RNA (1999), *5*:167–179. Cambridge University Press. Printed in the USA.

The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5′ splice site

JOSÉ L. REYES,[1] E. HILARY GUSTAFSON,[1] HONGBO R. LUO,[2] MELISSA J. MOORE,[2] and MARIA M. KONARSKA[1]

[1]The Rockefeller University, New York, New York 10021, USA
[2]Howard Hughes Medical Institute, Brandeis University, Department of Biochemistry, Waltham, Massachusetts 02454, USA

ABSTRACT

A U5 snRNP protein, hPrp8, forms a UV-induced crosslink with the 5′ splice site (5′SS) RNA within splicing complex B assembled in *trans*- as well as in *cis*-splicing reactions. Both yeast and human Prp8 interact with the 5′SS, branch site, polypyrimidine tract, and 3′SS during splicing. To begin to define functional domains in Prp8 we have mapped the site of the 5′SS crosslink within the hPrp8 protein. Immunoprecipitation analysis limited the site of crosslink to the C-terminal 50–60-kDa segment of hPrp8. In addition, size comparison of the crosslink-containing peptides generated with different proteolytic reagents with the pattern of fragments predicted from the hPrp8 sequence allowed for mapping of the crosslink to a stretch of five amino acids in the C-terminal portion of hPrp8 (positions 1894–1898). The site of the 5′SS:hPrp8 crosslink falls within a segment spanning the previously defined polypyrimidine tract recognition domain in yPrp8, suggesting that an overlapping region of Prp8 may be involved both in the 5′SS and polypyrimidine tract recognition events. In the context of other known interactions of Prp8, these results suggest that this protein may participate in formation of the catalytic center of the spliceosome.

Keywords: 5′ splice site; pre-mRNA splicing; Prp8; RNA:protein interaction; UV-crosslinking

INTRODUCTION

Removal of introns from eukaryotic pre-mRNA is catalyzed by a multicomponent complex termed the spliceosome. A set of small nuclear ribonucleoprotein (snRNP) particles U1, U2, and U4/U5/U6 snRNPs, accompanied by a number of protein factors, bind to pre-mRNA in a stepwise fashion (reviewed in Moore et al., 1993; Nilsen, 1994).

Correct splicing of mRNA precursors depends on the precise recognition of sequence elements that direct assembly of the spliceosome and represent the substrates for the reaction. One of them, the 5′ splice site (5′SS), defines the exon/intron junction and represents the substrate for the first step of splicing. The initial base pairing between U1 snRNA and the 5′SS formed in splicing complex E is subsequently disrupted and replaced by the interaction of the 5′SS with U4/U5/U6 snRNP within splicing complex B. A specific base-pairing interaction between nucleotides in the conserved ACAGAG sequence of U6 snRNA and the intron bases at the 5′SS was demonstrated both by genetic and biochemical analyses (Kandels-Lewis & Séraphin, 1993; Konforti et al., 1993; Lesser & Guthrie, 1993; Sontheimer & Steitz, 1993). Furthermore, exon sequences at the 5′SS are thought to interact with the conserved loop 1 in U5 snRNA (see Newman, 1997). In yeast, this interaction is dispensable for the first, but essential for the second catalytic step in vitro (O'Keefe et al., 1996). Finally, the most highly conserved element in the 5′SS sequence, the GU dinucleotide at the 5′ end of the intron, is implicated in interactions with U6 and U2 snRNAs (Sontheimer & Steitz, 1993; Kim & Abelson, 1996; Luukkonen & Séraphin, 1998a, 1998b). The U5 snRNP-specific yeast Prp8 protein and its mammalian homolog hPrp8 (also termed p220) interact with the 5′SS, branch site, polypyrimidine tract, and 3′SS regions (e.g., Wyatt et al., 1992; MacMillan et al., 1994; Teigelkamp et al., 1995a; Chiara et al., 1996; Reyes et al., 1996; Umen & Guthrie, 1996). Thus, Prp8 is the only splicing factor known to interact with all the important pre-mRNA sequences. Its role in the recognition of the 5′SS is the subject of the present study.

A simplified in vitro system in HeLa cell nuclear extracts was used to study interactions between the 5′SS and spliceosomal components. A short RNA oligonucleotide comprising the 5′SS consensus sequence (5′SS RNA, $A_5G/GUAAGUAdTdC_3$, where / represents the exon/intron junction), in the presence of a

Reprint requests to: Maria M. Konarska, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA; e-mail: konarsk@rockvax.rockefeller.edu.

longer RNA containing the branch site, polypyrimidine tract, and 3′SS (3′SS RNA), induces formation of splicing complex B and undergoes both steps of splicing *in trans* (Konforti & Konarska, 1995). Upon UV irradiation of these *trans*-splicing reactions, the 5′SS RNA becomes crosslinked to hPrp8 within splicing complex B. The site of crosslink maps to the highly conserved GU dinucleotide at the 5′ end of the intron (Reyes et al., 1996). This interaction is remarkably specific since substitution of the uridine residue within the GU with thymidine or 5-iodo-uridine diminishes the 5′SS RNA interaction with hPrp8. This effect is correlated with a decreased splicing efficiency of the modified 5′SS RNA in the *trans*-splicing assay (Reyes et al., 1996).

Here we show that the same crosslink can also be detected in *cis*-splicing reactions and that the 5′SS RNA:hPrp8 interaction occurs only in the context of the 5′SS sequence capable of other important interactions within the spliceosome. These data suggest a highly specific recognition of the GU dinucleotide by hPrp8 in the broader context of multiple interactions between the 5′SS and other spliceosomal components. Thus, to identify the region of hPrp8 involved in this interaction with the 5′SS, we have mapped the site of the 5′SS crosslink within the hPrp8 amino acid sequence. First, we have used an antibody against the C-terminal segment of hPrp8 to narrow down the location of the crosslink to the C-terminal 50–60-kDa region. Second, we have performed an extensive series of enzymatic and chemical proteolytic digestions of the 5′SS:hPrp8 crosslink. Size comparison of the resulting crosslink-containing peptides with the hPrp8 sequence allowed for a more precise mapping of the crosslink to a stretch of 5 amino acids (positions 1894–1898) in the C-terminal region of the protein. The site of the 5′SS:hPrp8 crosslink corresponds to a segment within the previously defined polypyrimidine tract recognition domain in yPrp8 (Umen & Guthrie, 1996), suggesting that an overlapping region of Prp8 may also be involved in the 5′SS recognition.

RESULTS

hPrp8 crosslinks to the 5′ splice site GU dinucleotide in *cis*-spliceosomes

To further characterize the interaction between hPrp8 and the GU dinucleotide at the 5′ end of the intron, we analyzed the time course of appearance of the 5′SS:hPrp8 crosslink in *trans*-splicing reactions (Fig. 1A). The 3 min irradiation with 254-nm UV light (as opposed to 15 min used previously; Reyes et al., 1996) yields lower levels of the 5′SS RNA:hPrp8 crosslink but more precisely reflects the selected time points of incubation. After irradiation, the crosslinked products were directly resolved in a 10% acrylamide/SDS gel. The 5′SS RNA:hPrp8 crosslink appears early in the reaction (5 min), reaches a maximal level at 10 min, and then slowly decreases over time (Fig. 1A). The major ~97-kDa crosslink detected in these reactions is not stably associated with spliceosomes because it is not detected in complex B resolved in native gels (Reyes et al., 1996) and it can be separated from splicing complexes upon centrifugation through glycerol cushions (data not shown and Sha et al., 1998). The 5′SS RNA:hPrp8 crosslinking profile parallels that of the splicing complex B formation (data not shown), consistent with the finding that interaction of hPrp8 with the 5′SS RNA occurs within complex B (Reyes et al., 1996). However, the efficiency of crosslinking decreases prior to the appearance of splicing intermediates and products (Fig. 1B), indicating that the 5′SS RNA:hPrp8 interaction is disrupted or changed at the later stages of splicing. Since even minor changes in the local physical environment may affect crosslinking, this loss of crosslink formation does not necessarily indicate that the 5′SS:hPrp8 interaction is discontinued after complex B formation. In fact, hPrp8 was shown to interact with the 5′ exon nucleotides even at the stage of splicing catalysis (Wyatt et al., 1992).

Because the *cis*- and *trans*-splicing systems may differ in the details of the 5′SS recognition, we tested whether the GU dinucleotide:hPrp8 interaction observed in *trans*-splicing reactions can also be detected in the context of a full-length pre-mRNA *cis*-splicing substrate. We have constructed a *cis*-splicing pre-mRNA based on the bimolecular substrates used in *trans*-splicing reactions. This substrate (213 nt) was generated by a DNA oligonucleotide-assisted ligation (Moore & Sharp, 1992) of the standard 5′SS RNA oligonucleotide (A_5G/GUAAGUAdT) ^{32}P-labeled at the 5′ end, with the 3′SS RNA transcript containing sequences derived from the second intron (145 nt) and third exon (54 nt) of the rabbit β-globin gene. Splicing reactions using such a 5′ end-labeled pre-mRNA were UV irradiated for 3 min and subjected to RNase A digestion prior to resolving the crosslinked products in a 10% acrylamide/SDS gel (Fig. 1C). Upon digestion with RNase A, only the crosslink products formed upstream of position U+2 (p*A_5G/GU) are detectable. A clearly visible 5′SS RNA:hPrp8 crosslink present in RNase A-treated reactions (Fig. 1C) is not detectable upon RNase T1 treatment (data not shown), effectively limiting the site of crosslinking within the RNA to the GU dinucleotide at the 5′ end of the intron. Splicing efficiency cannot be accurately determined using the 5′ end-labeled substrate since only formation of the ligated-exons product can be monitored. To follow the accumulation of lariat intermediates and products (Fig. 1D), a uniformly ^{32}P-labeled pre-mRNA transcript containing 20 nt 5′ exon, 191 nt intron, and 54 nt 3′ exon of β-globin was prepared by transcription (see Materials and Methods). In addition, a pre-mRNA generated by ligation of the 3′ end ^{32}P-labeled 5′SS RNA oligonu-

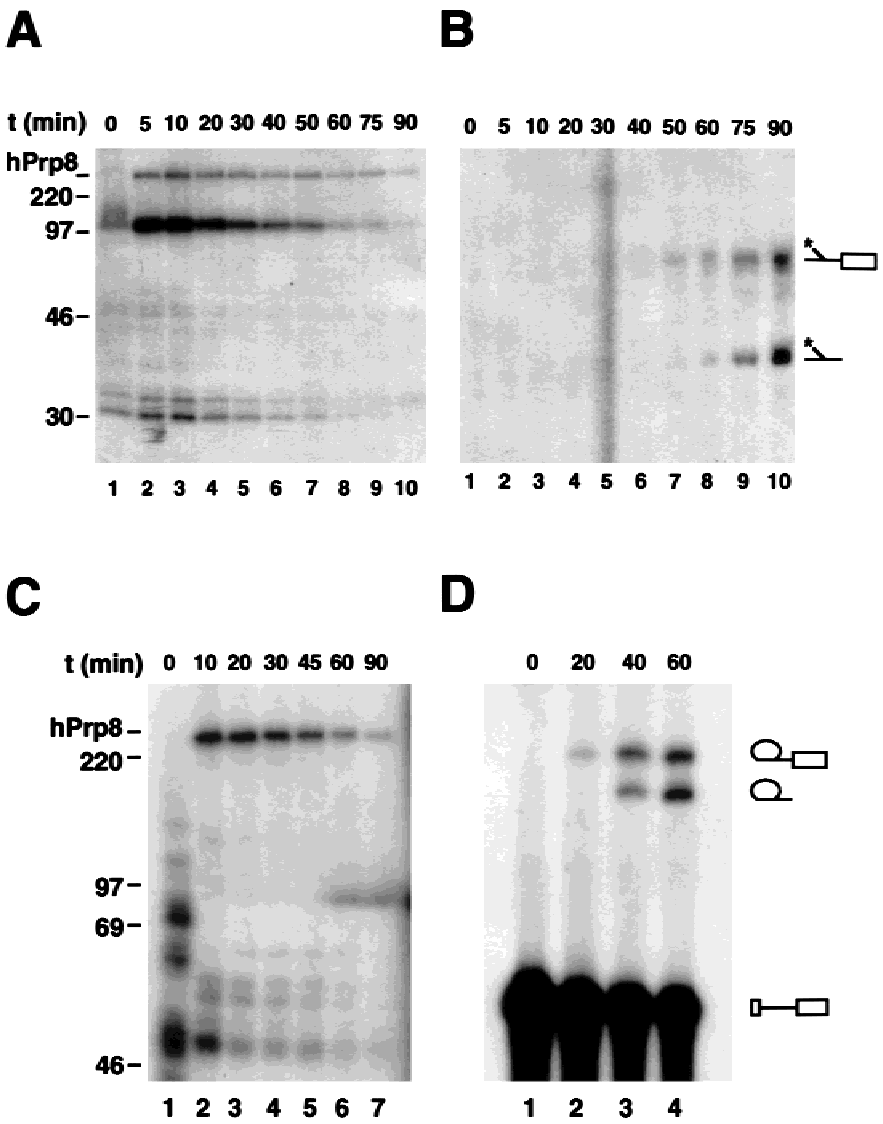

FIGURE 1. Kinetics of the 5′SS RNA: hPrp8 crosslink formation. *Trans*-splicing reactions, containing the 3′ end labeled 5′SS RNA (**A** and **B**) or *cis*-splicing reactions containing 5′ end labeled (**C**) or internally labeled (**D**) pre-mRNA were incubated for the times indicated and UV (254 nm) irradiated for 3 min. Samples were directly resolved in a 10% SDS gel (**A** and **C**). Alternatively, reactions were phenol extracted and RNA products of splicing resolved in a 10% polyacrylamide/8 M urea gel (**B** and **D**).

cleotide to a 3′SS RNA transcript, yielding 6 nt 5′ exon, 153 nt intron, and 54 nt 3′ exon RNA, has also been tested and shown to participate in splicing as efficiently as substrates generated by transcription (data not shown).

These results demonstrate that the interaction between hPrp8 and the GU dinucleotide at the 5′ end of the intron is detected both in *trans*-splicing (Reyes et al., 1996) and in *cis*-splicing reactions (Fig. 1). As in the case of *trans*-splicing, the appearance of the 5′SS RNA:hPrp8 crosslink in *cis*-splicing reactions coincides with complex B formation (data not shown) and peaks before the onset of splicing catalysis (Fig. 1C,D).

Crosslinking of hPrp8 to the 5′SS RNA within splicing complex B requires a splicing-competent 5′SS sequence

Crosslinking of the 5′SS RNA to hPrp8 within the spliceosome-like complexes was detected using the 5′SS consensus sequence. Mutations of the GU dinucleotide, as well as minor modifications of the U at position +2, interfere with complex formation and crosslinking to hPrp8 (Reyes et al., 1996). However, what elements of the functional 5′SS RNA are required, and whether the GU dinucleotide alone is sufficient to promote the interaction of the 5′SS RNA with hPrp8 has not been analyzed in detail. To determine whether crosslinking of hPrp8 to the 5′ end of the intron requires a full-length consensus 5′SS signal, we tested several derivatives of the 5′SS RNA in crosslinking experiments. The selected 5′SS RNA oligonucleotides contained either a shortened intron that exhibits a diminished potential for base pairing with U6 snRNA or a shortened 5′ exon sequence (Fig. 2). All 5′SS RNAs were incubated under *trans*-splicing conditions in the presence of the 3′SS RNA and allowed to form splicing complex B. Aliquots of the reactions were UV irradiated and resolved in a native gel to determine the efficiency of complex B formation (Fig. 2A) or in a 12% SDS gel

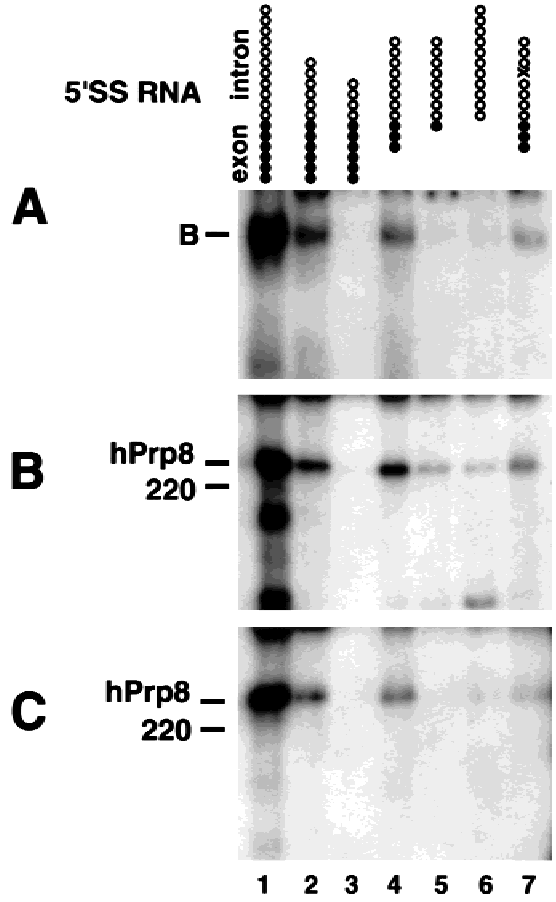

FIGURE 2. Crosslinking of hPrp8 to the 5′SS RNA within splicing complex B requires a splicing competent 5′SS sequence. *Trans*-splicing reactions were carried out in the presence of 5′SS RNAs of different lengths. The wt 5′SS RNA contained six exon nucleotides (A_5G, represented by black dots) and 11 intron nucleotides (open circles, GUAAGUAdT*dC_3; *: the P^{32} phosphate group, lane 1). Lanes 2–6 contained truncated versions of the wt 5′SS RNA, as indicated by their graphic representations. In lane 7, the 5′SS RNA contains a G+5 → U point mutation, represented by **x**. All 5′SS RNAs were incubated in the presence of HeLa nuclear extract under splicing conditions to form splicing complex B (**A**). Aliquots of the same reactions were UV irradiated, centrifuged through a 100 μL cushion of buffer D, and resolved in a 12% SDS gel either directly (**B**) or after treatment with RNase A (**C**).

to monitor crosslinking to hPrp8 (Fig. 2B). In parallel, aliquots of these reactions were treated with RNase A to confirm the location of the hPrp8 crosslink within the 5′SS RNA (Fig. 2C).

The full-length 5′SS RNA (A_5G/GUAAGUAdTdC_3), which efficiently assembles splicing complexes, participates in both catalytic steps, and crosslinks to hPrp8, was used as a reference for all subsequent experiments (Fig. 2A–C, lane 1). First, we analyzed the effect of shortening the intron portion of the 5′SS RNA. Intron positions +4 to +8 in the 5′SS have been shown to interact with U6 snRNA during splicing (Kandels-Lewis & Séraphin, 1993; Konforti et al., 1993; Lesser & Guthrie, 1993; Sontheimer & Steitz, 1993; Crispino & Sharp, 1995). Deletion of these nucleotides is expected to decrease the ability of such 5′SS RNAs to act as splicing substrates. In fact, as the potential base pairing of the 5′SS RNA to U6 snRNA is reduced, so is the efficiency of *trans*-splicing (data not shown), complex B formation, and crosslinking to hPrp8 (Fig. 2, lanes 2 and 3). Deletion analysis of the 5′ exon nucleotides showed a similar effect. Specifically, exon length reduction from 6 to 3 or 1 nt also resulted in a marked decrease in *trans*-splicing efficiency (data not shown and Konforti & Konarska, 1995), complex formation, and crosslinking to hPrp8 (Fig. 2, lanes 1, 4, and 5, respectively). Interestingly, a substrate with no exon nucleotides (intron only), but containing an 11 nt intron, exhibited a reduced level of complex B formation and crosslinking to hPrp8, comparable to that of the substrate with a single exon nucleotide and a shorter (8 nt) intron (compare Fig. 2, lanes 5 and 6). For comparison, we have also used a 5′SS RNA substrate containing the G+5 → U mutation, which is expected to interfere with base pairing to U6 snRNA. This mutation also results in a reduced level of splicing (data not shown), complex B formation, and crosslinking to hPrp8, but the effect is less dramatic than that observed with the above mentioned deletions (compare Fig. 2, lanes 5 and 6 to lane 7). Thus, the sole presence of a GU at the 5′ end of the intron does not suffice for its efficient interaction with hPrp8, as defined by the ability to form the crosslink. These results indicate that the interaction of hPrp8 with the GU dinucleotide has to occur in the context of a functional 5′SS capable of establishing other important interactions with the spliceosome components.

Immunoprecipitation with 70R Ab maps the 5′SS RNA crosslink to the C-terminal region of hPrp8

The analysis of the 5′SS:hPrp8 crosslink suggests that the interaction between these two elements reflects a specific recognition of the GU dinucleotide by hPrp8. To further define the hPrp8:GU dinucleotide interaction, we have used two different approaches to determine the location of the crosslink within the hPrp8 amino acid sequence. First, we carried out immunoprecipitation experiments using the 70R antibody raised against the C-terminus of hPrp8 (positions 1876–2335). We first generated the 5′SS RNA:hPrp8 crosslink by UV irradiation of the 5′ end-labeled 5′SS RNA incubated in the presence of HeLa cell nuclear extracts under splicing conditions. The crosslinked species was isolated from an 8% SDS gel and used to generate a partial digest with endoprotease Arg-C (specific for arginine residues, Fig. 3B, lane 1). The resulting fragments were then immunoprecipitated with either 70R Ab or with preimmune serum (Fig. 3B, lanes 3 and 4, respectively). Alternatively, serum was left out from the experiments (Fig. 3B, lane 5). Bound material was subsequently resolved in a 20% acrylamide/SDS gel

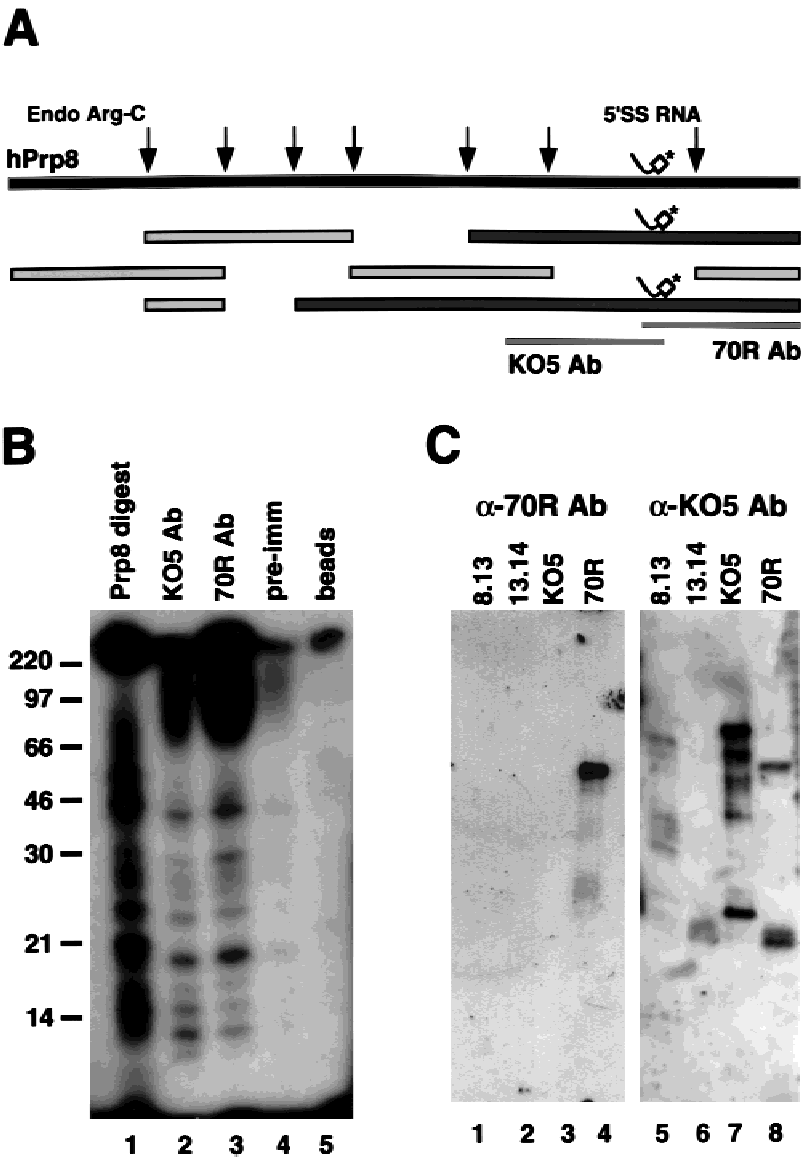

FIGURE 3. The 5′SS RNA:hPrp8 crosslink maps to the C-terminal end of the protein. **A**: The diagram represents the sequence of hPrp8 (black bar) and the region recognized by the antibodies used for immunoprecipitation experiments (70R and KO5 Abs). Arrows indicate sites of cleavage for endoprotease Arg-C. Digestion fragments containing the crosslink to the 5′SS RNA are highlighted. **B**: The 5′SS RNA:hPrp8 crosslink was digested with endoprotease Arg-C to generate a series of partial fragments (lane 1 and panel **A**) that were subsequently incubated with either KO5 antiserum (lane 2), 70R antiserum (lane 3), preimmune serum (lane 4), or protein-A beads (lane 5). The immunoprecipitated material was resolved in a 15% SDS gel and visualized by autoradiography. **C**: Western blot analysis of bacterially expressed hPrp8 fragments 8.13 (positions 761–1284, lanes 1 and 5), 13.14 (positions 1285–1408, lanes 2 and 6), KO5 (positions 1391–1902, lanes 3 and 7), and 70R (positions 1876–2335, lanes 4 and 8) was carried out to determine the specificity of the 70R Ab (lanes 1–4) and KO5 Ab (lanes 5–8). Molecular weight markers (in kDa) are indicated.

and visualized by autoradiography. Only immunoprecipitated fragments that still contain the crosslinked ^{32}P-labeled 5′SS RNA can be visualized in this way (Fig. 3A). As expected, a number of crosslinked hPrp8 fragments were specifically precipitated in the presence of 70R Ab (Fig. 3B, lane 3), whereas the addition of a preimmune serum (Fig. 3B, lane 4) or no serum (Fig. 3B, lane 5) did not generate a significant level of binding. In this experiment, the size of the smallest peptides recognized by the antibody defines the distance between the site of crosslinking and the epitope(s) recognized by the antibody (see Fig. 3A). The smallest fragments specifically immunoprecipitated with the 70R Ab and containing the 5′SS RNA crosslink fall in the range of 8 to 10 kDa (Fig. 3B, lane 3). The slight difference in migration of the immunoprecipitated fragments as compared to the input material (Fig. 3B, compare lanes 2 and 3 to lane 1, see also Fig. 4, lanes 2 and 7) is probably caused by a nuclease activity in the serum that reduces the size of the RNA component of the crosslink. Hence, the size of the immunoprecipitated peptide fragments could not be accurately determined in these experiments, because of the unknown contribution of the RNA moiety to migration of the crosslinked products. However, all immunoprecipitated peptide fragments must at least overlap with the C-terminal 50-kDa region (positions 1876–2335) used to generate the 70R antibody. This result effectively locates the site of the 5′SS RNA crosslink within the 50- to 60-kDa region at the C-terminus of hPrp8.

Similar immunoprecipitation experiments were also carried out with another antibody (KO5 Ab) raised against an overlapping region of hPrp8 (positions 1391–1902). A pattern similar to that seen with the 70R Ab was obtained (Fig. 3B, compare lanes 2 and 3); even small fragments (8–10 kDa) containing the crosslinked 5′SS RNA were immunoprecipitated with KO5 Ab. This suggests that the site of crosslink to the 5′SS RNA is

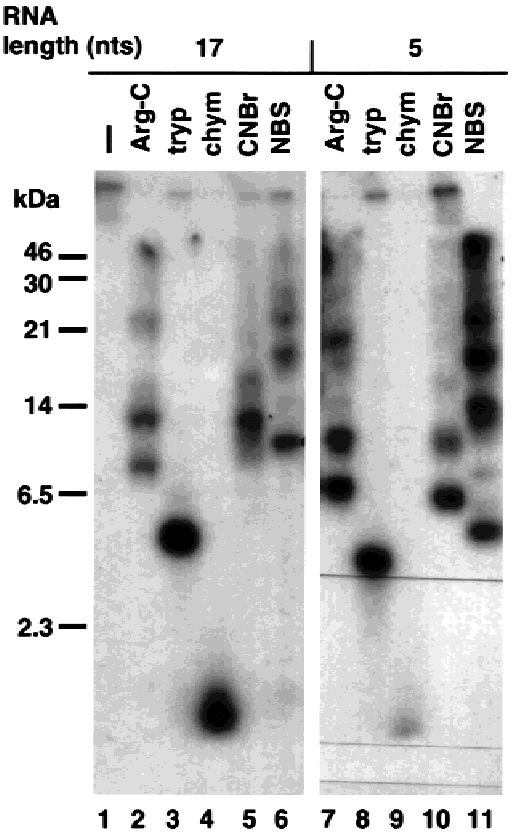

FIGURE 4. Effect of RNA length on migration of 5′SSRNA:hPrp8 crosslink peptide fragments in polyacrylamide/SDS gels. The 5′SS RNA:hPrp8 crosslink was gel purified and used for protease analysis. The 5′SS RNA in the crosslink was 17 nt (lanes 1–6) or 5 nt long (after RNase A digestion, lanes 7–11). The 5′SS RNA:hPrp8 crosslink was digested with the endoproteases Arg-C (lanes 2 and 7), trypsin (lanes 3 and 8), or chymotrypsin (lanes 4 and 9). Alternatively, the crosslink was treated with CNBr (lanes 5 and 10) or NBS (lanes 6 and 11). The digestion products were resolved in a 16% polyacrylamide/6 M urea Tris-Tricine SDS gel and visualized by autoradiography. Molecular weight markers (in kDa) are indicated.

located at, or close to, the overlapping region recognized by the two antibodies (positions 1876–1902). Control Western blot analyses showed that the KO5 Ab recognizes not only the corresponding KO5 antigen peptide, but also cross-reacts with the bacterially expressed 70R peptide fragment used to generate the 70R Ab (Fig. 3C, compare lanes 7 and 8, respectively). These results may be explained by a short stretch of overlapping sequence shared by the 70R and KO5 fragments (amino acids 1876–1902). Alternatively, the KO5 Ab may recognize a short N-terminal His-tag sequence present in both the 70R and KO5 peptides. Since two other fragments of hPrp8 (8.13, positions 761–1284, and 13.14, positions 1285–1408) also show a significantly lower, but detectable, level of reactivity with the KO5 Ab (Fig. 3C, lanes 5 and 6), a combination of both these interpretations may apply. In contrast, Western blot analysis using the 70R Ab and the same bacterially expressed fragments of hPrp8 showed only specific recognition of the 70R fragment originally used to raise this antibody (Fig. 3C, lane 4). No cross-reactivity with the other three fragments was observed (Fig. 3C, lanes 1–3), confirming the high specificity of the 70R Ab and its suitability for mapping experiments. Because of the uncertainty regarding the specificity of the KO5 serum, we did not rely on the results obtained with this antibody, although they are completely consistent with the proteolytic analysis of the 5′SS RNA:hPrp8 crosslink shown below. Unfortunately, all attempts to raise immunoprecipitation-competent antisera to other regions of hPrp8 (e.g., 8.13 and 13.14) have been unsuccessful to date (H.R. Luo, M.J. Moore, unpubl. data). Minimally, the immunoprecipitation results indicate that the site of the 5′SS RNA:hPrp8 crosslink is located within the C-terminus of the hPrp8 protein. The segment spanning the crosslink, a fragment of 8–10 kDa, at least overlaps with the region recognized by the 70R antibody, between positions 1876 and the C-terminus of the protein (position 2335).

Proteolytic analysis of the 5′SS RNA:hPrp8 crosslink

We have extended the analysis of the 5′SS RNA:hPrp8 crosslink to include a number of proteases and chemical reagents of distinct specificities. Based on the size of the fragments containing the 5′SS RNA crosslink and the known sequence of hPrp8 (GenBank accession number AF092565), we could precisely locate the site of crosslink within the protein sequence.

The 5′SS RNA:hPrp8 crosslink isolated from *trans*-splicing reactions was subjected to digestion with a variety of proteolytic reagents, and the generated products were resolved in a high percentage Tris-Tricine/SDS polyacrylamide gel that allows for a higher resolution of peptides in the size range of 2–10 kDa (Schägger & von Jagow, 1987). However, in the case of short peptides, the contribution of the 5′SS RNA to migration of the crosslink in SDS-PAGE becomes significant, precluding accurate size determination. This is demonstrated by the analysis of the 5′SS RNA:hPrp8 crosslink containing either a 17-nt-long 5′SS RNA (Fig. 4, lanes 1–6) or a 5-nt-long segment of the 5′SS RNA (Fig. 4, lanes 7–11), digested with a battery of reagents. The length of the 5′SS RNA most significantly affected the migration of fragments obtained with cyanogen bromide (CNBr; Fig. 4, lanes 5 and 10) and N-bromosuccinimide (NBS; Fig. 4, lanes 6 and 11). In contrast, the RNA size did not significantly affect migration of the trypsin peptide (Fig. 4, lanes 3 and 8). While we do not understand the reason for these reagent-specific effects, the relative position of the crosslink within the peptide affecting the overall topology of the molecule may contribute to this difference. In addition, local amino acid composition may also be involved in these effects. To more precisely determine the size of the crosslinked peptides, similar digestions

were performed with the 5′SS RNA:hPrp8 crosslink containing a single ^{32}P at position G+1 of the 5′SS (A_7G/p*GUAAGUAdTdC_3). Upon digestion of the crosslink with nuclease P1, a single p*G residue remains linked to hPrp8, thus minimizing the RNA contribution to the crosslink mobility in Tris-Tricine/SDS gels. In all subsequent analyses, the molecular weight of the crosslink-containing peptides generated by digestion with a variety of proteolytic reagents was calculated from crosslinks containing a single G nucleotide. Remarkably, digestion of the 5′SS RNA:hPrp8 crosslink with either trypsin or chymotrypsin generates single ^{32}P-labeled peptides (Fig. 4, lanes 3 and 4), indicating a high level of homogeneity of the crosslink with respect to the protein component. Together with the previous analysis that mapped the same crosslink to the GU dinucleotide within the 5′SS RNA, these results indicate a highly site-specific, localized physical interaction between the 5′SS and a unique segment in hPrp8.

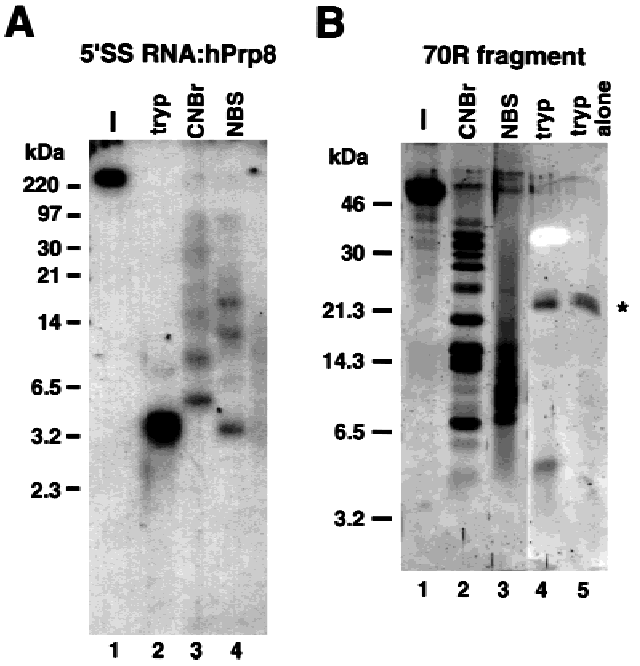

FIGURE 5. Protease analysis of the 5′SS RNA:hPrp8 crosslink. **A**: The 5′SS RNA:hPrp8 crosslink was generated using a 5′SS RNA ^{32}P-labeled at position G+1 (A_6G*pGUAAGUAdTdC_3), which upon digestion with nuclease P1 yields a single *pG nucleotide crosslinked to the protein (lane 1). The crosslink was subsequently digested with trypsin (lane 2), CNBr (lane 3), or NBS (lane 4), resolved in a 16% polyacrylamide/6 M urea Tris-Tricine SDS gel and visualized by autoradiography. **B**: The 70R fragment (50 kDa, positions 1876–2335) of hPrp8 (lane 1) was digested with CNBr (lane 2), NBS (lane 3), or trypsin (lane 4). As a control, a mock trypsin digestion was carried out in the absence of 70R peptide (lane 5). Digestion products were resolved in a 16% polyacrylamide/6 M urea Tris-Tricine SDS gel and visualized by silver staining. The migration position of trypsin is indicated by a *. Molecular weight markers (in kDa) are indicated.

Cyanogen bromide, N-bromosuccinimide and trypsin digestion of the 5′SS RNA:hPrp8 crosslink

Figure 5A shows digestion patterns of the gel-purified 5′SS RNA:hPrp8 crosslink (lane 1) after treatment with trypsin (lane 2) and chemical reagents cyanogen bromide (lane 3) and N-bromosuccinimide (lane 4), resolved in a 16% Tris-Tricine/SDS gel. As described above, the crosslinked peptides contain a single ^{32}P-labeled G nucleotide, which allows one to more accurately estimate the peptide size. Cyanogen bromide (CNBr) specifically cleaves proteins at the carboxy-terminus of methionine residues. The extent of its cleavage depends both on the conditions used and the amino acid context around the cleavage site (Gross, 1967). In the context of the 5′SS RNA:hPrp8 crosslink, we have not been able to obtain a complete CNBr digestion generating a single product (Fig. 5A, lane 3, and Fig. 4, lane 10). The gel-purified partial digestion products were not further cleaved with CNBr (data not shown), possibly due to modification of methionines. The most abundant cleavage product corresponds to the smallest fragment of 5.1 kDa, which may represent a complete digestion product that does not include any internal methionines. From the amino acid sequence of hPrp8, a number of CNBr fragments can be found in the range of 4–6 kDa (indicated in green in Fig. 7A). However, since the 5.1-kDa fragment may represent a partial CNBr digestion product, we have also considered peptides containing one internal methionine for further analysis. Together, 22 distinct CNBr fragments were considered.

N-bromosuccinimide cleaves proteins after tryptophan, histidine, and tyrosine residues (Shechter et al., 1976). As with CNBr, the sequence context affects the extent of its cleavage. Digestion of the 5′SS RNA:hPrp8 crosslink with NBS generated three predominant products of 4.1, 12, and 17 kDa (Fig. 5A, lane 4). While the gel-purified 12-kDa fragment generated the 4.1-kDa fragment upon further NBS treatment, the additional extensive NBS treatment of the 4.1-kDa peptide did not yield smaller digestion products (data not shown). These results strongly suggest that the 4.1-kDa peptide represents a complete cleavage product that does not contain any internal W, H, or Y residues. In the entire hPrp8 protein sequence, six fragments (positions 257–293, 364–394, 1274–1317, 1792–1823, 1840–1875, 1876–1920) are predicted to have a similar molecular size (3.5–5.1 kDa; Fig. 7A, shown in orange). While an additional 11 NBS–partial-cleavage fragments in the same size range can be identified in the entire hPrp8 sequence, none of them is compatible with the proteolytic results described here (data not shown).

Cleavage of the 5′SS RNA:hPrp8 crosslink with trypsin generates a single 4.0-kDa peptide containing the crosslink (Fig. 5A, lane 2). In the entire hPrp8 sequence, there is a single trypsin peptide of this size

(position 1684–1723); however, it does not overlap with any of the sequences identified with CNBr or NBS. This suggests that the trypsin fragment represents a partial cleavage product that contains at least one blocked internal trypsin site (a lysine or arginine residue). A protein may be resistant to trypsin cleavage either because of particular structural features or because the crosslinked RNA itself could block the cleavage site. In fact, when the 4-kDa trypsin fragment was gel-purified and subsequently digested with a high excess of trypsin, a low level of a smaller, 2.0-kDa fragment containing the crosslinked 5′SS RNA could be detected (data not shown).

To independently confirm that the trypsin fragment represents a partially digested peptide, we purified the 70R fragment of hPrp8 (position 1876–2335) expressed in *Escherichia coli* as a His-tagged protein. The gel-purified 70R fragment (Fig. 5B, lane 1) was digested in parallel with CNBr, NBS, or trypsin, and products were resolved in a 20% polyacrylamide/SDS gel and visualized by silver staining (Fig. 5B, lanes 2–5). The sequence of the 70R fragment predicts no trypsin products larger than 3.0 kDa. However, a single peptide with gel mobility of ~4.0 kDa was detected in this trypsin digestion, indicating that it must represent a partial digestion product (Fig. 5B, lane 4). This 4-kDa peptide originates from the 70R fragment since it was not detected in the absence of the 70R substrate (Fig. 5B, lane 5). Smaller digestion products generated in this reaction could not be detected, presumably because of the poor staining with silver under these conditions (Fig. 5B, lanes 2–5). Digestion with CNBr and NBS also yields a partial digestion of the 70R fragment, judging by the size of the detected products (Fig. 5B, lanes 2 and 3). Detection of the 4-kDa trypsin peptide of the 70R fragment suggests that a similar partial peptide in the 5′SS RNA:hPrp8 crosslink results from a local protein sequence or structure rather than from the presence of the crosslink at the trypsin cleavage site. Predicted trypsin fragments of hPrp8 in the range of 3.1 to 5.0 kDa (a total of 38 fragments) are represented by blue bars in Figure 7A. As with CNBr fragments, partial products are also included.

A combination of CNBr, NBS, and trypsin digestions eliminates a majority of the initial pool of the candidate sites and identifies only three possible locations consistent with the proteolysis results shown in this section (Fig. 7, compare A and B). These sequences are located within the trypsin-generated fragments at positions 228–261, 1263–1290, and 1867–1898 (Fig. 7C).

Chymotrypsin and Endo Glu-C protease analysis

In addition to proteolytic digestion of the full-length 5′SS RNA:hPrp8 crosslink described above, we have also performed secondary digestions of the selected, gel-isolated peptide fragments. By imposing additional constraints, this approach allowed us to further distinguish between the three possible locations of the crosslink defined above (Fig. 7C). Digestion of the 5′SS RNA:hPrp8 crosslink with chymotrypsin (which cleaves after tryptophan, tyrosine, and phenylalanine residues) generated a small, 2.5-kDa peptide containing the crosslink (Fig. 6, lane 11). However, secondary chymotrypsin digestion of the gel-purified 4.0-kDa trypsin fragment (Fig. 6, lane 1) generated a smaller, 0.7-kDa fragment (Fig. 6, lane 2), indicating that the 2.5-kDa chymotrypsin fragment is not completely contained within the 4.0-kDa trypsin fragment. All of the three candidate trypsin fragments defined above contain internal chy-

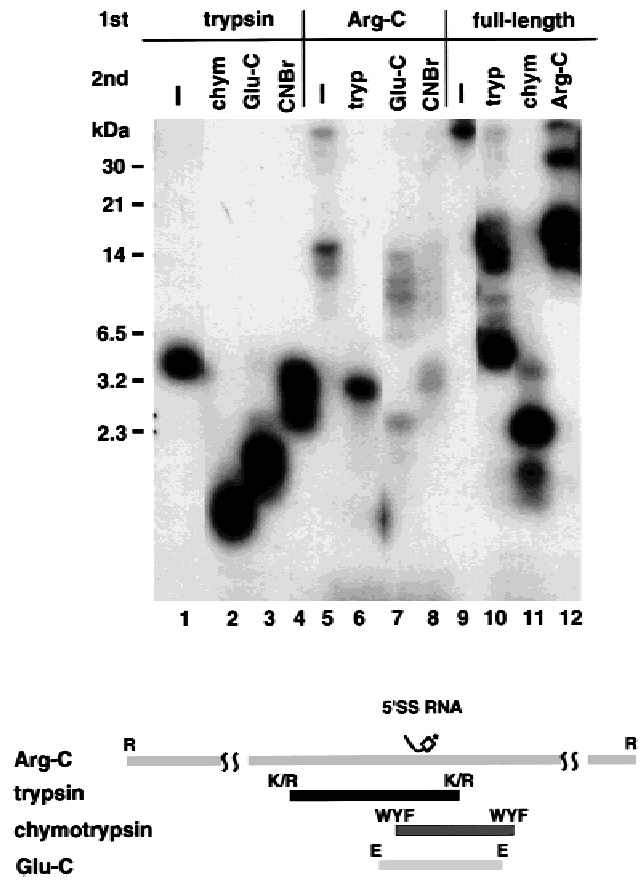

FIGURE 6. Protease analysis of the 5′SS RNA:hPrp8 crosslink: Secondary digestions. The 5′SS RNA:hPrp8 crosslink was generated using a 5′SS RNA ^{32}P-labeled at position G+1 (A_6G*pGUAA $GUAdTdC_3$) and digested with nuclease P1 to yield a single *pG nucleotide crosslinked to the protein (lane 9). The 5′SS RNA:hPrp8 crosslink was digested with trypsin (lane 10), chymotrypsin (lane 11), or endoprotease Arg-C (lane 12). The smallest fragments from trypsin and endo Arg-C digestion were further gel purified (4.0 kDa, lane 1 and 14 kDa, lane 5, respectively). The resulting fragments were subsequently treated with chymotrypsin (lane 2), endoprotease Glu-C (lanes 3 and 7), CNBr (lanes 4 and 8), or trypsin (lane 6). Digestion products were resolved in a 16% polyacrylamide/6 M urea Tris-Tricine SDS gel and visualized by autoradiography. Molecular weight markers (in kDa) are indicated. The bottom panel represents the location of fragments generated by digestion with endo Arg-C, trypsin, chymotrypsin, and endo Glu-C with respect to each other and to the site of crosslink to the 5′SS RNA. The protease specificities are indicated.

motrypsin sites (Fig. 7B,C). However, the trypsin fragment at position 1263–1290 overlaps with a partial chymotrypsin fragment, thus making it less likely as the crosslink site (Fig. 7C).

Digestion of the 5′SS:hPrp8 crosslink with endoprotease Arg-C generates a 14-kDa partial fragment (Fig. 6, lanes 5 and 12). The 14-kDa Arg-C fragment contains the entire 4-kDa trypsin fragment, as its secondary digestion with trypsin yields the same 4.0-kDa peptide (Fig. 6, lane 6). When the 14-kDa Arg-C peptide was further digested with endoprotease Glu-C, the smallest product containing the crosslink was an ~2.5-kDa peptide (Fig. 6, lane 7). Under the conditions used, Glu-C cleaves only after glutamic acid, and not aspartic acid residues (data not shown). The gel-purified 4.0-kDa trypsin fragment digested with Glu-C generates an ~1-kDa peptide, which is smaller than the 2.5-kDa Glu-C peptide obtained from digestion of the Arg-C fragment (Fig. 6, compare lanes 3 and 7). Thus, within the 14-kDa Arg-C fragment, the 4.0-kDa trypsin peptide containing the crosslink must overlap with both Glu-C and chymotrypsin fragments (Fig. 7C). Inspection of the three candidate locations identified by CNBr, NBS, and trypsin digestions identifies the C-terminal site (trypsin peptide position 1867–1898) as the location of the 5′SS RNA:hPrp8 crosslink. The fact that the N-terminal trypsin fragment (position 228–261) does not contain any internal glutamic acid residues, and that the second trypsin fragment (position 1869–1914) contains partial chymotrypsin and Glu-C peptides that do not overlap, eliminates both these sites as a possible location of the crosslink (Fig. 7C).

In summary, the data from digestion of the 5′SS RNA:hPrp8 crosslink with CNBr, NBS, and trypsin gen-

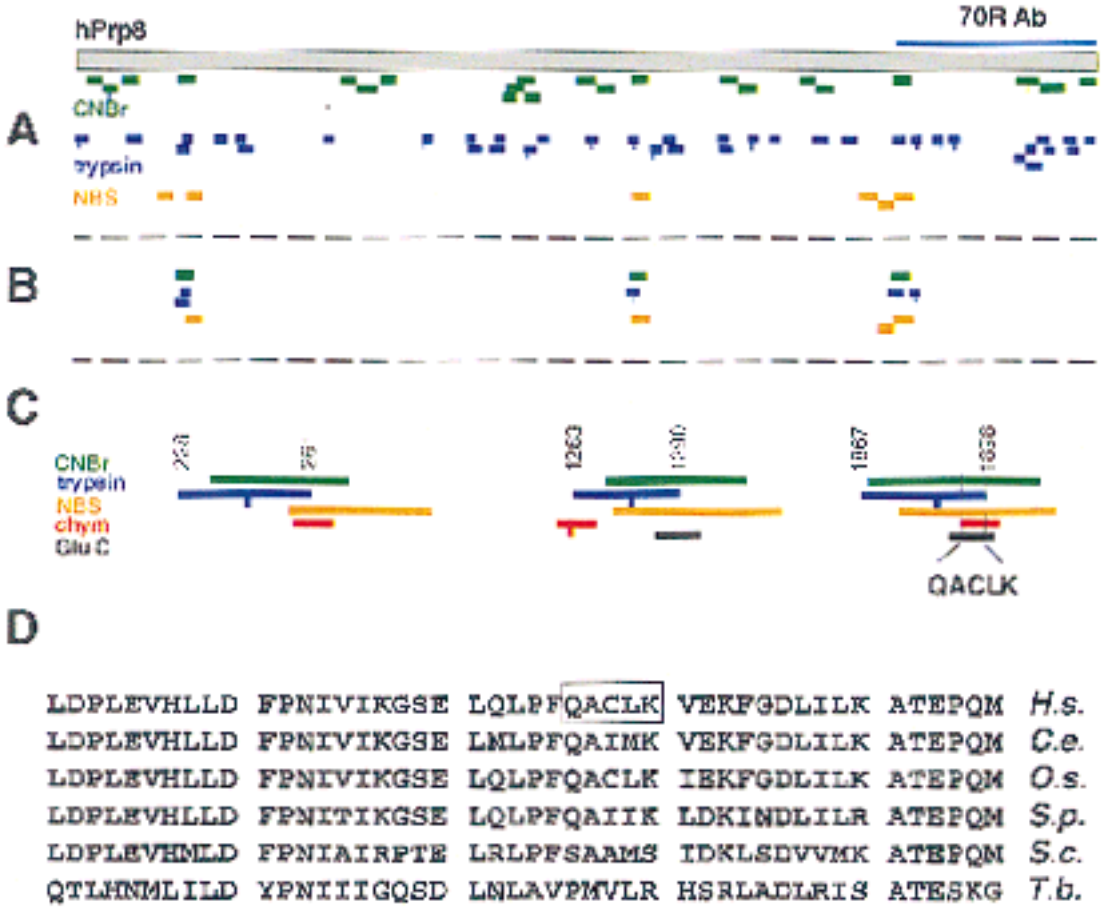

FIGURE 7. Summary of the mapping analysis of the 5′SS RNA:hPrp8 crosslink. **A**: The 5′SS RNA:hPrp8 crosslink (gray bar) was digested with the indicated proteases. Fragments predicted from the hPrp8 sequence are shown for CNBr, NBS, and trypsin digestions. The results from the immunoprecipitation experiments using the 70 Ab are also shown. The three candidate locations for the 5′SS RNA:hPrp8 crosslink shown in **B** are represented as the overlapping fragments obtained from CNBr, NBS, trypsin, endo Glu-C, and chymotrypsin digestion in **C**. In the deduced location of the crosslink, the overlapping region between the different proteolytic fragments is only five amino acids, QACLK (positions 1894–1898), in the hPrp8 amino acid sequence. **D**: Comparison of known Prp8 sequences from other organisms. The alignment shows the region encompassing the site of crosslinking to the 5′SS RNA in hPrp8 (H.s., GenBank accession number AF092565). The sequences from *Caenorhabditis elegans* (*C.e.*), *Oriza sativa* (*O.s.*), and *Saccharomyces cerevisiae* (*S.c.*) are from Hodges et al., 1995. The sequence of *Schizosaccharomyces pombe* (*S.p.*) was obtained from GenBank (accession number 2440191). The sequence of *Trypanosoma brucei* (*T.b.*) is from Lücke et al., 1997.

erated three candidate segments that may contain the site of crosslink (Fig. 7B,C). All three segments comply with the requirements of representing overlapping peptides generated by different proteolytic digestions and having a size compatible with the estimates from the high resolution Tris-Tricine/SDS gels. For the three candidate fragments, the trypsin peptide is considered to represent a partial digestion product with one internal trypsin site. Data from secondary digestions of the isolated trypsin peptide with chymotrypsin and endo Glu-C proteases were used to eliminate two of the three possible locations. Only the C-terminal candidate segment (trypsin peptide positions 1867–1898) is consistent with all the proteolytic analyses and, moreover, with the immunoprecipitation experiments (Fig. 7A,C). Based on the proteolysis results, the site of crosslink is limited by the chymotrypsin and trypsin sites (Fig. 7C), restricting it to only five amino acids in the hPrp8 sequence: QACLK (positions 1894–1898). The achieved resolution is remarkable, considering the total length of hPrp8: 2335 amino acids.

While the entire sequence of Prp8, including the region spanning the crosslink site, exhibits a high degree of conservation across eukaryotes, the amino acid sequence directly adjacent to the mapped crosslink site is significantly less well conserved. The human QACLK sequence (positions 1894–1898) corresponds to QAIMK in *Caenorhabditis elegans*, QAIIK in *Schizosaccharomyces pombe*, and SAAMS (positions 1966–1970) in *Saccharomyces cerevisiae* (Fig. 7D). It should be noted that in *S. cerevisiae*, the mapped site is located in the proximity of amino acid residues implicated in polypyrimidine tract recognition (positions 1834 and 1960; Umen & Guthrie, 1995, 1996), and in genetic interaction with U4 snRNA (Kuhn et al., 1999). Together, these results suggest that the region of Prp8 involved in interactions with the polypyrimidine tract and possibly with U4/U6 snRNAs, may also be involved in recognition of the 5′ splice site.

DISCUSSION

Recognition of the 5′SS by base pairing to the 5′ end of U1 snRNA represents one of the initial steps of spliceosome assembly. While this interaction controls the overall selection of the 5′ splice site, it does not specify the actual cleavage site (reviewed in Moore et al., 1993; Nilsen, 1994). Subsequently, the 5′SS:U1 snRNA basepairing must be disrupted to allow the 5′SS to interact with Prp8, among other components of the U4/U5/U6 snRNP (see Kandels-Lewis & Séraphin, 1993; Konforti et al., 1993; Lesser & Guthrie, 1993; Sontheimer & Steitz, 1993). Both human and yeast Prp8 have been shown to crosslink to the 5′SS region (Wyatt et al., 1992; Teigelkamp et al., 1995a, 1995b; Chiara et al., 1996; Reyes et al., 1996). We have previously shown that a highly specific GU dinucleotide:hPrp8 crosslink at the 5′SS can be detected within splicing complex B assembled in *trans*-splicing reactions, where the 5′SS is provided as a short RNA oligonucleotide (Reyes et al., 1996). The analogous 5′SS:hPrp8 crosslink can also be detected within complex B formed in *cis*-splicing reactions, using a unimolecular pre-mRNA substrate (Fig. 1). Also the kinetics of crosslinking indicate that in both *trans*- and *cis*-splicing, the 5′SS:hPrp8 interaction takes place early in the reaction, after formation of complex B, but before the appearance of splicing intermediates and products. The interaction of Prp8 with exon sequences at the 5′SS is maintained beyond the first step of splicing, as evidenced by the persistence of crosslinks in both yeast and mammalian systems (Wyatt et al., 1992; Teigelkamp et al., 1995b). In contrast, the GU:hPrp8 interaction described here occurs within spliceosome complex B, but crosslink formation is not detected at later stages in the reaction (Fig. 1). This crosslinking profile suggests that either the GU:hPrp8 interaction is disrupted at later stages of splicing, or that an alteration in the physical environment near the contact site may not permit UV crosslinking, even though the interaction itself may be maintained.

The close contact between the 5′SS and Prp8 extends beyond the GU dinucleotide at the 5′ end of the intron. Introduction of relatively small acetamide groups ($\sim$3 Å) attached through the ribose backbone at positions flanking the 5′SS junction (from positions -2 to $+3$) inhibits spliceosome formation, suggesting a steric hindrance in the interaction of Prp8 with the derivatized 5′SS RNA (Sha et al., 1998). Furthermore, photoreactive azidophenacyl groups attached to the 5′ exon (positions -3, -4) form crosslinks with hPrp8 (Sha et al., 1998), consistent with earlier reports of the 5′ exon: Prp8 crosslinks (Wyatt et al., 1992; Teigelkamp et al., 1995a). In addition to Prp8, U2 and U6 snRNAs have also been implicated in interactions with the GU dinucleotide at the 5′SS (Sontheimer & Steitz, 1993; Kim & Abelson, 1996; Luukkonen & Séraphin, 1998a, 1998b). Finally, the 5′SS intron nucleotides positions $+4$ to $+7$ were shown to interact with U6 snRNA (Kandels-Lewis & Séraphin, 1993; Konforti et al., 1993; Lesser & Guthrie, 1993; Sontheimer & Steitz, 1993) and a number of other spliceosomal proteins (Sha et al., 1998). The multiplicity of factors that contact the 5′SS region may explain why the GU:hPrp8 crosslinking is affected by both the exon and intron sequences flanking the 5′SS junction (Fig. 2). All these different interactions involving the 5′SS contribute to the overall recognition of the substrate, its joining with the spliceosome, and its proper positioning at the active site of the complex.

The 5′SS RNA:hPrp8 crosslink maps to a small segment in the C-terminal region of the protein

To gain more insight into the interaction between the 5′SS RNA and hPrp8, we identified the region of the

protein involved in contacting the GU dinucleotide. In the first approach, we carried out immunoprecipitation experiments using antibodies raised against two overlapping C-terminal regions of hPrp8 (70R and KO5 Abs; Fig. 4). From a mixture of proteolytic fragments containing the crosslink, peptides as small as ~8–10 kDa could be immunoprecipitated with these antibodies, indicating that the site of crosslink is located within the C-terminal portion of hPrp8. Formally, the crosslink should be located near the region recognized by both antibodies, that is, positions 1876–1902. Because of some cross-reactivity of the KO5 Ab, we only considered the results obtained with the highly specific 70R Ab, effectively defining the crosslink site to a 50–60-kDa region at the C-terminus of hPrp8. However, the results of the second mapping approach using a battery of endoproteolytic reagents to analyze the 5′SS RNA:hPrp8 crosslink were consistent with the immunoprecipitation experiments. The crosslink peptides obtained from digestions with a number of proteolytic reagents constituted a single product, indicating that the site of crosslinking is highly homogeneous within the hPrp8 sequence, and allowing for the high-resolution mapping of the crosslink. By comparing the size of the resulting crosslink-containing peptides with the digestion pattern predicted from the sequence of hPrp8, we could restrict the site of crosslink to a short, five-amino-acid stretch.

The use of just three different proteolytic reagents (CNBr, NBS, and trypsin) allowed us to identify three potential sites of crosslinking within the entire hPrp8 sequence. Digestions of the trypsin-generated fragment with chymotrypsin and endoprotease Glu-C allowed us to eliminate two of these candidate locations. Significantly, the third location is consistent with all the proteolysis as well as the immunoprecipitation results. The region of overlap between the different proteolytic fragments limits the site of crosslink to the amino acid sequence QACLK (positions 1894–1898) at the C-terminal region of hPrp8. While it was not possible to determine the exact amino acid that forms the photoadduct with the GU dinucleotide at the 5′SS, the cysteine residue at position 1896 is a likely candidate. In model DNA:amino acid crosslinking systems, cysteine is one of the most reactive residues, forming a number of different adducts with thymidine (Saito & Sugiyama, 1990), although other amino acids also have some ability to form crosslinks with DNA (Shetlar et al., 1984). Among the characterized DNA:protein complexes, a thymidine:cysteine crosslink has been characterized in the bacteriophage fd (Paradiso & Konigsberg, 1982), whereas uracil:methionine and uracil:tyrosine adducts have been described among ribosomal RNA-protein crosslinks (Zwieb & Brimacombe, 1978; Maly et al., 1980). Finally, UV crosslinking of thymidine to lysine residues has also been observed in the context of DNA:histone interactions (Saito & Sugiyama, 1990).

Prp8 interacts with multiple spliceosomal components

The mapped 5′SS:hPrp8 crosslink does not represent the sole interaction between the pre-mRNA and Prp8. Human Prp8 has been shown to crosslink to the branch site region within splicing complexes B and C (MacMillan et al., 1994) and to sequences around the 3′SS within splicing complex C, most likely after the first and before the second catalytic step (Chiara et al., 1997). Similarly, yeast Prp8 forms crosslinks with the branch site region and the 3′ exon after the first catalytic step (Teigelkamp et al., 1995b). The involvement of Prp8 in the 3′SS recognition is supported by the identification of the yeast *PRP8* allele (*prp8-101*) isolated as a suppressor of mutations in the polypyrimidine tract that selectively block the second step of splicing (Umen & Guthrie, 1995). A screen for additional mutations has defined two regions in the yPrp8 protein involved in fidelity of the 3′SS choice and polypyrimidine tract recognition (Umen & Guthrie, 1996). A comparison of human and yeast Prp8 amino acid sequences indicates that the polypyrimidine tract recognition domain is located in close proximity to the 5′SS RNA:hPrp8 crosslink. The *prp8-101* and *prp8-102* alleles implicated in polypyrimidine tract recognition represent mutations of E (position 1960) in yPrp8, while the 5′SS crosslink corresponds to positions 1966–1970 of the yeast protein. This feature, together with the biochemical and genetic data indicating similar roles for yeast and human Prp8, suggests that the homologous region in yPrp8 is involved in recognition of the 5′SS.

In addition to the above-discussed contacts with the pre-mRNA, Prp8 is expected to interact with a number of spliceosomal components. An allele of *PRP8* (*prp8-201*) was found to suppress a cold-sensitive allele of U4 snRNA. The mutation in *prp8-201* maps close to the region involved in polypyrimidine tract recognition (Kuhn et al., 1999; Umen & Guthrie, 1996). Prp40, a component of yeast U1 snRNP, physically interacts with the proline-rich domain at the N-terminus of yPrp8 (Abovich & Rosbash, 1997). Finally, Prp8 was shown to genetically interact with two putative helicases: DED1 (Jamieson et al., 1991) and Prp28 (Strauss & Guthrie, 1991). Interestingly, the human homolog of Prp28 (U5-100kD) is also an integral component of the U5 snRNP (Teigelkamp et al., 1997), suggesting that Prp8 may directly interact with Prp28 and possibly other U5 snRNP components, including U5 snRNA.

While the crosslinking studies have confirmed a close contact between Prp8 and pre-mRNA, Prp8 also has been shown to interact with a series of other spliceosomal components throughout the reaction. Thus, it is likely that Prp8 plays a central role in splicing, coordinating a series of events through interactions with several spliceosomal components, and perhaps participating in formation of the active site of the splice-

osome. This could be achieved either by providing a binding site for the splice sites, or possibly also by contributing to the chemical reaction in a more direct way. At present, the mechanistic details concerning the pre-mRNA splicing catalysis are not known. The catalytic center is likely to involve a combination of specific RNA:RNA interactions and other contacts that rely on amino acid residues. Because of the discussed close contacts between Prp8, pre-mRNA, and other spliceosomal components, as well as the established RNA:RNA interactions involving U2 and U6 snRNAs, it seems likely that Prp8, along with U2 and U6 snRNAs, participates in formation of the catalytic center.

MATERIALS AND METHODS

5′SS RNA oligonucleotides

Oligonucleotides were synthesized using an Applied Biosystems 390 synthesizer, and phosphoramidites were from Glen Research (Sterling, Virginia). The 3′ end labeling of oligonucleotides was carried out using the Klenow fragment of *E. coli* DNA polymerase (Boehringer Mannheim) and α^{32}P-dCTP (NEN) as described (Konforti & Konarska, 1995). The 5′ end labeling of 5′SS RNAs was carried out using T4 polynucleotide kinase (New England Biolabs) and γ^{32}P-ATP (NEN) as described (Konforti & Konarska, 1995).

3′SS RNAs

In most experiments the 3′SS RNA contained 83 nt of intron and 45 nt of exon and was prepared by transcription with T7 RNA polymerase from a pBSAd13 plasmid (Konarska, 1989) cut with *Sau*3A1. For *cis*-splicing experiments, the rabbit β-globin pre-mRNA was derived from the pAL4 plasmid (Lamond et al., 1987) transcribed with T7 RNA polymerase. For experiments shown in Figure 1C, the 3′SS RNA transcript contained 145 nt of intron sequence and 54 nts of 3′ exon sequence. This RNA was subsequently ligated to the 5′ end ^{32}P-labeled 5′SS RNA (A_5G/GUAAGUAdT) using T4 DNA ligase (Boehringer Mannheim) and a bridging DNA oligonucleotide (Moore & Sharp, 1992). For Figure 1D, the β-globin pre-mRNA was transcribed from a DNA template containing 20 nt of the second exon and 46 nt of the intron spanning the 5′ splice site region, attached to the sequences of the 3′SS RNA (145 nt of the intron and 54 nt of the third exon).

In vitro *trans*-splicing assays

The preparation of HeLa cell nuclear extracts and *trans*-splicing reactions (Konforti & Konarska, 1995) were carried out as described. To analyze snRNP complex formation, aliquots of the reactions were resolved in 4% polyacrylamide/50 mM Tris-glycine nondenaturing gels (Konarska, 1989). Alternatively, phenol extracted RNA products were resolved in a 10% polyacrylamide/8 M urea gel. In either case, electrophoresis was carried out such that the unbound or unreacted 5′SS RNA remained in the gel to allow for determination of complex formation or *trans*-splicing efficiency, respectively. All gels were quantified using a Molecular Dynamics phosphorimager.

Immunoprecipitations

Polyclonal rabbit and chicken antisera were raised against bacterially expressed His-tagged fragments of hPrp8 (70R, 50-kDa protein fragment positions 1876–2335) and KO5 (60-kDa fragment positions 1391–1902), respectively. Anti-hPrp8 (70R Ab or KO5 Ab) or preimmune sera were coupled to Protein A-trisacryl beads (Pierce) in IP100 buffer (100 mM NaCl, 2 mM $MgCl_2$, 50 mM Tris, pH 7.6, 0.5 mM DTT, 0.05% Nonidet P40) for 2 h on ice and washed twice with the same buffer. Samples were incubated with the antibody-coupled beads for 3 h on ice, and rinsed twice with IP150-1M urea buffer (150 mM NaCl, 2 mM $MgCl_2$, 50 mM Tris, pH 7.6, 0.5 mM DTT, 0.05% Nonidet P40, 1 M urea). The bound material was resolved in a 15% polyacrylamide SDS gel.

The 8.13 (positions 761–1284), 13.14 (positions 1285–1408), KO5, and 70R peptides were expressed in *E. coli* as His-tagged fusion proteins (H. Luo, M. Moore, unpubl. results), and additionally gel-purified from a 10% polyacrylamide SDS gel. Western blot analysis was carried out by blocking the membrane overnight at 4 °C in TBST buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.1% Tween 20), followed by incubation with 70R Ab or KO5 Ab (1:5,000 dilution) for 2 h at 4 °C in TBST buffer. After extensive washes in TBS buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl), the secondary antibody (anti-rabbit IgG for 70R Ab, anti-chicken IgY for KO5) conjugated to horseradish peroxidase was added and incubation continued for 2 h in TBST buffer at 4 °C. Visualization of the results was performed using the Renaissance reagents from NEN.

Proteolytic analysis of the 5′SS RNA:hPrp8 crosslink species

One hundred microliters *trans*-splicing reactions were UV irradiated for 4 × 5 min and centrifuged through a 100 μL cushion of buffer D for 30 min at 70,000 rpm in a Beckman TL-100 ultracentrifuge. When indicated, the resulting pellet was further treated with RNase A (1 μg) or nuclease P1 (0.5 U) for 30 min at 37 °C. The 5′SS RNA:hPrp8 crosslink was then purified from an 8% polyacrylamide SDS gel. The recovered material was dissolved directly in the digestion buffer for each proteolytic reagent. Incubations with endoprotease Arg-C (1 U, Sigma) were carried out in 0.1 M sodium acetate, pH 7.4. To generate partial digestion products used in Figure 3A, two parallel incubations with 0.05 U and 0.1 U of Arg-C were used. After digestion, fragments were directly diluted in IP100 buffer and used for immunoprecipitation. Under these conditions, Arg-C activity during the incubation with antibodies was negligible. Digestions with trypsin (0.1 U, Sigma) were carried out in 0.1 M Tris-HCl, pH 8.0, 20 mM $CaCl_2$, chymotrypsin (0.1 U, Sigma) in 0.1 M sodium acetate, pH 8.0, and endoprotease Glu-C (0.1 U, Sigma) in 0.1 M ammonium bicarbonate, pH 7.8. All digestions were incubated in 10–20 μL for 2 h at 37 °C and reactions were terminated by addition of the SDS sample buffer (2% SDS, 15 mM Tris-HCl, pH 6.8, 30% glycerol, 0.06% bromophenol blue).

Cyanogen bromide was adjusted to 40 mg/mL and 70% formic acid in the reaction and incubation was carried out overnight at room temperature in the dark (Gross, 1967). N-bromosuccinimide reactions contained 15 mM NBS in 40% acetic acid, 6.6 M urea (Mirfakhrai & Weiner, 1993), and in-

cubations were carried out for 1 h at room temperature. The CNBr and NBS treated samples were ethanol precipitated in the presence of 10 mM Tris base. Protease and chemical digestions were resolved in a Laemmli 22% polyacrylamide/SDS gel. Alternatively, a 16% polyacrylamide/6 M urea Tris-Tricine SDS gel was used (Schägger & von Jagow, 1987).

ACKNOWLEDGMENTS

J.L.R. was supported by a John Calvert Eistenstein Fellowship. This work was supported by the NIH grant GM49044 to M.M.K.

Received September 23, 1998; returned for revision October 20, 1998; revised manuscript received October 30, 1998

REFERENCES

Abovich N, Rosbash M. 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell 89*:403–412.

Chiara MD, Gozani O, Palandjian L, Reed R. 1996. Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. *Mol Cell Biol 15*:3317–3326.

Chiara MD, Palandjian L, Kramer RF, Reed R. 1997. Evidence that U5 snRNP recognizes the 3′ splice site for catalytic step II in mammals. *EMBO J 16*:4746–4759.

Crispino JD, Sharp PA. 1995. A U6 snRNA:pre-mRNA interaction can be rate-limiting for U1-independent splicing. *Genes & Dev 9*:2314–2323.

Gross E. 1967. The cyanogen bromide reaction. In: Hirs CHW, ed. *Methods in enzymology*. New York: Academic Press. pp 238–255.

Hodges PE, Jackson SP, Brown JD, Beggs JD. 1995. Extraordinary sequence conservation of the PRP8 splicing factor. *Yeast 11*:337–342.

Jamieson DJ, Rahe B, Pringle J, Beggs JD. 1991. A suppressor of a yeast splicing mutation (prp8-1) encodes a putative ATP-dependent RNA helicase. *Nature 349*:715–717.

Kandels-Lewis S, Séraphin B. 1993. Role of U6 snRNA in 5′ splice site selection. *Science 262*:2035–2039.

Kim CH, Abelson J. 1996. Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5′ splice site. *RNA 2*:995–1010.

Konarska MM. 1989. Analysis of splicing complexes and small nuclear ribonucleoprotein particles by native gel electrophoresis. In: Dahlberg JE, Abelson JN, eds. *Methods in enzymology*. San Diego, California: Academic Press. pp 442–453.

Konforti BB, Konarska MM. 1995. A short 5′ splice site RNA oligo can participate in both steps of splicing in mammalian extracts. *RNA 1*:815–827.

Konforti BB, Koziolkiewicz MJ, Konarska MM. 1993. Disruption of base pairing between the 5′ splice site and the 5′ end of U1 snRNA is required for spliceosome assembly. *Cell 75*:863–873.

Kuhn A, Li Z, Brow DA. 1999. Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. *Mol Cell*. In press.

Lamond AI, Konarska MM, Sharp PA. 1987. A mutational analysis of spliceosome assembly: Evidence for splice site collaboration during spliceosome formation. *Genes & Dev 1*:532–543.

Lesser CF, Guthrie C. 1993. Mutations in U6 snRNA that alter splice site specificity: Implications for the active site. *Science 262*:1982–1988.

Lücke S, Klöckner T, Palfi Z, Boshart M, Bindereif A. 1997. Trans mRNA splicing in trypanosomes: Cloning and analysis of a PRP8-homologous gene from Trypanosoma brucei provides evidence for a U5-analogous RNP. *EMBO J 16*:4433–4440.

Luukkonen BGM, Séraphin B. 1998a. Genetic interaction between U6 snRNA and the first intron nucleotide in *Saccharomyces cerevisiae*. *RNA 4*:167–180.

Luukkonen BGM, Séraphin B. 1998b. A role for U2/U6 helix Ib in 5′ splice site selection. *RNA 4*:915–927.

MacMillan AM, Query CC, Allerson CR, Chen S, Verdine GL, Sharp PA. 1994. Dynamic association of proteins with the pre-mRNA branch region. *Genes & Dev 8*:3008–3020.

Maly P, Rinke J, Ulmer E, Zwieb C, Brimacombe R. 1980. Precise localization of the site of cross-linking between protein L4 and 23S ribonucleic acid induced by mild ultraviolet irradiation of *Escherichia coli* 50S ribosomal subunits. *Biochemistry 19*:4179–4188.

Mirfakhrai M, Weiner AM. 1993. Chemical Cleveland mapping: A rapid technique for characterization of crosslinked nucleic acid–protein complexes. *Nucleic Acids Res 21*:3591–3592.

Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNA by the spliceosome. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–357.

Moore MJ, Sharp PA. 1992. Site-specific modification of pre-mRNA: The 2′-hydroxyl groups at the splice sites. *Science 256*:992–997.

Newman AJ. 1997. The role of U5 snRNP in pre-mRNA splicing. *EMBO J 16*:5797–5800.

Nilsen TW. 1994. RNA–RNA interactions in the spliceosome: Unraveling the ties that bind. *Cell 78*:1–4.

O'Keefe RT, Norman C, Newman AJ. 1996. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell 86*:679–689.

Paradiso PR, Konigsberg W. 1982. Photochemical cross-linking of the gene 5 protein.fd DNA complex from fd-infected cells. *J Biol Chem 257*:1462–1467.

Reyes JL, Kois P, Konforti BB, Konarska MM. 1996. The canonical GU dinucleotide at the 5′ splice site is recognized by p220 of the U5 snRNP within the spliceosome. *RNA 2*:213–225.

Saito I, Sugiyama H. 1990. Photoreactions of nucleic acids and their constituents with amino acids and related compounds. In: Morrison H, ed. *Bioorganic photochemistry*. New York: J Wiley & Sons. pp 317–340.

Schägger H, von Jagow G. 1987. Tricine-sodium dodecyl sufate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem 166*:368–379.

Sha M, Levy T, Kois P, Konarska MM. 1998. Probing of the spliceosome with site-specifically derivatized 5′ splice site RNA oligonucleotides. *RNA 4*:1069–1082.

Shechter Y, Patchornik A, Burstein Y. 1976. Selective chemical cleavage of tryptophanyl peptide bonds by oxidative chlorination with N-chlorosuccinimide. *Biochemistry 15*:5071–5075.

Shetlar MD, Christensen J, Hom K. 1984. Photochemical addition of amino acids and peptides to DNA. *Photochem Photobiol 39*:125–133.

Sontheimer EJ, Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science 262*:1989–1996.

Strauss EJ, Guthrie C. 1991. A cold-sensitive mRNA splicing mutant is a member of the RNA helicase gene family. *Genes & Dev 5*:629–641.

Teigelkamp S, Mundt C, Achsel T, Will CL, Lührmann R. 1997. The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. *RNA 3*:1313–1326.

Teigelkamp S, Newman AJ, Beggs JD. 1995a. Extensive interactions of PRP8 protein with the 5′ and 3′ splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO J 14*:2602–2612.

Teigelkamp S, Whittaker E, Beggs JD. 1995b. Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. *Nucleic Acids Res 23*:320–326.

Umen JG, Guthrie C. 1995. A novel role for a U5 snRNP protein in 3′ splice site selection. *Genes & Dev 9*:855–868.

Umen JG, Guthrie C. 1996. Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3′ splice site selection. *Genetics 143*:723–739.

Wyatt JR, Sontheimer EJ, Steitz JA. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. *Genes & Dev 6*:2542–2553.

Zwieb C, Brimacombe R. 1978. RNA–protein cross-linking in *Escherichia coli* 30S ribosomal subunits: A method for the direct analysis of the RNA regions involved in the cross-links. *Nucleic Acids Res 5*:1189–1206.