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"

M. E. Rimmele and J. G. Belasco

RNA 1998 4: 1386-1396

RNA (1998), *4*:1386–1396. Cambridge University Press. Printed in the USA.

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″

MARTINA E. RIMMELE and JOEL G. BELASCO

Skirball Institute of Biomolecular Medicine and Department of Microbiology, New York University School of Medicine, New York, New York 10016, USA

ABSTRACT

The spliceosomal proteins U1A and U2B″ each use a homologous RRM domain to bind specifically to their respective snRNA targets, U1hpII and U2hpIV, two stem-loops that are similar yet distinct in sequence. Previous studies have shown that binding of U2B″ to U2hpIV is facilitated by the ancillary protein U2A′, whereas specific binding of U1A to U1hpII requires no cofactor. Here we report that U2A′ enables U2B″ to distinguish the loop sequence of U2hpIV from that of U1hpII but plays no role in stem sequence discrimination. Although U2A′ can also promote heterospecific binding of U1A to U2hpIV, a much higher concentration of the ancillary protein is required due to the ~500-fold greater affinity of U2A′ for U2B″. Additional experiments have identified a single leucine residue in U1A (Leu-44) that is critical for the intrinsic specificity of this protein for the loop sequence of U1hpII in preference to that of U2hpIV. Our data suggest that most of the difference in RNA-binding specificity between U1A and U2B″ can be accounted for by this leucine residue and by the contribution of the ancillary protein U2A′ to the specificity of U2B″.

Keywords: RNA-binding proteins; RNA splicing; RRM domain; snRNP; U1A; U2B″

INTRODUCTION

RNA-binding proteins (RNA-BPs) play a key role in a variety of cellular regulatory processes. Crucial to the function of this important class of proteins is their RNA target specificity. Only through their ability to distinguish between closely related RNA elements can these proteins act with the requisite selectivity.

Most RNA-BPs can be categorized into families on the basis of shared sequence motifs in their RNA-binding domains (reviewed by Mattaj, 1993; Burd & Dreyfuss, 1994). The largest of these is the RRM (RNA recognition motif) family of RNA-BPs, also referred to as the RNP family. RRM proteins are characterized by the presence of one or more structurally related RNA-binding domains, each comprising 90–100 amino acid residues and containing two conserved sequence motifs (RNP-1 and RNP-2). Members of this protein family occur in all types of organisms and bind to RNA targets that vary in sequence and secondary structure.

A paradigm for the RRM family is the spliceosomal protein U1A, a component of the U1 small nuclear ribonucleoprotein complex (U1 snRNP). U1A uses its amino-terminal RRM domain to bind hairpin II of U1 snRNA (U1hpII) (Scherly et al., 1989; Lutz-Freyermuth et al., 1990). The three-dimensional structure of this U1A domain has been solved in both the absence and presence of bound RNA (Nagai et al., 1990; Hoffman et al., 1991; Howe et al., 1994; Oubridge et al., 1994). It consists of a four-stranded β-sheet (the RNA-binding surface) supported on one side by two α-helices. The conserved RNP motifs lie on the two central β-strands (β1 and β3; see Fig. 1). It has been proposed that these conserved sequences provide basal RNA-binding activity, while the variable regions surrounding them are thought to determine binding specificity (Scherly et al., 1990a; Bentley & Keene, 1991).

Much has been learned about the specificity of RRM proteins by comparing U1A to the closely related spliceosomal protein U2B″, a U2 snRNP component that binds

Reprint requests to: Joel G. Belasco, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, New York, New York 10016, USA; e-mail: belasco@saturn.med.nyu.edu.

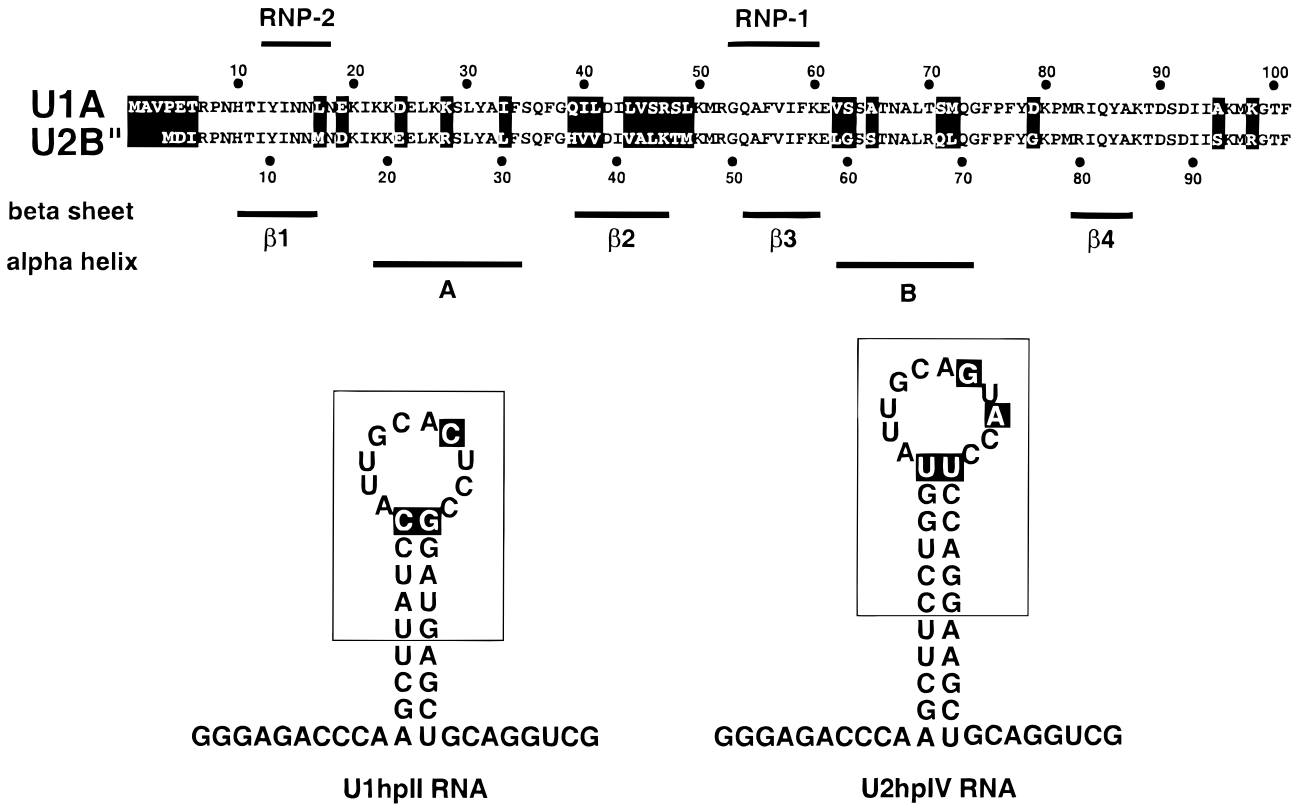


FIGURE 1. Comparison of human U1A and U2B″ protein sequences and U1hpII and U2hpIV RNA sequences. Above are shown the N-terminal domains of the two proteins. Lines mark α-helices, β-strands, and RNP consensus regions, as defined by Oubridge et al. (1994). Differences between the two proteins are highlighted. Below are the U1hpII and U2hpIV RNAs used for gel-retardation analysis. Portions identical to the human snRNA stem-loops are boxed. Differences between the two RNAs that are important for specificity are highlighted.

hairpin IV of U2 snRNA (U2hpIV). Despite being 75% identical and 94% similar in sequence (Fig. 1), the amino-terminal RRM domains of U1A and U2B″ bind their respective RNA targets selectively and with little cross-reactivity (Scherly et al., 1990a; Bentley & Keene, 1991; Scherly et al., 1991). However, in contrast to the high intrinsic affinity of U1A for U1hpII, U2B″ requires help from an ancillary snRNP protein, U2A′, to bind tightly to U2hpIV (Scherly et al., 1990a; Scherly et al., 1990b; Bentley & Keene, 1991; Boelens et al., 1991). The target RNAs U1hpII and U2hpIV also resemble one another (Fig. 1), but important sequence differences in both their loop and stem regions allow these hairpins to be distinguished by U1A and U2B″ (Scherly et al., 1990a; Hall & Stump, 1992). Presumably, a number of amino acid residues that are shared by U1A and U2B″ are important for the ability of each protein to recognize sequence features common to both U1hpII and U2hpIV, whereas differences between the two proteins and the contribution of U2A′ to RNA binding by U2B″ must account for the contrasting specificity of U1A and U2B″ in discriminating between U1hpII and U2hpIV.

Although most of the sequence differences between the amino-terminal domains of U1A and U2B″ are dispersed, nine are clustered in an 11 amino acid segment that forms the β2 strand and part of the β2/β3 loop of the RRM domain (Fig. 1). Swap experiments between U1A and U2B″ have shown that this cluster of nine amino-acid differences (U1A residues 39–41 and 44–49, U2B″ residues 36–38 and 41–46) plays a major role in distinguishing the RNA-binding specificity of these two proteins (Scherly et al., 1990b; Bentley & Keene, 1991; Jessen et al., 1991; Scherly et al., 1991). In the crystal structure of the U1A-U1hpII complex, some of these amino acids appear to contact the RNA directly (Oubridge et al., 1994). How these residues differentially interact with U1hpII and U2hpIV RNA and/or with other protein residues to enable each protein to preferentially bind its cognate RNA target and what their respective contributions are to binding specificity remains unclear.

The role of the ancillary protein U2A′ in RNA binding by U2B″ also is poorly understood. Mutational analysis suggests that U2A′ associates with the amino-terminal RRM domain of U2B″ by binding to the outside face of the first α-helix (helix A) (Fig. 1; Scherly et al., 1990b). This ancillary protein has little if any affinity for U2hpIV in the absence of U2B″ (Bentley & Keene, 1991; Boelens et al., 1991), and it is not clear whether it facilitates

RNA binding by reshaping the RNA-binding surface of U2B″ and/or by interacting directly with U2 RNA in the ternary complex. Nor is it clear why U2B″, but not U1A, requires an ancillary protein to bind its RNA target with high affinity.

Here we report mutational studies that address the structural origin of the difference in binding specificity between U1A and U2B″. Our data indicate that one β2 residue of U1A is particularly important for distinguishing its target specificity from that of U2B″. This U1A residue and the ancillary protein U2A′ both influence specificity by aiding discrimination between the loop sequences of U1hpII and U2hpIV. Together, they account for most of the difference in RNA-binding specificity between U1A and U2B″.

RESULTS

The identity of U1A residue 44 in U1A is critical for binding specificity

An earlier study (Laird-Offringa & Belasco, 1995) examined the change in U1A RNA-binding specificity that results from replacing the β2 strand and β2/β3 loop of the amino-terminal RRM domain of U1A ($U1A_N$) with the corresponding segment of U2B″. (For this and all subsequent derivatives of U1A and U2B″ described here, the subscript N indicates that the protein comprised only the amino-terminal RNA-binding domain.) The resulting hybrid protein (A/B''_N) discriminates between U1hpII and U2hpIV considerably less well than does $U1A_N$. Whereas $U1A_N$ binds 3,000 times more tightly to U1hpII than to U2hpIV, A/B''_N shows only a sevenfold preference for U1hpII (Laird-Offringa & Belasco, 1995; see also Table 1, as well as Bentley & Keene, 1991). The same study also examined the affinity of U2hpIV for several additional $U1A_N$ variants, each of which bore multiple random substitutions of U2B″ residues in the β2 strand and the β2/β3 loop. Those experiments revealed a possible correlation between increased affinity for U2hpIV and the presence of a substituted U2B″ valine residue in place of a leucine at position 44 of $U1A_N$. This finding suggested that the differential specificity of U1A and U2B″ might result in part from the identity of the amino acid residue at this position.

To test directly the importance of this valine residue for discrimination between U1hpII and U2hpIV, we constructed a $U1A_N$ variant ($U1A_N$-L44V) with a single leucine $\rightarrow$ valine substitution at position 44. $U1A_N$-L44V was purified to homogeneity, as were wild-type $U1A_N$ and A/B''_N. The affinity ($1/K_d$) of each of these polypeptides for U1hpII and U2hpIV was then compared by gel-retardation analysis by combining radiolabeled RNA with various concentrations of each purified protein (Fig. 2). In each case, a more slowly migrating band corresponding to a binary protein–RNA complex was observed, which grew in intensity as a function of protein concentration. The L44V substitution significantly increases the affinity of $U1A_N$ for U2hpIV (Fig. 2, Table 1) while impairing binding to U1hpII (Table 1). As a consequence, $U1A_N$-L44V binds much less selectively than wild-type $U1A_N$. By comparing the affinity of $U1A_N$, $U1A_N$-L44V, and A/B''_N for U1hpII and U2hpIV (Table 1), it is apparent that this single valine substitution reduces the ability of $U1A_N$ to discriminate between these two stem-loops by a factor of 14 and that replacing all of the other β2 residues as well reduces specificity by an additional factor of 35. Thus, a single amino acid residue is alone responsible for almost half of the difference in specificity between $U1A_N$ and A/B''_N.

To test the effect of the reciprocal mutation on RNA binding by U2B″, the corresponding valine residue of $U2B''_N$ was replaced with leucine ($U2B''_N$-V41L), and its affinity for U2hpIV was measured. As expected, this

TABLE 1. Binding specificity of $U1A_N$, $U1A_N$-L44V and A/B''_N.

		K_d (nM)		
		U1hpII	U2hpIV	Discrimination
$U1A_N$	Q I L . . L V S R S L	0.33 ± 0.02	1145 ± 167	3470
				14 x
$U1A_N$-L44V	Q I L . . V V S R S L	1.0 ± 0.2	243 ± 61	234
				35 x
A/B''_N	H V V . . V A L K T M	1.9 ± 0.1	13.3 ± 1.8	7.0
	39 40 41 . . 44 45 46 47 48 49			

The protein sequence corresponding to β2-region amino acids 39–49 of U1A is shown for each protein variant. U2B″-like residues are highlighted. The invariant amino acids at position 42 and 43 of U1A (Asp and Ile in both U1A and U2B″; see Fig. 1) are not shown. K_d values were determined by gel-retardation analysis (see Fig. 2). The ability of each protein to discriminate between U1hpII and U2hpIV (column on right) was calculated from the K_d ratio.

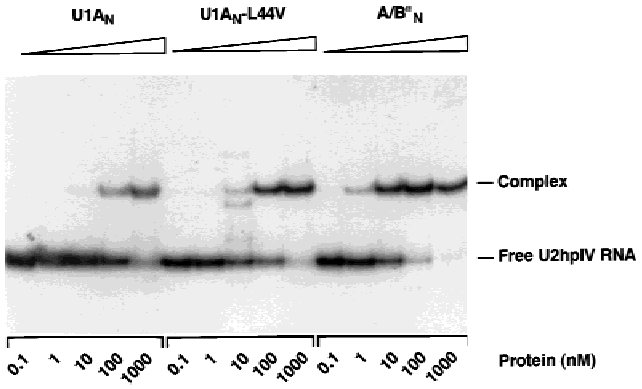


FIGURE 2. Gel-retardation analysis of U2hpIV binding to U1A variants U1A$_N$, U1A$_N$-L44V, and A/B″$_N$. Increasing amounts of purified protein (0.1 to 1,000 nM) were combined with radiolabeled U2hpIV RNA (10 pM). After 30 min at room temperature, the complexes were separated from free RNA by electrophoresis on a non-denaturing gel at 4 °C. The relative amounts of complexed RNA (binary complex with protein) and free RNA in each sample were quantified, and the K_d values (see Table 1) were determined according to Laird-Offringa and Belasco (1995). The origin of the additional bands of intermediate mobility in one U1A$_N$-L44V sample is not known.

valine → leucine mutation significantly reduces the affinity of U2B″$_N$ for U2hpIV (K_d of 27 ± 5 nM for U2B″$_N$-V41L versus 5 ± 1 nM for U2B″$_N$). Together, these findings indicate the importance of this β2 residue for distinguishing the binding specificity of U1A and U2B″.

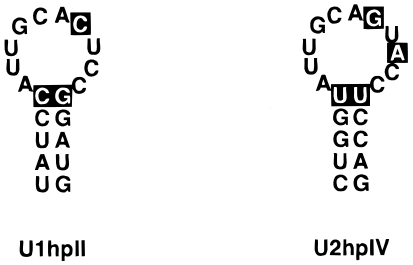
The L44V mutation enhances U1A$_N$ binding to U1hpII variants with loop mutations characteristic of U2hpIV

As a single valine substitution at position 44 of U1A$_N$ selectively increases the affinity of the protein for U2hpIV while reducing its affinity for U1hpII, we sought to identify the RNA sequence difference(s) between these two RNAs that accounts for this change in specificity. Three differences in the sequence of U2hpIV are responsible for its greater affinity for U2B″ than for U1A (Scherly et al., 1990a; Hall & Stump, 1992): a noncanonical U-U pair in place of a C-G pair at the top of the stem, a C → G substitution at the seventh nucleotide of the loop, and an additional adenylate nucleotide (+A) that increases the loop size (Table 2). To determine whether stem sequence, loop sequence, or loop size is responsible for the increased affinity for U2hpIV caused by the L44V substitution in U1A$_N$, we introduced each of these changes individually into U1hpII and determined their effect on the K_d of U1A$_N$ and U1A$_N$-L44V (Table 2).

Among the U1hpII variants tested, the relaxed specificity of the U1A$_N$-L44V mutant (versus U1A$_N$) was most pronounced with the U1hpII derivative bearing the C → G mutation. Whereas U1A$_N$ discriminates against this nucleotide change in U1hpII by a factor of 18, U1A$_N$-L44V discriminates against this change by only a factor of 3, making the L44V variant 6 times more tolerant of

TABLE 2. RNA sequence changes accommodated by the L44V mutation.

	U1A$_N$		U1A$_N$-L44V		
	K_d (nM)	Fold effect on K_d (versus U1hpII)	K_d (nM)	Fold effect on K_d (versus U1hpII)	Relative discrimination (versus U1hpII)
U1hpII	0.33 ± 0.02	1	1.0 ± 0.2	1	1
U1hpII C → G	6.0 ± 0.6	18	3.0 ± 0.4	3	6
U1hpII +A	0.62 ± 0.09	1.9	0.57 ± 0.09	0.57	3
U1hpII U–U	15 ± 2	45	24.5 ± 5.5	24.5	2
U2hpIV	1,145 ± 167	3,470	243 ± 61	243	11



The affinity of U1A$_N$ and U1A$_N$-L44V for U1hpII, U2hpIV, and three different U1hpII variants was determined by gel-retardation analysis. The U1hpII variants contained either a single nucleotide change in the loop (C → G substitution or +A insertion) or a base pair substitution at the top of the stem (U-U), each corresponding to one of the key sequence features characteristic of U2hpIV. The fold effect of each RNA mutation on K_d was calculated for each protein by dividing the K_d of that RNA by the K_d of U1hpII. The degree to which U1A$_N$ is better able than U1A$_N$-L44V to discriminate between U1hpII and each RNA variant (column on right) was calculated from the ratio of the fold effect on K_d of each RNA mutation for U1A$_N$ versus U1A$_N$-L44V. In other words, this calculated value indicates how well each RNA mutation is accommodated by the L44V amino acid substitution. Also shown are stem-loops U1hpII and U2hpIV, with their crucial nucleotide differences highlighted.

this loop mutation (Table 2). U1A$_N$-L44V also tolerates augmentation of the loop size (+A) 3 times better than does U1A$_N$. Indeed, this additional loop nucleotide actually enhances the affinity of the RNA for U1A$_N$-L44V, unlike U1A. On the other hand, the base-pair substitution at the top of the stem (C-G → U-U) significantly impairs binding by both proteins, with little (≤2-fold) additional tolerance observed for the L44V variant. Together, these data indicate that the primary effect of the L44V mutation is to render U1A$_N$ more tolerant of changes in the RNA loop.

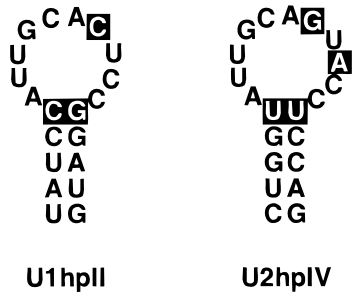
A truncated form of U2A′ enhances the affinity of U2B″$_N$ for U2hpIV

As the ability of U1A$_N$ to discriminate between U1hpII and U2hpIV is reduced when Leu-44 and other residues in the β2 strand and the β2/β3 loop are replaced with the corresponding U2B″ residues, we decided to examine the binding specificity of U2B″$_N$. Interestingly, we found that U2B″$_N$ by itself is unable to discriminate between U1hpII and U2hpIV, binding both of these stem-loops with a K_d of 5 ± 1 nM (Table 3; see also Bentley & Keene, 1991). However, in human cells U2B″ associates only with U2 snRNP (Bringmann & Luehrmann, 1986). A likely explanation for the difference in specificity observed in vivo and in vitro is suggested by previous immunoprecipitation studies, which indicated that, unlike U1A and U1hpII, U2B″ needs help from an ancillary factor, U2A′, in order to bind with high affinity to its snRNA target U2hpIV (Scherly et al., 1990a; Bentley & Keene, 1991).

Our initial efforts to examine how U2A′ achieves this effect on RNA binding were complicated by difficulty in obtaining significant amounts of recombinant U2A′ by expression in *Escherichia coli*. We solved this problem by instead overproducing and purifying a truncated form of U2A′ (U2A′$_N$) that comprised the first 203 amino acid residues of U2A′ fused to glutathione S-transferase (GST). Previously, a similar truncation mutant had been shown to bind efficiently to U2B″ despite the absence of 52 residues from the U2A′ carboxy terminus (Boelens et al., 1991). To test whether this truncated form of U2A′ can enhance the affinity of U2B″ for U2hpIV, a quantitative gel-retardation experiment was performed. When U2B″$_N$ was combined with U2hpIV in the presence of excess U2A′$_N$ (10 nM), we observed a ternary complex that migrated more slowly than the binary complex formed by U2B″$_N$ and U2hpIV in the absence of U2A′$_N$ (Fig. 3). By quantifying gel-retardation data obtained at various U2B″$_N$ concentrations, we determined that U2A′$_N$ increases the affinity of U2B″$_N$ for U2hpIV by a factor of 25. Only in the presence of this ancillary factor can U2B″$_N$ bind U2hpIV with a subnanomolar dissociation constant (0.20 ± 0.07 nM) comparable to that of the binary U1A$_N$-U1hpII complex (0.33 ± 0.02 nM). The enhancing effect of the ancillary factor on U2hpIV binding by U2B″$_N$ is half maximal at a U2A′$_N$ concentration of ~2 nM (Fig. 4). In a control experiment with GST added instead of U2A′$_N$, neither a ternary complex with U2B″$_N$ and U2hpIV nor an increase in RNA-binding affinity was observed (data not shown), thereby confirming that the U2A′ portion of the GST-U2A′ fusion protein is responsible for the chimera's

TABLE 3. RNA sequence features important for U2A′$_N$-assisted U2B″$_N$ binding.

	K_d (nM)		
	U2B″$_N$ alone	U2B″$_N$ + U2A′$_N$	Fold effect of U2A′$_N$ on binding affinity
U1hpII	5.0 ± 1.0	16 ± 3.0	0.31
U1hpII C → G	19 ± 2.0	17 ± 3.0	1.1
U1hpII +A	2.0 ± 0.5	2.0 ± 0.8	1.0
U1hpII C → G, +A	21 ± 4.0	1.0 ± 0.2	21
U1hpII U–U	333 ± 84	290 ± 49	1.1
U2hpIV	5.0 ± 1.0	0.20 ± 0.07	25



The affinity of U2B″$_N$ for various RNAs was measured by gel-retardation analysis in the presence or absence of a saturating concentration of U2A′$_N$. The fold effect of U2A′$_N$ on binding affinity was calculated from the ratio of K_d values. Also shown are stem-loops U1hpII and U2hpIV, with their crucial nucleotide differences highlighted.

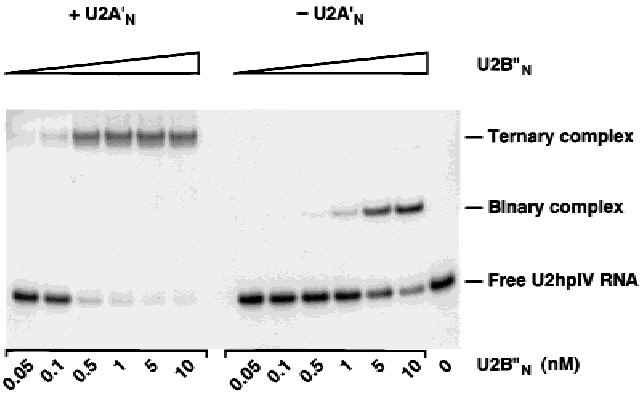


FIGURE 3. Gel-retardation analysis of U2B″$_N$ binding to U2hpIV RNA in the presence or absence of the ancillary protein U2A′$_N$. Increasing amounts of U2B″$_N$ (0.05 to 10 nM) were combined with radiolabeled U2hpIV RNA (10 pM) in the presence (left) or absence (right) of U2A′$_N$ (10 nM) and analyzed as described in Figure 2. The K_d values were calculated (Laird-Offringa & Belasco, 1995) by measuring the ratio of ternary complex to free RNA in the experiments with U2A′$_N$ or the ratio of binary complex to free RNA in the experiments without U2A′$_N$.

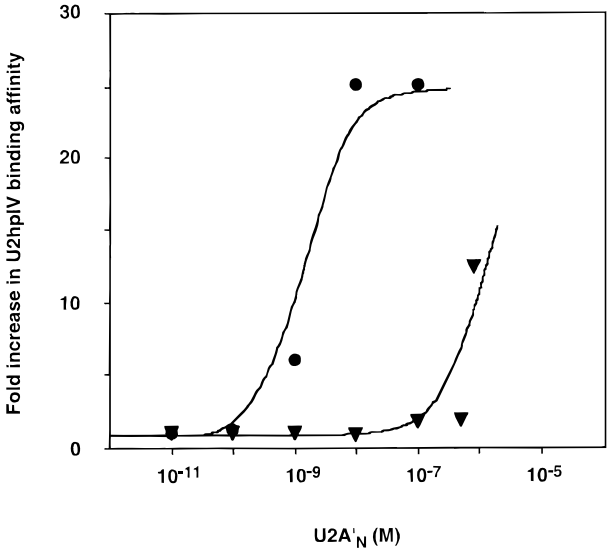


FIGURE 4. Effect of U2A′$_N$ concentration on the affinity of U2B″$_N$ and U1A$_N$ for U2hpIV RNA. The affinity ($1/K_d$) of U2B″$_N$ or U1A$_N$ for U2hpIV RNA was determined by measuring the ratio of total complexed RNA (binary + ternary complex) versus free RNA as a function of U2B″$_N$ or U1A$_N$ concentration in the absence of U2A′$_N$ and at various U2A′$_N$ concentrations. At each U2A′$_N$ concentration, the fold increase in U2hpIV binding affinity that results from the addition of U2A′$_N$ was determined for U2B″$_N$ (filled circles) and U1A$_N$ (filled triangles). Theoretical curves that best fit the data are shown. The dissociation constant of U2A′$_N$ from each ternary complex ($K_d^{U2A'_N/ternary}$) corresponds to the U2A′$_N$ concentration at which the increase in RNA-binding affinity is half maximal. For each binary protein complex, U2B″$_N$–U2A′$_N$ or U1A$_N$–U2A′$_N$, the dissociation constant of U2A′$_N$ ($K_d^{U2A'_N/binary}$) can be calculated by multiplying $K_d^{U2A'_N/ternary}$ by the maximum fold increase in RNA-binding affinity caused by U2A′$_N$. The highest U2A′$_N$ concentration that could be tested (1×10^{-6} M) was limited by the yield of this protein when overexpressed in *E. coli*.

enhancing effect. Alone, U2A′$_N$ does not bind U2hpIV detectably (data not shown). These findings show that the first 203 residues of U2A′ are sufficient to stimulate RNA binding by U2B″$_N$.

U2A′$_N$ reduces the affinity of U2B″$_N$ for U1hpII

The need for assistance from an ancillary protein to achieve high-affinity binding of U2B″$_N$ to U2hpIV might reflect an intrinsic mediocrity of U2B″$_N$ as an independent RNA-binding protein and/or features of U2hpIV that make it a difficult target for binding. To distinguish between these possibilities, we first tested whether U2A′$_N$ could assist U2B″$_N$ in binding to another stem-loop closely related to U2hpIV, namely U1hpII. Previous studies had differed as to whether U2A′$_N$ causes a modest increase or decrease in the affinity of U2B″$_N$ for U1hpII (Scherly et al., 1990b; Bentley & Keene, 1991), and it was not clear from those immunoprecipitation experiments whether sufficient U2A′$_N$ had been added to allow it to form a ternary complex with U2B″$_N$ and the noncognate U1hpII stem-loop. We therefore performed a gel-retardation experiment to determine whether such a ternary complex can form and, if so, to compare its dissociation constant with that of the binary U2B″$_N$-U1hpII complex.

At a U2A′$_N$ concentration of 1,000 nM, the predominant shifted species was a U2B″$_N$-U2A′$_N$–U1hpII ternary complex (Fig. 5). Furthermore, by varying the concentration of U2B″$_N$, the dissociation constant of this complex was determined to be 16 ± 3 nM, a value three times larger than the 5 ± 1 nM K_d of the binary U2B″$_N$-U1hpII complex. Thus, binding by U2A′$_N$ slightly reduces the affinity of U2B″$_N$ for U1hpII. The comparatively low affinity of U2A′$_N$ for the complex of U2B″$_N$ with U1hpII may help to explain why some binary complex was also observed in these experiments despite the presence of U2A′$_N$ (Fig. 5). We conclude that U2A′ is capable of joining with U2B″ to form a ternary complex with either U2hpIV or U1hpII but that the enhancing effect of U2A′ on RNA binding by U2B″ is specific for U2hpIV.

U1A and related proteins also bind better to U2hpIV in the presence of U2A′$_N$

We next investigated whether U2A′ can assist U2hpIV binding by other proteins related to U2B″. Previous studies had suggested that U1A cannot bind U2A′ detectably unless two amino acid residues in the first α-helix of U1A are exchanged for the corresponding U2B″ residues (Scherly et al., 1990b). Therefore, we introduced these two amino acid substitutions (D24E and K28R) into U1A$_N$ and U1A$_N$-L44V to create the variants U1A*$_N$ and U1A*$_N$-L44V, respectively. We then tested whether U2A′$_N$ affects binding of U2hpIV by U1A$_N$, U1A*$_N$, U1A$_N$-L44V, and U1A*$_N$-L44V.

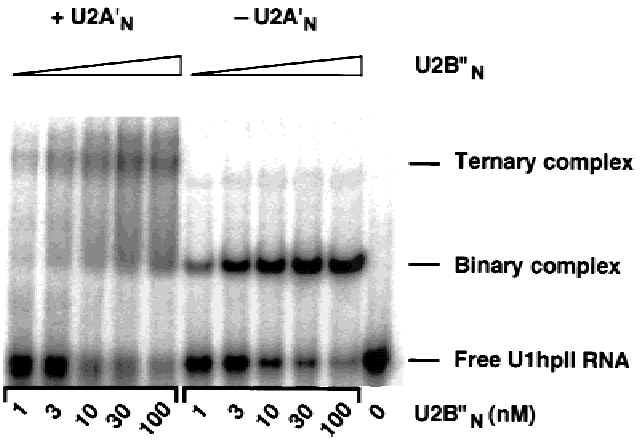


FIGURE 5. Gel-retardation analysis of U2B″$_N$ binding to U1hpII RNA in the presence or absence of U2A′$_N$. Experiments in the presence (left) or absence (right) of U2A′$_N$ were performed as described in Figure 3, except that U1hpII RNA and a higher U2A′$_N$ concentration (1,000 nM) were used. The background smear in the upper part of each lane containing U2A′$_N$ might be a consequence of partial dissociation of U2A′$_N$ from the ternary complex during electrophoresis.

Interestingly, U2A′$_N$ was found to enhance binding of U2hpIV by all four of these proteins, including U1A$_N$ (Fig. 6). Compared to gel-retardation experiments with U2B″$_N$, a 500-fold higher concentration of U2A′$_N$ (~1,000 nM) was necessary to overcome remaining impediments to the interaction of U2A′$_N$ with U1A$_N$ and its derivatives and to maximize the ability of U2A′$_N$ to influence RNA binding (Fig. 4). Under these conditions, the magnitude of the increase in U2hpIV binding (11–130-fold, depending on the U1A variant) was comparable to that observed with U2B″ (25-fold). Furthermore, as observed for U2B″$_N$, U2A′$_N$ impaired binding of U1A$_N$, U1A*$_N$, U1A$_N$-L44V, and U1A*$_N$-L44V to U1hpII (Fig. 6).

Surprisingly, the enhancing effect of U2A′$_N$ on U2hpIV binding was evident for shifted complexes of U2hpIV with U1A$_N$, U1A*$_N$, U1A$_N$-L44V, and U1A*$_N$-L44V that migrated at the position expected for a binary complex lacking U2A′$_N$, ternary complexes being barely detectable (Fig. 7). The absence of U2A′$_N$ in the predominant shifted species was confirmed by the failure of antibodies directed against the GST domain of the GST-U2A′$_N$ fusion protein to cause a supershift (data not shown). Nevertheless, U2A′$_N$ reduced the dissociation constant measured for this electrophoretic species (e.g., from 1145 ± 167 nM to 100 ± 50 nM for U1A$_N$). Apparently, the kinetic stability of these four ternary complexes is not sufficient to withstand gel electrophoresis without loss of U2A′$_N$, yet the remaining binary complexes persist and enter the gel before equilibrium can be reestablished, thereby allowing the effect of U2A′$_N$ to be "remembered." (There is ample precedent for the inability of some RNA–protein complexes to survive electrophoresis (see, e.g., Zillmann et al., 1988).)

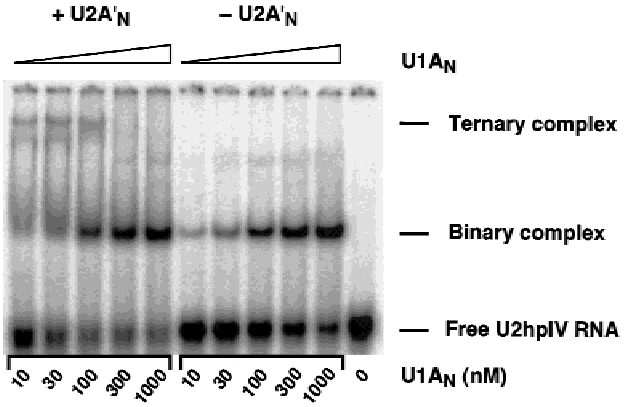


FIGURE 7. Gel-retardation analysis of U1A$_N$ binding to U2hpIV RNA in the presence or absence of U2A′$_N$. The experiments were performed as described in Figure 3, except that U1A$_N$ was used instead of U2B″$_N$ and U2A′$_N$ was added at a higher concentration (1,000 nM). Increasing amounts of U1A$_N$ (10 to 1,000 nM) were combined with radiolabeled U2hpIV RNA (10 pM) in the presence (left) or absence (right) of U2A′$_N$. Dissociation constants (K_d) were determined as described in Figures 2 and 3, using the ratio of total complexed RNA (binary + ternary complex) versus free RNA for the calculation in the presence of U2A′$_N$. The background smear in each lane containing U2A′$_N$ might be a consequence of partial dissociation of the RNA–protein complex during electrophoresis.

Together, these findings suggest that no feature unique to U2B″ makes this protein dependent on an ancillary factor for high-affinity binding to its U2 snRNA target, as U2A′$_N$ can assist even U1A$_N$ to bind U2hpIV. Instead, it appears that some aspect of the structure of U2hpIV makes it a difficult target for binding without the assistance of U2A′$_N$.

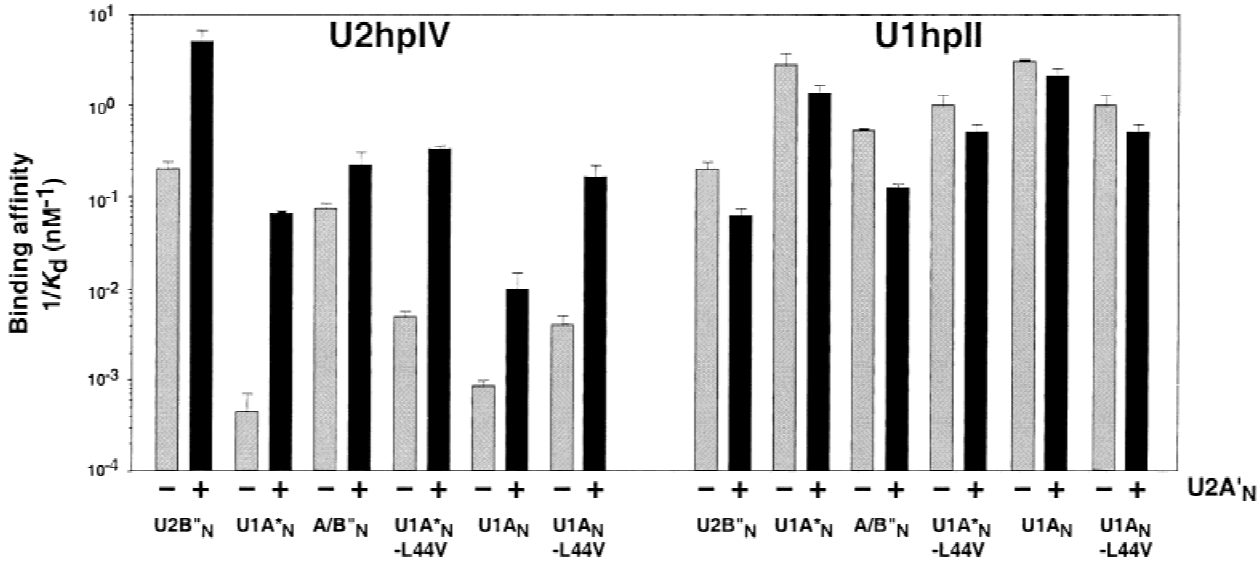


FIGURE 6. Effect of U2A′$_N$ on binding of U2hpIV or U1hpII by U2B″$_N$, U1A$_N$, and variants thereof. The bar graph shows binding affinities (1/K_d) for U2hpIV RNA (left) or U1hpII RNA (right) measured in the absence (light grey bars) or presence (black bars) of a saturating concentration of U2A′$_N$. Error estimates are indicated.

U2A′$_N$ enhances binding by accommodating sequence differences in the loop of U2hpIV

U2A′$_N$ enhances the affinity of U2B″$_N$ for U2hpIV but impairs binding to U1hpII, even though these two snRNA stem-loops are quite similar. As noted above, three key sequence differences distinguish U2hpIV from U1hpII: a noncanonical U-U pair at the top of the stem, a C → G substitution in the loop, and an additional adenosine nucleotide (+A) that augments the size of the loop (Scherly et al., 1990a). To determine which of these differences between U2hpIV and U1hpII accounts for their opposite response to U2A′$_N$, we determined the effect of U2A′$_N$ on U2B″$_N$ binding after introducing each of these three modifications individually into U1hpII (Table 3). Surprisingly, none of these individual mutations allowed U2A′$_N$ to enhance binding. However, when both loop mutations (C → G and +A) were introduced together into U1hpII, the enhancing effect of U2A′$_N$ on U2B″$_N$ binding was fully reconstituted (21-fold for the U1hpII double loop mutant, versus 25-fold for U2hpIV). These findings indicate that U2A′ facilitates U2B″ binding to U2hpIV by aiding recognition of otherwise deleterious features of the loop sequence of this RNA (C → G, +A) and that the noncanonical U-U pair in the stem of U2hpIV is not important for the contribution of U2A′ to binding.

DISCUSSION

U1A and U2B″ are closely related spliceosomal proteins that associate with different snRNPs by binding to snRNA stem-loop structures (U1hpII and U2hpIV) that are similar yet distinct. Binding by U1A is intrinsically specific, with more than a 3,000-fold preference for U1hpII over U2hpIV. Our data indicate that much of the ability of U1A to discriminate between these two stem-loops results from the presence of a leucine residue (Leu-44) at a critical position within the β2 strand of U1A. In contrast, U2B″ lacks an intrinsic ability to distinguish U2hpIV from U1hpII and instead depends entirely on an ancillary factor, U2A′, to target it preferentially to U2hpIV.

U2hpIV differs from U1hpII in three important ways: a C → G substitution at the seventh position of the loop, an additional adenylate nucleotide that increases the size of the loop to 12 nucleotides, and a noncanonical U-U base pair that replaces a C-G pair at the top of the stem (Fig. 1). Leucine-44 contributes to the binding specificity of U1A primarily by enhancing the ability of the protein to selectively recognize the loop sequence of U1hpII. Substituting the corresponding valine residue of U2B″ at this position of U1A (U1A$_N$-L44V) reduces the ability of the protein to discriminate between the loop sequences of U1hpII and U2hpIV, but does little to ameliorate the negative effect of substituting a U-U base pair at the top of the stem. Conversely, the affinity of U2B″ for U2hpIV is significantly reduced when the corresponding valine residue of that protein (Val-41) is replaced with leucine (U2B″$_N$-V41L).

Leucine-44 is particularly important for the ability of U1A to discriminate in favor of a cytidine residue at the seventh position of the U1hpII loop. Examination of the published crystal structure of U1A bound to U1hpII (Oubridge et al., 1994) reveals that Leu-44 lies at the periphery of the RNA loop and that its side chain appears to make a van der Waals contact with the ribose ring of this cytosine nucleotide. Replacement of Leu-44 with a smaller valine residue may relax the binding specificity of U1A by rendering it more sterically tolerant of RNA mutations that enlarge or distort the loop of U1hpII. Although crystallographic studies indicate that the cytosine base of the seventh loop nucleotide forms hydrogen bonds with the main-chain amide groups of U1A residues 90–92 (Oubridge et al., 1994), the identity of these three protein residues cannot explain the difference in specificity between U1A and U2B″ because they are shared by both proteins. The influence of Leu-44 on the binding specificity of U1A highlights the key role of non-hydrogen-bond interactions in sequence discrimination by RNA-binding proteins.

The importance of Leu-44 for target recognition by U1A is well illustrated by its conservation among U1A proteins in species as diverse as humans and potatoes, both of which share a very similar U1hpII loop sequence (Simpson et al., 1995). Further evidence for the biological significance of this protein residue comes from *Drosophila*, which contains a single protein, SNF/D25, that plays the role of both U1A and U2B″ (Polycarpou-Schwarz et al., 1996). This dual-purpose protein binds both to U1hpII and, with the help of a U2A′-like activity, to U2hpIV, even though these two snRNA stem-loops closely resemble their human counterparts. This property implies that SNF/D25 must have a relaxed RNA-binding specificity. Consistent with our finding that a valine residue at position 44 relaxes the specificity of U1A and facilitates binding to U2hpIV, SNF/D25 has a valine residue at this critical position.

In contrast to U1A, U2B″ alone is unable to discriminate between U2hpIV and U1hpII, binding both of these stem-loops with a K_d of 5 nM. Instead, U2B″ is entirely dependent on the ancillary protein U2A′ for its snRNA-binding specificity. U2A′$_N$ selectively enhances the affinity of U2B″$_N$ for U2hpIV by a factor of 25 while diminishing its cross-reactivity with U1hpII by a factor of 3. The increased binding specificity induced by U2A′$_N$ thus reflects a calculated 75-fold greater affinity of this ancillary protein for the complex of U2B″$_N$ with U2hpIV versus the mismatched U2B″$_N$-U1hpII complex. Only with the help of U2A′$_N$ does U2B″$_N$ attain an affinity for U2hpIV comparable to the affinity of U1A$_N$ for U1hpII. Our data show that U2A′ acts to improve the ability of U2B″ to recognize the loop sequence of U2hpIV in preference to that of

U1hpII. The two sequence differences between these loops (a guanosine substitution and an additional adenylate residue) are each located in the 3′ half of the loop (Fig. 1), and both must be present for U2A′ to enhance binding. The presence of a C-G or U-U pair at the top of the stem is inconsequential for the selective binding enhancement elicited by U2A′.

In principle, there are a number of ways in which U2A′ could aid U2B″ in recognizing U2hpIV. As this ancillary protein can bind U2B″ in the absence of RNA but has little, if any, intrinsic affinity for U2hpIV (Boelens et al., 1991; Bentley & Keene, 1991), it seems likely that, in the ternary complex, U2A′ either (1) contacts only U2B″, reshaping its RNA-binding site to better fit the loop of U2hpIV, or (2) contacts both U2B″ and U2hpIV, helping to clamp them together. In the former case, association of U2A′ with the outside face of the amino-proximal α-helix of U2B″ (via Glu-21 and Arg-25; Scherly et al., 1990b) might optimize the location of other U2B″ residues that contact the RNA loop. In the latter case, U2A′ might touch the loop nucleotides of U2hpIV directly, or it might instead aid indirectly in the recognition of the loop sequence by contacting another region of U2hpIV (e.g., the stem) whose spatial position and accessibility to interaction with U2A′ might depend on the loop sequence.

Our data show that U2A′ can assist both U2B″ and U1A to bind U2hpIV. As the magnitude of this binding enhancement is similar in both cases, the key to the selective effect of U2A′ derives instead from its much greater affinity for U2B″. From the titration data in Figure 4, we calculate that the affinity of U2A′$_N$ for U2B″$_N$ ($K_d \sim 2$ nM in the presence of bound U2hpIV and ~50 nM in the absence of an RNA ligand) is about 500-fold greater than its affinity for U1A$_N$ ($K_d \sim 1{,}000$ nM in the presence of bound U2hpIV and ~50,000 nM in its absence). Therefore, assuming that U2A′ is not present in cells in large excess, it should associate primarily with U2B″ and preferentially direct that spliceosomal protein to bind U2 snRNA.

Though unimportant for the enhancement in RNA-binding specificity elicited by U2A′, the noncanonical U-U base pair at the top of the stem of U2hpIV nevertheless plays a significant role in distinguishing the binding specificity of U1A and U2B″. Substituting a U-U pair for the C-G pair that is normally present at the top of the stem of U1hpII can markedly diminish binding by both U1A$_N$ and U2B″$_N$. In the case of U1A$_N$, the U-U stem substitution impairs RNA binding by a factor of 31–45 whether or not the two loop mutations characteristic of U2hpIV (C → G, +A) are also present (Table 2 and other data not shown). In contrast, this stem substitution impairs binding by U2B″$_N$ (by a similar factor of 18–67) only in the context of the U1hpII loop sequence; in the context of the U2hpIV loop sequence (C → G, +A), the U-U substitution instead enhances U2B″$_N$ binding 4–5-fold, whether or not the ancillary protein U2A′$_N$ is present (Table 3). Thus, the loop sequence of U2hpIV is crucial not only for the ability of U2A′ to enhance binding by U2B″ but also for the ability of U2B″ to tolerate the U-U base pair present in this stem-loop. The deleterious effect of the C-G → U-U substitution in the context of the U1hpII loop sequence may in part be a consequence of disrupting the interaction of the C-G base pair with an arginine residue present in the β3 strand of both proteins (Oubridge et al., 1994), whereas the beneficial effect of this stem substitution in the context of the U2hpIV loop sequence may be indicative of an altered loop conformation that allows U2B″ to interact productively with the noncanonical base pair. Because this beneficial effect on binding is a property of U2B″ not shared by U1A (due to the distinct sequence of its β2 strand and/or β2/β3 loop; data not shown), it constitutes another important distinguishing characteristic that helps to differentiate the RNA-binding specificity of these two proteins.

In light of these and previous data, how much of the difference in RNA-binding specificity between U1A and U2B″ can now be explained? U1A$_N$ binds U1hpII >3,000 times more tightly than U2hpIV, whereas U2B″$_N$ and U2A′$_N$ together have the opposite specificity, binding U2hpIV 80 times more tightly than U1hpII (Table 4). Thus, despite the structural similarity of U1A and U2B″ and their respective RNA targets, the RNA-binding specificity of U2B″ in combination with its ancillary factor differs from that of U1A by more than a factor of 2×10^5. In the presence of sufficient U2A′$_N$, we find that U1A*$_N$-L44V binds U1hpII and U2hpIV with nearly equal affinity (Table 4). These data indicate that the selective influence of U2A′ on RNA binding combined with the target discrimination imposed on U1A by Leu-44 accounts for most of the difference in snRNA-binding specificity between U1A and U2B″.

A variety of RNA-binding proteins function in post-transcriptional gene regulation by acting in conjunction with ancillary proteins with which they form a protein complex. With certain RNA-BPs (e.g. poly(A)-binding protein, Brown et al., 1996; Tarun et al., 1997), the identity of the ancillary protein(s) and even the regulatory mechanism can vary. The marked effect of U2A′ on the binding specificity of U2B″ is a reminder that

TABLE 4. Cumulative effect of mutations and U2A′$_N$ on specificity.

	K_d (nM)		
	U1hpII	U2hpIV	Specificity
U1A$_N$	0.33 ± 0.02	1,145 ± 167	3,470:1
U1A$_N$-L44V	1.0 ± 0.2	243 ± 61	243:1
U1A*$_N$-L44V + U2A′$_N$	2.0 ± 0.4	3.0 ± 0.3	1.5:1
U2B″$_N$ + U2A′$_N$	16 ± 3	0.20 ± 0.07	1:80

RNA binding affinities were determined by gel-retardation analysis. Values reported for U1A*$_N$-L44V and U2B″$_N$ were measured in the presence of a saturating concentration of the ancillary protein U2A′$_N$. The specificity of each protein complex (column on right) was calculated from the ratio of K_d values.

protein cofactors which seemingly lack intrinsic RNA-binding affinity may nonetheless influence the target specificity of their RNA-BP partner. In this manner, it would be possible for a single RNA-binding protein, working in combination with diverse ancillary proteins, to direct different RNAs along distinct regulatory pathways.

MATERIALS AND METHODS

Plasmids for protein and RNA synthesis

Plasmid pU1A$_N$ was constructed by subcloning a 1.34-kb *Nco* I–*Cla* I fragment of pDISPLAY-U1A101 (Laird-Offringa & Belasco, 1995) between the corresponding sites of pET3d (Studier et al., 1990). The resulting pU1A$_N$ plasmid encodes a protein fusion (U1A$_N$) comprising the amino-terminal RRM domain of human U1A (residues 1–101) joined at the carboxyl end to a (His)$_6$ affinity tag and a c-Myc epitope tag. Plasmids pU1A$_N$-L44V and pU1A*$_N$ are identical to pU1A$_N$ except for the substitution of either a valine (GTG) codon at position 44 of U1A or glutamate (GAA) and arginine (AGG) codons at positions 24 and 28, respectively. Plasmid pU1A*$_N$-L44V combines all three of these mutations.

To construct pU2B″$_N$, a segment of a human U2B″ cDNA (gift of W. Boelens; Habets et al., 1987) comprising codons 1–39 was amplified by PCR using the two primers 5′-TAACA CAACCATGGATATCAGACCAAATC-3′ (5′ primer) and 5′-GTCCACCACATGGCCAAACTGAGAAAACAG-3′ (3′ primer) under reaction conditions that favor replication fidelity (Mattila et al., 1991). The PCR product was cleaved with *Nco* I and *Msc* I and substituted for the corresponding segment of plasmid pDISPLAY-U1A/B″ (Laird-Offringa & Belasco, 1995). Inserting a 1.34-kb *Nco* I–*Cla* I fragment of the resulting plasmid between the corresponding sites of pET3d (Studier et al., 1990) yielded pU2B″$_N$. This chimera encodes the amino-terminal RRM domain of U2B″ (residues 1–98) with eight U1A-derived substitutions near the carboxyl end: L59V, G60S, S62A, Q68S, L69M, G76D, S92A, and R95K. Previous studies have indicated that these eight amino acid substitutions do not affect the RNA-binding specificity of U2B″ (Bentley & Keene, 1991). For unknown reasons, expression of an intact U2B″ cDNA lacking these substitutions was poor. Like U1A$_N$, U2B″$_N$ bears a (His)$_6$ affinity tag and a c-Myc epitope tag at the carboxyl end. Plasmid pU2B″$_N$-V41L is identical to pU2B″$_N$ except for the substitution of a leucine codon (CTA) at U2B″ position 41.

Plasmid pGST-U2A′203 encodes U2A′$_N$, a hybrid protein comprising GST fused to the first 203 residues of U2A′. To construct this plasmid, the corresponding segment of a human U2A′ cDNA clone (gift of W. Boelens; Sillekens et al., 1989) was amplified by PCR (Mattila et al., 1991), using the primers 5′-TCGGATCCGACCATGGTCAAGCTGACGGCG-3′ (5′ primer) and 5′-GCGCCCGAATTCAGTCAGTCAGCGG CCGCAGCACCTGGATTAAAAGTTTT-3′ (3′ primer). The 5′ primer created a *Bam*H I site 5′ to the first codon of U2A′ and the 3′ primer introduced a *Not* I site, a stop codon, and an *Eco*R I site 3′ to codon 203 of U2A′. After cleavage with *Bam*H I and *Eco*R I, the PCR product was cloned in-frame between the corresponding sites of pGEX-2T (Pharmacia) to generate pGST-U2A′203.

The in vitro transcription templates pEP-U1hpII and pEP-U2hpIV have been described previously (Laird-Offringa & Belasco, 1995). Variants of pEP-U1hpII bearing nucleotide substitutions or insertions were created by replacing the *Hin*d III–*Pst* I segment that encodes U1hpII with a synthetic oligonucleotide duplex of the desired sequence.

All plasmid constructs were confirmed by DNA sequencing. The pGST-U2A′203 construct was found to contain one difference from the published U2A′ sequence (Sillekens et al., 1989). A single nucleotide substitution changed Arg-193 (AGA) of U2A′$_N$ to a lysine residue (AAA). As this substitution was also present in the original U2A′ cDNA clone, it is possible that human U2A′ normally contains a lysine at this position.

Protein synthesis and purification

Plasmids for the overproduction of recombinant proteins were transformed into either BL21 (DE3) (Studier et al., 1990) containing the *lac* I^q plasmid pRJM386 (gift from D. Chattoraj, NCI) or into JM109. Protein synthesis was induced by adding IPTG (isopropyl β-D-thiogalactoside, 3 mM) to a log-phase culture growing at 30 °C. Three hours after induction, the cells were centrifuged, resuspended in sonication buffer (10 mM Tris/HCl pH 8, 250 mM NaCl, 0.5% Triton X-100), and lysed by sonication. Hexahistidine-tagged proteins were affinity-purified on Ni^{2+}-agarose beads (Qiagen), eluting with increasing imidazole concentrations in sonication buffer containing 10% glycerol, whereas U2A′$_N$ was affinity-purified on glutathione Sepharose 4B beads (Pharmacia), eluting with glutathione. The concentration of each purified protein was measured initially by Bradford assay (Bio-Rad) and then confirmed by SDS gel electrophoresis followed by either Coomassie staining and densitometry (Molecular Dynamics Personal Densitometer SI) or by Sypro Orange staining (Molecular Probes) and fluorimetry (Molecular Dynamics Fluorimeter 575).

RNA synthesis and purification

Plasmid templates for the in vitro synthesis of radiolabeled RNA were linearized with *Acc* I and then transcribed in a reaction mixture (15 μl) containing 40 mM Tris/HCl (pH 7.9), 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM DTT, RNAsin (20 units, Promega), the DNA template (2 μg), ATP, GTP, and UTP (0.3 mM each, Boehringer Mannheim), CTP (0.02 mM), [α-^{32}P] CTP (70 μCi = 88 pmol, ICN), and T7 RNA polymerase (15 units, Promega). After 16 h of in vitro transcription at 37 °C, the resulting RNA transcripts were each purified by gel electrophoresis and quantified, as described (Laird-Offringa & Belasco, 1996).

Gel-retardation analysis

RNA-binding reactions were performed in 10 μl TENT buffer (10 mM Tris/HCl, pH 7.4, 250 mM NaCl, 1 mM EDTA, 0.5% Triton X-100) containing the purified RRM-domain protein (0.1 nM to 10 μM, including concentrations above and below the measured K_d value), the radiolabeled and purified RNA target (0.1 fmol), and tRNA (10 μg). The effect of U2A′$_N$ on binding was assessed by adding this protein at a concentration of 10 to 1,000 nM, depending on the amount needed to

maximize its effect on RNA binding. In addition, to measure the K_d for dissociation of U2A′$_N$ from its ternary complex with U2hpIV and either U2B″$_N$ or U1A′$_N$, the affinity of U2B″$_N$ and U1A′$_N$ for U2hpIV RNA was measured at several different U2A′$_N$ concentrations (0.1–1,000 nM). Binding reactions were incubated for 30 min at room temperature, combined with 30% glycerol (2.5 μl), and loaded immediately onto cold, nondenaturing Tris-glycine/8% polyacrylamide gels that had been prerun for 30 min and were still running. After electrophoresis for 2 h at 150 V (4 °C), the gels were dried, and the relative amounts of complexed and free RNA in each sample were quantified on a Molecular Dynamics Phosphorimager. K_d values were calculated as previously described (Laird-Offringa & Belasco, 1995), with each reported value based on at least three independent experiments. Error estimates correspond to standard deviations.

ACKNOWLEDGMENTS

We are grateful to Ite Laird-Offringa for help in initiating these studies and for invaluable advice. We also thank Walther van Venrooij for antisera and helpful discussions. These studies were supported by a research grant (GM55624) to J.G.B. from the National Institutes of Health, by a research grant (NP-947) and a Faculty Research Award (FRA-419) to J.G.B. from the American Cancer Society, and by a postdoctoral fellowship (LT 0769/1995-M) to M.E.R. from the Human Frontier Science Program.

Received July 6, 1998; returned for revision July 20, 1998; revised manuscript received July 29, 1998

NOTE ADDED IN PROOF

Price et al. (*Nature 394*:645–650 (1998)) have recently reported the X-ray crystal structure of a ternary complex of U2hpIV bound to amino-terminal fragments of U2B″ and U2A′. They find that U2A′ contacts both U2B″ and the stem of U2hpIV. The spatial orientation of this stem, which differs from the orientation of the U1hpII stem in the U1A-U1hpII complex, allows U2A′ to contact U2hpIV. Our binding data suggest that this important difference in stem orientation is a consequence of the distinct loop sequences of these RNAs and not of the base pair at the top of the stem (U-U versus C-G).

REFERENCES

Bentley RC, Keene JD. 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.

Boelens W, Scherly D, Beijer RP, Jansen EJR, Dathan NA, Mattaj IW, van Venrooij WJ. 1991. A weak interaction between the U2A′ protein and U2 snRNA helps to stabilize their complex with the U2B″ protein. *Nucleic Acids Res 19*:455–460.

Bringmann P, Luehrmann R. 1986. Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. *EMBO J 5*:3509–3516.

Brown CE, Tarun SZ, Boeck R, Sachs AB. 1996. PAN3 encodes a subunit of the Pab1p-dependent Poly(A) nuclease in *Saccharomyces cerevisiae*. *Mol Cell Biol 16*:5744–5753.

Burd CG, Dreyfuss G. 1994. Conserved structures and diversity of functions of RNA-binding proteins. *Science 265*:615–621.

Habets WJ, Sillekens PTG, Hoet MH, Schalken JA, Roebroek AJM, Leunissen JAM, van de Ven WJM, van Venrooij WJ. 1987. Analysis of a cDNA clone expressing a human autoimmune antigen: Full-length sequence of the U2 small nuclear RNA-associated B″ antigen. *Proc Natl Acad Sci USA 84*:2421–425.

Hall KB, Stump WT. 1992. Interaction of N-terminal domain of U1A protein with an RNA stem/loop. *Nucleic Acids Res 20*:4283-4290.

Hoffman DW, Query CC, Golden BL, White SW, Keene JD. 1991. RNA-binding domain of the A protein component of the U1 small nuclear ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. *Proc Natl Acad Sci USA 88*:2495–2499.

Howe PW, Nagai K, Neuhaus D, Varani G. 1994. NMR studies of U1 snRNA recognition by the N-terminal RNP domain of the human U1A protein. *EMBO J 13*:3873–881.

Jessen TH, Oubridge C, Teo CH, Pritchard C, Nagai K. 1991. Identification of molecular contacts between the U1A small nuclear ribonucleoprotein and U1 RNA. *EMBO J 10*:3447-3456.

Laird-Offringa IA, Belasco JG. 1995. Analysis of RNA-binding proteins by in vitro genetic selection: Identification of an amino acid residue important for locking U1A onto its RNA target. *Proc Natl Acad Sci USA 92*:11859–11863.

Laird-Offringa IA, Belasco JG. 1996. In vitro genetic analysis of RNA-binding proteins using phage display libraries. *Methods Enzymol 267*:149–68.

Lutz-Freyermuth C, Query CC, Keene JD. 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 USA 87*:6393–6397.

Mattaj IW. 1993. RNA recognition: A family matter? *Cell 73*:837–840.

Mattila P, Korpela J, Tenkanen T, Pitkanen K. 1991. Fidelity of DNA synthesis by the *Thermococcus litoralis* DNA polymerase—extremely heat-stable enzyme with proofreading activity. *Nucleic Acids Res 19*:4967–973.

Nagai K, Oubridge C, Jessen TH, Li J, Evans PR. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature 348*:515–20.

Oubridge C, Ito N, Evans PR, Teo C-H, Nagai K. 1994. Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature 372*:432–438.

Polycarpou-Schwarz M, Gunderson SI, Kandels-Lewis S, Séraphin B, Mattaj IW. 1996. *Drosophila* SNF/D25 combines the functions of the two snRNP proteins U1A and U2B″ that are encoded separately in human, potato, and yeast. *RNA 2*:11–23.

Scherly D, Boelens W, Dathan NA, van Venrooij WJ, Mattaj IW. 1990a. Major determinants of the specificity of interaction between small nuclear ribonucleoproteins U1A and U2B″ and their cognate RNAs. *Nature 345*:502–506.

Scherly D, Boelens W, van Venrooij WJ, Dathan NA, Hamm J, Mattaj IW. 1989. Identification of the RNA binding segment of human U1A protein and definition of its binding site on U1 snRNA. *EMBO J 8*:4163–4170.

Scherly D, Dathan NA, Boelens W, van Venrooij WJ, Mattaj IW. 1990b. The U2B″ RNP motif as a site of protein–protein interaction. *EMBO J 9*:3675–3681.

Scherly D, Kambach C, Boelens W, van Venrooij WJ, Mattaj IW. 1991. Conserved amino acid residues within and outside of the N-terminal ribonucleoprotein motif of U1A small nuclear ribonucleoprotein involved in U1 RNA binding. *J Mol Biol 219*:577–584.

Sillekens PT, Beijer RP, Habets WJ, van Venrooij WJ. 1989. Molecular cloning of the cDNA for the human U2 snRNA-specific A′ protein. *Nucleic Acids Res 17*:1893–1906.

Simpson GG, Clark GP, Rothnie HM, Boelens W, van Venrooij W, Brown JWS. 1995. Molecular characterization of the spliceosomal proteins U1A and U2B″ from higher plants. *EMBO J 14*:4540–4550.

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol 185*:60–89.

Tarun SZ, Wells SE, Deardorff JA, Sachs AB. 1997. Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proc Natl Acad Sci USA 94*:9046–9051.

Zillmann M, Zapp ML, Berget SM. 1988. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol Cell Biol 8*:814–821.