2 snRNA is typically thought to be accompanied by the U2A' protein. The molecular mechanisms that lead to formation of the RNA/U2B''/U2A' complex and its localization to the U2 snRNP are investigated here, using a combination of *in vitro* RNA–protein and protein–protein fluorescence and isothermal titration calorimetry binding experiments. We find that U2A' protein binds to U2B'' with nanomolar affinity but binds to U1A with only micromolar affinity. In addition, there is RNA-dependent cooperativity (linkage) between protein–protein and protein–RNA binding. The unique combination of tight binding and cooperativity ensures that the U2A'/U2B'' complex is partitioned only to the U2 snRNP.



The spliceosomal snRNPs each consist of one snRNA bound by many proteins.¹ The U1 snRNP is the simplest snRNP, with only three unique proteins (in addition to the common Sm proteins), while the U2 snRNP not only contains more than a dozen unique proteins but also undergoes a dynamic rearrangement of its protein composition. Curiously, among the unique proteins in the U1 and U2 snRNPs of jawed vertebrates are two phylogenetically related proteins, U1A and U2B'', respectively. U1A and U2B'' are thought to uniquely localize to either snRNP, where they bind similar RNA sequences in the respective snRNAs. The U1A protein binds to U1 snRNA stemloop II (SLII), and U2B'' binds U2 snRNA stemloop IV (SLIV) (Figure 1). Protein localization occurs despite the homology and strong sequence similarity between U1A and U2B'', as well as the similarity between their RNA stemloop binding sites within the snRNAs.

U1A and U2B'' are related in sequence and structure; both have two RNA recognition motifs (RRMs) separated by an unstructured linker, and there is >70% sequence identity between RRM paralogs (Figure 1). Their N-terminal RRM1 structures are highly homologous (Figure 1), and they display conserved amino acids on their β -sheet surfaces where RNA binds. RNA binding by the U1A protein has been studied extensively^{2–5} and found to require only RRM1 for RNA recognition.^{6–8} Its C-terminal RRM2 does not bind RNA,⁹ and there are no data regarding any protein–protein interactions it may have. The N-terminal RRM (RRM1) of U2B'' has also been shown to be responsible for specific RNA binding.^{6,10} Biochemical experiments suggest that *in vitro*, the U2B'' linker does make a contribution to RNA binding affinity but that like U1A, its C-terminal RRM (RRM2) does not bind RNA.¹¹

Unlike U1A, U2B'' is bound not only to the U2 snRNA but also to the U2A' protein.^{6,10} During pre-mRNA splicing, the U2 snRNA undergoes significant conformational changes, and many U2 snRNP proteins are exchanged during these rearrangements. However, U2B'' and U2A' are found in the U2 snRNP throughout its tenure in the spliceosome.¹² Human U2A' is a modular protein with a 180-amino acid N-terminal leucine-rich repeat (LRR) domain and a C-terminus predicted to be mostly disordered (IUPRED).¹³ In a cocrystal of U2 snRNA stemloop IV (SLIV), U2B'' RRM1, and the U2A'-LRR domain,¹⁴ the RRM is sandwiched between the RNA on the surface of the β -sheet and the LRR that wraps around α 1 on the opposite face of the RRM (Figure 1).

In early studies, several reports^{15,16} concluded that U2B'' was unable to bind to the U2 snRNA specifically in the absence of U2A'. These results also showed a direct protein–protein interaction between U2B'' and U2A' and suggested that in spite of the similarity between the U1A and U2B'' protein sequences, U1A was compromised in its ability to bind U2A'.¹⁵ More quantitative experiments established that U2A' did appear to increase the apparent binding affinity of U2B'' for SLIV,^{17,18} supporting the existing hypothesis that U2A' function was to increase the affinity of U2B'' for SLIV. In more recent *in vitro* experiments with recombinant human U2B'', we showed that U2B'' does not discriminate between SLII and SLIV but that the binding affinity is still reasonably tight for both RNAs (K_D

Received: April 10, 2014

Revised: May 22, 2014

Published: May 27, 2014

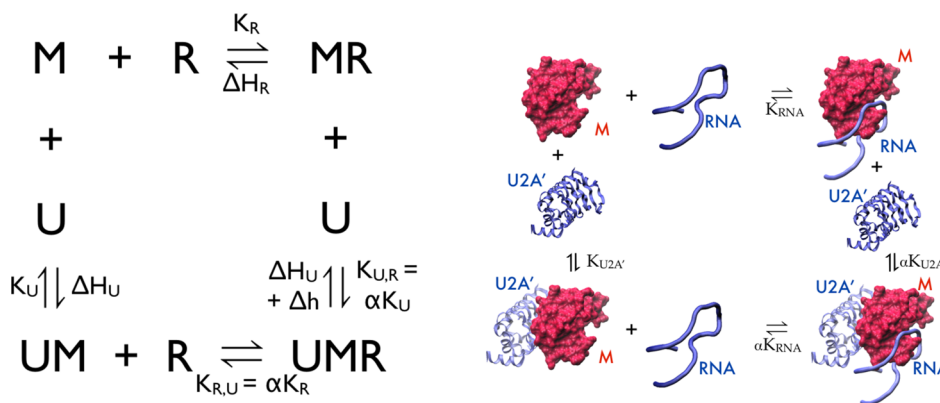


Figure 2. Schematics of binding model and thermodynamic cycles for ternary complex formation. R, U, and M represent the RNA, U2A', and RRM-containing protein (U1A or U2B''), respectively. α is the linkage parameter. K_R and K_U are the bimolecular binding constants for the RRM–RNA and RRM–U2A' interactions, respectively. The enthalpies associated with protein–protein and protein–RNA binding are indicated by ΔH_U and ΔH_R , respectively, and the enthalpy associated with cooperativity between the two binding events is indicated by Δh .

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Full-length human U1A and U2B'' proteins were purified as described previously.^{13,26} The full-length human U2A' protein was highly prone to aggregation and went entirely into inclusion bodies when it was overexpressed in *Escherichia coli*. A truncated form that included the first 180 amino acids was subcloned into our Ptac expression vector and transformed into BL-21(DE3) cells. The cells were grown in LB medium at 37 °C to an optical density of 0.6–0.8 and were induced with 0.1 mM IPTG overnight at 17 °C. Cells were harvested and stored at –70 °C until they were lysed or processed immediately. Cells were resuspended in 30 mM sodium acetate (pH 5.3), 200 mM NaCl, 2 mM EDTA, 8.5% sucrose, and 10 mM BME. PMSF, DNase II, and a protease inhibitor cocktail (Sigma) were added prior to French pressing the cells. The lysate was collected and spun down in an ultracentrifuge at 4 °C and 45000g. The supernatant was filtered through a 0.22 μm cellulose acetate membrane and loaded onto an SP Sepharose column pre-equilibrated in 50 mM Tris (pH 7.5). U2A' was eluted over 170 min, using a 50 to 375 mM NaCl gradient. All column buffers were sterile-filtered through 0.45 μm cellulose nitrate filters (Nalgene), and containers used in the purification were acid washed to remove RNases. Fractions containing U2A' were concentrated using a Vivaspin concentrator with a molecular mass cutoff of 10 kDa and buffer-exchanged into 100 mM arginine, 50 mM KCl, 10 mM cacodylate (pH 7), and 5 mM DTT; arginine was necessary to maintain protein solubility at high concentrations. Gel filtration of the protein with a Superdex 75 10/300 GL (GE) column was performed with a flow rate of 0.3 mL/min to remove impurities. The protein was eluted as a single symmetric peak. Clean fractions were collected and concentrated to ~100 μM for further use, and the final protein concentration was determined spectrophotometrically.

Fluorescently Labeled RNA Hairpins. For fluorescence binding experiments, we used chemically synthesized RNAs (IDT) with 5'-6-carboxyfluorescein (6-FAM): 5'-6-FAM-GG-GCCCGGCAUUGCACCUCGCCGGGUCC (SLII) and g5'-6-FAM-GGGCCCGGAUUGCAGUACCGCCGGGUCC (SLIV).

Loop nucleotides are underlined. To assess whether the 5'-fluorescein label affects RNA binding, these RNAs were 3'-end-labeled (using T4 RNA ligase) with [α -³²P]pCp (cytidine 3',5'-

bis-phosphate) for use in nitrocellulose filter binding experiments. FAM-RNA and RNAs transcribed with T7 RNA polymerase were bound with equal affinity by U1A and U2B'', so the FAM-RNAs were used in fluorescence experiments to measure binding affinity.

Fluorescence Titrations. U1A or U2B''/U2A' titrations were performed in 250 mM KCl, 10 mM potassium phosphate (pH 8), 1 mM MgCl₂, 40 μg/mL BSA, 5 mM DTT, and RNasin. Titrations were performed at 23 °C, with constant stirring. For a single titration of U1A/U2B'' or U1A/U2B''/U2A' into fluorescein-labeled RNA, the cuvette and titrant concentration of fluorescein-labeled RNA was held constant at 0.1 or 0.5 nM (the lower concentration was used for the highest-affinity interactions). The cuvette and titrant also contained identical concentrations of U2A'. The sample was excited at 490 nm, and the emission intensity at 520 nm was recorded (excitation and emission slit openings of 8 and 16 nm, respectively). U1A or U2B'' with or without U2A' was titrated into the RNA, and the fluorescence emission intensity was recorded for each addition of protein. The intensity data were converted to fluorescence enhancement and normalized to the maximal fluorescence enhancement to represent the fraction of bound RNA. Titrations were collected at multiple concentrations of U2A', and the data were globally fit in Scientist (Micromath) to eqs 1–4:

$$F_{M+UM} = \frac{1}{R_T} [K_R MR (1 + K_U U)] \tag{1}$$

$$M = \frac{M_T - F_{M+UM} R_T}{1 + K_U U} \tag{2}$$

$$R = \frac{R_T}{1 + K_R M + K_R K_U M U} \tag{3}$$

$$U = \frac{U_T}{1 + K_U M + K_R K_U M R} \tag{4}$$

where F_{M+UM} is the fraction of the total RNA, bound either to U1A/U2B'' (M) or to U2A':U1A/U2B'' (UM); R_T , U_T , and M_T are the total RNA, U2A', and U1A/U2B'' concentrations, respectively; R , U , and M are the concentrations of free RNA, U2A', and U1A or U2B'', respectively; α is the cooperativity parameter; and K_R and K_U are the bimolecular association constants for the SNF–RNA and SNF–U2A' interactions,

respectively. The schematic for data analysis in terms of a thermodynamic cycle of protein and RNA binding is illustrated in Figure 2.

Titration series were performed at least twice for each RNA. The parameter values represent the average of the series fits, with uncertainties that are the larger of either the propagated error or the standard deviation between measurements. For SLII, the difference in binding affinity with or without U2A' is small, such that U2A' binding affinity could not be extracted from these experiments. In fits of the binding data, the U1A–U2A' binding constant was fixed to the value obtained from experiments with SLIV. This allowed fitting of the linkage parameter α .

Partitioning surfaces were calculated in Scientist based on the model parameters determined in the fluorescence binding experiments. For these surfaces, SLII and SLIV were considered competitive ligands. The partitioning surfaces were plotted in MatLab.

ITC Experiments. RNAs for ITC experiments were transcribed *in vitro* using T7 RNA polymerase, purified via polyacrylamide gel electrophoresis, and reconstituted in water following ethanol precipitation. Concentrations were determined spectrophotometrically. The RNA was refolded by being heated to 95 °C for 3 min and then quenched on ice. Buffer was added to a total volume of 50 μ L, and this was placed in mini dialyzers [ThermoScientific, 2000 molecular weight cutoff (MWCO)] to dialyze against the final buffer [250 mM KCl, 10 mM potassium phosphate, and 1 mM MgCl₂ (pH 8)]. The RNA sequences used in ITC experiments were 5'-GGGCAC-AUUGCACCUCGUGUCCAGACUUCGGUC (SLII) and 5'-GGAGUUUCCAGGACGUAUUGCAGUACCUCGUCCUGG (SLIV). Loop nucleotides are underlined. The longer constructs were necessary to reduce the level of RNA dimerization at concentrations up to 1 mM: SLII includes a UUCG tetraloop at its 3'-end, and SLIV includes a poly-U tail at its 5'-end. Shorter constructs (like those used in the fluorescence binding experiments) were found to dimerize in a concentration-dependent manner, beginning at \sim 10 μ M (data not shown). The longer RNA constructs bind to the RRM proteins with affinities that are identical within error to those of shorter RNA constructs, when assayed by nitrocellulose filter binding (not shown).

Protein samples were diluted to 2 times their final concentration from stock solutions into 250 mM KCl, 10 mM potassium phosphate, and 1 mM MgCl₂ (pH 8) (buffer used in experiments) and dialyzed in minidialyzers (ThermoScientific, 2000 MWCO) against the experimental buffer. Final samples were prepared by diluting the RNA and/or protein samples (U1A, U2B'', U2A', or a mixture of these) with equal volumes of the final buffer supplemented with a final BME concentration of 5 mM. Samples were degassed prior to being loaded into the ITC injection syringe or cell. Titrations were performed on a NanoITC instrument (TAinstruments) and analyzed using the Triple Complex model in SedPhat.²⁷

RESULTS

RRMs, RNA, and U2A'. To determine the thermodynamic parameters of ternary complex formation, we performed titrations of full-length U1A and U2B'' into RNAs, with and without U2A' (we use only the LRR domain of human U2A'). For these experiments, the RNAs were labeled at the 5'-end with fluorescein (FAM). Protein binding by either U1A or U2B'' results in a 20% increase in FAM fluorescence upon

saturation by either protein (not shown). Addition of excess U2A' to the RNA did not change the RNA fluorescence, and the fluorescence of RNA/(U1A/U2B'') complexes was not altered by the presence of U2A'.

Binding of U2B'' and U1A to RNA was measured directly in the fluorescence experiments with or without U2A'. Representative data are shown in Figure 3, and data series were fit to a binding model that takes into account cooperativity between the protein–protein and protein–RNA interaction (see schematic and Figure 2). This provides estimates for the protein–protein and protein–RNA bimolecular binding constants, as well as the concentration-independent linkage parameter α (Tables 1 and 2).

U1A binds with subnanomolar affinity to SLII, and as Figure 3A shows, addition of U2A' modestly increases the affinity. However, these titrations show that addition of U2A' significantly increases the affinity of U1A for SLIV, and fitting these data to our binding model allowed us to estimate the affinity of U2A' for U1A [$K_{U2A',app} = (1.4 \pm 0.2) \times 10^{-6}$ M]. We used this value for fitting the cooperativity parameter α for the U2A'/U1A/SLII titrations; all of the binding parameters associated with U1A are summarized in Table 1. Rather surprisingly given the *in vivo* snRNA partitioning of U1A, the linkage parameter (α) for SLIV/U1A/U2A' binding is 89 ± 11 , corresponding to a substantial increase in the apparent U1A/SLIV binding affinity when U2A' is present.

As a short aside, full length (FL) U1A bound to SLIV with an affinity surprisingly high compared to values reported previously.²⁴ However, most studies to date have used RRM1 constructs to study RNA binding by U1A, and indeed, U1A RRM1 and FL U1A bind SLII with very similar affinities. In contrast, SLIV binding is substantially influenced by U1A construct length, with FL U1A binding more tightly to SLIV than U1A RRM1 (Figure 1 of the Supporting Information), accounting for the discrepancies in SLIV binding between this study and other studies. The U1A interdomain linker is highly positively charged, in particular at its N-terminus, which could contribute to the increased affinity of FL U1A for SLIV (the affinity of RRM1 for SLII is so tight that an effect would not be easily measured). Linker effects on binding affinity have been seen in other members of this protein family,²⁵ and we suspect that contributions of the interdomain linker to RNA binding may be a fairly general method of increasing the binding affinity of this family of proteins for RNA.

The most dramatic U2A'-dependent enhancement in RNA binding affinity was seen in the U2B''/SLIV titrations (Figure 3B and Table 2). Formation of the ternary complex is facilitated by thermodynamic linkage (cooperativity α of 140). The affinity of U2B'' for SLII is also enhanced by the presence of U2A'; however, the cooperativity parameter is smaller than that for the SLIV interaction by a factor of \sim 10 ($\alpha = 15$). The free energy associated with this cooperativity (ΔG) is -2.9 ± 0.2 kcal/mol for SLIV/U2B''/U2A' and -1.6 ± 0.1 kcal/mol for SLII/U2B''/U2A'. Linkage in the U2A'/U2B''/SLIV ternary complex leads to preferential stabilization of this species over the bimolecular species.

Localization of the Protein to the snRNPs. Partitioning of protein to the U1 and U2 snRNAs was modeled using the experimentally determined binding parameters. This is illustrated in simulations of the fractions of SLII and SLIV found in a bimolecular complex with U1A and ternary complex with U2B'' and U2A' (Figure 4; the populations of other species on SLII and SLIV are shown in Figures 2 and 3 of the

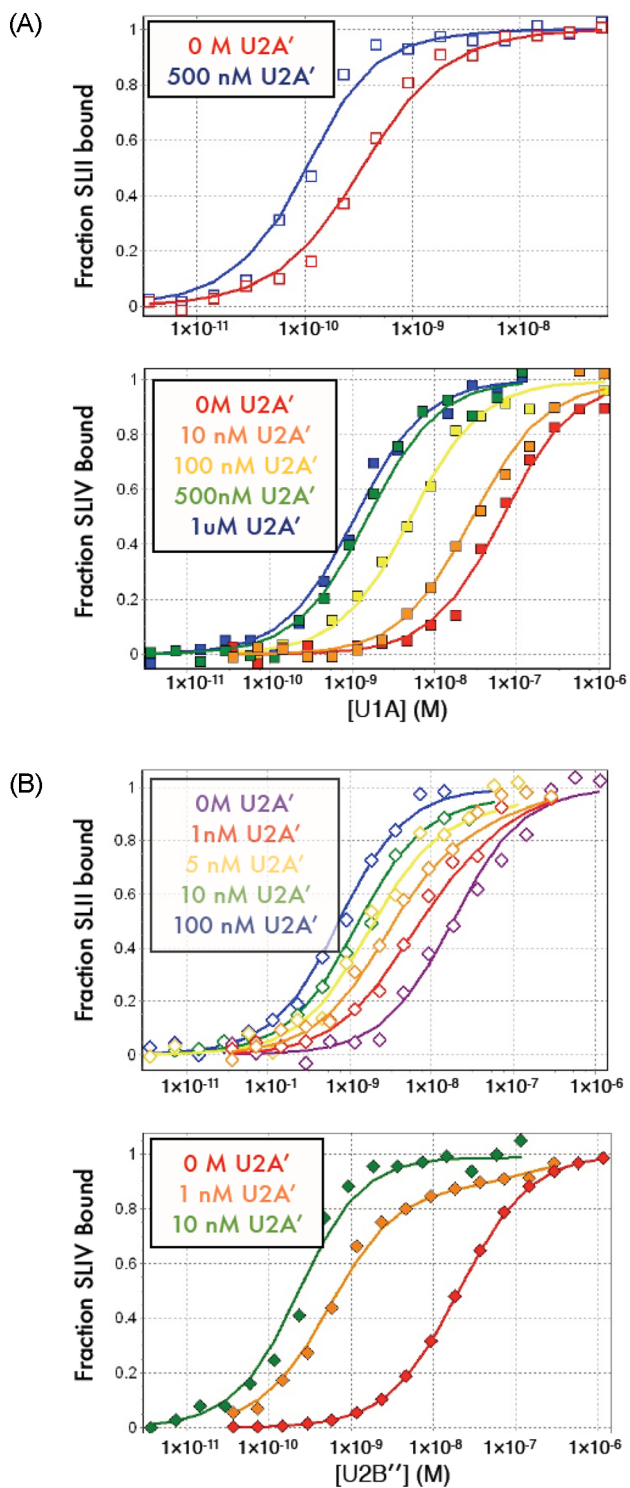


Figure 3. Binding experiments with FAM-RNA. Titrations of FAM-SLII (top) or -SLIV (bottom) into U1A (A) or U2B'' (B) were performed at different concentrations of U2A'. The U2A' concentration was kept constant for any given titration and is indicated in the figure. All fluorescence experiments were conducted in 250 mM KCl, 10 mM sodium phosphate, and 1 mM MgCl₂ (pH 8) at 22 °C. The linkage parameter for complex formation (α) is calculated to be 7.3 for U1A/SLII/U2A', 89 for U1A/SLIV/U2A', 16 for U2B''/SLII/U2A', and 140 for U2B''/SLIV/U2A'.

Supporting Information). The simulations were conducted assuming constant cellular concentrations of the RNAs: 3 μ M

Table 1. Binding Parameters from Global Fits of Titrations of U1A/U2A' into Fluorescein-Labeled SLII and SLIV^a

	SLII	SLIV
$K_{D,RNA,app}$ (M)	$(3.3 \pm 0.3) \times 10^{-10}$	$(5.8 \pm 0.4) \times 10^{-8}$
ΔG°_{RNA} (kcal/mol)	-12.8 ± 0.1	-9.8 ± 0.1
$K_{D,U2A',app}$ (M)	–	$(1.4 \pm 0.2) \times 10^{-6}$
$\Delta G^{\circ}_{U2A'}$ (kcal/mol)	–	-7.9 ± 0.1
α	7.3 ± 1.3	89 ± 11
Δg (kcal/mol)	-1.2 ± 0.1	-2.6 ± 0.1

^aDissociation constants are reported here. The affinity of U1A for U2A' was too weak to measure accurately in these experiments. Parameter values reflect the average values from at least two separate data series. The uncertainty represents the larger of either the standard deviation of the parameter values from different fits or the propagated error.

Table 2. Binding Parameters from Global Fits of Titrations of U2B''/U2A' into Fluorescein-Labeled SLII and SLIV^a

	SLII	SLIV
$K_{D,RNA,app}$ (M)	$(1.5 \pm 0.1) \times 10^{-8}$	$(1.7 \pm 0.3) \times 10^{-8}$
ΔG°_{RNA} (kcal/mol)	-10.6 ± 0.1	-10.5 ± 0.1
$K_{D,U2A',app}$ (M)	$(5.1 \pm 1.1) \times 10^{-9}$	$(4.4 \pm 2.4) \times 10^{-9}$
$\Delta G^{\circ}_{U2A'}$ (kcal/mol)	-11.2 ± 0.1	-11.3 ± 0.3
α	15.8 ± 1.6	139 ± 49
Δg (kcal/mol)	-1.6 ± 0.1	-2.9 ± 0.2

^aDissociation constants are used here. Parameter values reflect the average values from at least two separate data series. The uncertainty represents the larger of either the standard deviation of the parameter values from different fits or the propagated error.

for U1 SLII and 1.5 μ M for U2 SLIV.²⁶ It is also known that U1A is found at levels in the cell higher than those of U2B'', so simulations were performed assuming $[U1A] = 2[U2B'']$. [Results from additional simulations conducted at various U1A:U2B'' ratios of ≥ 1 showed overall results similar to those found with a 2:1 ratio (Figure 4 of the Supporting Information).]

Figure 4 shows that unless U2A' concentrations are in excess of U2B'', then U2B'' and U2A' are effectively excluded from binding U1 SLII, and the ternary complex with U2B'' readily forms on U2 SLIV (U1A is also effectively excluded from binding). The protein concentration ranges over which a bimolecular complex is formed on the U1 snRNA and a ternary complex is formed on the U2 snRNA indicate that the thermodynamics of the systems effectively partition U2A' to the U2 snRNP and prevent incorporation into the U1 snRNP.

Protein–Protein Interactions. The LRR domain of U2A' surrounds $\alpha 1$ of U2B'' as illustrated in Figure 1. At this interface, several charged residues from U2B'' (Arg25 and Glu22) form a polar patch that makes contact with U2A'. In U1A, the same polar patch comes from Lys28 and Glu25. Despite this conservation of charge, the difference in the binding affinities of U1A and U2B'' for U2A' is quite large (nearly 3 orders of magnitude). To further probe the thermodynamics of the interactions, we measured the protein–protein interactions directly by ITC. Calorimetric titrations of binding of U1A and U2B'' to U2A' are shown in Figure 5. Under our experimental solution conditions, the apparent enthalpy of U1A/U2A' binding is slightly unfavorable ($\Delta H^{\circ} = 2 \pm 1$ kcal/mol), so this association is entropically driven. In contrast, U2B'' binds to U2A' with a favorable apparent enthalpy of binding ($\Delta H^{\circ} = -85 \pm 2$ kcal/mol).

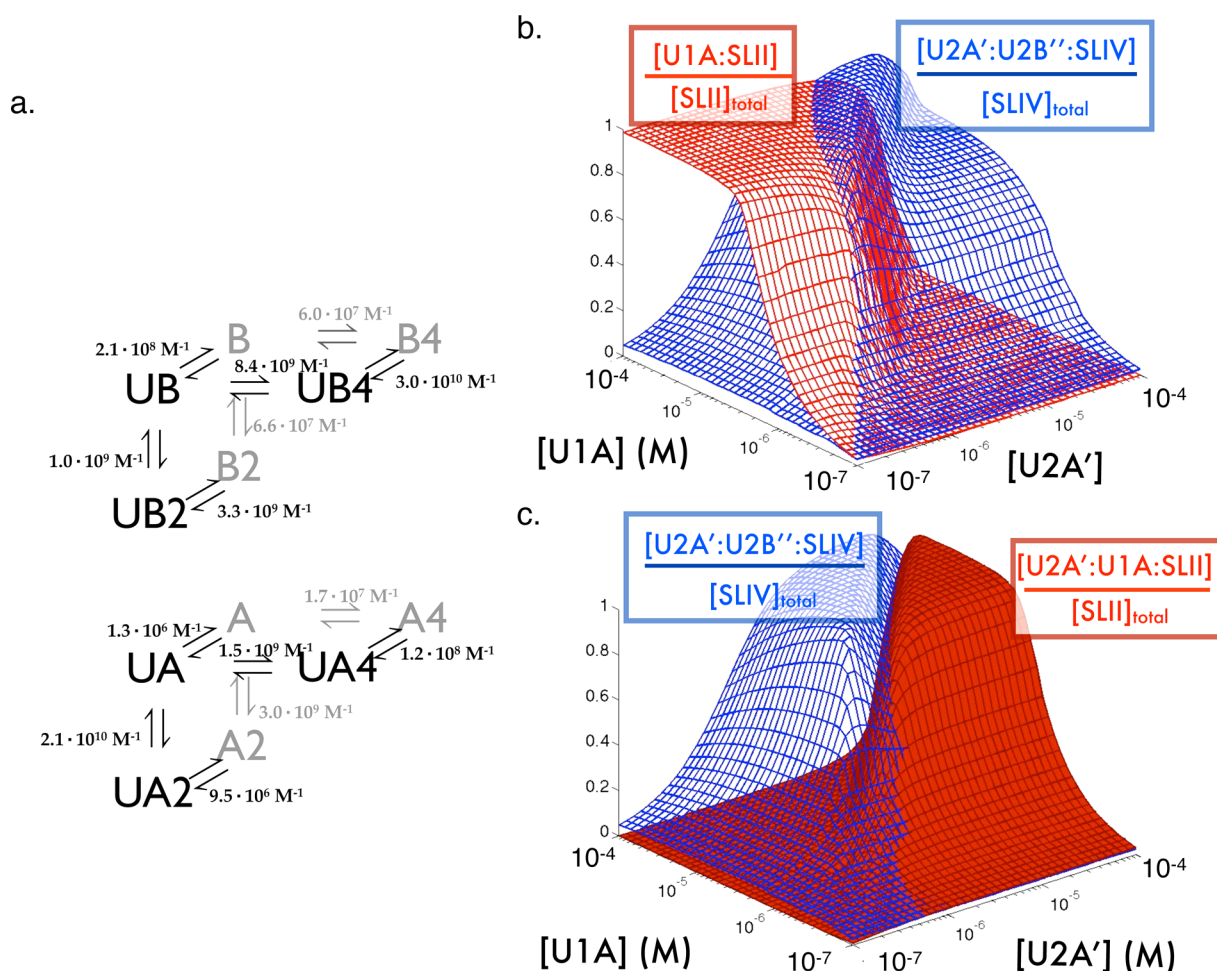


Figure 4. Protein partitioning simulations. (a) Thermodynamic model including all components (U1A, U2B^{''}, U2A', SLII, and SLIV) with all binding parameters obtained from fluorescence titrations. Protein partitioning surfaces were calculated over a wide range of U1A, U2B^{''}, and U2A' concentrations, given the thermodynamic parameters in panel a. (b) Fractions of SLII found in a bimolecular complex with U1A (red) and of SLIV in a ternary complex with U2B^{''} and U2A' (blue) are shown over a range of U1A, U2B^{''}, and U2A' concentrations. (c) Similar partitioning surface showing the fraction of SLII in a ternary complex with U1A and U2A' (red) and SLIV in a ternary complex with U2B^{''} and U2A' (blue). These simulations were conducted assuming [U1A] = 2[U2B^{''}].

Using the cocrystal structure¹⁴ of SLIV/U2B^{''}/U2A' (Figure 1) to estimate the surface areas, the binding enthalpy (at 22 °C) that can be expected from the buried surface area of the two proteins (ΔH°) is -15 kcal/mol.²⁷ Our experimental U2B^{''}/U2A' binding enthalpy is 6-fold larger and favorable. It could originate from conformational changes, linked protonation, or both, which are linked to binding. U2B^{''} uses several charged residues to contact the U2A' surface, which suggests electrostatic interactions also play a role at this interface.

Calorimetric Analysis of Ternary Complex Formation.

Having determined the calorimetric thermodynamic parameters for the binary protein–protein interactions, we determined the calorimetric parameters for ternary complex formation. Figure 6 shows results from ITC experiments based on titrations of U1A into U2A', RNA into U1A, and RNA into an equimolar mix of U1A and U2A'. Similar results are shown for U2B^{''} in Figure 7. These experiments were performed under the same solution conditions as in the fluorescence binding assays. Fitting parameters from global fits of the data are listed in Tables 3 and 4.

The calorimetry data were fit in SEDPHAT to a binding model that used the injection heats as input.^{28,29} This provides estimates for the binding enthalpies associated with both of the

bimolecular interactions (ΔH°_{U2A} and ΔH°_{RNA}) as well as the enthalpy associated with linkage (Δh). The calorimetric measurements are more limited in their sensitivity to tight binding than the fluorescence measurements. In particular, the apparent binding of U2B^{''} to SLIV in the presence of U2A' is very tight, producing an extremely steep transition in the titration (Figure 7, bottom right), making it impossible to estimate $\alpha_{U2A'/U2B''/SLIV}$ with any reliability. Similarly, binding of U1A to SLII is very tight (in the presence and absence of U2A'), making estimates of $K_{A,appU1A/SLII}$ and $\alpha_{U2A'/U1A/SLII}$ less reliable than those obtained from the fluorescence-based titrations. Within the limitations of the calorimetric data, agreement between spectroscopic and calorimetric data is reasonable, while the binding enthalpies (including the enthalpic contribution to linkage) are accurate and robust.

U2B^{''} binds to SLII and SLIV with similar apparent enthalpies of binding ($\Delta H^\circ \sim -14$ kcal/mol in these solution conditions). Binding is enthalpically driven, and the entropic cost is relatively small. Binding of U1A to SLIV has similar thermodynamic parameters. In contrast, the apparent binding enthalpy for U1A/SLII is much larger ($\Delta H^\circ = -24$ kcal/mol). This contributes to the higher binding affinity but is also

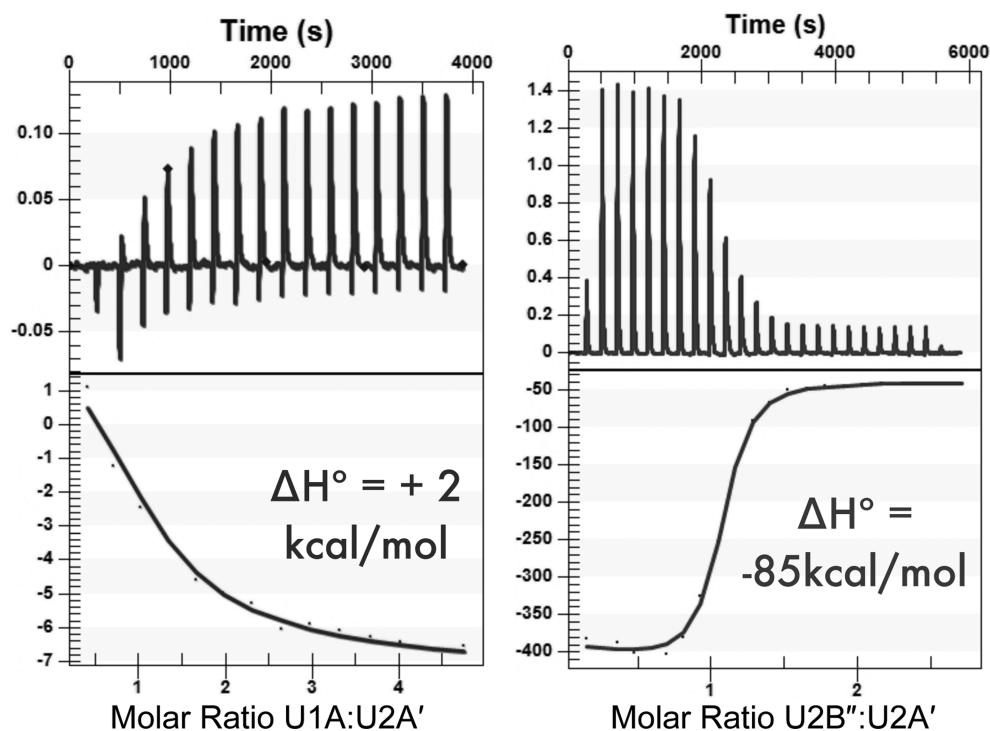


Figure 5. Protein–protein interactions. Calorimetric titrations of U1A (left) and U2B'' (right) into U2A' show very different thermodynamic signatures of binding. Calorimetric titrations were conducted in 100 mM arginine, 50 mM KCl, and 10 mM cacodylate (pH 7) at 22 °C.

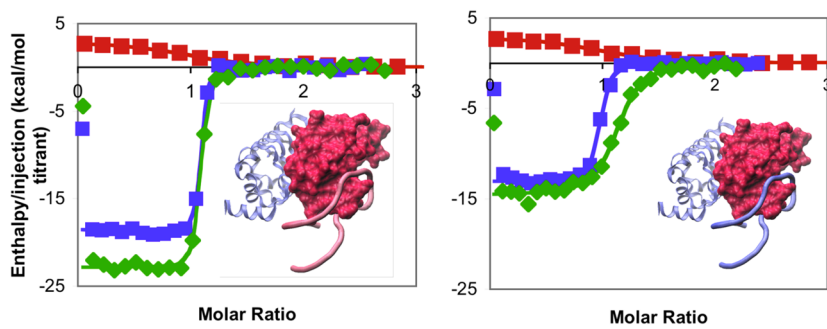


Figure 6. Calorimetric titrations and ternary complex formation for U1A. Results for SLII are shown on the left, and results for SLIV are shown on the right. Titrations of U1A into U2A' (red squares), RNA into U1A (purple squares), or RNA into an equimolar mix of U1A and U2A' (green diamonds) are shown, along with the results of globally fitting the data for each set of experiments (lines). Titrations were conducted in 250 mM KCl, 10 mM potassium phosphate (pH 8), and 1 mM MgCl_2 at 22.5 °C, and parameters from the fits are listed in Table 3.

accompanied by a larger entropic penalty under these solution conditions.

We have already noted that under the conditions studied, there is significant positive linkage (α) between U2A' and SLIV binding to U1A or U2B''. The calorimetric titrations show, however, that this is not the result of a net increase in the apparent enthalpy of binding. The enthalpy Δh associated with the linked equilibria is approximately twice as large for binding of protein to SLII ($\Delta h = 4.5$ kcal/mol) compared to binding to SLIV ($\Delta h = 2$ kcal/mol); both are unfavorable. Instead, the origin of the positive linkage is entropic. For these complexes, possible entropic contributions could come from water or ion release or the increased flexibility of a structural element.

Binding of U1A and U2B'' to SLII and SLIV is characterized by salt dependence, enthalpy–entropy compensation, and a negative heat capacity ($\Delta C_{p,\text{obs}}$).^{30,31} Previous van't Hoff determinations of U1A/SLII binding thermodynamics gave the following values: $\Delta H^\circ = -34$ kcal/mol and $\Delta S^\circ = -74$ eu

at 22 °C in 200 mM NaCl, 1 mM MgCl_2 , and 10 mM sodium cacodylate (pH 6), with a heat capacity $\Delta C_{p,\text{obs}}$ of -3.1 ± 0.4 kcal mol⁻¹ K⁻¹. Binding of SLIV to U2B'' in 250 mM KCl could be fit to a linear van't Hoff equation to give the following binding thermodynamics: $\Delta H^\circ = -16 \pm 1$ kcal/mol and $\Delta S^\circ = -21.2 \pm 3.5$ cal mol⁻¹ K⁻¹. These van't Hoff values for the enthalpy are in excellent agreement with our new calorimetric determinations at the same temperature at nearly identical salt concentrations. The effect of U2A' on the temperature dependence of RNA binding thermodynamics is still to be determined.

DISCUSSION

Human U1A and U2B'' proteins have very different affinities and specificities for U1 snRNA SLII and U2 snRNA SLIV^{11,32} and so segregate to the U1 and U2 snRNP.^{15,12} We propose, however, that the localization of U2A' to the U2 snRNP is the

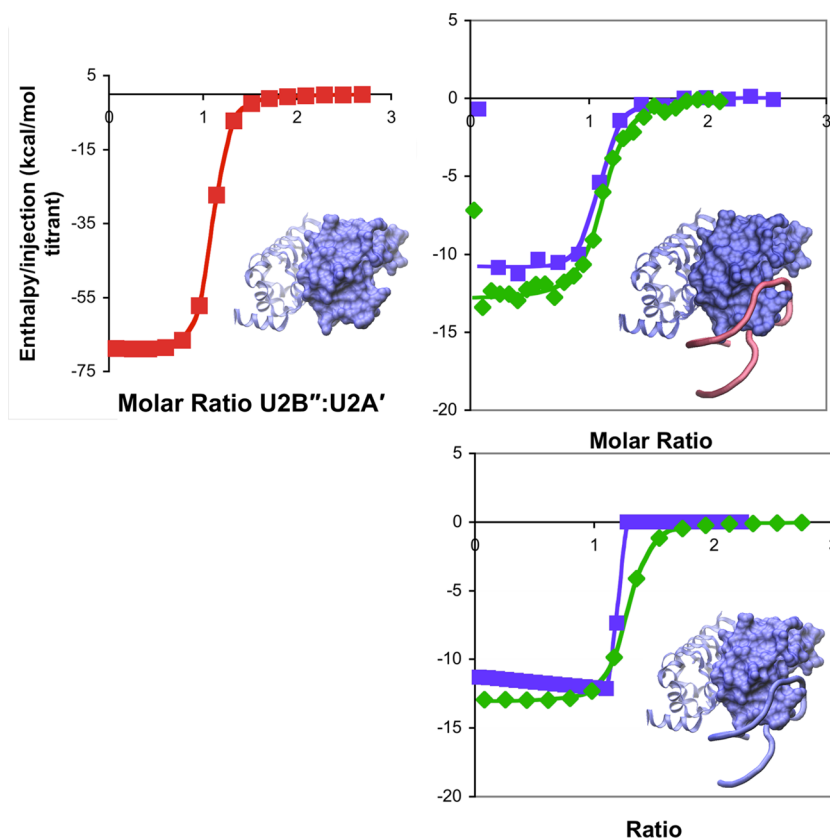


Figure 7. Calorimetric titrations and ternary complex formation for U2B''. A titration of U2B'' into U2A' is shown on the left (red squares), along with fits from global analyses of either SLII (top right) or SLIV (bottom right). RNA was titrated into U2B'' (purple squares) or into an equimolar mix of U2B'' and U2A' (green diamonds). The results of globally fitting the data for each set of experiments (lines) are also shown. Titrations were conducted in 250 mM KCl, 10 mM potassium phosphate, and 1 mM MgCl₂ (pH 8) at 22.5 °C, and parameters from the fits are listed in Table 4.

Table 3. Parameters for Global Fits of Calorimetric Titrations for U1A-Related Thermodynamic Parameters^a

	SLII	SLIV
$K_{D,U2A'}$ (M)	$(1.0 \pm 0.9) \times 10^{-6}$	$(8.3 \pm 2.2) \times 10^{-7}$
$\Delta G^{\circ}_{U2A'}$ (kcal/mol)	-8.1 ± 1.5	-8.2 ± 0.9
$\Delta H^{\circ}_{U2A'}$ (kcal/mol)	2.9 ± 1.2	2.8 ± 0.6
$K_{D,RNA}$ (M)	$<7 \times 10^{-9}$	$(1.0 \pm 1.2) \times 10^{-7}$
ΔG°_{RNA} (kcal/mol)	<-11	-9.1 ± 0.4
ΔH°_{RNA} (kcal/mol)	-23.5 ± 0.7	-14.8 ± 1.7
α	1.4 ± 7.9	34.0 ± 17.0
Δg (kcal/mol)	-0.2 ± 1.4	-2.1 ± 0.6
Δh (kcal/mol)	4.6 ± 2.5	1.5 ± 0.6

^aDissociation constants are used here. U1A binds too tightly to SLII to accurately measure affinity by ITC. Parameter values reflect the average values from at least two separate data series. The uncertainty represents the larger of either the standard deviation of the parameter values from different fits or the propagated error.

primary *raison d'être* for the combination of binding affinities and cooperativity among these RNAs and proteins.

In humans, U2A' localization is largely accomplished by different intrinsic affinities of the protein for U1A and U2B'' and by stronger linkage among SLIV, U2B'', and U2A' than among SLII, U2B'', and U2A'. The strong intrinsic affinity of U1A for SLII and its relative abundance are also important, as the binding affinity of U1A for SLII is still approximately 5-fold tighter than the affinity of U2B''/U2A' for SLII. These factors, together with the relative paucity of U2A', are sufficient to

Table 4. Parameters for Global Fits of Calorimetric Titrations for U2B''-Related Thermodynamic Parameters^a

	SLII	SLIV
$K_{D,U2A'}$ (M)	$(5.9 \pm 1.1) \times 10^{-9}$	$(6.0 \pm 1.1) \times 10^{-9}$
$\Delta G^{\circ}_{U2A'}$ (kcal/mol)	-11.1 ± 0.1	-11.1 ± 0.1
$\Delta H^{\circ}_{U2A'}$ (kcal/mol)	-73.2 ± 1.4	-73.2 ± 1.3
$K_{D,RNA}$ (M)	$(9.2 \pm 3.9) \times 10^{-8}$	$(2.3 \pm 0.5) \times 10^{-9}$
ΔG°_{RNA} (kcal/mol)	-9.5 ± 0.5	-10.3 ± 0.2
ΔH°_{RNA} (kcal/mol)	-13.3 ± 1.0	-14.9 ± 0.2
α	7.4 ± 6.8	>30
Δg (kcal/mol)	-1.2 ± 0.5	<-2
Δh (kcal/mol)	4.5 ± 0.8	2.1 ± 0.5

^aDissociation constants are reported here. Values of cooperativity (α) could not be accurately determined in ITC experiments with SLIV. Parameter values reflect the average values from at least two experiments. The uncertainty represents the larger of either the standard deviation of the parameter values from different fits or the propagated error.

maintain U1A localization to the U1 snRNP and restrict U2B''/U2A' binding.

The molecular mechanism of the cooperativity we observe in the interaction of U2B'' with SLIV and U2A' remains unknown, and we can only speculate about its origins. Ternary complex formation that includes SLIV is facilitated by cooperativity, so the SLIV sequence certainly contributes to the binding mechanism. SLII and SLIV differ in their loop-closing base pairs, in the identity of the seventh loop nucleotide (C in SLII and G in SLIV), and in an A inserted on the 3'-side

of the SLIV loop. How U2B'' interacts with these sites on SLIV could determine how it responds to U2A' binding, resulting in the cooperativity we observe. However, in the reciprocal pathway for ternary complex formation, U2B'' first interacts with U2A' and also leads to cooperative binding by SLIV. We suspect that loop 3 of U2B'' interacts with SLII and SLIV very differently, specifically near the loop-closing base pair, and that its interactions and flexibility in either binary complex enhance its subsequent interactions in the ternary complex. Loop 3 is a notable site of amino acid variation in this family of RRMs,³² particularly at its N-terminus, and the unique combination of U2B'' loop 3 amino acids, SLIV, and (uncharacterized) U2A' amino acids could lead to cooperativity in forming the SLIV/U2B''/U2A' ternary complex.

Linkage in the U1A/U2B''/SNF Family. The binding parameters that we have determined for this system are sufficient to explain the *in vivo* localization of the different protein components to the U1 and U2 snRNPs. Their reliance on intrinsic differences in binding affinity for SLII and SLIV as a dominant mechanism of localization is strikingly different from that of SNF, the single *Drosophila* protein that binds to both snRNAs. In the protein phylogeny, U1A and U2B'' emerged after a relatively recent gene duplication, while SNF is more closely related to the single ancestral protein.³² U1A, U2B'', and SNF share many properties, and among these is strong linkage with SLIV binding. All three proteins use an RNA-dependent cooperative binding mechanism to guide their snRNP localization, but while these properties of linked equilibria are fundamental to SNF segregation and contribute significantly to U2B'' ternary complex formation, they would appear to be vestigial for U1A function.

The localization of U2A' to the U2 snRNP is accomplished through distinct mechanisms in humans and *Drosophila*, and it is of interest to compare the two systems. SNF binds to both U1 snRNA SLII and U2 snRNA SLIV, but only when it is in the U2 snRNP does it form the ternary complex with RNA and dmU2A' (dm is *Drosophila*). SNF's affinity for dmU2A' is modest, but binding is very strongly coupled to SLIV binding, increasing the apparent affinity for the RNA by a factor of 350.³¹ This effect is a reciprocal one, so SLIV binding increases the apparent binding affinity of the SNF–dmU2A' interaction by 350-fold. In contrast, dmU2A' binding has almost no effect on SNF–SLII interactions. The result is that the very large difference in cooperativity is effective in partitioning U2A' to the U2 snRNP.

Following the gene duplication in an ancestor of jawed vertebrates, one protein evolved to bind U1 SLII with very high affinity and specificity, eventually becoming human U1A. In contrast, the second protein evolved to lose specificity for SLII.³³ The system also evolved a large difference in the intrinsic binding affinities of the two RRM paralogs for U2A'. Given the single protein origin of the phylogeny, it is instructive to think about how the proteins would be expected to partition if a single protein with the characteristics of either U1A or U2B'' were present in humans. Results from simulations similar to those shown in Figure 4 (maintaining the U1A and U2B'' binding parameters but considering only U1A or U2B'' to be present) are shown in Figure 5 of the Supporting Information. If U2B'' were to be lost from human cells, the partitioning of U2A' to the U2 snRNP would be significantly compromised over a substantial range of protein concentrations. If U1A were lost, the population of U2A' on the U2 snRNP would be much more significant.

Reports of U2A' function suggest that its principal biochemical role is to increase the binding affinity of U2B'' for SLIV. While linkage analysis shows that this effect undoubtedly occurs, it seems unlikely that this is the purpose of the protein. Most metazoans function with a single SLII/SLIV binding protein (as in flies), for which a duplicate high-affinity RNA target on the U2 snRNA could easily circumvent the need for an auxiliary protein. In systems with separate U1A and U2B'' proteins, the system would again be more parsimonious with two high-affinity, high-specificity RNAs determining protein partitioning. Therefore, it seems more likely that both systems evolved to exclude U2A' from the U1 snRNP and localize it to the U2 snRNP.

Biological Necessity of U2B''/U2A'. In experiments with *Xenopus* oocytes in which endogenous U2 snRNA was inactivated, pre-mRNA splicing could be rescued when exogenous U2 snRNA was expressed.^{20,21} This system was used to investigate the regions of U2 snRNA that are required for spliceosome formation and splicing. In independent experiments, these investigators deleted SLIV in exogenous U2 snRNA and observed that splicing was impaired but not inactivated. Examination of spliceosome assembly revealed that complex A (containing U1 snRNP and U2AF protein) was present in large amounts but complex B (where U2 snRNP is added) and complex C were not detected. Because splicing was observed (complex C is the active spliceosome), Hamm et al.²⁰ concluded that with the truncated U2 snRNA and the consequent absence of U2B''/U2A', these complexes were unstable and did not survive purification. Pan and Prives²¹ came to a similar conclusion, as they monitored splicing of endogenously transcribed SV40 pre-mRNA in *Xenopus* oocytes. They concluded that U2 snRNA lacking SLIV (and therefore also U2B''/U2A') was unstable.

Yeast (*S. cerevisiae*) contain both U2B'' (Yib9p or YU2B'')³⁴ and U2A' (Lea1p).³⁵ Cells lacking Lea1p, Yip9p, or both spliced at greatly reduced levels had slow growth, and levels of U2 snRNA were low. Those yeast cells accumulated Commitment Complex 2 (lacking U2 snRNP), and no prespliceosome was detected. In other studies, the yeast U2 snRNA SLIV was deleted; the result was that pre-mRNA splicing was inhibited but not abolished.^{36,37} As Caspary and Seraphin concluded,³⁵ both Yip9p and Lea1p are essential for efficient prespliceosome formation. Yeast U2B'' has only one RRM, and unlike U2A', the C-terminal domain of Lea1p is not predicted to be disordered. These different features of the proteins did not seem to preclude their binding to human U2 snRNA when it replaced the yeast snRNA³⁷ *in vivo*. In another report, the association of a GST-YU2B'' with human SLIV by an electrophoretic mobility shift assay could not be detected, although the protein did bind to human U1 SLII.³⁴

The necessity for sequestering U2B''/U2A' to the U2 snRNP may have its origin in the (unknown) function of the C-terminal tail of U2A'. This region of U2A' is predicted to be mostly disordered, although there is a putative helical region in human and *Drosophila* proteins.¹³ There are many proteins associated with the U2 snRNP, and a large proportion of them are transiently bound.¹ On the basis of the *Xenopus* and yeast results, it has been suggested that U2B''/U2A' may be essential for stable spliceosome formation. Our investigations have provided a mechanism that explains how the SLIV/U2B''/U2A' ternary complex is localized to the U2 snRNP. We propose that U2A' provides protein–protein interactions that stabilize the

U2 snRNP and the prespliceosome and that U2B'' is the scaffold that anchors U2A' to the snRNP.

■ ASSOCIATED CONTENT

● Supporting Information

Binding isotherms and simulations of protein distributions on snRNAs. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8231, St. Louis, MO 63110. E-mail: kathleenhal@gmail.com. Phone: (314) 362-4196.

Funding

This work was funded by National Institutes of Health Grants 5 R01 GM096444 to K.B.H. and F31-GM089576 to S.G.W.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Michael Rau for his help finding RNA constructs that were not prone to dimerization. He also transcribed and purified some of the RNAs used in this work. We thank Dr. Kim Delaney for cloning the human U2A' protein.

■ ABBREVIATIONS

RNP, ribonucleoprotein; RRM, RNA recognition motif; LRR, leucine-rich repeat; ITC, isothermal titration calorimetry; snRNP, small nuclear ribonucleoprotein.

■ REFERENCES

- (1) Will, C. L., and Lührmann, R. (2011) Spliceosome structure and function. *Cold Spring Harbor Perspect. Biol.* 3, 1–23.
- (2) Williams, D. J., and Hall, K. B. (1996) Thermodynamic comparison of the salt dependence of natural RNA hairpins and RNA hairpins with non-nucleotide spacers. *Biochemistry* 35, 14665–14670.
- (3) Law, M. J., Chambers, E. J., Katsamba, P. S., Haworth, I. S., and Laird-Offringa, I. A. (2005) Kinetic analysis of the role of the tyrosine 13, phenylalanine 56 and glutamine 54 network in the U1A/U1 hairpin II interaction. *Nucleic Acids Res.* 33, 2917–2928.
- (4) Kormos, B. L., Baranger, A. M., and Beveridge, D. L. (2007) A study of collective atomic fluctuations and cooperativity in the U1A-RNA complex based on molecular dynamics simulations. *J. Struct. Biol.* 157, 500–513.
- (5) Hall, K. B., and Stump, W. T. (1992) Interaction of N-terminal domain of U1A protein with an RNA stem/loop. *Nucleic Acids Res.* 20, 4283–4290.
- (6) Scherly, D., Boelens, W., van Venrooij, W. J., Dathan, N. A., Hamm, J., and Mattaj, I. W. (1989) Identification of the RNA binding segment of human U1 A protein and definition of its binding site on U1 snRNA. *EMBO J.* 8, 4163–4170.
- (7) Boelens, W. C., Scherly, D., Mattaj, I. W., and van Venrooij, W. J. (1990) Analysis of the in vitro binding of A and B'' snRNP proteins to U1 and U2 snRNAs using mutated A and B'' proteins. *Mol. Biol. Rep.* 14, 159.
- (8) Lutz-Freyermuth, C., Query, C. C., and Keene, J. D. (1990) Quantitative determination that one of two potential RNA-binding domains of the A protein component of the U1 small nuclear ribonucleoprotein complex binds with high affinity to stem-loop II of U1 RNA. *Proc. Natl. Acad. Sci. U.S.A.* 87, 6393–6397.
- (9) Lu, J., and Hall, K. B. (1995) An RBD that does not bind RNA: NMR secondary structure determination and biochemical properties

of the C-terminal RNA binding domain from the human U1A protein. *J. Mol. Biol.* 247, 739–752.

(10) Scherly, D., Boelens, W., Dathan, N. A., Kambach, C., van Venrooij, W. J., and Mattaj, I. W. (1990) Binding specificity determinants of U1A and U2B'' proteins. *Mol. Biol. Rep.* 14, 181–182.

(11) Williams, S. G., and Hall, K. B. (2011) Human U2B'' protein binding to snRNA stemloops. *Biophys. Chem.* 159, 82–89.

(12) Wahl, M. C., Will, C. L., and Lührmann, R. (2009) The spliceosome: Design principles of a dynamic RNP machine. *Cell* 136, 701–718.

(13) Dosztányi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005) IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21, 3433–3434.

(14) Price, S. R., Evans, P. R., and Nagai, K. (1998) Crystal structure of the spliceosomal U2B''-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* 394, 645–650.

(15) Scherly, D., Dathan, N. A., Boelens, W., van Venrooij, W. J., and Mattaj, I. W. (1990) The U2B'' RNP motif as a site of protein-protein interaction. *EMBO J.* 9, 3675–3681.

(16) Fresco, L. D., Harper, D. S., and Keene, J. D. (1991) Leucine periodicity of U2 small nuclear ribonucleoprotein particle (snRNP) A' protein is implicated in snRNP assembly via protein-protein interactions. *Mol. Cell. Biol.* 11, 1578–1589.

(17) Bentley, R. C., and Keene, J. D. (1991) Recognition of U1 and U2 small nuclear RNAs can be altered by a 5-amino-acid segment in the U2 small nuclear ribonucleoprotein particle (snRNP) B'' protein and through interactions with U2 snRNP-A' protein. *Mol. Cell. Biol.* 11, 1829–1839.

(18) Rimmele, M. E., and Belasco, J. G. (1998) Target discrimination by RNA-binding proteins: Role of the ancillary protein U2A' and a critical leucine residue in differentiating the RNA-binding specificity of spliceosomal proteins U1A and U2B''. *RNA* 4, 1386–1396.

(19) Flickinger, T. W., and Salz, H. K. (1994) The *Drosophila* sex determination gene *snf* encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein. *Genes Dev.* 8, 914–925.

(20) Hamm, J., Dathan, N. A., and Mattaj, I. W. (1989) Functional analysis of mutant *Xenopus* U2 snRNAs. *Cell* 59, 159–169.

(21) Pan, Z. Q., and Prives, C. (1989) U2 snRNA sequences that bind U2-specific proteins are dispensable for the function of U2 snRNP in splicing. *Genes Dev.* 3, 1887–1898.

(22) Nagengast, A. A., and Salz, H. K. (2001) The *Drosophila* U2 snRNP protein U2A' has an essential function that is SNF/U2B'' independent. *Nucleic Acids Res.* 29, 3841–3847.

(23) Will, C. L., Rümpler, S., Klein Gunnewiek, J., van Venrooij, W. J., and Lührmann, R. (1996) In vitro reconstitution of mammalian U1 snRNPs active in splicing: The U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic Acids Res.* 24, 4614–4623.

(24) Hall, K. B. (1994) Interaction of RNA hairpins with the human U1A N-terminal RNA binding domain. *Biochemistry* 33, 10076–10088.

(25) Williams, S. G., and Hall, K. B. (2010) Coevolution of *Drosophila* *snf* protein and its snRNA targets. *Biochemistry* 49, 4571–4582.

(26) Tycowski, K. T., Kolev, N. G., Conrad, N. K., Fok, V., and Steitz, J. A. (2006) The evergrowing world of small nuclear ribonucleoproteins. In *The RNA World*, 3rd ed., pp 327–368, Cold Spring Harbor Monograph Archive, Cold Spring Harbor Laboratory Press, Plainview, NY.

(27) Baker, B. M., and Murphy, K. P. (1998) Prediction of binding energetics from structure using empirical parameterization. *Methods Enzymol.* 295, 294–315.

(28) Houtman, J. C. D., Brown, P. H., Bowden, B., Yamaguchi, H., Appella, E., Samelson, L. E., and Schuck, P. (2007) Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: Application to adaptor protein complexes in cell signaling. *Protein Sci.* 16, 30–42.

- (29) Martinez-Julvez, M., Abian, O., Vega, S., Medina, M., and Velazquez-Campoy, A. (2012) Studying the allosteric energy cycle by isothermal titration calorimetry. *Methods Mol. Biol.* 796, 53–70.
- (30) Williams, D. J., and Hall, K. B. (1996) RNA hairpins with non-nucleotide spacers bind efficiently to the human U1A protein. *J. Mol. Biol.* 257, 265–275.
- (31) Williams, S. G., and Hall, K. B. (2014) Linkage and allostery in snRNP protein:RNA complexes. *Biochemistry* 53, 3529–3539.
- (32) Williams, S. G., Harms, M. J., and Hall, K. B. (2013) Resurrection of an Urbilaterian U1A/U2B^{''}/SNF protein. *J. Mol. Biol.* 425, 3846–3862.
- (33) Delaney, K. J., Williams, S. G., Lawler, M., and Hall, K. B. (2014) Climbing the vertebrate branch of U1A/U2B^{''} protein evolution. *RNA*, 10.1261/rna.044255.114.
- (34) Tang, J., Abovich, N., and Rosbash, M. (1996) Identification and characterization of a yeast gene encoding the U2 small nuclear ribonucleoprotein particle B^{''} protein. *Mol. Cell. Biol.* 16, 2787–2795.
- (35) Caspary, F., and Séraphin, B. (1998) The yeast U2A'/U2B complex is required for pre-spliceosome formation. *EMBO J.* 17, 6348–6358.
- (36) McPheeters, D. S., Fabrizio, P., and Abelson, J. (1989) In vitro reconstitution of functional yeast U2 snRNPs. *Genes Dev.* 3, 2124–2136.
- (37) Shuster, E. O., and Guthrie, C. (1990) Human U2 snRNA can function in pre-mRNA splicing in yeast. *Nature* 345, 270–273.
- (38) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual Molecular Dynamics. *J. Mol. Graphics* 14.1, 33–38.