Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

Distinct functions of the closely related tandem RNA-recognition motifs of hnRNP A1.

A Mayeda, S H Munroe, R M Xu, et al.

RNA 1998 4: 1111-1123

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *RNA* go to:
http://rnajournal.cshlp.org/subscriptions

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

RNA (1998), *4*:1111–1123. Cambridge University Press. Printed in the USA.
Copyright © 1998 RNA Society.

Distinct functions of the closely related tandem RNA-recognition motifs of hnRNP A1

AKILA MAYEDA,[1] STEPHEN H. MUNROE,[2] RUI-MING XU,[1] and ADRIAN R. KRAINER[1]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA
[2]Department of Biology, Marquette University, Milwaukee, Wisconsin 53201, USA

ABSTRACT

hnRNP A1 regulates alternative splicing by antagonizing SR proteins. It consists of two closely related, tandem RNA-recognition motifs (RRMs), followed by a glycine-rich domain. Analysis of variant proteins with duplications, deletions, or swaps of the RRMs showed that although both RRMs are required for alternative splicing function, each RRM plays distinct roles, and their relative position is important. Surprisingly, RRM2 but not RRM1 could support this function when duplicated, despite their very similar structure. Specific RNA binding and annealing are not sufficient for hnRNP A1 alternative splicing function. These observations, together with phylogenetic and structural data, suggest that the two RRMs are quasi-symmetric but functionally nonequivalent modules that evolved as components of a single bipartite domain.

Keywords: alternative splicing; hnRNP A1; RNA binding; RNA-recognition motif (RRM); RNA–RNA annealing

INTRODUCTION

Following transcription in the cell nucleus, nascent mRNA precursors (pre-mRNAs) rapidly associate with a set of abundant RNA-binding proteins, known as heterogenous nuclear ribonucleoproteins, or hnRNPs. It is in association with hnRNPs and other less abundant proteins that pre-mRNA molecules undergo processing and subsequent transport from the nucleus to the cytoplasm. More than 20 proteins, designated hnRNP A to U, have been identified by immunoprecipitation as components of hnRNP complexes in vivo (reviewed in Dreyfuss et al., 1993).

The ubiquitous association of hnRNP proteins with pre-mRNA suggests their direct involvement in many different aspects of posttranscriptional nuclear RNA metabolism. Indeed, hnRNP A1, one of the most abundant core hnRNPs, plays an active role in nuclear-cytoplasmic mRNA transport and in alternative pre-mRNA splicing. This protein has been shown to shuttle continuously between the nucleus and the cytoplasm (Piñol-Roma & Dreyfuss, 1992). Shuttling of hnRNP A1 requires a short C-terminal peptide sequence, termed M9 (Michael et al., 1995; Siomi & Dreyfuss, 1995; Weighardt et al., 1995), which interacts specifically with transportin, a component of a receptor-mediated protein import pathway (Pollard et al., 1996).

In addition to its role in mRNA transport, hnRNP A1 also plays a role in pre-mRNA alternative splicing. In transcripts containing alternative 5′ splice sites, increasing concentrations of the splicing factor SF2/ASF promote the use of the proximal 5′ splice site (Ge & Manley, 1990; Krainer et al., 1990). hnRNP A1 counteracts this activity, thereby promoting use of the distal 5′ splice site in vitro and in vivo (Mayeda & Krainer, 1992; Cáceres et al., 1994; Yang et al., 1994). SF2/ASF is the founding member of a highly conserved set of proteins, termed SR proteins, which are general splicing factors widely distributed among eukaryotes (reviewed in Cáceres & Krainer, 1997). hnRNP A1 is the prototype of a structurally related set of proteins, termed the hnRNP A/B family, which also includes hnRNP $A1^B$, A2, and B1 (reviewed in Dreyfuss et al., 1993). The antagonistic activities in the selection of alternative 5′ splice sites are also present to varying degrees in the other members of the SR and hnRNP A/B protein families (Fu et al., 1992; Mayeda & Krainer, 1992; Mayeda et al., 1993, 1994; Yang et al., 1994; Shen et al., 1995). Both SF2/ASF and hnRNP A1, as prototypes of the SR and hnRNP A/B proteins, have been well characterized with respect to their RNA-binding properties and

Reprint requests to: Adrian R. Krainer, Cold Spring Harbor Laboratory, P.O. Box 100, 1 Bungtown Road, Cold Spring Harbor, New York 11724-2208, USA; e-mail: krainer@cshl.org; or to Stephen H. Munroe, Department of Biology, Marquette University, P.O. Box 1881, Milwaukee, Wisconsin 53201, USA; e-mail: munroes@vms.csd.mu.edu.

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

their association with pre-mRNA splicing complexes. However, the mechanisms by which hnRNP A1 and SF2/ASF modulate the use of alternative 5′ splice sites remain poorly understood.

The N-terminal two-thirds of hnRNP A1 comprises two tandem copies of an RNA-recognition motif (RRM), also known as an RNA-binding domain (RBD) or a ribonucleoprotein consensus (RNP-CS), which is found in a very large number of RNA-binding proteins (reviewed in Kenan et al., 1991; Birney et al., 1993). A glycine-rich C-terminal domain (G domain) follows the second RRM. This arrangement of structural domains, RRM1-RRM2-G, is characteristic of all hnRNP A/B proteins (reviewed in Dreyfuss et al., 1993). Recently, the crystal structure of UP1, the N-terminal portion of hnRNP A1 including both RRMs, was determined (Shamoo et al., 1997; Xu et al., 1997). Each RRM of hnRNP A1 is folded into a compact globular domain, which consists of six strongly conserved secondary structure elements, β_1-α_A-β_2-β_3-α_B-β_4. The four β-strands form a single antiparallel β-sheet and the two α-helices pack tightly on one side of it. Two conserved solvent-exposed Phe residues, found at the center of the β-sheet in each RRM, have been shown to contact RNA directly (Merrill et al., 1988; Mayeda et al., 1994).

We previously analyzed the role of the individual domains of hnRNP A1 in mediating alternative splicing, facilitating the rapid annealing of complementary RNA molecules, and binding RNA (Mayeda et al., 1994). Deletion of the G domain resulted in a complete loss of alternative splicing activity and a striking decrease in RNA binding and annealing. To determine whether each of the RRMs contributes equally to these properties, each RRM was altered by mutation of two conserved Phe residues in β_3 (Mayeda et al., 1994). The replacement of these residues in either RRM resulted in the complete loss of alternative splicing, demonstrating that both RRMs are essential for alternative splicing activity. In contrast, the RNA binding and RNA annealing activity of hnRNP A1 decreased only slightly upon substitution of the Phe residues in either RRM, and only upon simultaneous mutation of both RRMs were these activities strongly affected.

One explanation for the different effects of substitutions within individual RRMs on splicing, compared with RNA binding or annealing, is that splicing requires the maintenance of precise and stringent interactions with both RRMs of hnRNP A1, but that the two RRMs are functionally redundant with respect to RNA binding and annealing. To determine whether the two RRMs of hnRNP A1 are in fact functionally redundant, and to analyze how the different structural features of the protein contribute to its role in alternative splicing, we have now examined the functional properties of a series of variant hnRNP A1 proteins in which each RRM is replaced with a duplicate copy of the other. The results revealed an unexpected functional difference between RRM1 and RRM2. These observations are discussed in the context of RRM sequences and structures, and are relevant to numerous other proteins that possess multiple RRMs.

RESULTS

Design of hnRNP A1 variants

The domain structure of hnRNP A1 is shown in Figure 1A. The overall structures of the two RRMs in hnRNP A1 are remarkably similar, both in length and amino acid sequence. Their sequences can be aligned without gaps and are 35% identical and 60% similar (Kenan et al., 1991; Matunis et al., 1992; Xu et al., 1997). This similarity is much higher than found in pairwise comparisons between RRMs from unrelated proteins (Birney et al., 1993). To determine whether the two RRMs of hnRNP A1 differ in their ability to promote RNA binding, RNA annealing, as well as changes in alternative splicing, three variant proteins were constructed in which one RRM segment of hnRNP A1 was precisely replaced with a duplicate copy of the other (Fig. 1B). The portion of each RRM selected for duplication was based on a careful analysis of the alignment of related RRM sequences (Birney et al., 1993). End points for the duplicated portions of the protein were selected to fall within conserved sequences at either end of the RRMs. In the variant protein A1-D(RRM2), the N-terminal RRM (RRM1), extending from residues 15 to 92, was replaced with the corresponding portion of the second RRM (RRM2), residues 106 to 183, to produce a protein in which two copies of RRM2 are separated by a short unaltered linker (RRM2-RRM2-G). Similarly, in the variant protein A1-D(RRM1), RRM2 was replaced with an exact copy of RRM1, to produce a protein with two copies of RRM1 (RRM1-RRM1-G). The third variant, A1-S(RRM2,1), contains a single copy of RRM1 and RRM2, but with their order reversed (RRM2-RRM1-G).

Three additional variants were constructed by deleting specific segments within the tandem RRMs. A1-ΔRRM1 lacks the N-terminal sequence (residues 2–92) comprising RRM1, and A1-ΔRRM2 lacks residues 93–183, which comprise RRM2. The inter-RRM linker (IRL) consists of a 13-amino acid peptide that connects RRM1 to RRM2 (Fig. 1A). A deletion mutant, A1-ΔIRL, was constructed in which this IRL (residues 93–105) is replaced by a single alanine. All the recombinant proteins were overexpressed in *Escherichia coli* without any tags, and were purified to apparent homogeneity (Fig. 1B).

General and sequence-specific RNA-binding properties of hnRNP A1 variant proteins

The RNA–protein interactions of wild-type and variant hnRNP A1 proteins were compared using two different binding assays: one for specific high-affinity binding

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

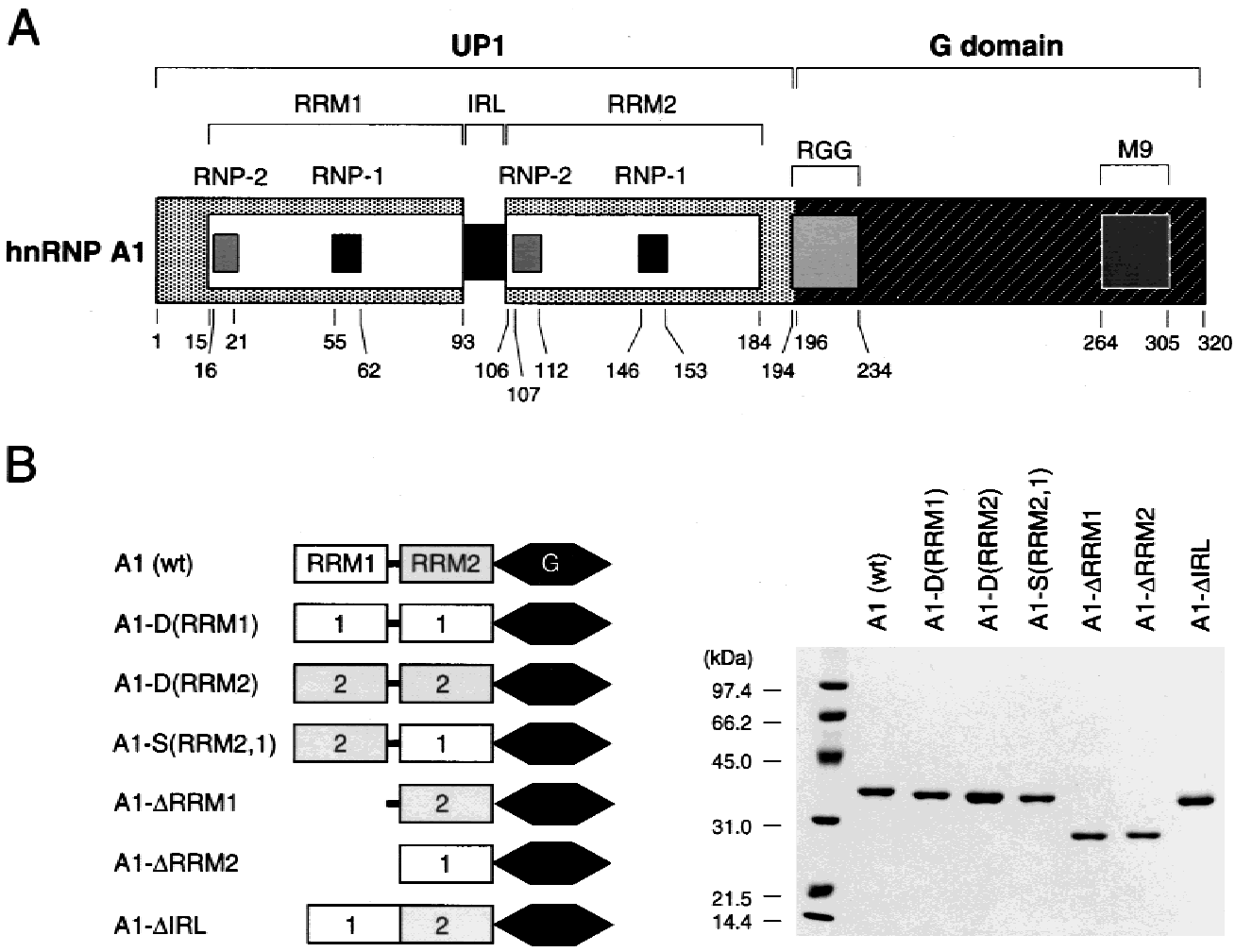

FIGURE 1. A: Schematic domain structure of human hnRNP A1. hnRNP A1 consists of two domains; the N-terminal UP1, which is a proteolytic product of hnRNP A1, and the C-terminal Gly-rich domain (G domain). UP1 consists of two copies of canonical RRMs (RRM1 and RRM2), each of which possesses the conserved RNP-2 and RNP-1 submotifs. The G domain comprises an RGG box (Kiledjian & Dreyfuss, 1992) and the M9 nuclear localization motif (Siomi & Dreyfuss, 1995; Weighardt et al., 1995) in 12 imperfect Gly-rich repeats (Cobianchi et al., 1988). Amino acid residues are numbered from the initiation codon Met 1. **B:** Purified recombinant hnRNP A1 wild-type and variant proteins. All the proteins are untagged; their structures are shown schematically at left. Proteins (60 pmol each) were analyzed by 12% SDS-PAGE and Coomassie Blue staining.

and the other for nonspecific general binding. Specific RNA binding was measured by a filter-binding assay with a short RNA probe that binds to hnRNP A1 with high affinity. This probe, designated Rd 6-1, was selected by six rounds of iterative in vitro RNA-binding selection (SELEX; see Materials and Methods). Figure 2A compares the binding of wild-type hnRNP A1 with that of three variant proteins, A1-D(RRM1), A1-D(RRM2), and A1-S(RRM2,1). Wild-type protein bound to the high-affinity probe with markedly higher affinity ($K_d \sim 10$ nM) than any of the variant proteins. The specificity of hnRNP A1 binding to Rd 6-1 is apparent from the very low level of binding observed with the control Rd 0-1 probe, which was cloned from the original pool of random RNA sequences prior to selection. Although Rd 6-1 and Rd 0-1 are identical over almost two-thirds of their sequence, they differed by at least 100-fold in their affinity for hnRNP A1.

Of the three variant proteins, A1-D(RRM2) bound significantly more strongly than A1-S(RRM2,1), and A1-D(RRM1) bound with the lowest affinity (Fig. 2A). The three deletion mutants examined also displayed reduced binding, compared with wild-type protein (Fig. 2B). Deletion of either RRM, in A1-Δ(RRM1) and A1-Δ(RRM2), led to a significant loss in binding to the Rd 6-1 probe, especially in the case of the RRM2 deletion. These comparisons suggest that the two-RRM structure is important for optimal RNA-binding activity, but that the intrinsic RNA-binding activity of RRM1 may be lower than that of RRM2. A1-ΔIRL had an intermediate affinity for Rd 6-1, which was very similar to that of A1-S(RRM2,1).

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

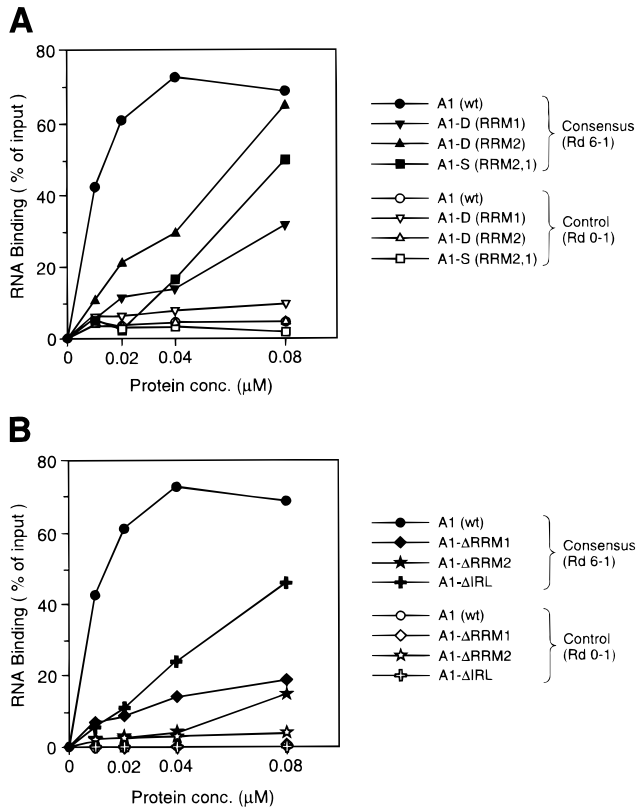

FIGURE 2. Sequence-specific RNA-binding properties of wild-type and variant hnRNP A1 proteins measured by a filter-binding assay. ^{32}P-labeled short RNAs (20 fmol) comprising the SELEX winner consensus sequence (Rd 6-1 probe) or a control sequence (Rd 0-1) were incubated under splicing conditions with the indicated concentrations of each protein. Radioactivity of protein-bound RNA retained on the filters was measured and RNA binding is expressed as percentage of the input RNA radioactivity.

General or nonspecific RNA binding by hnRNP A1 and the six variant proteins was also assayed using a 501-nt β-globin pre-mRNA splicing substrate (data not shown). Although the binding affinity of wild-type protein was about 10-fold lower than for binding to the high-affinity Rd6-1 probe, the relative affinities of wild-type and variant proteins were nearly identical with the two kinds of RNA probes. However, binding by A1-D(RRM2) was more similar to that of wild type when the β-globin probe was used.

RNA annealing properties of hnRNP A1 variant proteins

An interesting property of hnRNP A1 is its ability to promote the rapid base pairing of complementary RNA strands (Kumar & Wilson, 1990; Pontius & Berg, 1990; Munroe & Dong, 1992; Mayeda et al., 1994; Portman & Dreyfuss, 1994). RNA annealing reactions are thought to be involved in structural rearrangements of the spliceosome during splicing catalysis (reviewed in Staley & Guthrie, 1998). We therefore assayed RNA annealing of complementary single-stranded RNA molecules with hnRNP A1 variant proteins (Fig. 3). The annealing of a β-globin pre-mRNA, which was used in the general binding assay described above, was measured with an antisense RNA complementary to 200 nt at the 5′ end of the pre-mRNA. Although the relative ability of the variant proteins to facilitate annealing of complementary RNAs correlated with their specific and nonspecific RNA affinities, the A1-D(RRM2) variant was as active as wild-type A1, and even more active than the wild type at protein concentrations in the 0.3–0.6 μM range (Fig. 3A and data not shown). This increased annealing activity of A1-D(RRM2) at high protein concentrations may reflect a perturbation in the balance between duplex unwinding and strand annealing activities. Thus, the increase in annealing activity of A1-D(RRM2) may reflect its significantly decreased affinity for binding to single-stranded RNA (Fig. 2A). The other variants tested displayed annealing activities that closely paralleled their relative activities in the specific binding assays (Fig. 2A,B). A1-ΔIRL and A1-S(RRM2,1) had

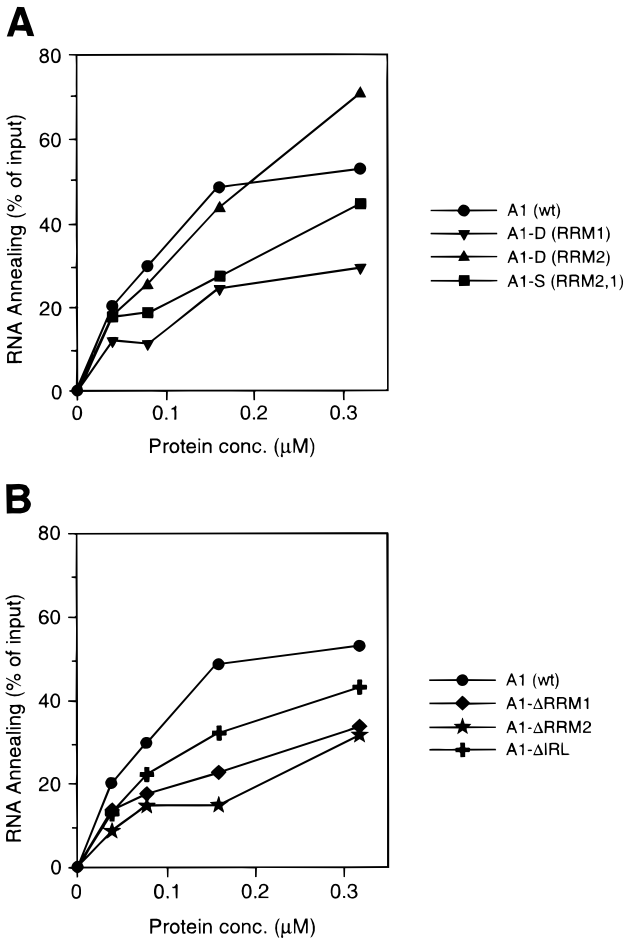

FIGURE 3. RNA annealing activity of wild-type and variant hnRNP A1 proteins. Annealing activity is expressed as the percentage of input ^{32}P-labeled pre-mRNA resistant to RNase T1 digestion, after correcting for the length of the base paired region.

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

similar levels of annealing activity, which were greater than those observed for A1-D(RRM1), A1-ΔRRM1, and A1-ΔRRM2, all of which showed similarly reduced activity in the assay.

Alternative splicing activity of hnRNP A1 variant proteins

The hnRNP A1 variant proteins were assayed for alternative splicing activity using a model β-globin pre-mRNA with a duplicated 5′ splice site (Fig. 4; Reed & Maniatis, 1986; Krainer et al., 1990). Because HeLa cell cytosolic S100 extract contains ~10-fold less hnRNP A1 than nuclear extract (Mayeda et al., 1993), the proximal 5′ splice site is almost exclusively selected in splicing reactions performed in the S100 extract complemented with SF2/ASF (Krainer et al., 1990). Therefore, this complemented extract system was used to obtain maximal switching from the proximal to the distal 5′ splice site upon addition of hnRNP A1 wild-type or variant proteins (Fu et al., 1992; Mayeda & Krainer, 1992; Mayeda et al., 1994).

Using the variant recombinant proteins with duplicated or deleted RRMs, we investigated whether the

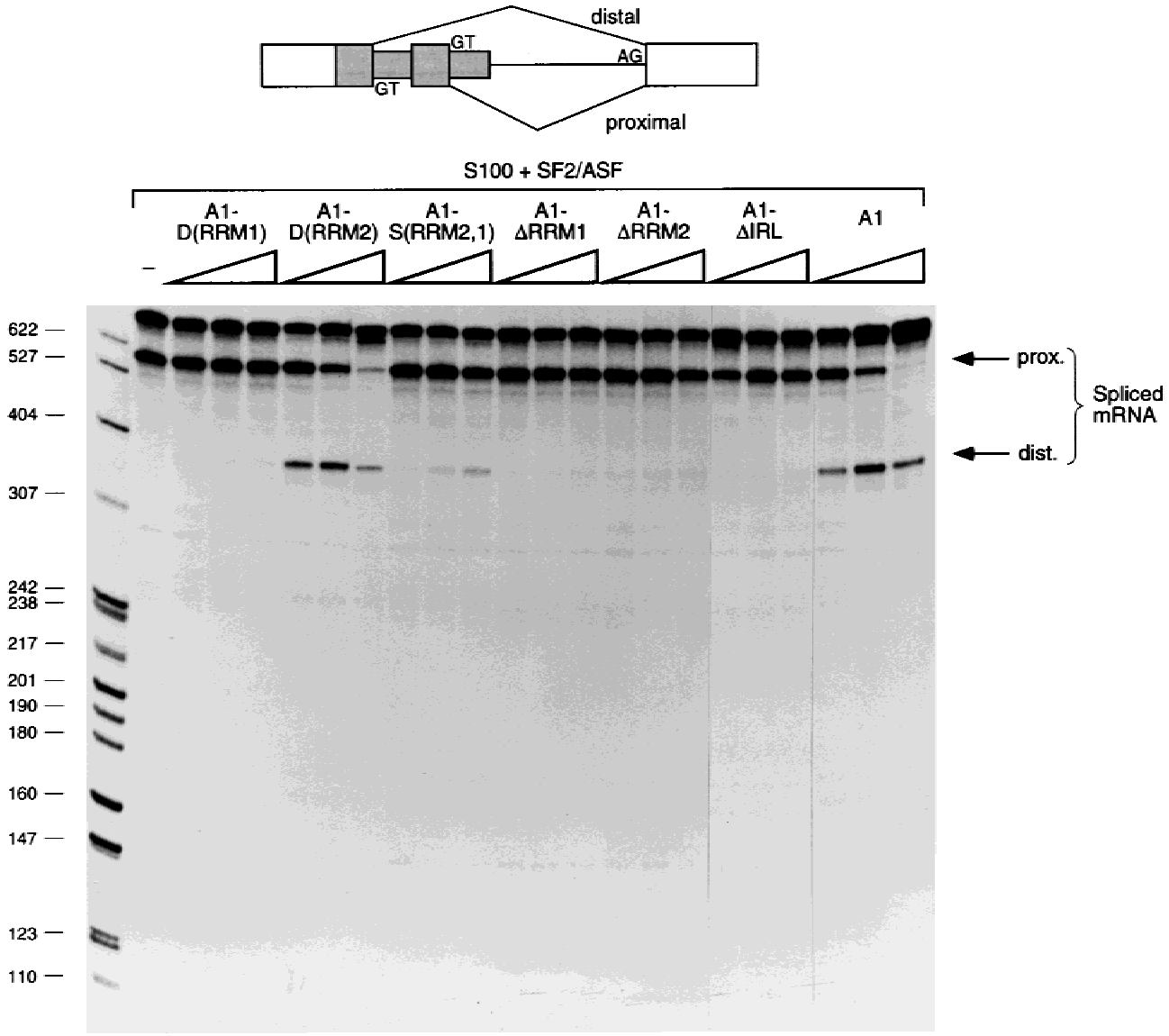

FIGURE 4. Alternative 5′ splice-site switching activity of wild-type and variant hnRNP A1 proteins. Structure of the model β-globin pre-mRNA with duplicated 5′ splice sites and the two possible splicing paths are shown schematically at the top. Splicing reactions contained 6 μl of HeLa cell S100 extract complemented with 0.4 μM purified SF2/ASF plus increasing amounts of the indicated hnRNP A1 wild-type and variant proteins: 0 (–), 0.3, 0.6, 1.2 μM final protein concentration, respectively. Positions of the spliced mRNAs generated by selection of proximal or distal 5′ splice sites are indicated. pBR322/*Hpa* II DNA markers are shown in the first lane with their sizes indicated at left.

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

individual RRMs of hnRNP A1 have similar properties with respect to promoting distal alternative 5′ splice-site selection by antagonizing the activity of SF2/ASF (Fig. 4). Surprisingly, the hnRNP variant comprising a duplicated RRM1, A1-D(RRM1), completely lacked alternative splicing activity, whereas the variant including a duplicated RRM2, A1-D(RRM2), was nearly as active as the wild type. This striking difference may indicate that RRM2 has a more significant role than RRM1 in distal 5′ splice-site selection. However, an intact two-RRM structure is necessary, because deletion of either RRM, in the A1-ΔRRM1 or A1-ΔRRM2 variants, completely abolished the splice-site switching activity. Upon swapping the positions of RRM1 and RRM2, in the A1-S(RRM2,1) variant, we observed very weak splice-site switching activity. We conclude that the presence of at least one copy of RRM2 is essential, and the relative position of RRM1 and RRM2 is important, for efficient activity. The A1-ΔIRL variant, which has a shortened inter-RRM linker connecting RRM1 to RRM2, was also tested for alternative splicing activity. This mutant was completely inactive in the alternative splicing assay. Thus, the IRL is likely to play an important role, either directly by participating in RNA contacts, or indirectly by allowing proper positioning of RRM1 and RRM2.

Several lines of evidence indicate that the variant proteins did not lose activity because of misfolding. First, there is strong sequence and structural similarity between RRM1 and RRM2. Second, all the recombinant proteins were expressed in soluble form in *E. coli*. Third, the variant proteins that were inactive in the alternative splicing assay retained significant levels of RNA binding and RNA annealing activities (Figs. 2, 3).

The variant proteins that were inactive in the splice-site switching assay did not have dominant-negative effects on splicing, because efficient splicing via the proximal 5′ splice site continued to be observed even at the highest levels of added protein (Fig. 4). Figure 5 summarizes the RNA binding, annealing, and alternative splicing activities obtained with the different proteins.

Phylogenetic analysis of hnRNP A1-like proteins

The different properties observed for the variants A1-D(RRM1) and A1-D(RRM2) demonstrated that the two RRMs of hnRNP A1 are functionally distinct. To obtain insights into the differences between these domains, we compared the sequences of individual RRMs present in proteins closely related to human hnRNP A1. A total of 17 sequences was identified by multiple BLAST searches of current databases using query sequences from each RRM of hnRNP A1. An alignment of 22 RRMs from seven vertebrate and four invertebrate representative proteins is shown in Figure 6A. This alignment extends a similar sequence analysis of *Drosophila* hnRNP A/B-related proteins (Matunis et al., 1992). Additional vertebrate sequences with more than 95% identity to those shown were omitted, as were isoforms arising from alternative splicing, because the isoform-specific sequences lie outside of their RRMs or IRLs.

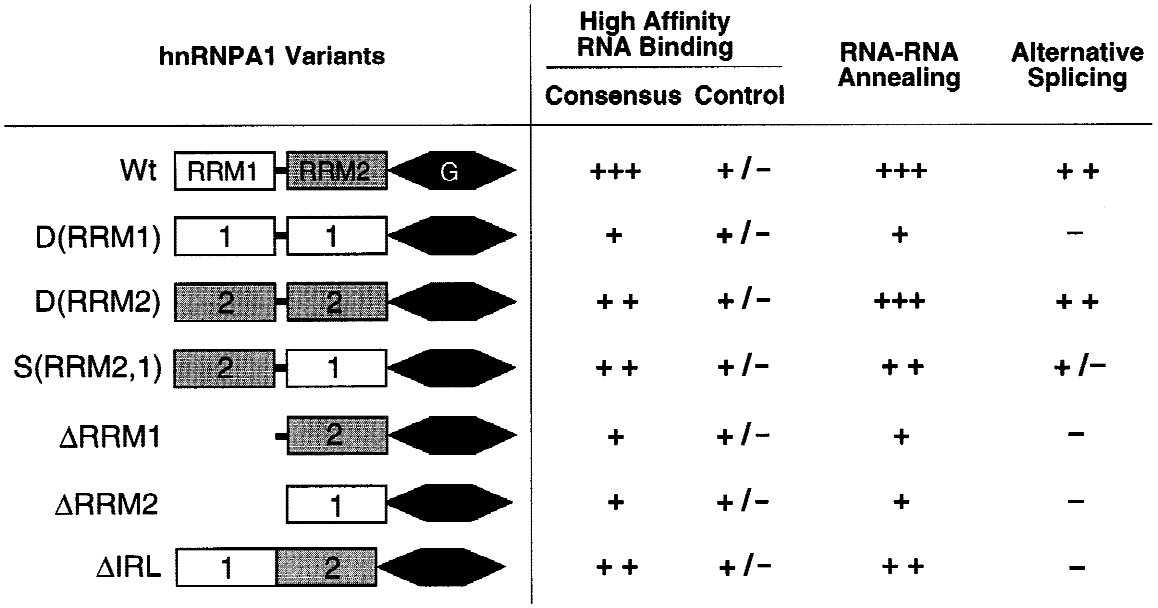

FIGURE 5. Summary of the activities of hnRNP A1 variant proteins. Structures of the wild-type (Wt) and variant proteins are shown schematically at left. The RRMs and G domain are indicated by boxes and hexagons, respectively. Relative specific activities for high-affinity RNA binding, RNA annealing, and alternative 5′ splice-site switching are indicated by + and − signs. See Figures 2, 3, and 4 for data used in scoring these values.

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

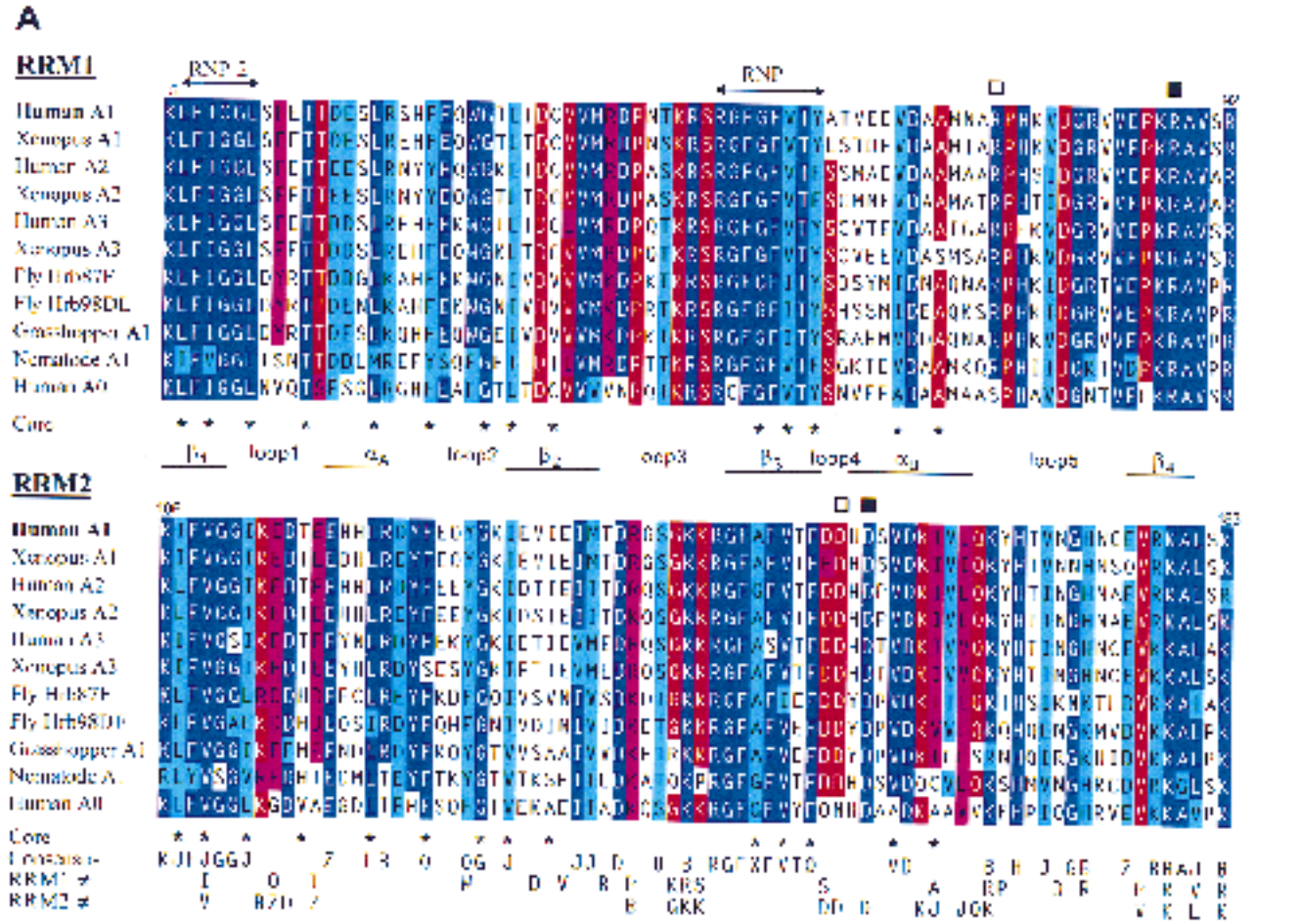

FIGURE 6. Alignment of sequences from the tandem RRMs of 11 proteins similar to hnRNP A1. Standard one-letter abbreviations are used for amino acids. Six groups of similar amino acids are indicated as follows: B = H, K, R; J = I, L, M, V; O = F, W, Y; U = S, T; X = A, G; Z = D, E. **A:** Alignment of RRM1 and RRM2. Shading indicates regions that are conserved within RRM1 or RRM2 in at least 9/11 sequences. Blue shading indicates positions that consist of a single conserved residue (white letters shaded dark blue) or two or more similar residues (black letters shaded light blue) present in both RRMs. Red shading indicates positions that differ between the two RRMs, i.e., positions that consist of a single conserved residue (white letters shaded dark red) or a single type of residue (black letters shaded light red) within a given RRM, but for which that residue is present no more than once in the sequences of the other RRM. Locations of the conserved RNP-1 and RNP-2 submotifs are indicated at the top. Conserved secondary structure elements are indicated between the RRM1 and RRM2 alignments. Asterisks indicate the position of residues located within the hydrophobic core (Birney et al., 1993). Two pairs of Arg and Asp residues involved in inter-RRM salt bridges (Shamoo et al., 1997; Xu et al., 1997) are indicated with open and filled squares. Summary lines at the bottom indicate residues that are tightly conserved in at least 18/22 sequences representing both RRMs (Consensus), residues conserved in RRM1 but that differ in RRM2 (RRM1≠), and residues conserved in RRM2 that differ in RRM1 (RRM2≠). **B:** Alignment of inter-RRM linker (IRL) segments. Identical and similar amino acids are shown on dark blue and light blue backgrounds, respectively. Consensus sequence is indicated below the alignment. Accession numbers for these sequences are as follows: human A1 (gb:X06747), A2 (gb:M29065), A3 (SW:P51991), and A0 (gb:U23803); *Xenopus laevis* A1 (gb:M31041), A2 (gb:L02954), and A3 (gb:L02956); *Drosophila* HRB87F (gb:X54803) and HRB98DE (gb:M15766); grasshopper A1 (gb:X54670); *C. elegans* A1 (gb:D10877).

The sequences shown are 48–92% identical. Each of the RRMs forms an ungapped alignment with RRM1 and RRM2 of hnRNP A1.

These 11 hnRNP A1-like proteins display a number of common features. All contain two and only two closely spaced RRMs. Each protein contains a short N-terminal region (7–35 amino acid residues long) and a longer C-terminal domain (129–193 residues long), which is extremely (37–53%) Gly-rich. In each of these proteins, the two RRMs are connected by an IRL of exactly

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

13 amino acids, which is also conserved in sequence (Fig. 6B). The aligned sequences in Figure 6A are highlighted to indicate positions that are highly conserved within each RRM (see Fig. 6 legend). The red highlighting indicates conserved positions that differ between the two RRMs, and are thus most likely to be associated with the different behaviors of RRM1 and RRM2 in the domain duplication and swap variants. Some positions highlighted in blue, indicating conserved positions that are similar or identical between the two RRMs, may also contribute to functional differences between RRM1 and RRM2. Specifically, positions 18, 37, 53, 75, 82, 88, and 90 in RRM1 of human hnRNP A1 correspond to sites with similar but not identical conserved residues in each domain.

To examine the evolutionary relationships among the RRMs in the hnRNP A1-like proteins, a phylogenetic tree was constructed with the above sequences. The result of bootstrap analysis of a neighbor-joining tree is displayed in Figure 7. The single RRM of an 83-amino acid cyanobacterial RNA-binding protein was used as an outgroup RRM to construct the tree. This RRM is similar in overall length and sequence (30 ± 5% identity) to the RRMs of the hnRNP A1-like proteins, and thus it can be unambiguously aligned, with only a single amino acid insertion in loop 5.

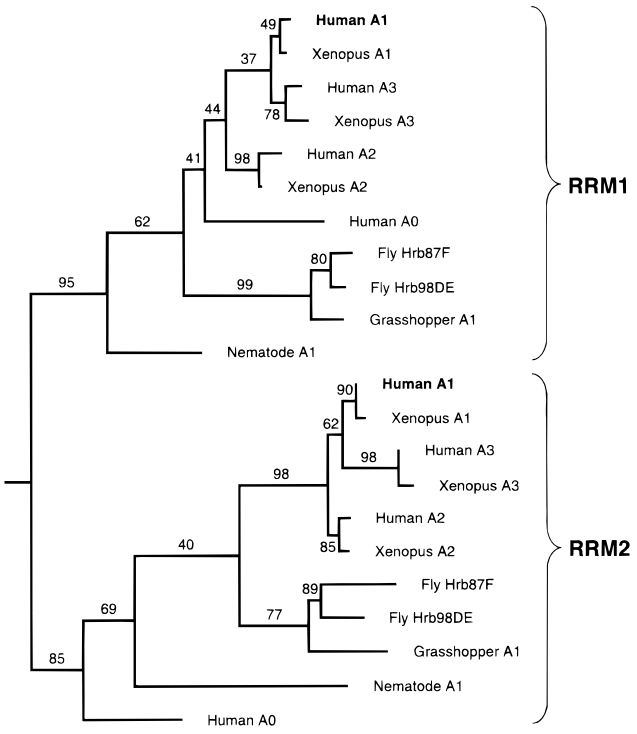

FIGURE 7. Phylogenetic tree of the RRMs from hnRNP A1-like proteins. Bootstrap analysis of 500 replicas of the RRM sequences shown in Figure 6 was performed as described in Materials and Methods. The RRM of a cyanobacterial protein (PIR accession number S77263) was used as an outgroup RRM to construct the tree. The percent of 500 bootstrap replicas that match the consensus tree is indicated by the number on each internal branch. Lengths of branches are proportional to the distances calculated by PROTDIST using the Categories model (Felsenstein, 1993).

The sequences of RRM1 and RRM2 cluster in two separate groupings with bootstrap confidence intervals of 85% and 95%. The RRM1 sequences are somewhat more tightly conserved (>60% identity with human A1) than are the RRM2 sequences (>48% identity). Strikingly, the patterns of branching for RRM1 and RRM2 sequences are almost identical, except for the placement of the branch representing the minor human variant hnRNP A0 (Myer & Steitz, 1995). Further analysis of the sequence of hnRNP A0, including sequences at the C-terminal domain (data not shown) and the IRL (Fig. 6B), indicates that this protein groups most closely with the other vertebrate sequences, as observed in the branching for the more highly conserved RRM1 sequences. The nearly identical pattern of branching for RRM1 and RRM2 strongly suggests that the two RRMs have evolved in parallel.

Aside from two divergent proteins, hnRNP A0 (Myer & Steitz, 1995) and the *Caenorhabditis elegans* protein (Iwasaki et al., 1992), whose RRMs show 50–69% identity to those of human hnRNP A1, the hnRNP A1-like proteins cluster into two groups representing insect and vertebrate proteins. The division of insect and vertebrate RRM sequences on two separate branches most likely represents independent duplications of an ancestral hnRNP A1-like protein, because each of the insect proteins is almost equally distant from each of the vertebrate proteins. The two *Drosophila* proteins are functionally similar to human A1: overexpression of Hrb87F and Hrb98DE has been shown to promote exon skipping in vivo (Shen et al., 1995; Zu et al., 1996), an effect that may involve preferential use of distal splice sites.

DISCUSSION

The striking differences observed in the activities of three hnRNP A1 variants in which the RRMs were duplicated or swapped clearly demonstrate that the tandem RRMs of hnRNP A1 are neither redundant nor functionally equivalent, in spite of their similar sequence and overall structure. Analysis of the sequences of closely related proteins from nematodes, insects, and mammals revealed many tightly conserved positions that are either similar in both RRM1 and RRM2 or that differ between the two RRMs. Such comparisons suggest that the different properties of RRM1 and RRM2 reflect the contributions of a limited number of structural differences between the two domains. The strong conservation of strikingly different residues at certain positions in RRM1 and RRM2 and absolute conservation in the length of the short IRL peptide connecting the two RRMs suggest that the tandem RRMs of the hnRNP A1-like proteins comprise a single integral functioning unit that has been tightly conserved throughout metazoan evolution.

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

Correlations between RNA-binding/annealing activities and alternative splicing

We showed that an hnRNP A1 variant with two identical copies of RRM1 was completely inactive in alternative splicing, whereas the reciprocal variant with a duplication of RRM2 was fully active. The different contributions of RRM1 and RRM2 toward RNA binding were previously suggested by the fact that hnRNP A1 variants lacking one or the other RRM yielded different high-affinity RNA consensus sequences in SELEX experiments (Burd & Dreyfuss, 1994).

In our experiments, the relative strengths of sequence-specific RNA binding and RNA annealing activities showed an approximate correlation with the alternative splicing activity of the mutants tested (Fig. 5). The RRM2 duplication variant was active in alternative splicing and it showed higher RNA annealing and specific RNA binding activities than the RRM1 duplication variant, which was inactive in the splicing assay. In general, the requirement for structural integrity appeared to be more stringent in the case of the alternative splicing function. Several hnRNP A1 variants that displayed a partial loss of RNA binding and annealing activities completely lost the ability to modulate alternative splicing. The decreased RNA binding and annealing activities of the two deletion mutants, which contain a single copy of each RRM, reflect a perturbation in cooperative interactions between adjacent RRMs. The IRL deletion mutant protein was relatively active in the specific RNA binding and annealing assays, but was inactive in splice-site switching. Taken together, these results suggest that sequence-specific RNA binding and RNA annealing are necessary but not sufficient for the alternative splicing activity of hnRNP A1.

Nonequivalent roles of each RRM in alternative splicing

The contrasting properties of the RRM-duplication variants in alternative splicing can be interpreted in terms of three different models. In the first model, the activity of A1-D(RRM2) may reflect some essential, intrinsic property of RRM2 that is essential for alternative splicing but is lacking in RRM1. According to this model, RRM1 cannot carry out the specific function of RRM2, but plays a less-specific role in its interactions with RNA. For example, RRM2 may make sequence-specific RNA contacts that are important for function, whereas RRM1 may contribute to the overall free energy of binding through nonspecific RNA contacts. It is also possible that RRM2 is involved in important protein contacts, e.g., via the α-helices, to form higher-order hnRNP assemblies or to interact with spliceosomal components. The residual activity of the domain-swap variant A1-S(RRM2,1) is consistent with this model, i.e., RRM2 can still participate in a specific, required interaction, although the altered RRM position is somehow suboptimal.

In the second model, the lack of activity of A1-D(RRM1) may reflect the fact that RRM1 has evolved for a distinct role compatible only with its position near the N-terminus of hnRNP A1. Thus, RRM1 may include some specialized features that make it unable to function in the central domain of hnRNP A1. According to this model, RRM2 can replace RRM1 at the N-terminal position simply because it is more versatile. Comparative analysis of related sequences offers two points to support this second model: (1) RRM1 is more tightly conserved than RRM2, suggesting that its sequence has been constrained during evolution, perhaps reflecting its role in functions other than alternative splicing; (2) RRM1, but not RRM2, contains three highly conserved Pro residues, located in loop 3, loop 5, and β_4. The Pro residues in the two loops are responsible for two of the most significant differences in the polypeptide backbone conformation when the structures of RRM1 and RRM2 are superimposed, as discussed below (Xu et al., 1997).

A third possibility is that the activity of wild-type hnRNP A1 and the variant A1-D(RRM2) reflects requirements for interactions between the two domains, rather than the intrinsic properties of either RRM1 or RRM2. Indeed, an intact two-RRM structure is required, because deletion of RRM1 or mutation of two surface Phe residues in the RNP-1 submotif of RRM1 abolish alternative splicing function completely (Fig. 4; Mayeda et al., 1994). Furthermore, because the RRM-swap variant had greatly reduced alternative splicing activity, the relative position of the individual RRMs may affect the interactions between them. This model is also consistent with the conservation of the length and sequence of the IRL and the conservation of two salt bridges between RRM1 and RRM2 within the hnRNP A1-like proteins, as discussed below. A number of multi-RRM proteins show synergy between two or more of their RRMs for binding to RNA (Burd et al., 1991; Ye & Sugiura, 1992; Zamore et al., 1992; Cáceres & Krainer, 1993; Mayeda et al., 1994; Kuhn & Pieler, 1996; Deardorff & Sachs, 1997), and it will be interesting to learn to what extent the structural arrangement and functional modularity of RRMs seen in hnRNP A1 apply to other proteins with two or more RRMs.

Comparison of RRM1 and RRM2 of hnRNP A1-like proteins

The specific distribution of conserved and variable residues within and between different RRMs is informative with respect to the structural and functional constraints on specific domains within these hnRNP A1-like proteins. Comparisons of RRM sequences from many RNA-binding proteins have revealed that the most highly conserved sequences in RRMs are those in the RNP-1

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

and RNP-2 submotifs (Kenan et al., 1991; Birney et al., 1993), which correspond to the two β-strands, β_3 and β_1, at the center of the RNA-binding surface. The most conserved positions independent of these submotifs correspond precisely to residues that form the hydrophobic core, as noted in the U1-A protein previously (Nagai et al., 1990; Kenan et al., 1991; Birney et al., 1993). Many other positions within the alignment of RRM1 and RRM2 of hnRNP A1-like proteins are also tightly conserved, particularly those residues located in loops 2, 3, and 5 connecting adjacent elements of secondary structure, as well as those found on the interior face of helix α_A and in the center of helix α_B. These residues contribute to the similarities noted between RRM1 and RRM2 of this family of proteins.

Other conserved positions within each individual RRM that differ between the two RRMs are especially interesting with respect to the functional differences of the domain-duplication and domain-swap variants. Of particular interest is the presence of three conserved Pro residues located at positions 49, 76, and 86 within loop 3, loop 5, and β_4 of RRM1, respectively. Each of these prolines is replaced in RRM2 with a very different, but also highly conserved residue, namely Arg 140, Tyr 167, and Val 177. The presence of the conserved prolines in three segments of RRM1 and their absence in RRM2 can account for the differences in the structure and conformational flexibility between these RRMs. Many other divergent substitutions are apparent in comparing the most conserved positions of RRM1 and RRM2. Fully 67 of 78 positions ($\sim$86%) of the sites in RRM1 or RRM2 are tightly conserved in at least one domain. Of these conserved positions, 25 of 67 positions ($\sim$37%) are different in RRM1 and RRM2 (see bottom two consensus lines in Fig. 6A). This number includes 19 positions (shaded red), where there is a significant difference in the type of residues, plus 6 positions (shaded blue), where two different but similar residues are strictly conserved in each domain. The most striking conserved differences are clustered in loop 1, β_2, loop 3, loop 4, α_B, and loop 5. Other, more subtle differences are found in β_1 and β_4. Among these conserved differences, those located on the presumptive RNA-binding surface of the RRMs in loops 1 and 3 and β_2 are particularly likely to alter the association of RNA with the RRM directly.

Importance of the salt bridges between RRM1 and RRM2

The recently determined crystal structure of UP1, which comprises both RRMs of hnRNP A1, provides important insights into the phylogenetically conserved differences between RRM1 and RRM2 (Shamoo et al., 1997; Xu et al., 1997). Loop 3 and loop 5 show the most pronounced backbone deviations when the two RRMs are superimposed. The independently folded RRMs are held together in a fixed geometry, such that the two RRMs can probably function as a single entity in RNA binding. The two RRMs are antiparallel and held in close contact mainly by two pairs of salt bridges, i.e., Arg 75–Asp 155 and Arg 88–Asp 157 (Fig. 6A). Significantly, these charged residues are invariant among the hnRNP A1-like proteins (Fig. 6A), with only two exceptions: hnRNP A0 and the monkey hnRNP A1 γ-isoform (not shown), which are both unusual forms of hnRNP A1 (An & Wu, 1993; Myer & Steitz, 1995). Also, it is striking that these four salt-bridge residues are conserved only in proteins possessing a 13-amino acid IRL (data not shown).

Comparison of the sequences of RRM1- and RRM2-duplicated variants suggests that the two salt bridges are fully maintained in the active RRM2 duplication (Lys 75–Asp 155 and Lys 88–Asp 157). In contrast, only one salt bridge (Arg 88–Glu 157) can be reconstructed in the inactive RRM1-duplication, because the second salt bridge is disrupted by replacement of Asp 155 with Thr 155. Likewise, only one salt bridge (Lys 88–Glu 157) can be formed in the weakly active RRM-swap variant. To test whether the lack of activity of the RRM1-duplication is solely due to the loss of one of the salt bridges, we reconstructed the salt bridge by a Thr 155 to Asp 155 mutation in the context of the A1-D(RRM1) variant. However, this new variant remained inactive in alternative splicing (data not shown). There may be a requirement for additional interactions at the interface between the two RRMs, as suggested by the conservation of several sequences adjacent to Asp 155 and Asp 157 in RRM2. Alternatively, or in addition, unique RRM2 residues may be required for proper interaction with RNA in the context of alternative splicing. The phylogenetic analysis and the conservation of residues of appropriate charge in the active variant proteins strongly suggest that the two salt bridges observed in the UP1 structure are an important structural feature of the hnRNP A/B family, rather than a result of crystal packing forces.

Phylogenetic conservation of tandem RRMs and IRL

The conservation of the tandem RRMs in hnRNP A1-like proteins from nematodes, insects, and vertebrates is impressive. The RRMs are conserved much more tightly than adjacent N-terminal and C-terminal sequences in all these closely related proteins, suggesting that the tandem RRMs have a very specific function within the context of the whole protein. Our phylogenetic study of hnRNP A1-like proteins revealed that RRM1 and RRM2 are clearly separated as distinct groups (Fig. 7). The observed phylogenetic pattern is consistent with the idea that each RRM evolved independently after the ancestral duplication event (Bandziulis et al., 1989; Birney et al., 1993).

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

The IRL is also highly conserved in both length and sequence among related hnRNP A/B proteins (Fig. 6B). Because the hnRNP A1 IRL is located between the RRM1 and RRM2 β-sheets on the same side as their RNA-binding surfaces, it is likely to contact bound RNA directly (Xu et al., 1997). In addition, the IRL is responsible, in conjunction with the two salt bridges, for the precise spatial arrangement of the two RRMs. Consistent with these observations, alternative splicing activity was completely abolished by shortening the 13-amino acid hnRNP A1 IRL to a single alanine (Figs. 4, 5). IRL lengths are highly variable in other proteins (Shamoo et al., 1995), and a single-alanine IRL separates the two RRMs in p54^{nrb}, PSF, and related proteins (Dong et al., 1992). The oligo-glycine IRL that separates the two RRMs of SF2/ASF can be shortened with essentially no effect on in vitro splicing activity or function in an in vivo viability assay (Zuo & Manley, 1993; Wang et al., 1996). These observations suggest that IRLs have very different structural roles in divergent multi-RRM proteins.

In conclusion, the tandem RRM1-RRM2 structure of hnRNP A1 constitutes a highly conserved ancient motif. The two RRMs probably function in a cooperative or coordinated manner to achieve functional RNA binding, RNA annealing, and alternative splicing. The recently elucidated structure of the tandem RRMs strongly suggested that they bind RNA as a single unit. However, an unexpected finding in the present study that was not apparent from the structure is that, although the two-RRM structure is essential, only one of the two RRMs appears to play a specific role in the function of the protein in alternative splicing. The molecular basis for this striking difference between RRM1 and RRM2 may be elucidated through functional analysis of further hnRNP A1 mutants designed in the context of structural and sequence conservation data.

MATERIALS AND METHODS

Construction of variant hnRNP A1 expression plasmids

The expression plasmids were constructed by one or more rounds of overlap-extension PCR with appropriate primers and/or deletion of restriction fragments, followed by replacement of the desired fragments by subcloning into the parent pET9d-hnRNP A1 plasmid (Mayeda & Krainer, 1992; Mayeda et al., 1994). Plasmid pET9d-A1-D(RRM1) has a duplication of RRM1 residues Lys 15–Arg 92 in place of RRM2 residues Lys 106–Lys 183, whereas plasmid pET9d-A1-D(RRM2) has the reciprocal duplication and replacement. In plasmid pET9d-A1-S(RRM2,1), the above RRM1 and RRM2 sequences were swapped. Plasmid pET9d-A1-ΔRRM1 encodes the N-terminal sequence Met-Ala followed by residues Asp 94–Phe 320 of hnRNP A1. Plasmid pET9d-A1-ΔRRM2 has a deletion of RRM2 and the 13 residues preceding it (Glu 93–Lys 183). In plasmid pET9d-A1-ΔIRL, the 13 residues (Glu 93–Lys 105) that connect RRM1 and RRM2 are replaced by a single Ala. All the constructions were verified by DNA sequencing.

Preparation of *E. coli*-expressed recombinant proteins

Expression of all the recombinant proteins in the *E. coli* strain BL21(DE3)pLysS and purification of recombinant A1-D(RRM1), A1-D(RRM2), and A1-ΔRRM1 proteins were performed essentially as described for wild-type hnRNP A1 (Mayeda & Krainer, 1992).

A1-ΔRRM2, A1-S(RRM2,1), and A1-ΔIRL were recovered from the precipitate formed during dialysis after fractionation by CsCl density gradient centrifugation (Mayeda et al., 1994). These proteins were denatured in buffer containing 6 M urea and purified by Mono S chromatography under denaturing conditions, essentially as described (Mayeda et al., 1994; Screaton et al., 1995).

After purification, all recombinant proteins were dialyzed into buffer C (Mayeda et al., 1994) containing 20% (v/v) glycerol, followed by a final centrifugation step to remove insoluble material. The renaturation protocol used for some of the variants was shown previously to have no effect on the splicing activity of the wild-type protein (Mayeda et al., 1994). The final protein concentrations were estimated by the dye-binding method (Bio-Rad), standardized with bovine serum albumin. Each protein was adjusted to the desired concentration (5 pmol/μl) before use.

Filter-binding assay

Specific RNA binding was assayed using a 67-nt RNA selected by six rounds of SELEX for binding to the hnRNP A2 protein (I. Watakabe & A.R. Krainer, unpubl.). This oligoribonucleotide, Rd 6-1, consists of the 24-nt sequence UAGG UUAGGAAUAGGGAAUUAAGG, flanked by the constant sequences 5′-GGGCCACCAACGACAUUU and UUGAUAU AAAUAGUGCCCAUGGAUC-3′. It bound to the three closely related hnRNP A/B proteins A1, A2, and B1 with high affinity ($K_d < 10$ nM; unpubl. data). The control oligoribonucleotide, Rd 0-1, had the same flanking sequences surrounding the 19-nt sequence AUGUUUGGAGGCCACGCGC. It was isolated prior to selection for binding and showed no significant binding at protein concentrations below 50 nM. Rd 6-1 contains one perfect copy of the hnRNP A1 high-affinity binding site consensus, UAGGG(A/U) (Burd & Dreyfuss, 1994), plus three copies with matches at 5/6 positions.

Labeled RNA substrates were prepared by in vitro runoff transcription of linearized plasmid templates in the presence of [α-^{32}P] UTP. The filter-binding assay was performed as described (Mayeda et al., 1994).

RNA annealing assay

RNA–RNA annealing was measured with an RNase T1 protection/urea-PAGE assay, as described (Mayeda et al., 1994). Dried gels were scanned using an Ambis 100 radioanalytical imager. Annealing was determined by comparison with undigested controls. Background radioactivity deter-

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

mined in samples incubated in the absence of protein (<2% of input) was subtracted from each experimental point. Annealing is expressed as percentage of input RNA, corrected for content of labeled ribonucleotides.

In vitro splicing assays

The plasmid pSP64-5′D16X (Reed & Maniatis, 1986; Krainer et al., 1990), containing a duplicated 5′ splice site, was linearized with *Bam*H I and used as a template for runoff transcription with SP6 RNA polymerase (Mayeda & Krainer, 1998a). HeLa cell S100 extract and purified SF2/ASF were prepared as described (Mayeda & Krainer, 1998b; Mayeda et al., 1993). In vitro splicing reactions in 25 μl with the indicated amounts of S100 extract, purified SF2/ASF, the appropriate wild-type or variant recombinant hnRNP A1, and 20 fmol ^{32}P-labeled pre-mRNA substrate were incubated at 30 °C for 4 h as described (Mayeda & Krainer, 1998a). RNA products were analyzed by electrophoresis on a 5.5% polyacrylamide/7 M urea gel followed by autoradiography with an intensifying screen at −70 °C.

Comparative sequence analysis

The database of hnRNP A1-related proteins was assembled by multiple BLAST searches (Altschul et al., 1990) of the nonredundant protein database maintained by the NCBI. Query sequences included the individual RRM1 and RRM2 of human hnRNP A1 and the region extending from RNP-1 of RRM1 through RNP-2 of RRM2. The database of hnRNP A1-related proteins was extended by further searches using the same regions of diverse hnRNP A1-like proteins retrieved by prior BLAST searches. Retrieved sequences were aligned with the GCG Pileup program using the individual RRMs, with boundaries as defined by Birney et al. (1993), and pairwise alignments were evaluated by the GCG Gap program. Bootstrapped distance trees of RRMs from hnRNP A1-like proteins were constructed using the neighbor-joining method (Saitou & Nei, 1987), as implemented by the PHYLIP package using Categories in PROTDIST to calculate the distances (Felsenstein, 1993).

ACKNOWLEDGMENTS

We thank I. Watakabe and L. Snell for technical assistance, and Dr. E. Waters for helpful discussions of phylogenetic methods. A.M. and A.R.K. were supported in part by grant CA13106 from the NCI; R.-M.X. was supported by grant GM55874 from the NIH and by the W.M. Keck foundation; S.H.M. was supported by grants DK40834 and GM55922 from the NIH.

Received June 2, 1998; returned for revision June 9, 1998; revised manuscript received June 19, 1998

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol 215*:403–410.

An G, Wu R. 1993. cDNA cloning of a hnRNP A1 isoform and its regulation by retinol in monkey tracheobronchial epithelial cells. *Biochim Biophys Acta 1172*:292–300.

Bandziulis RJ, Swanson MS, Dreyfuss G. 1989. RNA-binding proteins as developmental regulators. *Genes & Dev 3*:431–437.

Birney E, Kumar S, Krainer AR. 1993. Analysis of the RNA-recognition motif and RS and RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res 21*:5803–5816.

Burd CG, Dreyfuss G. 1994. RNA binding specificity of hnRNP A1: Significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J 13*:1197–1204.

Burd CG, Matunis EL, Dreyfuss G. 1991. The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. *Mol Cell Biol 11*:3419–3424.

Cáceres JF, Krainer AR. 1993. Functional analysis of pre-mRNA splicing factor SF2/ASF structural domains. *EMBO J 12*:4715–4726.

Cáceres JF, Stamm S, Helfman DM, Krainer AR. 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science 265*:1706–1709.

Cáceres JF, Krainer AR. 1997. Mammalian pre-mRNA splicing factors. In: Krainer AR, ed. *Eukaryotic mRNA processing*. Oxford, UK: IRL Press. pp 174–212.

Cobianchi F, Karpel RL, Williams KR, Notario V, Wilson SH. 1988. Mammalian heterogeneous nuclear ribonucleoprotein complex protein A1. Large-scale overproduction in *Escherichia coli* and cooperative binding to single-stranded nucleic acids. *J Biol Chem 263*:1063–1071.

Deardorff JA, Sachs AB. 1997. Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast poly(A)-binding protein. *J Mol Biol 269*:67–81.

Dong B, Horowitz DS, Kobayashi R, Krainer AR. 1992. Purification and cDNA cloning of HeLa cell p54^{nrb}, a nuclear protein with two RNA recognition motifs and extensive homology to human splicing factor PSF and *Drosophila* NONA/BJ6. *Nucleic Acids Res 21*:4085–4092.

Dreyfuss G, Matunis MJ, Piñol-Roma S, Burd CG. 1993. hnRNP proteins and the biogenesis of mRNA. *Annu Rev Biochem 62*:289–321.

Felsenstein J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Seattle, Washington: University of Washington. Department of Genetics.

Fu XD, Mayeda A, Maniatis T, Krainer AR. 1992. General splicing factors SF2 and SC35 have equivalent activities in vitro and both affect alternative 5′ and 3′ splice site selection. *Proc Natl Acad Sci USA 89*:11224–11228.

Ge H, Manley JL. 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell 62*:25–34.

Iwasaki M, Okumura K, Kondo Y, Tanaka T, Igarashi H. 1992. cDNA cloning of a novel heterogeneous nuclear ribonucleoprotein gene homologue in *Caenorhabditis elegans* using hamster prion protein cDNA as a hybridization probe. *Nucleic Acids Res 20*:4001–4007.

Kenan DJ, Query CC, Keene JD. 1991. RNA recognition: Towards identifying determinants of specificity. *Trends Biochem Sci 16*:214–220.

Kiledjian M, Dreyfuss G. 1992. Primary structure and binding activity of the hnRNP U protein: Binding RNA through RGG box. *EMBO J 11*:2655–2664.

Krainer AR, Conway GC, Kozak D. 1990. The essential pre-mRNA splicing factor SF2 influences 5′ splice site selection by activating proximal sites. *Cell 62*:35–42.

Kumar A, Wilson SH. 1990. Studies of the strand-annealing activity of mammalian hnRNP complex protein A1. *Biochemistry 29*:10717–10722.

Kuhn U, Pieler T. 1996. *Xenopus* poly(A) binding protein: Functional domains in RNA binding and protein–protein interaction. *J Mol Biol 256*:20–30.

Matunis EL, Matunis MJ, Dreyfuss G. 1992. Characterization of the major hnRNP proteins from *Drosophila melanogaster*. *J Cell Biol 116*:257–269.

Mayeda A, Helfman DM, Krainer AR. 1993. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Mol Cell Biol 13*:2993–3001.

Mayeda A, Krainer AR. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell 68*:365–375.

Downloaded from rnajournal.cshlp.org on December 11, 2013 - Published by Cold Spring Harbor Laboratory Press

Mayeda A, Krainer AR. 1998a. Mammalian in vitro splicing assays. In: Haynes SR, ed. *Methods in molecular biology, RNA–protein interaction protocols*. Totowa, New Jersey: Humana Press Inc. Forthcoming.

Mayeda A, Krainer AR. 1998b. Preparation of HeLa cell nuclear and cytosolic S100 extracts for in vitro splicing. In: Haynes SR, ed. *Methods in molecular biology, RNA–protein interaction protocols*. Totowa, New Jersey: Humana Press Inc. Forthcoming.

Mayeda A, Munroe SH, Cáceres JF, Krainer AR. 1994. Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J 13*:5483–5495.

Merrill BM, Stone KL, Cobianchi F, Wilson SH, Williams KR. 1988. Phenylalanines that are conserved among several RNA-binding proteins form part of a nucleic acid-binding pocket in the A1 heterogeneous nuclear ribonucleoprotein. *J Biol Chem 263*:3307–3313.

Michael WM, Choi M, Dreyfuss G. 1995. A nuclear export signal in hnRNP A1: A signal-mediated, temperature-dependent nuclear protein export pathway. *Cell 83*:415–422.

Munroe SH, Dong X. 1992. Heterogeneous nuclear ribonucleoprotein A1 catalyzes RNA–RNA annealing. *Proc Natl Acad Sci USA 89*:895–899.

Myer VE, Steitz JA. 1995. Isolation and characterization of a novel, low abundance hnRNP protein: A0. *RNA 1*:171–182.

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*:515–520.

Piñol-Roma S, Dreyfuss G. 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature 355*:730–732.

Pollard VW, Michael WM, Nakielny S, Siomi MC, Wang F, Dreyfuss G. 1996. A novel receptor-mediated nuclear protein import pathway. *Cell 86*:985–994.

Pontius BW, Berg P. 1990. Renaturation of complementary DNA strands mediated by purified mammalian heterogeneous nuclear ribonucleoprotein A1 protein: Implications for a mechanism for rapid molecular assembly. *Proc Natl Acad Sci USA 87*:8403–8407.

Portman DS, Dreyfuss G. 1994. RNA annealing activities in HeLa nuclei. *EMBO J 13*:213–221.

Reed R, Maniatis T. 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell 46*:681–690.

Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol 4*:406–425.

Screaton GR, Cáceres JF, Mayeda A, Bell MV, Plebanski M, Jackson DG, Bell JI, Krainer AR. 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. *EMBO J 14*:4336–4349.

Shamoo Y, Abdul-Manan N, Williams KR. 1995. Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res 23*:725–728.

Shamoo Y, Krueger U, Rice LM, Willams KR, Steitz TA. 1997. Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 Å resolution. *Nature Struct Biol 4*:215–222.

Shen J, Zu K, Cass CL, Beyer AL, Hirsh J. 1995. Exon skipping by overexpression of a *Drosophila* heterogeneous nuclear ribonucleoprotein in vivo. *Proc Natl Acad Sci USA 92*:1822–1825.

Siomi H, Dreyfuss G. 1995. A nuclear localization domain in the hnRNP A1 protein. *J Cell Biol 129*:551–560.

Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell 92*:315–326.

Wang J, Takagaki Y, Manley JL. 1996. Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. *Genes & Dev 10*:2588–2599.

Weighardt F, Biamonti G, Riva S. 1995. Nucleo-cytoplasmic distribution of human hnRNP proteins: A search for the targeting domains in hnRNP A1. *J Cell Sci 108*:545–555.

Xu RM, Jokhan L, Cheng X, Mayeda A, Krainer AR. 1997. Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure 5*:559–570.

Yang X, Bani MR, Lu SJ, Rowan S, Ben-David Y, Chabot B. 1994. The A1 and A1^{B} proteins of heterogeneous nuclear ribonucleoparticles modulate 5′ splice site selection in vivo. *Proc Natl Acad Sci USA 91*:6924–6928.

Ye L, Sugiura M. 1992. Domains required for nucleic acid binding activities in chloroplast ribonucleoproteins. *Nucleic Acids Res 20*:6275–6279.

Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature 355*:609–614.

Zu K, Sikes ML, Haynes SR, Beyer AL. 1996. Altered levels of the *Drosophila* HRB87F/hrp36 hnRNP protein have limited effects on alternative splicing in vivo. *Mol Biol Cell 7*:1059–1073.

Zuo P, Manley JL. 1993. Functional domains of the human splicing factor ASF/SF2. *EMBO J 12*:4727–4737.