Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells

Tetsuro Hirose and Joan A. Steitz*

Department of Molecular Biophysics and Biochemistry, Boyer Center for Molecular Medicine, Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, CT 06536

Contributed by Joan A. Steitz, September 18, 2001

In mammalian cells, all small nucleolar RNAs (snoRNAs) that guide rRNA modification are encoded within the introns of host genes. A database analysis of human box C/D snoRNAs revealed conservation of their intronic location, with a preference for 70–80 nt upstream of the 3′ splice site. Transfection experiments showed that synthesis of *gas5*-encoded U75 and U76 snoRNAs dropped significantly for mutant constructs possessing longer or shorter spacers between the snoRNA and the 3′ splice site. However, the position of the snoRNA did not affect splicing of the host intron. Substitution mutations within the spacer indicated that the length, but not the specific sequence, is important. A *in vitro* system that couples pre-mRNA splicing and processing of U75 has been developed. U75 synthesis *in vitro* depends on its box C and D sequences and requires an appropriate spacer length. Further mutational analyses both *in vivo* and *in vitro*, with subsequent mapping of the branch points, revealed that the critical distance is from the snoRNA coding region to the branch point, suggesting synergy between splicing and snoRNA release.

The nucleoli of eukaryotic cells contain large numbers of small nucleolar ribonucleoprotein (snoRNP) particles that function in the nucleolytic processing and nucleotide modification of precursor rRNAs. Currently, more than 150 small nucleolar RNA (snoRNA) species have been identified. SnoRNAs can be divided into two classes: those that possess boxes C (RUGAUGA) and D (CUGA), required for association with an abundant nucleolar protein fibrillarin, and those that possess boxes H (ANANNA) and ACA, which mediate the binding of Gar1 protein (reviewed in refs. 1 and 2). Box C/D snoRNAs and box H/ACA snoRNAs target specific sites in pre-rRNA for 2′-*O*-methylation and pseudouridylation, respectively (3–8). The methylation reaction is guided by an extensive region (10–21 nt) of complementarity between the box C/D snoRNA and rRNA sequences flanking the modification site (3, 5, 9–11).

Vertebrate snoRNAs are encoded within the introns of snoRNA host genes, which can be either protein coding or noncoding (reviewed in refs. 12 and 13). A common feature of mammalian snoRNA host gene transcripts is a 5′ terminal oligopyrimidine sequence, whose precise function with respect to snoRNA synthesis is unknown (14, 15). In most cases, it appears that snoRNAs are released from excised, debranched introns by exonucleolytic trimming (16–18); a minor pathway involves endonucleolytic cleavage of flanking intron sequences (19, 20). Many yeast snoRNAs are transcribed as monocistronic or polycistronic precursors from independent transcription units, and proteins involved in their processing have been characterized (21–23). On the other hand, no factors involved in snoRNA release have been identified in mammalian cells.

The processing of intronic snoRNAs is directed by elements residing within the snoRNA coding region. Exonucleolytic trimming and accumulation of vertebrate box C/D snoRNAs depend on the C and D boxes and an adjacent 4- to 5-nt helix that brings together the 5′ and 3′ ends of the snoRNA (24, 25). This terminal structure functions as a binding site for snoRNP proteins identified both in vertebrates and yeast: fibrillarin (Nop1p), Nop58p (Nop5p), Nop56p, and 15.5-kDa protein (Snu13p) (26–32). In mammalian cells, the catabolism of the excised intron is very rapid; introns linearized by debranching are degraded within a few seconds (33), suggesting that the assembly of snoRNP proteins onto the snoRNA-coding region occurs before debranching.

Here, we report that human box C/D snoRNAs are preferentially located at a conserved distance, about 70 nt upstream of the 3′ splice site of the host intron. We have developed an *in vitro* splicing-snoRNA processing system from HeLa nuclear extracts. Using both *in vivo* transfection and the *in vitro* system, we demonstrate that alterations in the distance between the snoRNA coding region and the branch point affect snoRNA release. Our results suggest the existence of molecular interactions between the splicing machinery and snoRNP proteins or snoRNP processing components during snoRNP biogenesis.

Materials and Methods

Plasmid Construction. A part of the mouse *gas5* gene containing its promoter region (−308 bp from the transcription start) and exons 1–5 (34) was amplified by PCR from mouse genomic DNA and cloned into the *Hin*dIII and *Sal*I sites of the pGEM-3Z plasmid. The 3′ untranslated region from the bovine growth hormone gene with its polyadenylation signal (Invitrogen) was inserted into the *Xba*I and *Kpn*I sites downstream of the *gas5* fragment. Site-directed mutagenesis was carried out by using the QuikChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. DNA sequences of all mutant constructs were verified by analyses performed at the Keck Facility at Yale University.

Transient Transfection. Human embryonic kidney cells (HEK293) were grown in DMEM (GIBCO/BRL) containing 10% FBS (GIBCO/BRL) in 35-mm dishes and transiently transfected with 1 μg plasmid by using Lipofectamine reagent (GIBCO/BRL) according to the manufacturer's directions. Twenty four hours after transfection, the growth medium was removed and the cells were washed once with PBS. Total RNA was extracted with Trizol (GIBCO/BRL) according to the manufacturer's directions.

Primer Extension and RNase Protection Assays. Primer extension was carried out according to Tycowski *et al.* (16). Briefly, primers U76PE (positions 47–67 in the mouse U76 snoRNA), U75PE (positions 43–62 in the mouse U75 snoRNA) (14, 34), and U15PE (positions 131–147 in human U15 snoRNA) (16) were 5′ end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP (6,000 Ci/mmol; NEN). The three labeled primers (1 pmol each) were then annealed at 95°C for 3 min with 5 μg total RNAs prepared from the transfected HEK293 cells, followed by cooling to 42°C. Reverse transcription (RT) was carried out by using

Abbreviations: snoRNA, small nucleolar RNA; snoRNP, small nucleolar ribonucleoprotein; RT, reverse transcription; WT, wild type.

*To whom reprint requests should be addressed. E-mail: joan.steitz@yale.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

www.pnas.org/cgi/doi/10.1073/pnas.231490998

12 units of avian myeloblastosis virus reverse transcriptase (Roche Molecular Biomedical) in RT buffer (100 mM Tris·HCl, pH 8.3/140 mM KCl/10 mM $MgCl_2$/0.025 mM EDTA/3.75 mM DTT/0.5 mM dNTPs) for 1 h at 42°C, and the products were separated by 8% PAGE containing 7 M urea and visualized by autoradiography.

RNase protection assays were performed by using standard procedures (35). The probes were internally ^{32}P-labeled antisense *gas5* RNAs extending either from the first nucleotide of exon 1 to 1 nt upstream of the U76 coding region excluding introns 1 and 2 (probe A) or from the first nucleotide of exon 2 to 1 nt upstream of the U75 coding region (probe B).

***In Vitro* Splicing-snoRNA Processing.** The splicing-snoRNA processing substrate contained exon 2, intron 2 (including U75), and exon 3 of the mouse *gas5* gene, transcribed *in vitro* with [α-^{32}P]UTP (3,000 Ci/mmol; Amersham Pharmacia) as described (35). Standard *in vitro* assays contained 12 μl of HeLa cell nuclear extract (60% final concentration) in buffer D (36) (4.8 mM $MgCl_2$/0.5 mM ATP/20 mM creatine phosphate/2% polyvinyl alcohol/10 fmol pre-mRNA substrate) (5×10^5 cpm) incubated for 4 h at 30°C. RNAs were treated with proteinase K, extracted with phenol/$CHCl_3$/isoamyl alcohol (50:49:1), and precipitated with ethanol. Samples were resolved on a 5% denaturing gel.

Branch Point Mapping. Branch point mapping was carried out according to Vogel *et al.* (37). cDNA synthesis was performed with Superscript II RNase H$^-$ reverse transcriptase (GIBCO/BRL), 1 pmol of RT primer (see Fig. 5*D*), and half the total RNA prepared from a 10-μl *in vitro* splicing reaction (data not shown). The subsequent PCR used the PCR1 and PCR2 primer pair to amplify the synthesized cDNA, followed by cloning into pGEM-3Z. Sequencing was performed at the Keck Facility at Yale University.

Results

The Spacing Between Human snoRNA-Coding Regions and 3′ Splice Sites Is Conserved. Initially, we noted that most snoRNAs encoded within two multiple snoRNA host genes in both mouse and human *gas5* and *UHG* (10, 14) are located ≈70 nt upstream of the 3′ splice site (data not shown). To investigate the generality of this observation, a database analysis of the location of known box C/D snoRNA-coding regions in the human genome was carried out. For many snoRNAs isolated by RT-PCR (3), the host genes had not been identified. Thus, genomic sequences surrounding snoRNAs were compared with cDNA sequences in the human expressed sequence tag database, allowing definition of snoRNA location within host introns. When combined with snoRNAs whose genomic organization had been established, the total number of box C/D snoRNAs analyzed was 57.

Fig. 1 shows a plot of the distribution of distances between human box C/D snoRNAs and 3′ splice sites (black bars), with the upstream spacer length (the length between the 5′ splice site and the snoRNA coding region) shown for comparison (gray bars). Strikingly, all box C/D snoRNAs but one are encoded more than 65 nt upstream of the 3′ splice site. The single exception is U16, which appears 56 nt upstream of the 3′ splice site of a minor class (U12-dependent) intron, and has been reported to be processed via endonucleolytic cleavage, as well as by a splicing-dependent pathway (20). The distribution peaks sharply between 71 and 80 nt for the downstream spacer, whereas no such peak is evident for the upstream spacer. This observation suggests that the positioning of box C/D snoRNAs relative to the 3′ splice site of the host intron facilitates efficient processing.

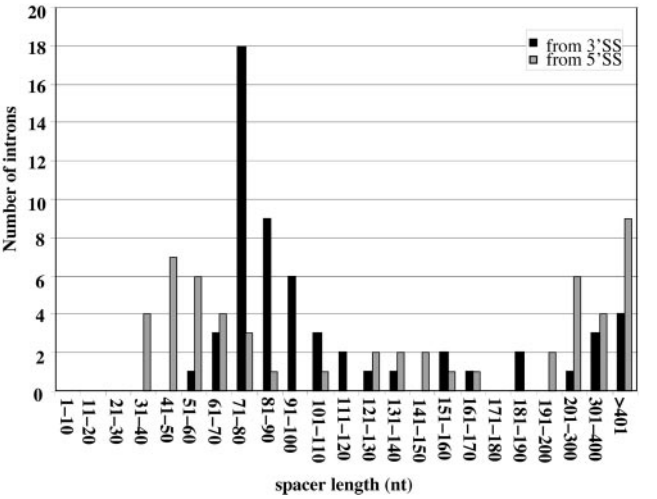

Fig. 1. Distribution of the lengths of spacer sequences for 57 human box C/D snoRNAs. The black and gray bars represent distances from the 5′ and 3′ splice sites, respectively.

Alterations of Downstream Spacer Length Decrease the Synthesis of U75 and U76 snoRNAs. Box C/D snoRNAs U75 and U76 are located 71 and 69 nt from the 3′ end of their respective introns in the mouse *gas5* gene (14). To investigate the significance of this distance for snoRNA synthesis, a series of length mutants of the downstream spacer were constructed (Fig. 2*A* and Fig. 6*A*, which is published as supporting information on the PNAS web site, www.pnas.org). *In vivo* analyses were carried out by transfecting a mouse *gas5* minigene, a genomic fragment containing the promoter region (from −308 bp) through exon 5 (Fig. 2*A*), into human embryonic kidney cells (HEK293). The production of mouse U75 and U76 snoRNAs was followed by primer extension analysis, whereas pre-mRNA splicing of the host introns (the second and third introns of *gas5*) was monitored by RNase protection.

Deletions within the downstream spacer proved to have the greatest impact on snoRNA synthesis. Specifically, whereas U76 was moderately decreased with a downstream spacer length of 65 nt (construct 3–65) [Fig. 2*B*; 29% of wild type (WT)], its appearance was nearly abolished with spacers of 60 nt or less (constructs 3–60, 3–55, 3–50, and 3–45 in Fig. 2*B*). Similarly, the synthesis of U75 from the second intron dropped to 6% with a 65-nt spacer and decreased further with spacers of 60 or 50 nt (see Fig. 6*B*). In contrast, levels of both U76 and U75 were only somewhat diminished when the downstream spacer was made longer. For U76, synthesis slowly tapered off (Fig. 2*C Left*) to 17% at 219 nt and was barely detectable at 379 nt (10%). A similar pattern was observed for U75 (see Fig. 6*C*). Meanwhile, the splicing of the host *gas5* introns was unperturbed and the splice site was unaltered by either deletions or insertions (Fig. 2 *B* and *C Right* and data not shown). We conclude that downstream spacer length strongly affects snoRNA processing but not splicing of snoRNA host introns.

The Length But Not the Specific Sequence of the Downstream Spacer Is Important for snoRNA Synthesis. We next asked whether a specific sequence within the downstream spacer or simply the length of the spacer is important for snoRNA synthesis. Additional mutants (Fig. 3*A*) were constructed and transfected into cultured cells. Whereas the results in Fig. 2*B* might have suggested that the 5 nt removed in 3–60 relative to 3–65 were important, mutant 3–65-2, in which these 5 nt were substituted (Fig. 3*A*), produced U76 at the same level as 3–65 (Fig. 3*C*). Likewise, U76 synthesis was abolished by deleting 9 nt in the

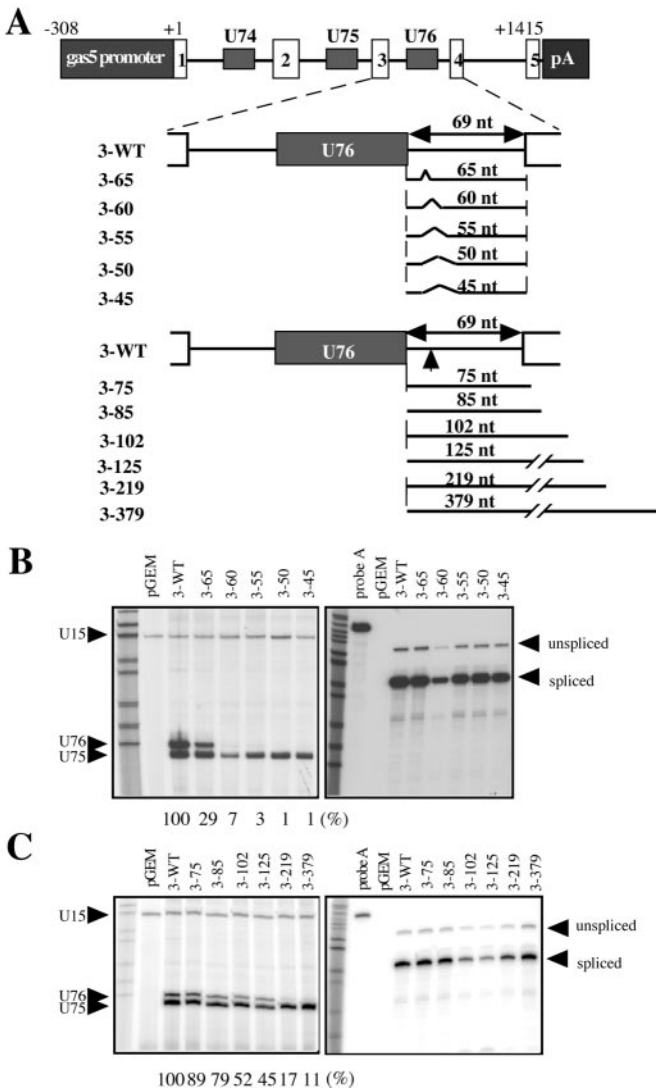

Fig. 2. Influence of downstream spacer length on the synthesis of U76 snoRNA. (*A*) Mutant constructs used for transfection analyses. All constructs were based on a mouse *gas5* minigene containing the *gas5* promoter region (gray box), exons 1–5 (open boxes with numbers), introns 1–4 (lines between the exons), snoRNA-coding regions for U74, U75, and U76 (boxes in the intron), and a polyadenylation signal (filled box with pA). Numbers above the gene structure are relative to the *gas5* transcription start site (nt). Deletion and insertion mutations made in the downstream spacer of the third intron are shown. (*B* and *C*) *gas5* minigenes with mutations in the downstream spacer of the third intron were transfected into HEK293 cells and U76, U75, and control endogenous U15 levels were assessed by primer extension (*Left*). Splicing of the third intron monitored by RNase protection is shown (*Right*). The intensities of the primer extension bands were quantitated by PhosphorImager (Molecular Dynamics) with the relative levels of U76 normalized to the internal control (U75).

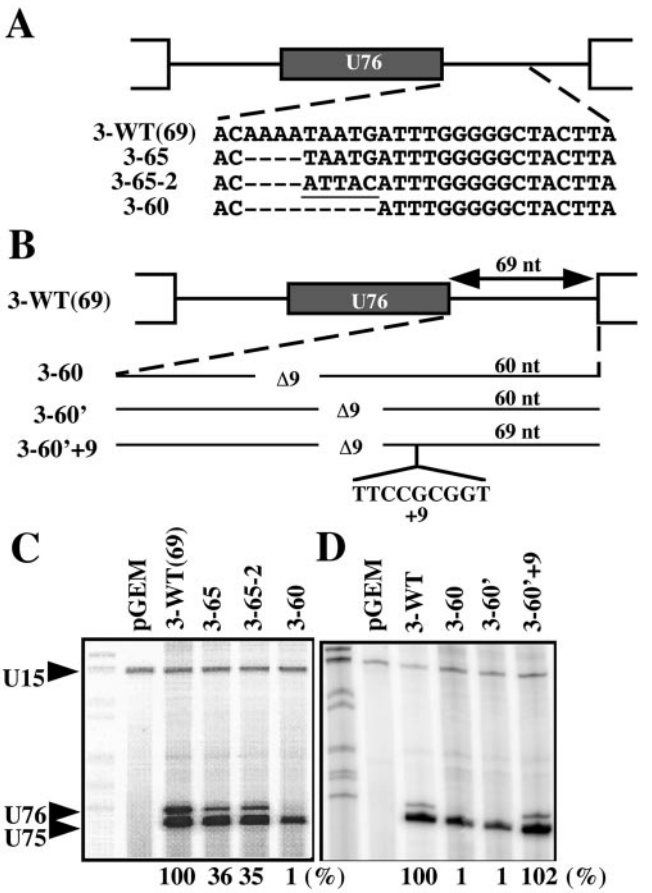

Fig. 3. Significance of the length but not specific sequence of the downstream spacer. (*A* and *B*) Mutant constructs used for transfection. Sequences of the region downstream of the U76 snoRNA in WT and mutants 3–65, 3–65-2, and 3–60 are shown. The substituted sequence in 3–65-2 is underlined (*A*). The 9-nt deletion (Δ9) in 3–60, 3–60′, and 3–60′+9, and the site of a 9-nt insertion (+9) in 3–60′+9 are shown. The total length of the downstream spacer in each construct is indicated above (*B*). (*C* and *D*) *gas5* minigene constructs were transfected and assayed as in Fig. 2.

downstream spacer (3–60′ in Fig. 3*B*) different from the 9 nt deleted in 3–60 (Fig. 3*D*). Conversely, the insertion of 9 unrelated nt just downstream of the deletion site (3–60′+9 in Fig. 3*B*) restored U76 synthesis (Fig. 3*D*). Together, these results indicate that U76 synthesis requires a downstream spacer longer than 65 nt, but suggest that no specific sequence is necessary.

An *in Vitro* System that Couples pre-mRNA Splicing and snoRNA Processing. Most previous *in vitro* studies analyzing the processing of mammalian snoRNAs have focused on the exonucleolytic trimming of linear intron substrates (16–18). To address the detailed mechanism of snoRNA production from the host intron, an *in vitro* system where both splicing and snoRNA processing occurs has been developed from HeLa nuclear extracts. Of three mouse *gas5* introns examined as substrates, only the second intron encoding U75 snoRNA was efficiently excised under standard *in vitro* splicing conditions. By adjusting the $MgCl_2$ concentration (to 4.8 mM) and adding 2% polyvinyl alcohol, we obtained detectable U75 after 3-h incubation (Fig. 4*A*). RNase H analysis confirmed that the band at 65 nt is the U75 snoRNA; it disappeared in the presence of an oligonucleotide complementary to U75 (oligo-U75), but not with an oligonucleotide complementary to the second exon of *gas5* (oligo-EX2) (see Fig. 7*A*, which is published as supporting information). The production of U75 in the coupled *in vitro* system was completely abolished if splicing of the host intron was inhibited by addition of either 2′-*O*-methyl oligonucleotide U2b, which prevents the interaction between U2 snRNA and the branch point (Fig. 4*B*) or of two other 2′-*O*-methyl oligonucleotides, which inhibit splicing at distinct steps (data not shown). These data indicate that the *in vitro* synthesis of U75 depends on the splicing of the host intron, consistent with *in vivo* observations (18).

Next, we investigated the importance of two *in vivo* requirements for box C/D snoRNA synthesis in our *in vitro* system. One

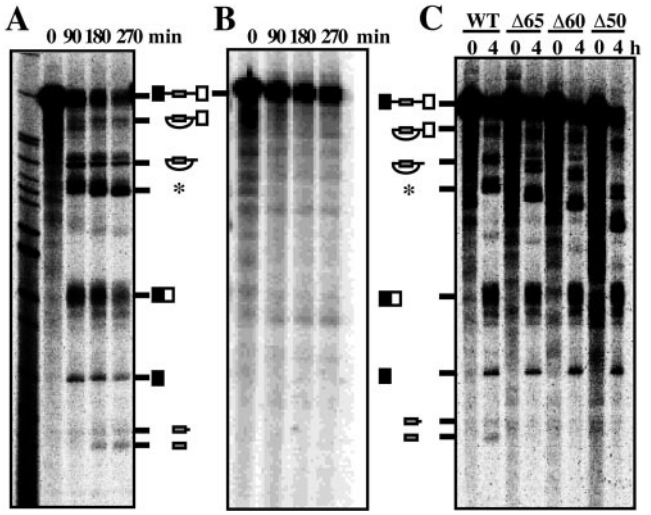

Fig. 4. Splicing and snoRNA processing *in vitro*. (*A*) Time course of the *in vitro* synthesis of U75 snoRNA from the mouse *gas5* intron 2 substrate. (*B*) Inhibition of splicing by 4 μM U2b 2′-*O*-methyl oligonucleotide (46). (*C*) The products and intermediates of *in vitro* splicing-snoRNA processing with mutant substrates in which the spacer length of the host intron is altered are shown. The identities of the intermediates and products [deduced from RNase H mapping [see Fig. 7*A* or mutant analyses (data not shown)] are shown on the sides. The * may represent a lariat intron lacking the sequence downstream of the branch point produced by exonucleolytic trimming during the longer incubation.

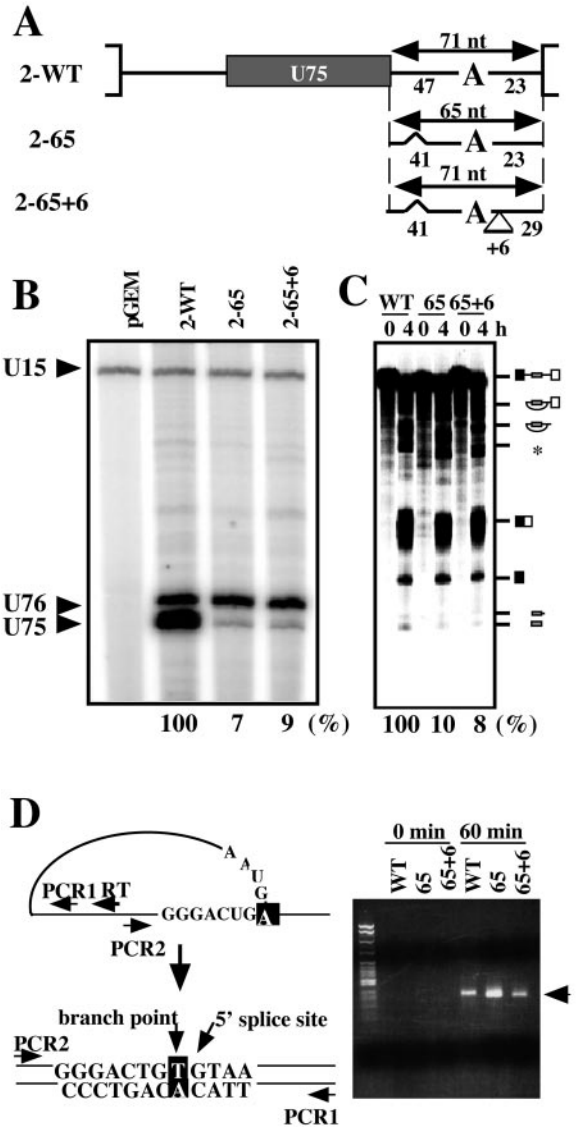

Fig. 5. Importance of the distance between the snoRNA-coding region and the branch point. (*A*) Mutant constructs used for transfection. The lengths of downstream spacers are shown above, with the distances between the U75 snoRNA-coding region and the branch point, and between the branch point and 3′ splice site indicated below. (*B*) Mutants were transfected and analyzed as in Fig. 2. (*C*) *In vitro* splicing-snoRNA processing of mouse *gas5* intron 2 substrates. The three splicing substrates, WT, 65, and 65 + 6, containing introns 2-WT, 2–65, and 2–65 + 6, respectively, were incubated for 4 h. The identities of the intermediates and products are shown on the sides. Quantitations of the relative levels of U75 are shown below. (*D*) Branch point mapping of introns excised *in vitro*. The RT-PCR strategy for the specific amplification of cDNAs derived from the lariat intron used the indicated primers for RT and PCR (PCR1 and PCR2). The sequence expected surrounding the branch point-5′ splice site ligation is shown. A 2% agarose gel fractionating the RT-PCR products is shown on the right.

is the box C/D element, and the other is the length of the downstream spacer described above. We first confirmed that the synthesis of U75 is abolished *in vivo* upon mutation of either box C or D by transfecting mutant constructs (data not shown). Box C or D mutations (see Fig. 7*B*) likewise abolish the *in vitro* synthesis of U75, whereas pre-mRNA splicing occurred with the same efficiency as with WT substrate. Most importantly, the synthesis of U75 was abolished by deletions in the downstream spacer without affecting splicing (Fig. 4*C*), consistent with the *in vivo* observations in Fig. 2*B*. Together these results indicate that our *in vitro* system reproduces the *in vivo* requirements for the production of the U75 snoRNA from its host intron.

The Distance to the Branch Point Is the Critical Determinant for snoRNA Synthesis. Two elements essential for splicing of the host intron, the branch point and the 3′ splice site, reside in the spacer downstream of box C/D snoRNAs. Both sequences interact with factors (including the U2 snRNP) whose binding profile changes dynamically as splicing proceeds. To clarify whether the distance from the snoRNA coding region to the branch point or the 3′ splice site is important, mutational analyses designed to alter the position of the branch point were carried out. The effects of these mutations were assessed *in vivo*. However, branch point identification *in vivo* is difficult because the excised lariat intron does not accumulate and mutations of the authentic branch point adenosine often shift utilization to another close-by adenosine residue (38). To map the branch point experimentally, we therefore used our *in vitro* system.

Because shortening the downstream spacer of the second intron of mouse *gas5* from 71 nt (WT) to 65 nt (construct 2–65) led to a dramatic decrease in U75 *in vivo* (Fig. 2*B*), we inserted 6 nt into a site downstream of the putative branch point of 2–65. Thus, in 2–65 + 6, the length between the snoRNA coding region and the branch point is 41 nt, as it is in 2–65, but the whole length of the downstream spacer is 71 nt, as it is in WT (Fig. 5*A*). If the whole length of the downstream spacer is important for snoRNA synthesis, the 2–65 + 6 should produce U75 snoRNA with the same efficiency as WT. If instead, the distance to the branch point is critical, U75 synthesis should be low. Transfection of these three constructs followed by snoRNA analysis indicated that the level of U75 snoRNA synthesized from 2–65 + 6 was less than 10% that of WT, the same efficiency as observed for 2–65 (Fig. 5*B*). This finding suggested that snoRNA distance from the branch point determines its level of production.

To investigate whether alteration of the downstream spacer has the same effect in our *in vitro* system, three substrates

containing the WT second intron, 2–65, and 2–65 + 6 mutants were constructed. All spliced accurately and with the same efficiency (Fig. 5*C*), consistent with the *in vivo* splicing results (data not shown). Compared with WT, the levels of U75 produced at 4 h from the 2–65 and the 2–65 + 6 mutants were about 10% (Fig. 5*C*), also consistent with *in vivo* observations (Fig. 5*B*).

Branch point mapping was carried out by using an RT-PCR method (37) on RNA prepared from the *in vitro* splicing reaction after 1 h (Fig. 5*D Left*). Molony murine leukemia virus RNase H^- reverse transcriptase has been observed to read through a 2′-5′ phosphodiester linkage by misincorporating adenosine (instead of inserting thymidine) at the branch point itself (Fig. 5*D*). We therefore synthesized cDNA from the lariat intron by using the RT primer and then specifically amplified it by using two other primers. Fig. 5*D Right* shows that RT-PCR products with the expected sizes (WT: 104 nt; 2–65 and 2–65 + 6: 98 nt) were obtained from the *in vitro*-spliced RNA, but not from the splicing substrates. Cloning and sequencing of these PCR fragments revealed that the expected branch point was used for splicing in all three constructs (data not shown). These data argue that what is important is the length between the snoRNA-coding region and the branch point. For the constructs tested, a distance of 47 nt yields efficient snoRNA synthesis; when shorter than 41 nt, snoRNA production drastically decreases.

Discussion

Here, we describe conservation of the position of box C/D snoRNAs within their host introns. The preferred location is ≈70 nt upstream of the 3′ splice site. Transfection experiments using the *gas5* snoRNA host gene verify that a downstream spacer length between 65 and 85 nt is optimal for snoRNA synthesis. The minimum length is particularly rigid: if it is shorter than 65 nt, synthesis of the snoRNA is abolished. The only known exception to the rule that human box C/D snoRNAs are all located farther than 65 nt from the 3′ splice site is the U16 snoRNA, encoded in the first intron of the ribosomal protein L1 gene. This host intron is a minor class intron (AC-AT intron) excised by the U12-dependent spliceosome (39). Conservation of the downstream spacer length in snoRNA host introns has also been observed in the *Drosophila UHG* gene (40), suggesting that the length constraint is general for metazoan organisms.

Our analyses argue that the important element within the downstream spacer is not a specific sequence, but the distance between the snoRNA-coding region and the branch point. U75 synthesis is reduced to less than 10% when the WT distance (47 nt) is shortened to 41 nt. Because the length between the branch point and 3′ splice site is relatively constant in mammals (20–30 nt), the length of the entire downstream spacer appears conserved. Interestingly, in the minor class intron encoding U16 snoRNA, the distance to the branch site is 43 nt, even though the distance to the 3′ splice site is short (56 nt, ref. 20), because the more highly defined branch point consensus sequence of minor class introns is located only about 12 nt from the 3′ splice site (39).

Fig. 1 shows that no snoRNAs are located closer than 30 nt to the 5′ splice site. This finding suggests that those components of the splicing machinery that interact with the 5′ splice site (e.g., the U1 or U6 snRNP) probably also interfere with snoRNP assembly.

A number of snoRNAs are located farther than 200 nt from the 3′ splice site. It is not known whether these snoRNAs are expressed at lower levels relative to those ideally situated within their host introns. It is possible that snoRNAs located in exceptional positions form more stable structures and thus do not require splicing to facilitate protein binding and processing. For instance, in the mouse *gas5* gene, U74 is located 289 nt upstream of the 3′ splice site of the first intron. It has been observed that substantial levels of U74 accumulate (14); furthermore, U74 was produced from a transfected *gas5* minigene at equivalent levels to U76 (data not shown). The mouse U74 possesses canonical C and D boxes, a terminal stem structure, and, in addition, a long external stem structure of 12 bp. In contrast, no stable external stem is predicted to flank U75 or U76. An external stem structure has been reported to be important for expression of box C/D snoRNAs lacking a terminal stem structure (41). Thus, the presence of a stable external stem structure may compensate for suboptimal location of the snoRNA within the host intron.

In contrast to the effect of 3′ spacer region length on snoRNA synthesis, it is surprising that moving the snoRNA-coding region closer to the 3′ splice site does not affect splicing of the host intron. This finding indicates that the assembly of spliceosomal components either obligatorily precedes or can displace snoRNP proteins. Although the failure of snoRNA synthesis from mutants with short downstream spacers can therefore be explained, our observations that snoRNA synthesis is also negatively impacted in mutants with longer downstream spacers suggests a synergy between spliceosome and snoRNP assembly. Indeed, incubation of the linear RNA corresponding to the second intron of mouse *gas5* in our coupled *in vitro* system failed to produce mature U75 snoRNA (data not shown), consistent with the idea that splicing components may be essential for snoRNA release.

We propose that in mammalian cells snoRNP protein(s) responsible for accurate snoRNA processing bind the snoRNA-coding region at a particular step of spliceosome assembly or function. The branch point region is an important element for spliceosome assembly, as well as catalysis, and is bound by several essential factors during the splicing reaction. At the stage of spliceosomal E complex formation, six proteins associated with the U2 snRNP (SF3a and SF3b proteins) tightly bind to the "anchor region," which comprises the 25 nt upstream of the branch point (42). These interactions are required for subsequent A complex formation, involving base-pairing of the U2 snRNP with the pre-mRNA at the branch point. Thus, the SF3a/b proteins are good candidates for splicing factors that positively (or negatively) affect snoRNA release from host introns. Of the four box C/D snoRNP proteins identified so far (fibrillarin, Nop58p, Nop56p, and 15.5-kDa protein), only Nop56p is dispensable for snoRNA accumulation (31). The 15.5-kDa protein specifically binds to the box C/D core structure (28), making it a strong candidate for interaction with splicing components. Because of the presence of this protein in the U4 snRNP, it is conceivable that the 15.5-kDa protein relocates to the nascent box C/D snoRNA during the splicing reaction.

Another possibility is that failure of snoRNA synthesis from mutants with short downstream spacers occurs at the postsplicing stage. After the splicing reaction, ligated exons are released from the spliceosome, whereas the excised lariat intron remains bound to snRNPs (at least *in vitro*) (33). It is believed that only after the dissociation of snRNPs can the debranching enzyme access the 2′-5′ phosphodiester bond at the branch point (33). Thereby, an ongoing debranching reaction also could interfere with the binding of snoRNP proteins close to the branch point.

A few yeast snoRNAs are encoded within introns (12). For these, release partially depends on pre-mRNA splicing and subsequent debranching of the excised lariat intron by the debranching enzyme (Dbr1). Yet, intronic snoRNAs can be produced in the dbr1-deficient yeast mutant, suggesting involvement of an endonuclease that linearizes lariat introns (21). In both pathways, exonuclease trimming forms the correct 5′ and 3′ termini of the snoRNA. The yeast Rat1p and Xrn1p 5′→3′ exonucleases play an essential role in the 5′ end formation of both intronic and polycistronic snoRNAs (22). Maturation of some yeast snoRNAs by the trimming of short 3′ terminal trailer

sequences specifically requires the Rrp6p 3′→5′ exonuclease, which is a component of the yeast nuclear exosome (23).

We have developed a mammalian *in vitro* system for coupled splicing-snoRNA processing. The synthesis of U75 snoRNA *in vitro* depends on the splicing of its host intron, the integrity of both boxes C and D, and the length of the downstream spacer. Several intronic snoRNAs in vertebrates have been reported to be processed exclusively via a splicing-dependent pathway (18), and our *in vitro* system supports these observations. In contrast, *in vitro* processing of U15, which is located at an exceptional position within its host intron (155 nt upstream of the 3′ splice site), did not require splicing (16); it possesses a stable internal stem structure, perhaps folding it into the correct snoRNA structure without synergy from splicing. We have failed to produce U75 (which does not possess a stable stem structure) from the linearized intron under the same *in vitro* conditions that give processing of U15 (data not shown). We observed a low yield of U75 in comparison to other splicing products (Figs. 4 and 5). One possibility is that the amounts of snoRNP proteins are limiting in our nuclear extracts, resulting in snoRNA decay. Alternatively, inefficient debranching of the lariat intron (seen in Figs. 4 and 5, where lariat intron lacking the sequence downstream of the branch point accumulates after longer incubation, shown by asterisk) may be responsible. Other extract preparations containing higher debranching activity yielded more snoRNA processing intermediate and a reduction of the truncated lariat (data not shown). We observed the production of a band at ≈70 nt, which by RNase H analysis (see Fig. 7*A*) is a processing intermediate of U75, presumably with a mature 5′ end and a short 3′ terminal trailer (Figs. 4 and 5*C*). Production of similar intermediates has been observed during the *in vitro* processing of U15 and U17 snoRNAs (16, 17). The human homologue of yeast Rrp6p has been identified as the PM-Scl100 protein (43, 44), which is the target of autoimmune antibodies in patients suffering from polymyositis-scleroderma overlap syndrome (45). It is possible that the concentration of PM-Scl100 protein is limited in our *in vitro* system, leading the accumulation of the 3′ extended precursor U75. Our system will be useful for pursuing the detailed molecular mechanism of the biogenesis of the intron encoded-snoRNAs in mammalian cells and identifying the factors involved.

We thank Brenda Bass, Kazio Tycowski, Lara Weinstein Szewczak, Niamh Cahill, Jens Lykke-Andersen, and other members of the Steitz laboratory for stimulating discussions and critical comments on the manuscript. This work was supported by a Human Frontier Science Program long-term fellowship (to T.H.) and National Institutes of Health Grant GM26154 (to J.A.S.). J.A.S. is an Investigator of the Howard Hughes Medical Institute.

1. Smith, C. M. & Steitz, J. A. (1997) *Cell* **89,** 669–672.
2. Tollervey, D. & Kiss, T. (1997) *Curr. Opin. Cell Biol.* **9,** 337–342.
3. Kiss-Laszlo, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M. & Kiss, T. (1996) *Cell* **85,** 1077–1088.
4. Nicoloso, M., Qu, L. H., Michot, B. & Bachellerie, J. P. (1996) *J. Mol. Biol.* **260,** 178–195.
5. Cavaille, J., Nicoloso, M. & Bachellerie, J. P. (1996) *Nature (London)* **383,** 732–735.
6. Tycowski, K. T., Smith, C. M., Shu, M. D. & Steitz, J. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 14480–14485.
7. Ganot, P., Bortolin, M. L. & Kiss, T. (1997) *Cell* **89,** 799–809.
8. Ni, J., Tien, A. L. & Fournier, M. J. (1997) *Cell* **89,** 565–573.
9. Kiss-Laszlo, Z., Henry, Y. & Kiss, T. (1998) *EMBO J.* **17,** 797–807.
10. Tycowski, K. T., Shu, M. D. & Steitz, J. A. (1996) *Nature (London)* **379,** 464–466.
11. Cavaille, J. & Bachellerie, J. P. (1998) *Nucleic Acids Res.* **26,** 1576–1587.
12. Maxwell, E. S. & Fournier, M. J. (1995) *Annu. Rev. Biochem.* **35,** 897–933.
13. Weinstein, L. B. & Steitz, J. A. (1999) *Curr. Opin. Cell. Biol.* **11,** 378–384.
14. Smith, C. M. & Steitz, J. A. (1998) *Mol. Cell. Biol.* **18,** 6897–6909.
15. Pelczar, P. & Filipowicz, W. (1998) *Mol. Cell. Biol.* **18,** 4509–4518.
16. Tycowski, K. T., Shu, M. D. & Steitz, J. A. (1993) *Genes Dev.* **7,** 1176–1190.
17. Kiss, T. & Filipowicz, W. (1993) *EMBO J.* **12,** 2913–2920.
18. Kiss, T. & Filipowicz, W. (1995) *Genes Dev.* **9,** 1411–1424.
19. Fragapane, P., Prislei, S., Michienzi, A., Caffarelli, E. & Bozzoni, I. (1993) *EMBO J.* **12,** 2921–2928.
20. Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P. & Bozzoni, I. (1996) *EMBO J.* **15,** 1121–1131.
21. Ooi, S. L., Samarsky, D. A., Fournier, M. J. & Boeke, J. D. (1998) *RNA* **4,** 1096–1110.
22. Petfalski, E., Dandekar, T., Henry, Y. & Tollervey, D. (1998) *Mol. Cell. Biol.* **18,** 1181–1189.
23. Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. & Tollervey, D. (1999) *EMBO J.* **18,** 5399–5410.
24. Watkins, N. J., Leverette, R. D., Xia, L., Andrews, M. T. & Maxwell, E. S. (1996) *RNA* **2,** 118–133.
25. Xia, L., Watkins, N. J. & Maxwell, E. S. (1997) *RNA* **3,** 17–26.
26. Tyc, K. & Steitz, J. A. (1989) *EMBO J.* **8,** 3113–3119.
27. Watkins, N. J., Newman, D. R., Kuhn, J. F. & Maxwell, E. S. (1998) *RNA* **4,** 582–593.
28. Watkins, N. J., Ségault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C. & Lührmann, R. (2000) *Cell* **103,** 457–466.
29. Wu, P., Brockenbrough, J., Metcalfe, A., Chen, S. & Aris, J. (1998) *J. Biol. Chem.* **273,** 16453–16463.
30. Lafontaine, D. L. & Tollervey, D. (1999) *RNA* **5,** 455–467.
31. Lafontaine, D. L. & Tollervey, D. (2000) *Mol. Cell. Biol.* **20,** 2650–2659.
32. Lyman, S. K., Gerace, L. & Baserga, S. J. (1999) *RNA* **5,** 1597–1604.
33. Moore, M. J., Query, C. C. & Sharp, P. A. (1993) in *The RNA World*, eds. Gesteland, R. F. & Atkins, J. F. (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 303–358.
34. Coccia, E. M., Cicala, C., Charlesworth, A., Ciccarelli, C., Rossi, G. B., Philipson, L. & Sorrentino, V. (1992) *Mol. Cell. Biol.* **12,** 3514–3521.
35. Tymms, M. J. (1995) in *Methods in Molecular Biology*, ed. Tymms, M. J. (Humana, Clifton, NJ), Vol. 37, pp. 31–46.
36. Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
37. Vogel, J., Hess, W. R. & Börner, T. (1997) *Nucleic Acids Res.* **25,** 2030–2031.
38. Ruskin, B., Greene, J. M. & Green, M. R. (1985) *Cell* **41,** 833–844.
39. Burge, C. B., Padgett, R. A. & Sharp, P. A. (1998) *Mol. Cell* **2,** 773–785.
40. Tycowski, K. T. & Steitz, J. A. (2001) *Eur. J. Cell. Biol.* **80,** 119–125.
41. Darzacq, X. & Kiss, T. (2000) *Mol. Cell. Biol.* **20,** 4522–4531.
42. Gozani, O., Feld, R. & Reed, R. (1996) *Genes Dev.* **10,** 233–243.
43. Briggs, M. W., Burkard, K. T. & Butler, J. S. (1998) *J. Biol. Chem.* **273,** 13255–13263.
44. Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D. & Mitchell, P. (1999) *Genes Dev.* **13,** 2148–2158.
45. Bluthner, M. & Bautz, F. A. (1992) *J. Exp. Med.* **176,** 973–980.
46. Tarn, W. Y. & Steitz, J. A. (1996) *Cell* **84,** 801–811.