Molecular Cell, Vol. 7, 319–329, February, 2001, Copyright ©2001 by Cell Press

A Biochemical Function for the Sm Complex

Dong Zhang, Nadja Abovich, and Michael Rosbash*
Howard Hughes Medical Institute
Department of Biology
Brandeis University
Waltham, Massachusetts 02254

Summary

Within the yeast commitment complex, SmB, SmD1, and SmD3 make direct contact with the pre-mRNA substrate, close to the 5′ splice site. Only these three Sm proteins have long and highly charged C-terminal tails, in metazoa as well as in yeast. We replaced these proteins with tail-truncated versions. Genetic assays demonstrate that the tails contribute to similar and overlapping functions, and cross-linking assays show that the tails make direct contact with the pre-mRNA in a largely sequence-independent manner. Other biochemical assays indicate that they function at least in part to stabilize the U1 snRNP-pre-mRNA interaction. We speculate that this role may be general, and may have even evolved to aid weak intermolecular nucleic acid interactions of only a few base pairs.

Introduction

In eukaryotes, a large group of small nuclear ribonucleoprotein particles (snRNPs) contains a set of common proteins, called Sm proteins. These Sm-containing snRNPs participate in many different RNA-processing reactions and include U1, U2, U4, and U5 snRNPs, which are involved in the splicing of U2-dependent introns (Yu et al., 1999). U7 snRNP is also an Sm snRNP, and is essential for the 3′ end formation of histone mRNA (Grimm et al., 1993). It has recently been demonstrated that even the budding yeast telomerase is an Sm-containing snRNP (Seto et al., 1999). All of these snRNPs share certain features. The snRNA component is an RNA polymerase II transcript whose 5′ end is commonly capped and then further modified by hypermethylation (Reddy and Busch, 1988). The snRNAs all contain a conserved sequence, the Sm site, and the consensus sequence is $PuAU_{4-6}GPu$, where the Sm protein complex binds (Guthrie and Patterson, 1988). There are seven prominent Sm proteins: SmB/B′, SmD1, SmD2, SmD3, SmE, SmF, and SmG. All seven contain two conserved motifs, Sm1 and Sm2. Recent structural data show that these Sm domains mediate protein–protein interactions in a proposed seven-member ring/doughnut structure, whose positively charged interior is thought to interact with the Sm site of the snRNAs (Kambach et al., 1999). The Sm complex has been shown to influence the hypermethylation of U snRNAs as well as the assembly and nuclear localization of snRNPs in metazoans (Mattaj, 1986; Hamm et al., 1987, 1990; Fischer et al., 1993; Nelissen et al., 1994). It is also known that detailed features of the Sm complex and its positioning within an snRNP are critical for snRNP function (Maroney et al., 1990; Nelissen et al., 1994). But the primary function(s) of the Sm complex and that of individual Sm proteins remains largely unknown.

During the earliest step of pre-mRNA splicing, U1 snRNP interacts with the pre-mRNA substrate to form the commitment complex in yeast or the E (early) complex in mammals (Séraphin and Rosbash, 1989; Michaud and Reed, 1991; Séraphin and Rosbash, 1991). Within this complex there is an important RNA–RNA interaction between U1 snRNA and the 5′ splice site (5′ ss) region (Zhuang and Weiner, 1986; Séraphin et al., 1988; Siliciano and Guthrie, 1988). We previously used a 4-thiouridine cross-linking assay to show that there are also direct protein–RNA contacts between eight commitment complex proteins and the pre-mRNA substrate (Zhang and Rosbash, 1999). Three of these are Sm proteins, SmB, SmD1, and SmD3, which make direct contacts with the 5′ exon region close to the 5′ ss. According to biochemical, two-hybrid, and crystal structure data, the three cross-linked Sm proteins are on the same side of the ring/doughnut structure in the order of D1-B-D3 (Lehmeier et al., 1994; Hermann et al., 1995; Raker et al., 1996; Kambach et al., 1999). Interestingly, of the seven Sm proteins, only these three have long and positively charged C-terminal tails, and the three tails are highly disordered in the crystal structures of Sm subcomplexes. We reasoned that the tails of these adjacent Sm proteins might be responsible for the Sm protein–RNA interactions, and therefore important for splicing and other processing events. Indeed, genetic and biochemical assays of different tail-truncated Sm-containing strains and extracts indicate that these positively charged C-terminal tails directly contact the pre-mRNA substrate within the commitment complex, and that these contacts contribute to complex stability.

Results

C-Terminal Tails Are Responsible for Protein–RNA Contacts in the Commitment Complex

To study the function of the Sm tails, we constructed viable yeast strains in which inviable Sm gene deletions were rescued by C-terminal truncated versions (truncation sites are indicated by arrows in Figure 1A). Comparable shortened human versions of the three proteins were used in structural studies (Kambach et al., 1999). The truncated Sm proteins are stable in yeast, but with somewhat variable expression levels, similar to their wild-type counterparts (Figure 1B; Roy et al., 1995). To test whether the truncated proteins are successfully assembled into snRNPs, we immunoprecipitated with antibodies specific for the tagged, truncated proteins, and assayed U1 and U2 snRNAs by primer extension. Both snRNAs were recovered from all of the single and double mutant strains (Figure 1C) except for the SmD3ΔC single truncation strain (Figure 1C, lane 5). We interpret this

* To whom correspondence should be addressed (e-mail: rosbash@brandeis.edu).

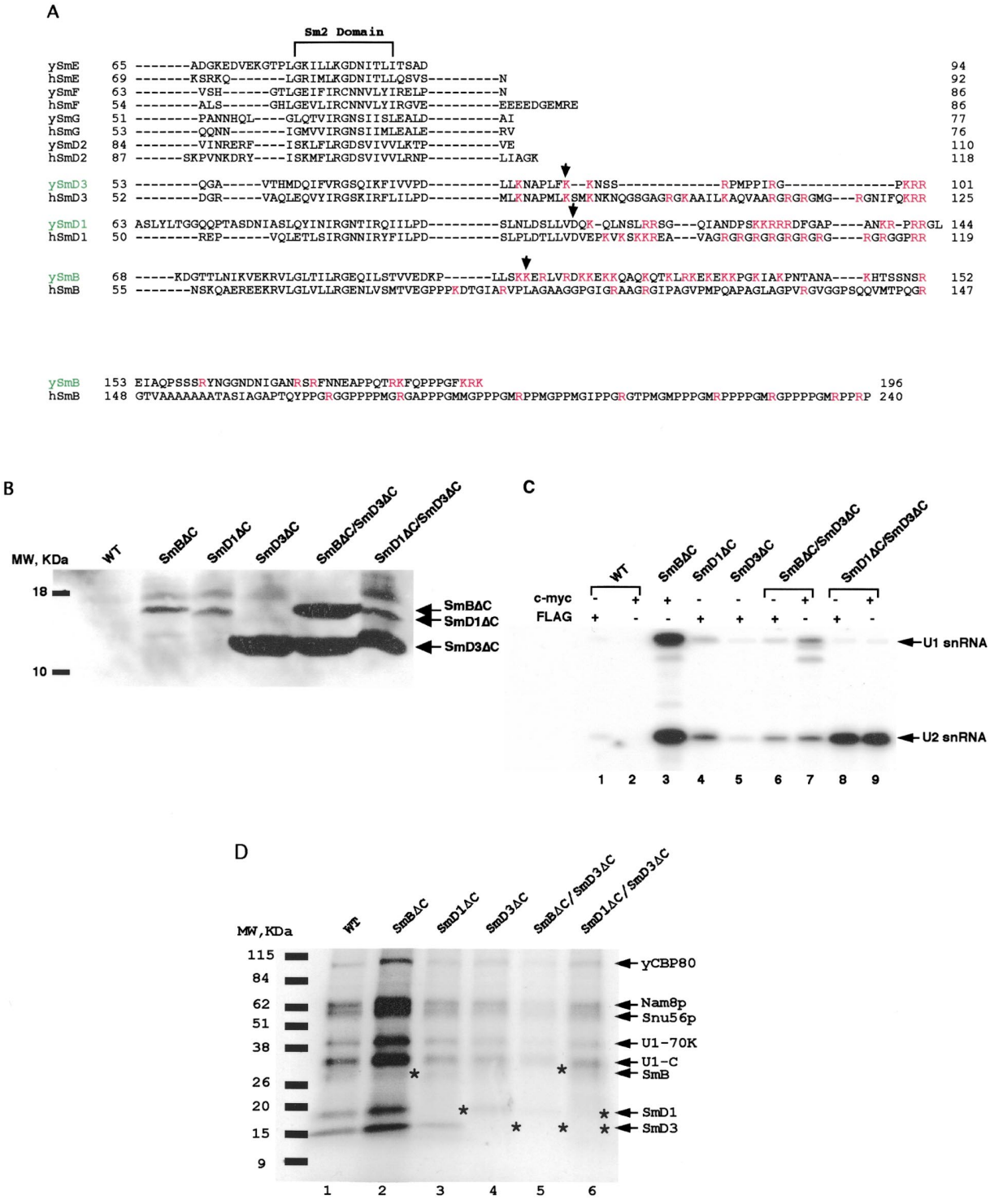


Figure 1. Cross-Linking Profile of Sm Tail Truncation Mutants

(A) Sequence alignment of yeast and human Sm proteins. Only the C-terminal regions that include the Sm2 domains and the C-terminal tails are shown. Basic residues in the C-terminal regions are colored in red. Arrows indicate the truncation point of yeast SmB, SmD1, and SmD3 proteins used in this study, and of the human SmB, SmD1, and SmD3 used in structural studies (Kambach et al., 1999).

(B) Western blot of tail truncation splicing extracts. Details of strain construction are described in Experimental Procedures. 5 μl of splicing extract was resolved by 15% SDS-PAGE. Since the tail-truncated Sm proteins were tagged with either a FLAG or a c-myc epitope at their C termini, a mixture of monoclonal antibodies against FLAG and c-myc (Sigma) was used for Western blotting. The bands corresponding to the truncated Sm proteins are indicated by arrows. The tags in each strain are: SmBΔC-c-myc, SmD1ΔC-FLAG, SmD3ΔC-FLAG, SmBΔC-c-myc/SmD3ΔC-FLAG, and SmD1ΔC-c-myc/SmD3ΔC-FLAG.

(C) Primer extension of immunoprecipitated U1 snRNA and U2 snRNA from tail truncation extracts. The antibody used for immunoprecipitation

failure to epitope inaccessibility in this strain. U1 snRNA was also just barely detectable from the SmD1ΔC/SmD3ΔC double truncation strain (Figure 1C, lanes 8 and 9).

We then made splicing extracts from these strains and assayed them by cross-linking the protein-pre-mRNA contacts in the commitment complex. As indicated by asterisks in Figure 1D, all protein-pre-mRNA contacts were present, except those corresponding to the particular truncated Sm protein(s). For example, SmD1 is the only missing protein in the cross-linking pattern using the SmD1ΔC extract, and there is no new band that corresponds to the size of the truncated SmD1 protein (Figure 1D, compare lanes 1 and 3). Similar results were obtained for the SmBΔC and SmD3ΔC extracts as well as for the two viable double-truncated strains, SmBΔC/SmD3ΔC and SmD1ΔC/SmD3ΔC; we note, however, that the assignment of the SmB protein band is somewhat uncertain as previously described (Zhang and Rosbash, 1999). These results are consistent with the notion that the positively charged C-terminal tails of SmB, SmD1, and SmD3 make direct contacts with the pre-mRNA within the commitment complex. The band intensities in SmBΔC were somewhat variable (Figure 1D, lane 2 and data not shown). In contrast, the band intensities in SmBΔC/SmD3ΔC were consistently lower than those in the other extracts (Figure 1D, lane 5); this suggests that the formation or stability of the commitment complex is compromised in the SmBΔC/SmD3ΔC extract (see Figures 3 and 4). We also found that all of the cross-linked proteins in the tail-truncated extracts bind to the same regions of the pre-mRNA substrate (data not shown; Zhang and Rosbash, 1999), which indicates that there is no dramatic distortion of the commitment complex after removal of the tail(s).

C-Terminal Tails Contribute to Shared or Overlapping Functions and Affect In Vivo Pre-mRNA Splicing Efficiency

Although these Sm tails are inessential for cell viability, we examined the growth rates of the truncation strains more carefully (Figures 2A and 2B). The absence of the SmB tail has no effect on growth rate, and the absence of the SmD3 tail has only a mild effect. In contrast, the SmD1 tail deletion strain grows poorly. We also observed strong synthetic interactions, such as that the double SmB and SmD1 truncation is lethal, and that the two other double mutant combinations are viable but grow poorly and show temperature-sensitive phenotypes. These synthetic interactions indicate that the three Sm tails contribute to shared or overlapping functions.

To examine the effect on in vivo splicing, we assayed the splicing efficiency of precursor U3 small nucleolar RNA (pre–U3 snoRNA) in the tail truncation strains by primer extension (Figure 2C; quantitation of pre–U3 snoRNA is shown in Figure 2D). Whereas the wild-type strain accumulates no detectable pre–U3 snoRNA, all Sm tail-truncated cells show a substantial increase in pre–U3 snoRNA levels. U2 snRNA was used as an internal control. (U snRNA levels are not affected by Sm protein tail truncations; data not shown.) The pre–U3 snoRNA accumulation assay correlates well with the growth rate assay, as the SmB and SmD3 truncations have the mildest effects by both criteria, whereas the SmD1 truncation and the two double mutants have severe effects. The SmD1 truncation strain reproducibly accumulates more pre–U3 snoRNA than other strains (Figures 2C and 2D); we have no salient explanation for this observation. However, the overall results suggest that the Sm tails contribute directly to shared or overlapping functions of U snRNPs during pre-mRNA splicing.

Contacts of Tails with Pre-mRNA Contribute to Commitment Complex Stability

To test directly for effects on an individual snRNP, we examined in vitro commitment complex formation using snRNA/pre-mRNA psoralen cross-linking (see Experimental Procedures for details). Consistent with the robust protein-pre-mRNA cross-links (Figure 1B), the tail truncations had little effect on RNA–RNA cross-linking efficiency with a wild-type capped pre-mRNA substrate (Figure 3A, upper panel). The different extracts also formed different levels of commitment complex by native gel electrophoresis (data not shown), but the effects were modest and not easily attributed to the tail truncations. Because the pre-mRNA cap has been previously shown to contribute to commitment complex formation and/or stability, we also used an uncapped substrate for psoralen cross-linking (Izaurralde et al., 1994; Colot et al., 1996; Lewis et al., 1996a, 1996b). The strategy was to present a less favorable substrate (Puig et al., 1999). Whereas commitment complex formation as assayed by psoralen cross-linking is almost insensitive to pre-mRNA cap status in the wild-type extract, very little complex formation with uncapped pre-mRNA occurs in all Sm tail-truncated extracts (Figure 3A, lower panel). The results suggest that the Sm tails make a contribution to U1 snRNP function during commitment complex formation, which is revealed in the absence of the pre-

is shown at the top of each lane. After immunoprecipitation, total RNA was extracted with phenol, precipitated with ethanol, and used for primer extension. Bands corresponding to U1 and U2 snRNAs are indicated by arrows. Wild-type (WT) non-tagged extract was used as a negative control (lanes 1 and 2).

(D) Cross-linking profiles of yeast commitment complex proteins to pre-mRNA using extracts from wild-type, single, and double Sm tail truncations. Cross-linking was performed as described previously (Zhang and Rosbash, 1999). Briefly, ^{32}P-radiolabeled and 4-thioU-substituted WT-72, a 5′ splice site containing RNA, were incubated in yeast splicing extract for 30 min at 25°C and then irradiated with 365 nm UV light for 5 min at 4°C. The commitment complex was immunoprecipitated using an antibody against Prp40, a U1 snRNP–specific protein. After immunoprecipitation, RNase A/T1 was used to degrade RNA, and labeled proteins were visualized by 4%–20% linear gradient SDS-PAGE and autoradiography. The migration of a molecular weight standard is indicated on the left. The protein concentration of the different extracts varied by less than 30% by Bradford assay. All eight cross-linked proteins, indicated by arrows on the right, have been identified previously (Zhang and Rosbash, 1999). The location of proteins missing in the cross-linking profile is marked by asterisks.

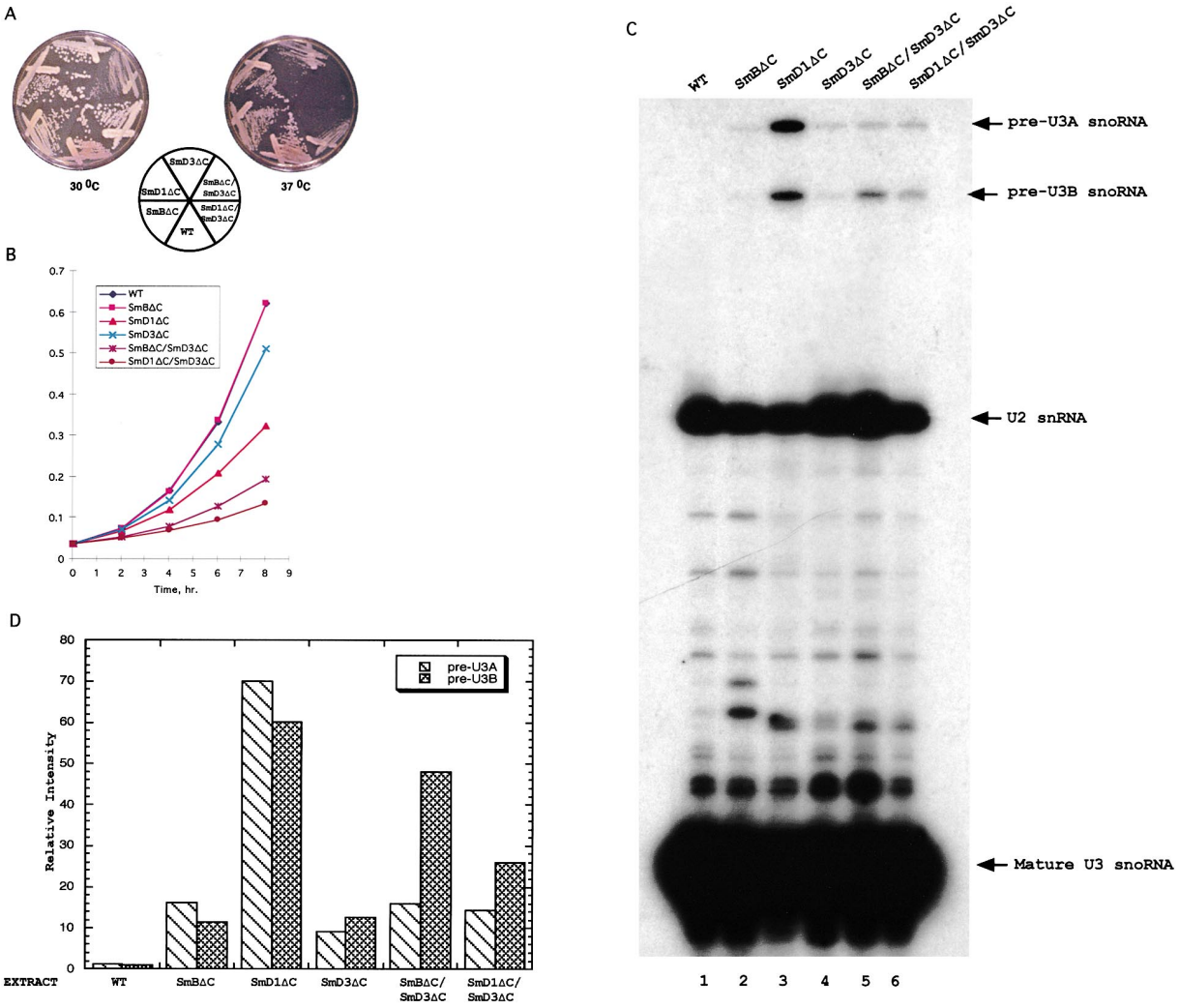


Figure 2. Growth and In Vivo Splicing Defects of Sm Tail Truncation Mutants

(A) Growth of Sm truncation strains at 30°C and 37°C. Single Sm truncation strains show no temperature-sensitive phenotype. SmBΔC and SmD1ΔC are synthetically lethal, and the SmBΔC/SmD3ΔC and SmD1ΔC/SmD3ΔC double mutants show different degrees of temperature-sensitive growth at 37°C.

(B) Growth rate in rich media at 30°C. The doubling time for individual strains at 30°C is 1.9 hr for wild type and SmBΔC, 2.2 hr for SmD3ΔC, 2.7 hr for SmD1ΔC, 3.2 hr for SmBΔC/SmD3ΔC, and 4.1 hr for SmD1ΔC/SmD3ΔC.

(C) Primer extension to determine precursor U3 snoRNA accumulation in Sm truncation strains. Primer extension was performed using two oligonucleotides: DT1967 (CCAAGTTGGATTCAGTGGCTC), specific for U3 snoRNA, and a U2-specific oligonucleotide DT58 (GCCAAAAAATGT GTATTGTAA) as an internal control. Primer extension reactions were resolved by 5% denaturing PAGE and bands were visualized by autoradiography. The positions of pre–U3A and pre–U3B snoRNA, and U2 snRNA and mature U3 snoRNA are indicated by arrows.

(D) Quantitation of pre–U3A and pre–U3B snoRNA in (C), normalized against U2 snRNA.

mRNA cap. They also suggest that the pre-mRNA cap and the Sm tails both contribute to the same aspect of commitment complex formation.

To further test for an Sm tail effect on U1 snRNP function, we compared extracts in a simple immunoprecipitation (IP) assay, using an antibody against a U1 snRNP–specific protein to precipitate radioactive pre-mRNA without cross-linking (Figure 3B). In this assay, the amount of U1 snRNP–associated pre-mRNA was substantially decreased in all truncation strain extracts. The data were normalized to the commitment complex formation capacity of each extract, as assayed by native gel electrophoresis (data not shown). The severity of the effects was similar, although not identical, to that observed in the two in vivo assays (Figure 2): the mildest effect was in the SmBΔC strain, and the most severe effects were in the two viable double mutant strains. These same truncation extracts, however, had little or no effect by psoralen cross-linking with capped pre-mRNA (Figure 3A, upper panel), indicating a difference between the two assays. This is particularly striking in the case of the SmBΔC/SmD3ΔC double truncation extract. One simple explanation is that the C-terminal tails contribute to complex stability and dissociate during the relatively stringent IP washes, namely, the complexes are less stable in the truncation extracts. Pso-

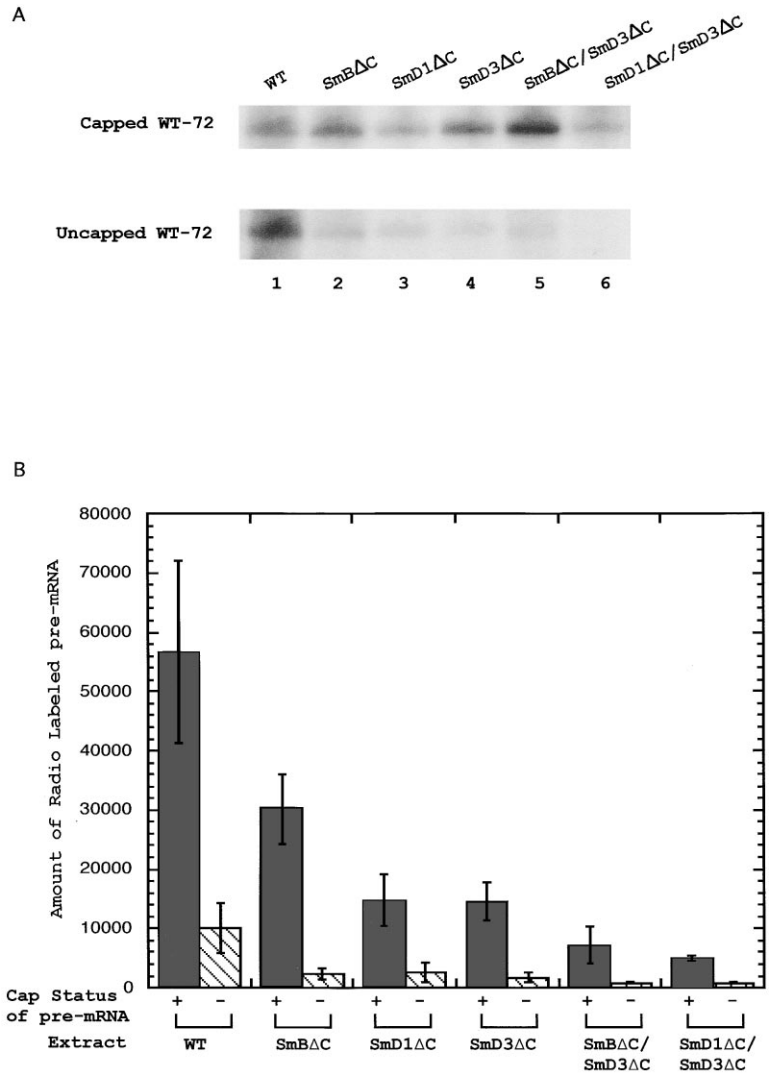


Figure 3. Commitment Complex Formation in Sm Truncation Extracts

(A) Psoralen cross-linking assay. 2×10^5 cpm of ^{32}P-radiolabeled WT-72, either capped (upper panel) or uncapped (lower panel), was incubated with 4 μl of splicing extract in a 10 μl standard commitment complex reaction as described previously (Abovich et al., 1990). After incubation, psoralen was added to the reaction and the mixture was irradiated for 10 min at 4°C to cross-link double-stranded RNA formed between U1 snRNA and WT-72. Samples were deproteinized by proteinase K digestion and phenol extraction. The RNA was ethanol precipitated, resolved by 6% denaturing PAGE, and bands were visualized by autoradiography. The intensity of the band indicates the amount of WT-72 that is cross-linked to U1 snRNA, which is a reflection of the amount of commitment complex formed.

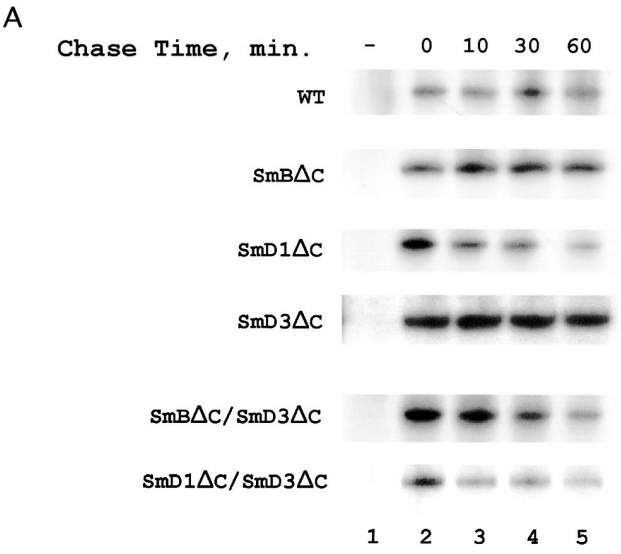
(B) Immunoprecipitation assay for commitment complex stability. 10^6 cpm of ^{32}P-radiolabeled WT-72, either capped (+) or uncapped (−), was incubated in 20 μl of splicing extract in a 50 μl standard commitment complex formation reaction for 20 min at 25°C. Complexes were immunoprecipitated with an anti-Prp40 antibody. Immunoprecipitated WT-72 was quantitated with a scintillation counter as a reflection of the stable commitment complex formed. The amount of active U1 snRNP in the different extracts was examined by native gel electrophoresis (data not shown) and used for normalization. The variability in U1 snRNP concentration was no more than 30%.

ralen as well as 4-thioU cross-linking (Figure 3A), however, traps these less stable complexes. Consistent with this interpretation, uncapped pre-mRNA gave rise to a much weaker signal in the IP assay, even in the wild-type extract. In all of the deletion extracts, near background levels of uncapped pre-mRNA were recovered (Figure 3B).

To directly address commitment complex stability, we performed chase experiments with the psoralen cross-linking assay. After commitment complex formation, a large excess of cold pre-mRNA was added and the incubation continued prior to addition of psoralen (Figure 4A). In the wild-type extract, there is very little change in the amount of complex after a 60 min incubation with excess cold pre-mRNA, consistent with the previously reported stability of the commitment complex (Legrain et al., 1988; Puig et al., 1999). Similar stabilities were observed in extracts from the SmBΔC and SmD3ΔC strains. However, the complexes were clearly less stable in the SmD1ΔC strain and in the two double mutant strains (Figures 4A and 4B). Taken together with the genetic interactions between the three truncation strains, the in vitro results suggest that the C-terminal tails of SmB, SmD1, and SmD3 contribute to commitment complex stability.

Replacement of the SmD3 Tail with Other RNA Binding Modules and Chemical Mapping

The large sequence differences between the yeast and mammalian C-terminal tails (Figure 1) suggest that they might function in a relatively sequence-independent manner. To address this possibility, the SmD3 C-terminal tail was replaced with two different RNA binding domains: a basic peptide derived from the HIV-1 nucleocapsid protein (SmD3ΔC-NC), or an arginine–serine peptide of 14 amino acids (SmD3ΔC-RS7). This strategy and these same two peptides have been previously used to influence the affinity of a yeast splicing protein for RNA (Berglund et al., 1998). These two fusion proteins partially rescued the slow growth phenotype of SmBΔC/SmD3ΔC at 30°C (Figure 5A). Although these strains were still temperature sensitive at 37°C (data not shown), the improved growth indicates that these novel tails contribute to Sm function. We also made splicing extracts from these strains and used them in commitment complex immunoprecipitation assays (Figure 5B). Strikingly, addition of either the NC or the RS7 tail to the SmD3ΔC protein rescued about 70% of the activity present in the SmBΔC single truncation strain (Figure 5B).



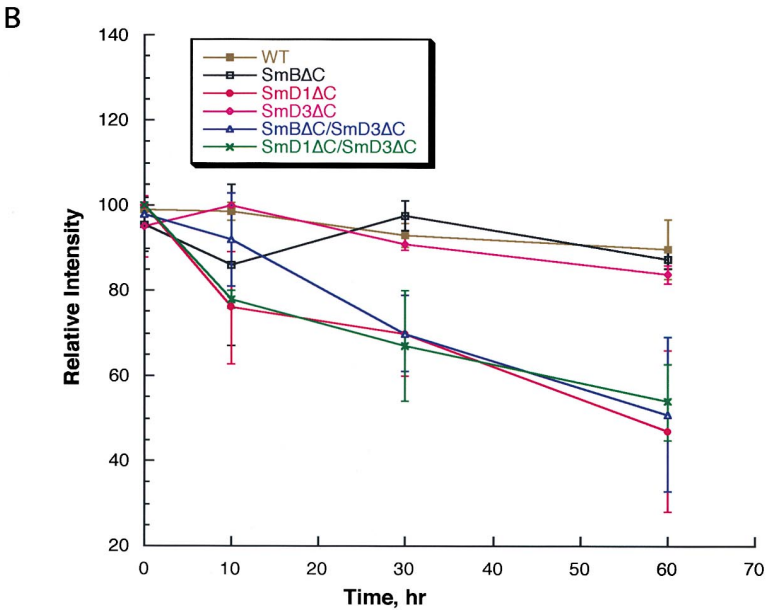


Figure 4. Commitment Complex Stability of Sm Tail Truncation Extracts Assayed by Psoralen Cross-Linking

(A) 2×10^5 cpm of ^{32}P-radiolabeled wild-type WT-72 was incubated with 4 μl of splicing extract in a 10 μl standard commitment complex reaction for 20 min at 25°C. A 500-fold molar excess of nonradioactive WT-72 was added and the incubation continued at 25°C for various times prior to psoralen addition (lanes 2–5). In lane 1, indicated by "−," excess cold WT-72 was added at the same time as the ^{32}P-radiolabeled WT-72.

(B) Quantitation of (A). Relative intensity of each chase time point to the zero time point from two independent experiments is plotted against time.

The SmD3ΔC-NC protein also restored direct cross-linking to the pre-mRNA substrate, suggesting that the pre-mRNA Sm contacts take place through the C-terminal basic tails (Figure 5C, lane 3). The SmD3ΔC-RS7 protein, however, did not restore detectable cross-linking (Figure 5C, lane 4). As this protein was equally effective in the two other assays (Figures 5A and 5B), it is likely that this failure is due to some subtle feature of this tail, such as the lack of appropriate residues in a proper orientation with the pre-mRNA.

To directly address the region of SmD3 that makes direct contact with the pre-mRNA, we used chemical cleavage by NCS (N-chlorosuccinimide) and cyanogen bromide (CNBr), which cleave only after tryptophan and methionine, respectively (Figure 6). Since there are no tryptophans in the wild-type SmD3 protein, we engineered one between Leu-83 and Phe-84, which is at the beginning of the predicted SmD3 tail (Figure 1A). This strain and U1 snRNP were unaffected by this single amino acid insertion (Figure 6 and data not shown). After NCS cleavage, there should be only a 9.1 kDa N-terminal fragment and a 3.1 kDa C-terminal fragment. After cross-linking, purification, and NCS cleavage only one band was visible, at approximately 5 kDa. This band is much smaller than 9 kDa, and therefore must be derived from the smaller C-terminal tail fragment. Moreover, we have shown previously that SmD3 binds primarily at or very close to the –2 U of the pre-mRNA 5′ exon. As the nucleotide sequence at the 5′ exon and 5′ splice site junction is AC^{-7} $AAAAU^{-2}G^{-1}$/GUAUGU (the 5′ splice site is underlined; Zhang and Rosbash, 1999), RNase A and T1 digestion should liberate the fragment $AAAAU^{-2}G^{-1}$, with a size of about 2 kDa. Subtracting this from the size of the radiolabeled peptide yields a predicted size of about 3 kDa, which matches the predicted size of the C-terminal tail very well (Figure 6, top).

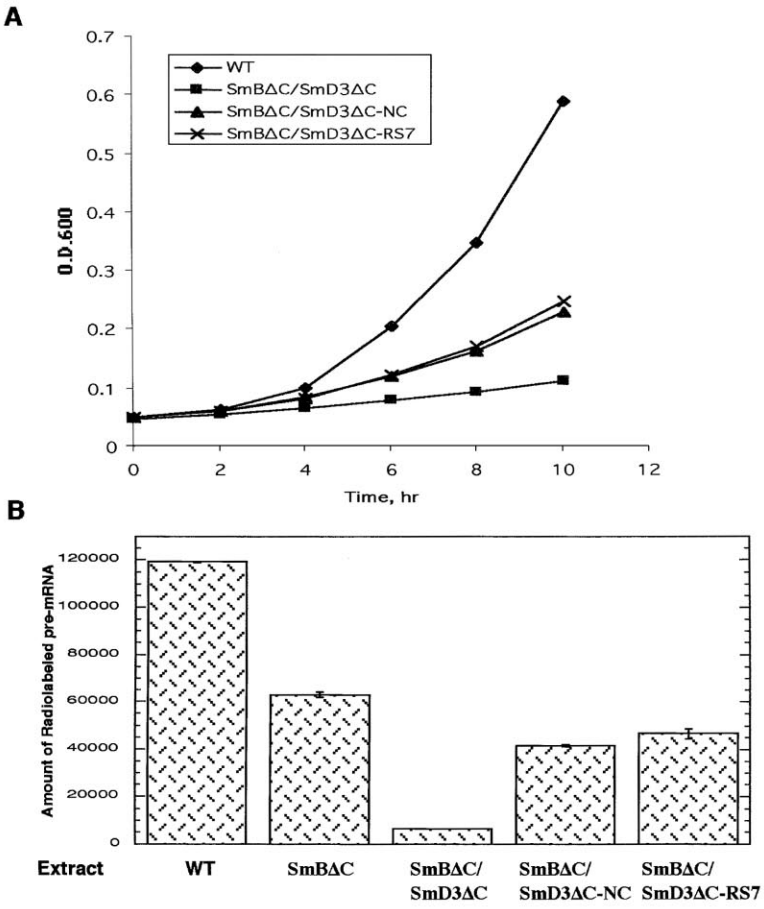


Figure 5. Partial Rescue of SmD3 Tail Truncation Phenotypes by Substitution of the SmD3 Tail with Either NC or RS7 RNA Binding Modules

The sequences of the NC and RS tails are VKGGRAPRKKGGERQANFLGKIWPSYKGR and RSRSRSRSRSRSRS, respectively.

(A) Growth rates of wild-type (WT), SmBΔC/SmD3ΔC, SmBΔC/SmD3ΔC-NC, and SmBΔC/SmD3ΔC-RS7 strains in Trp^-/Ura^- medium at 30°C.

(B) Immunoprecipitation assay for commitment complex stability as described in Figure 3B.

(C) 4-thioU cross-linking assays as described in Figure 1D using tail-substituted extracts. Wild-type and NC-substituted SmD3 are indicated by arrows on the left.

We independently cleaved wild-type SmD3 and SmD3ΔC-NC with CNBr. There are a total of five methionines in the wild-type SmD3, and the predicted size for the C-terminal fragment of wild-type SmD3 is 1.1 kDa. Since there are no methionines in the NC tail and since the NC tail is twice as big as the wild-type SmD3 tail, the predicted size for its C-terminal fragment is 6.2 kDa (Figure 6, top). The full-length SmD3ΔC-NC is about 1 kDa larger than the wild-type SmD3, consistent with the mobility difference between the two intact, cross-linked proteins (Figure 6, compare lanes 3 and 5). As we also observed a mobility shift for the two cleaved fragments (Figure 6, compare lanes 4 and 6), they cannot be the N-terminal two fragments of SmD3 and SmD3ΔC-NC, which are exactly the same. For SmD3ΔC-NC, this leaves only the 6.2 kDa C-terminal fragment. Subtracting

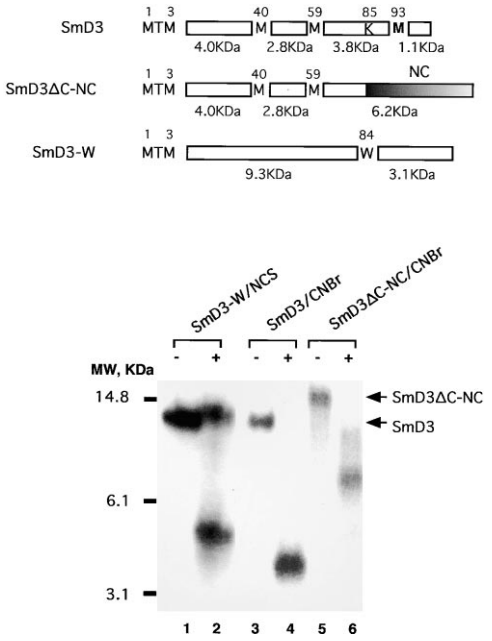


Figure 6. Chemical Cleavage Mapping of the 4-thioU Cross-Linking Region of SmD3 to WT-72

For details of the mapping protocol, see Experimental Procedures. Briefly, 4-thioU cross-linked commitment complex proteins were purified by 18% SDS-PAGE and the band corresponding to SmD3 was excised. 2000 cpm of gel-purified SmD3 was chemically cleaved and the cleaved products were visualized by 20% Tris-tricine-PAGE and autoradiography. The top panel shows a schematic diagram of wild-type SmD3, SmD3ΔC-NC, and SmD3-W, in which a tryptophan has been introduced. The shaded region in SmD3ΔC-NC indicates the NC tail. SmD3-W has a FLAG tag (DYDDDDK) at its C terminus and is approximately 1 kDa larger than the wild-type SmD3. The positions of methionine (M) or tryptophan (W) are shown with the amino acid numbered, as is the size of each predicted fragment after chemical cleavage. The cleavage pattern by either N-chlorosuccinimide (NCS), which cleaves after tryptophan, or cyanogen bromide (CNBr), which cleaves after methionine, is shown in the bottom panel. Lanes marked "+" and "−" are with and without chemical cleavage, respectively. Migration of a molecular weight standard is shown on the left, and the full-length proteins are indicated by arrows.

the predicted 2 kDa size of the cross-linked nucleic acid, the size of the SmD3 cleaved fragment would be about 1 kDa, in excellent agreement with its predicted C-terminal fragment. Taken together, these cleavage data indicate that the cross-linking site in SmD3 is between Met-92 and Arg-101 of its C-terminal tail.

In sum, these in vivo and in vitro assays strongly suggest that the tails of SmB, SmD1, and SmD3 contact pre-mRNA directly in a largely sequence-independent manner. On the other hand, the failure to rescue the growth phenotypes completely by SmD3ΔC-NC and SmD3-RS7 suggests that tail identity still plays some role in substrate recognition and/or snRNP structure and function.

Discussion

We have shown that removal of a single Sm tail has growth rate effects, suggesting that some feature of Sm complex function is compromised. There are also strong synthetic interactions when two tail truncations are combined, suggesting that all three tails contribute to overlapping functions. Of course, we do not know which Sm-dependent functions are sufficiently compromised to be rate limiting for growth. Although pre-mRNA splicing efficiency is depressed (Figure 2C), a nonsplicing snRNP may be the rate-limiting particle for growth. Suppressor and enhancer genetics might help identify the limiting snRNP(s).

The only proposed function for the Sm complex is in snRNP assembly and nuclear transport (Hamm et al., 1987, 1990; Fischer et al., 1993; Nelissen et al., 1994). Therefore, the tailless strains may be deficient in snRNP activity, due to insufficient quantities per cell or to the fact that much of some snRNP is in the wrong subcellular compartment. There is no evidence that yeast nuclear snRNPs ever exist in the cytoplasm (see below). However, R. Bordonné recently reported that the basic tails of both yeast SmB and SmD1, but not that of SmD3, have nuclear localization functions, which might serve to import the Sm complex or subcomplexes (Bordonné, 2000). Although U1 snRNP levels are normal or near normal in these strains, some snRNPs might still be present in limiting quantities. Our data suggest that splicing snRNP activity is also deficient in the truncation strains.

For U1 snRNP, several criteria indicate that the Sm tails contribute directly to commitment complex stability. An effect on stability, and perhaps also on formation, fits with the notion that basic residues within the three Sm tails make direct contact with the pre-mRNA within the 5′ splice site region (Zhang and Rosbash, 1999; Figure 1D). We imagine that these contacts complement the other RNA–RNA and RNA–protein contacts between the snRNP and protein factors on the one hand and the pre-mRNA substrate on the other. A much different view is that the tails affect some other region of the Sm protein and therefore interact only indirectly with the pre-mRNA substrate. However, addition of the NC module to the SmD3ΔC protein rescues cross-linking, and we are able to map the cross-linking site explicitly to the SmD3 tail region. This makes the less straightforward possibility unlikely, at least for the SmD3 protein. A similar argument posits that the tail truncation effects on complex stability are more complicated: U1 snRNP structure might be subtly altered by the tail deletions, which would affect complex stability in a more indirect manner. Although this possibility is difficult to exclude (and is not exclusive with the direct contact stabilization hypothesis), there are two indications that tail truncation snRNPs are otherwise normal or nearly so: the relative cross-linking signal intensities of the other proteins are unaffected by the tail truncations (Figure 1D), and the contact regions of these proteins on the pre-mRNA are unchanged from the wild-type snRNP (data not shown). Moreover, addition of the NC and RS7 RNA binding modules to the SmD3ΔC protein improved complex formation substantially. Although the effects of the modules on formation could be through complicated and subtle modulations of U1 snRNP structure, the unrelated nature of these modules to the natural SmD3 C-terminal tail sequence makes this possibility unlikely. We therefore favor the simple view that the Sm tails make direct

contact with the pre-mRNA, and that this helps stabilize the U1 snRNP-pre-mRNA complex. We note the importance of structural information, which will be required to extend this proposal of nucleic acid binding. As the Sm tails were disordered in the crystal structures (Kambach et al., 1999), substrate RNA or DNA might be required for a single, stable conformation.

Our previous results indicated the importance of multiple protein–RNA interactions to complex formation. For both the yeast and mammalian branchpoint sequence binding proteins (BBP and mBBP/SF1), accessory RNA binding modules can increase RNA binding affinity of a protein, and of a protein complex, for an RNA substrate without necessarily providing any sequence specificity (Berglund et al., 1997, 1998). These previous conclusions were based on in vitro binding experiments comparing a synthetic fusion protein with the wild-type protein. In this study, the fusion proteins were used to generate in vivo snRNPs that had replaced the wild-type SmD3 C-terminal tail with synthetic tails. They rescued an in vivo growth rate phenotype as well as in vitro phenotypes of the tail truncation snRNPs.

Although the nature of the Sm tail-pre-mRNA interactions is not known, it is possible that the basic residues of the tails contact the RNA phosphate backbone. Since the cross-linking is to residues very close to and even within the 5′ splice site (Zhang and Rosbash, 1999), the basic residues may even help stabilize the RNA–RNA interaction between the 5′ end of U1 snRNA and the 5′ splice site region. This is analogous to the proposed role of the U2AF65 RS domain in base pairing between U2 snRNA and the branchpoint region (Valcárcel et al., 1996).

In metazoa, the C-terminal tails of SmB, SmD1, and SmD3 may have additional functions. They are not only basic but also have two features not shared by their yeast counterparts. First, they all contain multiple (G)RG repeats, which have been suggested to be a key epitope of anti-Sm autoantibodies (Hirakata et al., 1993). Second, the C terminus of SmB contains impressive proline-rich regions that might be involved in protein–protein interactions. Both sequence features might contribute to another biological process that does not occur in yeast. Nuclear import of U snRNPs is a likely candidate. As mentioned above, there are no data in yeast systems that support a cytoplasmic phase for yeast nuclear snRNPs, and there is no obvious snurportin candidate in the yeast database (Huber et al., 1998). Another possibility is the recruitment of additional metazoan proteins during the assembly of Sm-containing snRNPs. As preliminary cross-linking data indicate that there are Sm protein-pre-mRNA contacts in the mammalian E complex (H. Du, personal communication), we suggest that metazoan Sm tails also share functional characteristics with yeast tails.

If the results for U1 snRNP are of general significance, then the Sm tails in other snRNPs will also increase affinity for their substrates. In the case of the other splicing snRNPs with the canonical set of Sm proteins (U2, U4, and U5 snRNPs), we predict that the affinity will be increased, either for different regions of the pre-mRNA substrate or for other snRNAs, such as U6 snRNA in the case of U4 snRNP tails. Consistent with these predictions, in vitro splicing is very inefficient in extracts from some of the tailless strains. Moreover, the effects are distributed among the strains differently from the effects on U1 snRNP (data not shown), suggesting that the activity of at least one other splicing snRNP is decreased and is rate limiting for in vitro splicing efficiency. It is possible that even the budding yeast telomerase might use the tails of these three Sm snRNPs for a similar purpose, to increase affinity for the DNA substrate or to aid function as a processive enzyme.

We suggest that the proposed stabilization functions might have evolved to aid short intermolecular base pairing, which is too unstable without protein assistance. The Sm complex might even have contributed to a putative transition from long to short intermolecular duplexes, thereby providing more flexibility for the processing machinery to select a broader range of substrates.

Experimental Procedures

Plasmid Constructions

Plasmids were constructed using standard techniques. All Sm expression plasmids are under the control of a galactose-inducible promoter(s), with the exception of pDZ4.

pDZ4 (wild-type SmB): a PCR product of yeast SmB genomic DNA (−503 to +816) was cloned into XbaI/XhoI of pRS316 (New England Biolabs). pDZ1 (SmBΔC): the nucleotide sequence encoding amino acids 1–106 of SmB was cloned into BamHI/SalI of pESC-TRP (Stratagene). pDZ7 (SmD1ΔC): the nucleotide sequence encoding amino acids 1–125 of SmD1 was cloned into EcoRI/ClaI of pESC-TRP. pDZ8 (SmD3ΔC): the nucleotide sequence encoding amino acids 1–85 of SmD3 was cloned into EcoRI/ClaI of pESC-TRP. pDZ3 (SmBΔC/SmD3ΔC): the nucleotide sequence encoding amino acids 1–85 of SmD3 was cloned into EcoRI/ClaI of pDZ1. pDZ12 (SmD1ΔC/SmD3ΔC): the nucleotide sequences encoding amino acids 1–125 of yeast SmD1 and 1–85 of yeast SmD3 were cloned into BamHI/SalI and EcoRI/ClaI of pESC-HIS (Stratagene), respectively. pSmD3ΔC-NC and pSmD3ΔC-RS7: a cloning strategy similar to that previously reported (Berglund et al., 1998) was used to fuse an NC or RS7 tail to the C terminus of SmD3ΔC, except that the coding sequences of the fusion proteins were cloned into BamHI/HindIII of pESC-URA (Stratagene). pDZ20 (SmD3-W): SmD3 was PCR amplified using an oligonucleotide that inserts a tryptophan between Leu-83 and Phe-84 of SmD3, and the PCR product was cloned into EcoRI/ClaI of pESC-TRP.

The pESC vectors contain a stretch of nucleotides that encode a c-myc epitope after SalI or a FLAG epitope after ClaI. Proteins cloned into BamHI/SalI have a c-myc tag at their C termini, and those cloned into EcoRI/ClaI have a FLAG tag at their C termini.

Yeast Strain Constructions

The wild-type strain used in this study is MGD353-13D (*MATa trp1-289 ura3-52 arg4 leu2-3, -112 ade2*). DZY2 (SmB): one copy of the wild-type gene encoding SmB in a diploid was replaced by a kanamycin resistance cassette (KAN^r). The strain was transformed with pDZ4 and sporulated on potassium acetate (KAc) media. Dissected haploid spores were tested for kanamycin resistance and growth on Ura$^-$ medium to select for DZY2 (*MATa trp1-289 leu2-3, -112 ura-3-52 smb::KAN* + pDZ4). DZY6 (SmBΔC): DZY2 was transformed with pDZ1 and grown on 5-FOA plates to force the loss of pDZ4. DZY19 (SmD1ΔC): BRY202 (*MATα trp1-289 ura3-52 leu2-3, -112 his3-Δ1, smd1::LEU2* + pBM150 [*URA3 GAL::SMD1*]), kindly provided by B.C. Rymond, was transformed with pDZ7 and grown on 5-FOA plates to force the loss of pBM150. DZY20 (SmD3ΔC): JWY2445 (*MATa his1 leu2-1 trp1-Δ101 ura3-52 smd3Δ1::LEU2* + pYCp50-*SMD3*), also kindly provided by B. C. Rymond, was transformed with pDZ8 and grown on 5-FOA plates to force the loss of pYCp50-*SMD3*. DZY21 (SmBΔC/SmD3ΔC): DZY2 was mated with JWY2445. After growth on 5-FOA plates, cells were transformed with pDZ3, sporulated, and spores tested for kanamycin resistance

and growth on Leu$^-$ and Trp$^-$ media to select for DZY21 (*MATa trp1 ura3-52 leu2 smb::KAN smd3Δ1::LEU2* + pDZ3). DZY11-A: GRY24104 (*MATa/α ura3D1 leu2D0 his3D1 smd3::KAN*; Research Genetics) was transformed with pYCp50-SMD3, sporulated, and dissected haploid spores tested for kanamycin resistance and growth on Ura$^-$ dropout plates to select for DZY11-A (*MATa ura3D1 leu2D0 his3D1 smd3::KAN* + pYCp50-SMD3). DZY12 was generated by mating DZY11-A with BRY202 and subsequent growth on 5-FOA plates. DZY13 (SmD1ΔC/SmD3ΔC): DZY12 was transformed with pDZ12, sporulated, and dissected haploid spores tested for kanamycin resistance and growth on Leu$^-$ and His$^-$ dropout media to select for DZY13 (*MATα leu2 ura3 his3 smd1::LEU, smd3::KAN* + pDZ12). SmBΔC/SmD3ΔC-NC and SmBΔC/SmD3ΔC-RS7: DZY21 was transformed with pSmBΔC/SmD3ΔC-NC or pSmBΔC/SmD3ΔC-RS7, and transformants selected on a Trp$^-$/Ura$^-$ dropout medium. DZY23 (SmD3-W): JWY2445 was transformed with pDZ20 and transformants plated on 5-FOA media to force the loss of pYCp50-SMD3 to select for DZY23 (*MATa his1 leu2-1 trp1-Δ101 ura3-52 smd3Δ1::LEU2* + pDZ20).

Psoralen Cross-Linking Assay

2×10^5 of ^{32}P-radiolabeled WT-72 (Zhang and Rosbash, 1999) was added to 10 μl of a standard commitment complex reaction (Abovich et al., 1990). After incubation for 20 min at 25°C, 1.1 μl of 62.5 μg/ml AMT-psoralen (HRI associates) was added and the samples were irradiated with 365 nm UV light for 10 min at 4°C. 50 μl of proteinase K buffer (50 mM Tris-HCl [pH 7.5], 12 mM EDTA, and 1% SDS) and 2.5 μl of 20 μg/μl proteinase K were added, and the incubation was continued for 20 min at 50°C. After phenol extraction, 60 μl of supernatant was removed and the RNA was ethanol precipitated prior to denaturing PAGE.

UV Cross-Linking of Commitment Complex and Immunoprecipitation Assay

All yeast strains were grown in a galactose-containing medium. Preparation of splicing extracts was performed as described previously (Abovich et al., 1990). Protein UV cross-linking was performed as described previously (Zhang and Rosbash, 1999). For the immunoprecipitation assay, 10^6 cpm of ^{32}P-radiolabeled WT-72 was incubated with 4 μl of splicing extract in a 10 μl standard commitment complex reaction for 20 min at 25°C. Commitment complexes were immunoprecipitated with an anti-Prp40 antibody. Immunoprecipitated WT-72 was quantitated with a scintillation counter.

Chemical Cleavage Mapping of the Cross-Linked Region on SmD3

Protein UV cross-linking was performed as described previously (Zhang and Rosbash, 1999), except that 1 ml of splicing extract was used. After immunoprecipitation and ribonuclease A/T1 digestion, samples were purified by an 18% preparative SDS-PAGE (Bio-Rad), and the bands corresponding to SmD3 were excised. The protein was eluted with volatile buffer (0.1% SDS and 50 mM ammonium bicarbonate) using an Electro-Eluter (Bio-Rad), and lyophilized. For cleavage of methionine with cyanogen bromide (Sigma), lyophilized proteins were resuspended in 50 μl of 70% formic acid, and 5 μl of 400 mg/ml CNBr was added (Reyes et al., 1999). The mixture was incubated at 25°C overnight. Samples were precipitated by adding 10 μl of 100 mM Tris (pH 8.3), 50 μg of BSA, and 800 μl of ethanol, and incubated on dry ice for 1.5 hr. Pellets were resuspended in 20 μl of SDS-PAGE loading buffer and the bands were visualized by 20% Tris-tricine-PAGE and autoradiography. For cleavage of tryptophan with N-chlorosuccinimide (Aldrich), a protocol was kindly provided by Dr. Ravinder Singh (University of Colorado) that is slightly modified from that previously reported (Mirfakhrai and Weiner, 1993).

Acknowledgments

We thank B. Rymond for kindly providing yeast strains (BRY202 and JWY2445), H. Du, T.H. Jensen, and K. Dower for helpful discussions, K. Dower, T.H. Jensen, I.W. Mattaj, R. Luhrmann, T.W. Nilsen, and B. Séraphin for excellent comments on the manuscript, and other members of Rosbash laboratory for suggestions and advice during the course of this work.

Received June 15, 2000; revised November 28, 2000.

References

Abovich, N., Legrain, P., and Rosbash, M. (1990). The yeast PRP6 gene encodes a U4/U6 small nuclear ribonucleoprotein particle (snRNP) protein, and the PRP9 gene encodes a protein required for U2 snRNP binding. Mol. Cell. Biol. *10*, 6417–6425.

Berglund, J.A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997). The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell *89*, 781–787.

Berglund, J.A., Fleming, M.L., and Rosbash, M. (1998). The KH domain of the branchpoint sequence binding protein determines specificity for the pre-mRNA branchpoint sequence. RNA *4*, 998–1006.

Bordonné, R. (2000). Functional characterization of nuclear localization signals in yeast Sm proteins. Mol. Cell. Biol. *20*, 7943–7954.

Colot, H.V., Stutz, F., and Rosbash, M. (1996). The yeast splicing factor MUD13 is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. Genes Dev. *10*, 1699–1708.

Fischer, U., Sumpter, V., Sekine, M., Satoh, T., and Luhrmann, R. (1993). Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. EMBO J. *12*, 573–583.

Grimm, C., Stefanovic, B., and Schümperli, D. (1993). The low abundance of U7 snRNA is partly determined by its Sm-binding site. EMBO J. *12*, 1229–1238.

Guthrie, C., and Patterson, B. (1988). Spliceosomal RNAs. Annu. Rev. Genet. *22*, 387–419.

Hamm, J., Kazmaier, M., and Mattaj, I.W. (1987). In vitro assembly of U1 snRNPs. EMBO J. *6*, 3479–3485.

Hamm, J., Darzynkiewicz, E., Tahara, S.M., and Mattaj, I.W. (1990). The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. Cell *62*, 569–577.

Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H., and Luhrmann, R. (1995). snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. EMBO J. *14*, 2076–2088.

Hirakata, M., Craft, J., and Hardin, J.A. (1993). Autoantigenic epitopes of the B and D polypeptides of the U1 snRNP. Analysis of domains recognized by Y12 monoclonal anti-Sm antibody and by patient sera. J. Immunol. *150*, 3592–3601.

Huber, J., Cronshagen, U., Kadokura, M., Marshallasay, C., Wada, T., Sekine, M., and Luhrmann, R. (1998). Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. EMBO J. *17*, 4114–4126.

Izaurralde, E., Lewis, J.D., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657–668.

Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V.A., Luhrmann, R., Li, J., and Nagai, K. (1999). Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell *96*, 375–387.

Legrain, P., Séraphin, B., and Rosbash, M. (1988). Early commitment of yeast pre-mRNA to the spliceosome pathway. Mol. Cell. Biol. *8*, 3755–3760.

Lehmeier, T., Raker, V., Hermann, H., and Luhrmann, R. (1994). cDNA cloning of the Sm proteins D2 and D3 from human small nuclear ribonucleoproteins: evidence for a direct D1-D2 interaction. Proc. Natl. Acad. Sci. USA *91*, 12317–12321.

Lewis, J.D., Gorlich, D., and Mattaj, I.W. (1996a). A yeast cap binding protein complex (yCBC) acts at an early step in pre-mRNA splicing. Nucleic Acids Res. *24*, 3332–3336.

Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I.W. (1996b). A nuclear cap-binding complex facilitates associa-

tion of U1 snRNP with the cap-proximal 5′ splice site. Genes Dev. *10*, 1683–1698.

Maroney, P.A., Hannon, G.J., Denker, J.A., and Nilsen, T.W. (1990). The nematode spliced leader RNA participates in trans-splicing as an Sm snRNP. EMBO J. *9*, 3667–3673.

Mattaj, I.W. (1986). Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. Cell *46*, 905–911.

Michaud, S., and Reed, R. (1991). An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. Genes Dev. *5*, 2534–2546.

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

Nelissen, R.L., Will, C.L., van Venrooij, W.J., and Luhrmann, R. (1994). The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. EMBO J. *13*, 4113–4125.

Puig, O., Gottschalk, A., Fabrizio, P., and Seraphin, B. (1999). Interaction of the U1 snRNP with nonconserved intronic sequences affects 5′ splice site selection. Genes Dev. *13*, 569–580.

Raker, V.A., Plessel, G., and Luhrmann, R. (1996). The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle in vitro. EMBO J. *15*, 2256–2269.

Reddy, R., and Busch, H. (1988). Small nuclear RNAs: RNA sequences, structure and modifications. In Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles, M.L. Birnstiel, ed. (Berlin: Springer-Verlag), pp. 1–37.

Reyes, J.L., Gustafson, E.H., Luo, H.R., Moore, M.J., and Konarska, M.M. (1999). The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5′ splice site. RNA *5*, 167–179.

Roy, J., Zheng, B., Rymond, B.C., and Woolford, J.L., Jr. (1995). Structurally related but functionally distinct yeast Sm D core small nuclear ribonucleoprotein particle proteins. Mol. Cell. Biol. *15*, 445–455.

Seto, A., Zaug, A., Sobel, S., Wolin, S., and Cech, T. (1999). *Saccharomyces cervisiae* telomerase is an Sm small nuclear ribonucleoprotein particle. Nature *401*, 177–178.

Séraphin, B., and Rosbash, M. (1989). Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. Cell *59*, 349–358.

Séraphin, B., and Rosbash, M. (1991). The yeast branchpoint sequence is not required for the formation of a stable U1 snRNP-pre-mRNA complex and is recognized in the absence of U2 snRNA. EMBO J. *10*, 1209–1216.

Séraphin, B., Kretzner, L., and Rosbash, M. (1988). A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5′ cleavage site. EMBO J. *7*, 2533–2538.

Siliciano, P.G., and Guthrie, C. (1988). 5′ splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev. *2*, 1258–1267.

Valcárcel, J., Gaur, R.K., Singh, R., and Green, M.R. (1996). Interaction of $U2AF^{65}$ RS region with pre-mRNA of branch point and promotion base pairing with U2 snRNA. Science *273*, 1706–1709.

Yu, Y.-T., Scharl, E.C., Smith, C., and Steitz, J.A. (1999). The growing world of small nuclear ribonucleoproteins. In The RNA World, R. Gesteland, T. Cech, and J. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 487–524.

Zhang, D., and Rosbash, M. (1999). Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. Genes Dev. *13*, 581–592.

Zhuang, Y., and Weiner, A.M. (1986). A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell *46*, 827–835.