J. Biochem. 120, 677-683 (1996)

Spliceosomal Introns in Conserved Sequences of U1 and U5 Small Nuclear RNA Genes in Yeast *Rhodotorula hasegawae*[1]

Yuzi Takahashi, Tokio Tani, and Yasumi Ohshima[2]

Department of Biology, Faculty of Science, Kyushu University, Hakozaki, Higashi-ku, Fukuoka 812-81

Received for publication, June 11, 1996

U1, U2, U4, U5, and U6 small nuclear RNAs are essential for precursor mRNA splicing. We previously found one and four mRNA-type, or spliceosomal introns, in the U2 and U6 genes, respectively, of yeast *Rhodotorula hasegawae* (*Erythrobasidium hasegawianum*). We proposed that U2 and U6 RNAs form a catalytic core for precursor mRNA splicing and that the introns in those small nuclear RNA genes may have been acquired through reverse splicing of an intron from a precursor mRNA into a catalytic site in small nuclear RNAs. In the present study, we analyzed U1, U4, and U5 genes in *R. hasegawae*. One spliceosomal intron was found in the U1 region forming base-pairs with a 5′ splice site of a precursor mRNA. The U5 gene has two spliceosomal introns in the region that interacts with 5′ and 3′ splice sites. In contrast, the gene for U4 RNA, which is released from the spliceosome prior to the first step of the splicing reaction, has no intron. These results lend a further support to the proposed relation between presence and position of an intron in an small nuclear RNA gene and the function of the encoded small nuclear RNA.

Key words: pre-mRNA splicing, spliceosomal intron, U1 snRNA, U4 snRNA, U5 snRNA.

A class of small nuclear RNAs (snRNAs) is present in the nuclei of eukaryotes. Of these, U1, U2, U4, U5, and U6 snRNAs are involved in pre-mRNA splicing. They form ribonucleoprotein particles and assemble on a pre-mRNA to make up a spliceosome, the machinery for pre-mRNA splicing. Several lines of experiments showed that U1 and U2 snRNAs base-pair with the 5′ splice site and the branch site of a pre-mRNA, respectively (*1-6*). U1 snRNA was also reported to base pair with the conserved 3′ splice site sequence, at least, in *Schizosaccharomyces pombe* (*7*). By genetic and biochemical experiments, U5 snRNA was proven to interact with both exon sequences adjacent to the 5′ or 3′ splice sites (*8-13*). U6 snRNA also determines specificity of the 5′ splice site recognition (*14, 15*). It was proposed that U4 snRNA functions as a negative regulator of U6 snRNA (*16*), although the exact function of U4 is unclear.

Catalysis of pre-mRNA splicing is presumed to be carried out by a machinery constructed with RNAs, in which U2, U5, and U6 snRNAs appear to play key roles (*17*). U1 and U4 snRNAs seem to be dispensable for catalytic steps and involved only in spliceosome assembly and commitment to the splicing pathway (*18-21*). U6 snRNA is the most conserved spliceosomal snRNA. When a spliceosomal intron was found in the U6 gene of *S. pombe* (*22*), it was proposed that U6 snRNA served as a catalytic element in pre-mRNA splicing and that the U6 intron had been generated through reverse splicing of an intron excised from a pre-mRNA into a catalytic domain of U6 snRNA, followed by reverse transcription and homologous recombination of the resultant complementary DNA with the U6 gene in the genome (*23*). Mutational analyses also suggested the catalytic function of U6 snRNA (for a review, see Ref. *24*). We found spliceosomal introns in U6 genes of two more yeasts, *Rhodosporidium dacryoidum* and *Rhodotorula hasegawae* (*Erythrobasidium hasegawianum*) (*25*). The U6 introns found in *S. pombe* and *R. dacryoidum* are flanked by sequences that affected the second step of pre-mRNA splicing *in vitro*, when mutated in *Saccharomyces cerevisiae* (*26*).

Some mutational experiments in *S. cerevisiae* and mammals suggested that U2 snRNA is also involved in catalytic steps of pre-mRNA splicing (*27-29*). We previously found a spliceosomal intron in the U2 gene of *R. hasegawae*, in which the U6 gene has four spliceosomal introns (*30*). The U2 intron was present in the conserved sequence downstream of the region encoding a branch site recognition sequence. U2 and U6 snRNAs form base-pairings essential for pre-mRNA splicing (helices I, II, and III) (*28-34*) and these pairings, together with the U2-branch site base-pairing and the U6-5′ splice site base-pairing, would bring both the branch point and the 5′ splice site of a pre-mRNA close to the putative catalytic domain formed by U2 and U6 snRNAs (*30, 34*). Catalytic core models of U2 and U6 snRNAs, which resemble a group II self-splicing intron or a hairpin ribozyme, were proposed (*28, 34*). U5 snRNA, which interacts with both splice sites, may specify

[1] This work was supported in part by a grant from a Research Fellowship of the Japan Society for the Promotion of Science for Young Scientists to Y. Takahashi, and grants from the Ministry of Education, Science, Sports and Culture of Japan. The sequence data reported in this paper have been entered in the DDBJ/EMBL/GenBank databases under the following accession numbers: D63681 for *R. hasegawae* U1 small nuclear RNA gene; D63682 for *R. hasegawae* U4 small nuclear RNA gene; D63683 for *R. hasegawae* U5 small nuclear RNA gene.

[2] To whom correspondence should be addressed. Tel: +81-92-642-2626, FAX: +81-92-642-2645, E-mail: yohshscb@mbox.nc.kyushu-u.ac.jp

Abbreviations: pre-mRNA, precursor mRNA; snRNA, small nuclear RNA.

the cleavage sites and align the two exons (*9, 10, 13*).

As mentioned above, one and four spliceosomal introns are present in the U2 and U6 genes of *R. hasegawae*, respectively, and their positions seem to be restricted to the RNA regions within or close to the presumed catalytic core in the spliceosome. For further clarification of the relation between presence and position of an intron in an snRNA gene and function of the snRNA, we analyzed U1, U4, and U5 genes of *R. hasegawae* in this study.

EXPERIMENTAL PROCEDURES

Oligonucleotides—Oligonucleotides used in this study are as follows.

Xho-T20, 5′-GCCTCGAGTTTTTTTTTTTTTTTTTTT-3′
Eco-U1LPA, 5′-CTCGAATTCA(G/A)(T/C/A)GGAGAT-CA(A/T)(G/T)AA-3′
Eco-U4, 5′-CTGAATTCACGA(T/G)C(T/C)TT(G/A)(T/C)GCA(G/T)(G/T)GG-3′
Eco-U5, 5′-GCGAATTCTCCT(C/T)GCCTTT(C/T)A(C/T)(C/T)A(A/G)A-3′
RHU1IUP, 5′-GTCGCCGTGCAATGCGAAGT-3′
RHU1IDWN, 5′-CTTTGGTCTTGTTGGCTTCG-3′
RHU1UP, 5′-ATACTTAGTTAGTTTTCCTC-3′
RHU1DWN, 5′-GGGTCTCGAGGGCCGAAGCC-3′
RHU4IUP, 5′-TGCAACCAGCAATTAGATGG-3′
RHU4IDWN, 5′-TCTTTGGGAATTTCTGCCCG-3′
RHU4UP, 5′-TCTTTGTGCGTCGGGCACGC-3′
RHU4DWN, 5′-ACGCCGTTAGGCGGGCAGAA-3′
RHU5S, 5′-GCGAGAAACATTGAATAAAC-3′
RHU5IUP, 5′-AAGCAAGAAAGCGAGAACCG-3′
RHU5IDWN, 5′-AATGTTTCTCGCCGCTCTGC-3′
RHU5UP, 5′-AAGCGCTAGTTCAGCGAACA-3′
RHU5DWN, 5′-ACCGCAGAGCGGCGAGAAAC-3′

Preparation of Total DNA and RNA—*R. hasegawae* was cultured at 23°C. Total DNA and RNA were prepared as described in Ref. *30*.

Determination of RNA Sequences—Sequences of snRNAs were determined by the protocol of Frank *et al.* (*35*), with some modifications. *R. hasegawae* total RNA was poly(A) tailed with poly(A) polymerase for 30 min at 37°C. Then cDNAs were synthesized using the oligonucleotide Xho-T20 as a primer and reverse transcriptase. Second-strand synthesis was carried out by two step PCR amplifications using oligonucleotides Xho-T20 and a sequence-specific primer for each snRNA (Eco-U1LPA, Eco-U4, and Eco-U5 for U1, U4, and U5 cDNA synthesis, respectively). In the first step PCR, annealing temperature was 35°C for the first 10 cycles and then raised to 45°C for additional 20 cycles. One microliter of resultant solution was applied to the second step PCR with annealing temperature at 55°C for 30 cycles. Each PCR sample was ethanol precipitated and analyzed on a 5% polyacrylamide gel. Amplified DNAs were cloned into the pCR™ II vector (Invitrogen) and were sequenced with a DNA sequencer (Applied Biosystems 373A).

Each snRNA region upstream of the determined sequences was sequenced using the dideoxy method and a ^{32}P-labeled sequence-specific primer for each snRNA (RHU1IUP, RHU4IUP, and RHU5S for U1, U4, and U5, respectively) (*25*). A 5′ end nucleotide of each snRNA was predicted in reference to a corresponding genomic sequence determined by inverse PCR.

Determination of Genomic Sequences—According to the RNA sequences, oligonucleotide primers specific for each snRNA (RHU1UP and RHU1DWN for U1, RHU4UP and RHU4DWN for U4, and RHU5UP and RHU5DWN for U5) were designed. Using these primers, each snRNA gene was amplified by PCR of 30 cycles each consisting of denaturation for 40 s at 94°C, annealing for 40 s at 55°C, and extension for 90 s at 72°C. Nucleotide sequences of each snRNA gene upstream and downstream of the determined sequences were analyzed by inverse PCR (*25*) with primers specific for each snRNA (RHU1IUP and RHU1IDWN for U1, RHU4IUP and RHU4IDWN for U4, and RHU5IUP and RHU5IDWN for U5).

RESULTS

Analysis of the U1 snRNA Gene—We determined *R. hasegawae* U1 snRNA sequence using the oligonucleotide Eco-U1LPA that has a conserved sequence of loop A of U1 snRNA. The result shows that the *R. hasegawae* U1 snRNA has candidates for a 5′ splice site recognition sequence, loop B, an Sm binding site, or loop A (Fig. 1A). *R. hasegawae* U1 snRNA can form a secondary structure conserved in higher eukaryotes and *S. pombe* (*36*), except for the presence of an additional stem-loop sticking out from the loop B.

We determined the sequence of a gene for U1 snRNA, as described in "MATERIALS AND METHODS." In the course of cloning the U1 gene, inverse PCR gives a single band (data not shown). It indicates that the U1 gene is present as a single copy in the *R. hasegawae* genome. Comparison of the RNA and genomic sequences of the *R. hasegawae* U1 revealed a spliceosomal intron, which has consensus 5′ and 3′ splice site sequences (Figs. 2A and 3A). This is the first intron to be found in the U1 gene, although U1 genes of more than 30 species have been cloned. The U1 intron is present in the sequence shown to base-pair with the 5′ splice site of a pre-mRNA in yeast and mammals (*1-3*).

Analysis of the U4 snRNA Gene—*R. hasegawae* U4 snRNA sequence was determined using the oligonucleotide Eco-U4 that has a conserved sequence for stem II which base-pairs with U6 snRNA. The resultant sequence (Fig. 1C) has an Sm binding site-like sequence and sequences that can base-pair with U6 snRNA of this yeast (*25*) forming stems I and II. A possible secondary structure of the U4/U6 snRNAs (Fig. 1C) resembles those of mammals. Figure 2B shows the nucleotide sequence of the U4 gene. The *R. hasegawae* U4 gene has no intron. Inverse PCR for the U4 gene gives a single band (data not shown), indicating that a single copy U4 gene is present in the *R. hasegawae* genome.

Analysis of the U5 snRNA Gene—We determined *R. hasegawae* U5 snRNA sequence using the oligonucleotide Eco-U5 that has a conserved sequence in loop I. The cloned U5 has sequences conserved among seven species of yeasts (*35*) with some exceptions. Figure 1B shows a presumed secondary structure of the *R. hasegawae* U5 snRNA. The sequence of the *R. hasegawae* U5 loop I matches better with the consensus sequence for plants and protozoa than with that for yeasts. U and C residues in the consensus sequence for yeasts are replaced with C and U residues in *R. hasegawae* (positions 37 and 47), respectively. In the stem I in *R. hasegawae*, yeast consensus G-C pair is

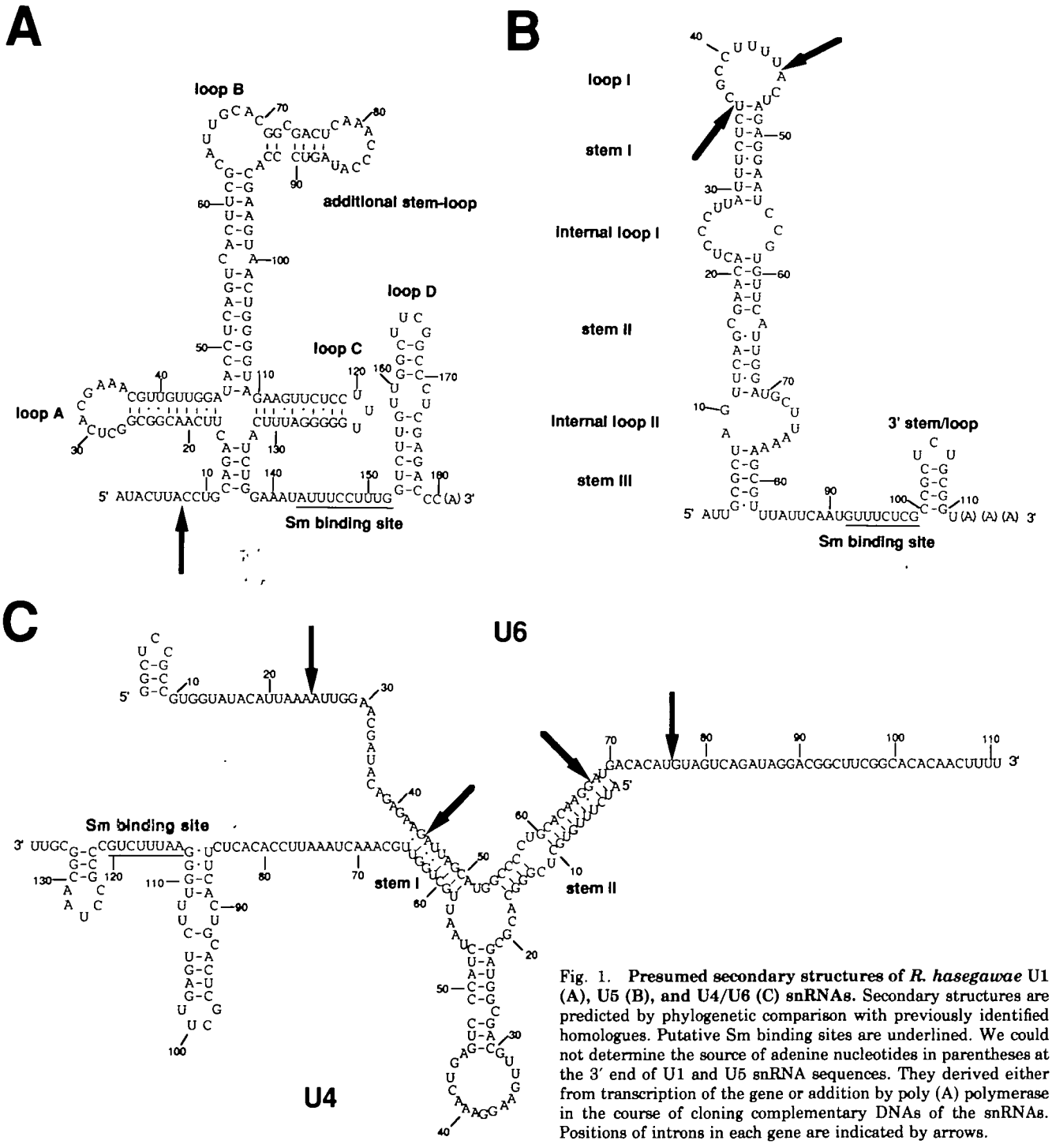

Fig. 1. **Presumed secondary structures of *R. hasegawae* U1 (A), U5 (B), and U4/U6 (C) snRNAs.** Secondary structures are predicted by phylogenetic comparison with previously identified homologues. Putative Sm binding sites are underlined. We could not determine the source of adenine nucleotides in parentheses at the 3′ end of U1 and U5 snRNA sequences. They derived either from transcription of the gene or addition by poly (A) polymerase in the course of cloning complementary DNAs of the snRNAs. Positions of introns in each gene are indicated by arrows.

replaced with U-A pair (positions 31 and 53). CCG is present in the 3′ side of the internal loop I (positions 56-58), but CG dinucleotides are not present in the 5′ side of the loop in *R. hasegawae* in contrast to the yeast consensus. Presumed Sm binding site of *R. hasegawae* contains two extra C residues in contrast to the consensus. Presumed Sm binding sites of U1 and U4 snRNAs in this yeast also have two and one extra C residues, respectively. We previously reported the presence of extra three C residues in the Sm binding site of U2 snRNA in this yeast (*30*). *R. hasegawae* snRNAs seem to contain extra C residues in the Sm binding sites. More than half of yeast U5 snRNAs, of which the sequences have been determined so far, have a variable stem-loop immediately upstream of the stem I, but the *R. hasegawae* U5 does not. The *R. hasegawae* U5 has a 3′ stem-loop that is present in less than half of the yeast species. A short version of U5 that lacks this stem-loop, as found in *S. cerevisiae* (*37*), was not detected in *R. hasegawae* in the course of this study.

Inverse PCR for determining the U5 genomic sequence

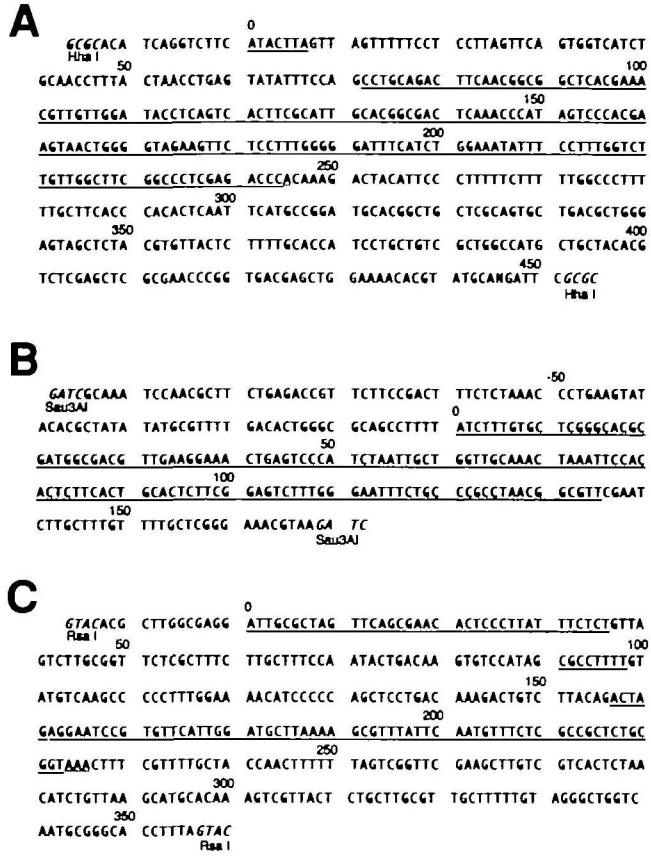

Fig. 2. **Nucleotide sequences of the *R. hasegawae* U1 (A), U4 (B), and U5 (C) snRNA genes.** Nucleotides in mature transcripts are underlined. Transcription of dot-underlined adenine nucleotides are not confirmed (see the legend of Fig. 1). Nucleotides in italics are cutting sites of indicated restriction enzymes used in inverse PCR. Intron sequences are indicated by outlined letters.

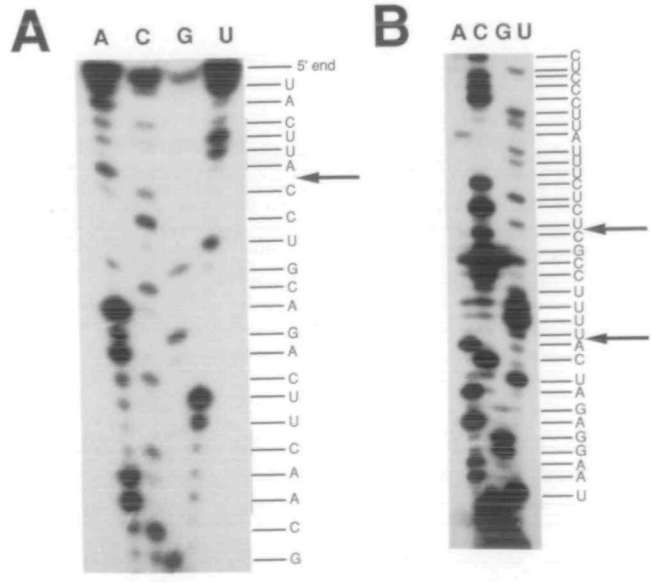

Fig. 3. **Partial nucleotide sequences near the intron positions of *R. hasegawae* U1 snRNA (A) and U5 snRNA (B).** Ten micrograms of total RNA from *R. hasegawae* was used for sequencing of each RNA. The sequence provided by the complementary DNA is given on the right. Arrows show positions of the introns.

	5' splice sites		putative branch sites		3' splice sites
U1	TTA/GTTAGT	——	TACTAAC	——	TTTCCAG/CCT
U2	CTT/GTACGT	——	GTCTGAC	——	CCTACAG/AAT
U5 1st	TCT/GTTAGT	——	TACTGAC	——	TCCATAG/CGC
U5 2nd	TTT/GTATGT	——	TCCTGAC	——	CTTACAG/ACT
U6 1st	AAA/GTGCGT	——	AGCTTAC	——	TTCTTAG/ATT
U6 2nd	AAG/GTAAAT	——	GTCTAAT	——	TCTACAG/ATT
U6 3rd	AGG/GTTTGT	——	AGCTTAC	——	TCTATAG/ATG
U6 4th	CAT/GTTTGT	——	AGTTAAC	——	TTTCCAG/GTA
consensus	wwt/GTWwGT	——	wNCTrAC	——	YYYaYAG/att

Fig. 4. **Comparison of the 5′ and 3′ splice sites and the putative branch site sequences of the snRNA genes in *R. hasegawae*.** The exon/intron boundaries and the putative branch sites of all the eight introns found in *R. hasegawae* spliceosomal snRNA genes are compared. Sequences of the U2 and U6 introns are adopted from Refs. *30* and *25*, respectively. Sequences of the U1 and U5 introns were determined in this study. In consensus sequences, uppercase letters indicate nucleotides conserved in seven or eight cases and lowercase letters designate those conserved in four to six cases. W, Y, and R are abbreviations for A or T, pyrimidine and purine, respectively.

gave a single band (data not shown), indicating that the U5 gene is present as a single copy in *R. hasegawae* genome. When we compared the sequence of the *R. hasegawae* U5 gene with that of U5 snRNA, we found two spliceosomal introns (Figs. 2C and 3B). These introns are present within the sequence corresponding to loop I, which contains a sequence known to interact with exon sequences adjacent to 5′ and 3′ splice sites of a pre-mRNA. The length of the exon between the two introns is only eight nucleotides. These are the first introns to be found in the U5 gene.

Splice Site Sequences of the Spliceosomal snRNA Genes in R. hasegawae—Figure 4 shows a comparison of sequences at 5′ and 3′ splice sites and putative branch sites of the snRNA genes in *R. hasegawae*. Consensus sequences of the 5′ and 3′ splice sites and the putative branch sites are wwt/GTWwGT, YYYaYAG/att, and wNCTrAC, respectively (see the legend of Fig. 4 for notations). All eight introns found in snRNA genes in this yeast start with GT and end with AG, which are typical for spliceosomal introns.

DISCUSSION

R. hasegawae U1 and U5 Introns Are Present in the Conserved Regions That Encode Sequences Interacting with Splice Sites of a Pre-mRNA—In this study, we analyzed U1, U4, and U5 genes of *R. hasegawae*, and found one and two spliceosomal introns in the U1 and U5 genes, respectively (Fig. 1, A and B). The U1 intron is present in the sequence shown to base-pair with the 5′ splice site of a pre-mRNA (*1-3*). The U1 sequence near the intron was also reported to base-pair with the 3′ splice site, at least in *S. pombe* (*7*). The U5 introns are in the loop I region (Fig. 1B), which contains a sequence that affected selection of both the 5′ and 3′ splice sites when mutated in *S. cerevisiae* (*8, 9*). This sequence was also shown to interact with both the splice sites by chemical cross-linking (*10, 11, 13*). In conclusion, the U1 and U5 introns are located within or immediately adjacent to the sequences involved in splice site selection (Fig. 5). The U4 gene of *R. hasegawae* has no

intron, and U4 snRNA leaves the spliceosome prior to catalytic steps of the splicing reaction (*20, 21*).

Origin of Introns in the snRNA Genes—Introns heretofore identified in snRNA genes are restricted to the RNA region that likely has an important role in pre-mRNA splicing. One possible explanation for this restriction is based on a hypothesis that these introns were generated through reverse splicing events. When a spliceosomal intron was found in the U6 gene of *S. pombe*, Brow and Guthrie (*23*) proposed that the U6 intron arose from a mishap during pre-mRNA splicing, that is to say, a spliceosomal intron excised from a pre-mRNA was inserted into a U6 snRNA molecule in the spliceosome. If such is the case, the site of intron insertion would be located in the region of U6 snRNA within or close to the catalytic center in the spliceosome. Similarly, a region of another snRNA, in which an intron is present, would reflect its proximity to the catalytic center of pre-mRNA splicing. The U1 and U5 introns are within or close to the sequences interacting with the splice sites in a pre-mRNA. This means that these snRNAs would locate close to the pre-mRNA in the spliceosome. Cross-linking experiments showed that U1 and U5 snRNAs were close to pre-mRNA and that U5 snRNA was close to an excised lariat intron (*10, 11, 13*). The cross-linking sites locate very close to the sites where the introns are found. Reverse splicing of a spliceosomal intron has not been detected directly. Reverse splicing reactions of group I and group II self-splicing introns were reported (*38-40*), although there are significant differences in the mechanisms of splicing between a spliceosomal intron and the self-splicing intron. If the above scenario on the origin of snRNA introns is correct, an snRNA intron that will be identified hereafter is expected to locate in a region close to a catalytic center of pre-mRNA splicing.

An alternative explanation for the presence of snRNA introns in the region important for pre-mRNA splicing is that only introns present in a functionally important region of the snRNA gene had been retained. Irrespective of whether introns were present in an ancient snRNA gene or whether they were occasionally inserted into snRNA genes, most of them might have been removed in the course of evolution and only introns present in functionally important regions might have been retained. In the genes for which an RNA is a final product, introns could be removed without causing a frame shift of neighboring sequences, in contrast to protein coding genes. Introns in RNA genes could generally be removed even if the removal generates a deletion, an insertion, or a base substitution in the neighboring sequences. However, introns in sequences important for the RNA function would be difficult to be removed without affecting the function. A spliceosomal intron in the *S. cerevisiae* U3 snRNA gene, which is the only intron found so far in the gene of a non-spliceosomal snRNA, also exists in a conserved sequence (*41*). This U3 intron could not be explained by the catalytic center hypothesis described before.

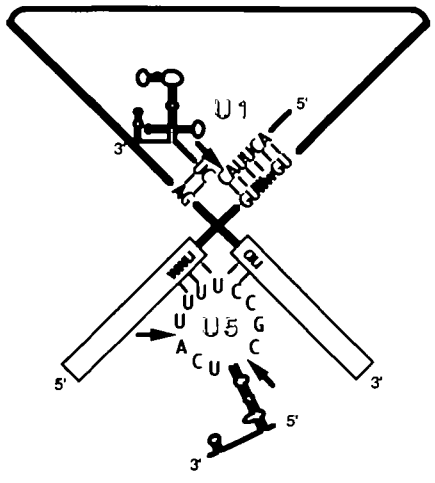

Fig. 5. **U1 and U5 snRNAs in a Holliday-like structure.** Interactions between U1 or U5 snRNAs and both 5′ and 3′ splice sites of a pre-mRNA lead to a Holliday-like structure (*44*). Structures of U1 and U5 snRNAs are drawn with *R. hasegawae* sequences. Sequences of 5′ and 3′ splice sites of the pre-mRNA are the consensus for *R. hasegawae* with upper case letters for highly conserved nucleotides. W represents A or U. U1-pre-mRNA and U5-pre-mRNA interactions are taken from those in *S. cerevisiae*, *S. pombe*, or a mammal (for a review, see Ref. *44*). Arrows indicate positions of the U1 and U5 introns.

We have found that *R. hasegawae* has a U6 gene containing five introns (Takahashi, Y. *et al.*, unpublished results), in addition to the U6 gene with four introns (*25*). So, in *R. hasegawae*, as many as thirteen introns are present in the spliceosomal snRNA genes, the coding regions of which amount to approximately 836 bp. We also analyzed an entire gene for triosephosphate isomerase and a 3′ half of elongation factor 1α gene from *R. hasegawae*, as genes encoding proteins. These regions have five and three spliceosomal introns, respectively (Takahashi, Y. *et al.*, unpublished results). In total, eight introns are present in the region coding for 519 amino acids that correspond to 1,557 bp DNA. Therefore, in *R. hasegawae*, spliceosomal snRNA genes may have more introns than protein coding genes.

Additional Stem-Loop in R. hasegawae U1 snRNA—A part of the loop B in the *R. hasegawae* U1 snRNA can form an extra stem-loop structure (Fig. 1A). This stem-loop is not present in the proposed secondary structures of U1 snRNAs in higher eukaryotes and *S. pombe*. In *S. cerevisiae* and *Kluyveromyces lactis*, U1 snRNA has an extremely long sequence instead of loop B (*42*). The sequence specific for *S. cerevisiae* and *K. lactis* might be an "expansion segment." Alternatively, the U1 snRNA of higher eukaryotes might have lost the "extra" region from a larger ancestral U1 snRNA (*42*). Whichever is the case, the *R. hasegawae* loop B and additional stem-loop might be an evolutionary intermediate between the long sequence in *S. cerevisiae* and *K. lactis* U1 snRNAs and the loop B of other species.

Presumed Expression or Processing Signals for the snRNAs in R. hasegawae—We compared the sequences of the 5′ and 3′ flanking regions of the snRNA genes in *R. hasegawae* with those of mammals or of *S. pombe*. The 5′ flanking regions of *R. hasegawae* U2, U4, and U6 genes have common sequences containing a TATA motif (GCTATATNNNNG, see Fig. 6A). The *S. pombe* U1, U2, U3, U4, and U5 genes have similar sequences in similar positions, however the conserved G in *R. hasegawae* (−32 in U2 gene) is A in *S. pombe* (*36*).

A 3′ end signal similar to the mammalian consensus is not found in *R. hasegawae*, nor is it found in *S. pombe* (*36, 43*).

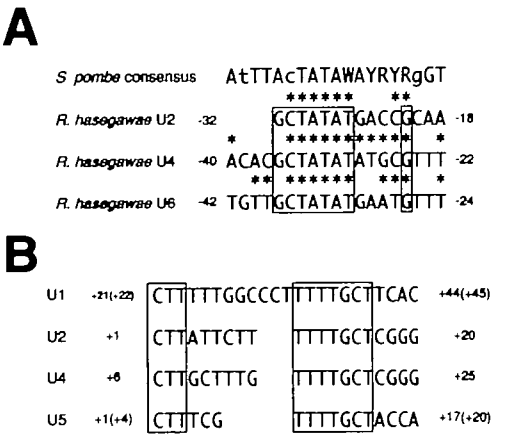

Fig. 6. **Comparison of the 5′ and 3′ flanking sequences of the *R. hasegawae* snRNA genes.** (A) Comparison of the 5′ flanking sequences. Nucleotides sequences of the *R. hasegawae* U2 (*30*), U4 and U6 (*25*) genes are compared. Numbering is relative to the 5′ end of the RNA. The sequence of the U1 and U5 genes around this region and the U2 gene upstream of nucleotides −32 were not determined. Boxes indicate sequences conserved among the U2, U4, and U6 genes. The *S. pombe* consensus sequence was adopted from Porter *et al.* (*36*), and is based on a comparison of the *S. pombe* U1, U2, U3, U4, and U5 genes. In the consensus sequences, uppercase letters indicate conserved nucleotides among all five genes and lower letters designate those conserved in the majority of the genes. W, Y, and R are abbreviations for A or T, pyrimidine, and purine, respectively. Asterisks above U2, U4, and U6 genes indicate the nucleotides consistent with the *S. pombe* consensus. (B) Comparison of the 3′ flanking sequence of the *R. hasegawae* U1, U2, U4, and U5 genes. The first nucleotide downstream of the 3′ end of each snRNA is numbered as +1. Numbers in parentheses are for the case in which all A residues dot-underlined in Fig. 2 are present in the 3′ end of each snRNA (see the legend of Fig. 1). Boxes indicate conserved sequences among the four genes.

Instead, the 3′ flanking regions of the *R. hasegawae* U1, U2, U4, and U5 genes have greatly T-rich regions containing common sequences ($CTTN_{3-9}TTTTGCT$) (Fig. 6B). These sequences may possibly function for correct formation of 3′ end of snRNA precursor. The *R. hasegawae* U6 gene lacks this sequence. As in other organisms (*43*), the *R. hasegawae* U6 gene may be transcribed by RNA polymerase III, and if so, U6 gene expression will not require these sequences that are required for transcription by RNA polymerase II.

REFERENCES

1. Séraphin, B., Kretzner, L., and Rosbash, M. (1988) A U1 snRNA: pre-mRNA base pairing interaction is required early in yeast soliceosome assembly but does not uniquely define the 5′ cleavage site. *EMBO J.* **7**, 2533-2538
2. 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
3. Zhuang, Y. and Weiner, A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell* **46**, 827-835
4. Parker, R., Siliciano, P.G., and Guthrie, C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**, 229-239
5. Wu, J. and Manley, J.L. (1989) Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev.* **3**, 1553-1561
6. Zhuang, Y. and Weiner, A.M. (1989) A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.* **3**, 1545-1552
7. Reich, C.I., VanHoy, R.W., Porter, G.L., and Wise, J.A. (1992) Mutations at the 3′ splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. *Cell* **69**, 1159-1169
8. Newman, A. and Norman, C. (1991) Mutations in yeast U5 snRNA alter the specificity of 5′ splice-site cleavage. *Cell* **65**, 115-123
9. Newman, A.J. and Norman, C. (1992) U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. *Cell* **68**, 743-754
10. Wassarman, D.A. and Steitz, J.A. (1992) Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. *Science* **257**, 1918-1925
11. Wyatt, J.R., Sontheimer, E.J., and Steitz, J.A. (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
12. Cortes, J.J., Sontheimer, E.J., Seiwert, S.D., and Steitz, J.A. (1993) Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5′ splice sites *in vitro*. *EMBO J.* **12**, 5181-5189
13. Sontheimer, E.J. and Steitz, J.A. (1993) The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* **262**, 1989-1996
14. Kandels-Lewis, S. and Séraphin, B. (1993) Role of U6 snRNA in 5′ splice site selection. *Science* **262**, 2035-2039
15. Lesser, C.F. and Guthrie, C. (1993) Mutations in U6 snRNA that alter splice site specificity: Implications for the active site. *Science* **262**, 1982-1988
16. Guthrie, C. and Patterson, B. (1988) Spliceosomal snRNAs. *Annu. Rev. Genet.* **22**, 387-419
17. Nilsen, T.W. (1994) RNA-RNA interactions in the spliceosome: Unraveling the ties that bind. *Cell* **78**, 1-4
18. Crispino, J.D., Blencowe, B.J., and Sharp, P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science* **265**, 1866-1869
19. Tarn, W.-Y. and Steitz, J.A. (1994) SR proteins can compensate for the loss of U1 snRNP functions *in vitro*. *Genes Dev.* **8**, 2704-2717
20. Lamond, A.I., Konarska, M.M., Grabowski, P.J., and Sharp, P.A. (1988) Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. *Proc. Natl. Acad. Sci. USA* **85**, 411-415
21. Yean, S.L. and Lin, R.J. (1991) U4 small nuclear RNA dissociates from a yeast spliceosome and does not participate in the subsequent splicing reaction. *Mol. Cell. Biol.* **11**, 5571-5577
22. Tani, T. and Ohshima, Y. (1989) The gene for the U6 small nuclear RNA in fission yeast has an intron. *Nature* **337**, 87-90
23. Brow, D.A. and Guthrie, C. (1989) Splicing a spliceosomal RNA. *Nature* **337**, 14-15
24. Madhani, H.D. and Guthrie, C. (1994) Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* **28**, 1-26
25. Tani, T. and Ohshima, Y. (1991) mRNA-type introns in U6 small nuclear RNA genes: Implications for the catalysis in pre-mRNA splicing. *Genes Dev.* **5**, 1022-1031
26. Fabrizio, P. and Abelson, J. (1990) Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* **250**, 404-409
27. McPheeters, D.S. and Abelson, J. (1992) Mutational analysis of the yeast U2 snRNA suggests a structual similarity to the catalytic core of group I introns. *Cell* **71**, 819-831
28. Madhani, H.D. and Guthrie, C. (1992) A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71**, 803-817
29. Madhani, H.D. and Guthrie, C. (1994) Randomization-selection analysis of snRNAs *in vitro*: Evidence for a tertiary interaction in the spliceosome. *Genes Dev.* **8**, 1071-1086
30. Takahashi, Y., Tani, T., and Ohshima, Y. (1993) An mRNA-type intron is present in the *Rhodotorula hasegawae* U2 small nuclear RNA gene. *Mol. Cell. Biol.* **13**, 5613-5619
31. Hausner, T.-P., Giglio, L.M., and Weiner, A.M. (1990) Evidence

for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes Dev.* **4**, 2146-2156
32. Datta, B. and Weiner, A.M. (1991) Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature* **352**, 821-824
33. Wu, J. and Manley, J.L. (1991) Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature* **352**, 818-821
34. Sun, J.-S. and Manley, J.L. (1995) A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev.* **9**, 843-854
35. Frank, D.N., Roiha, H., and Guthrie, C. (1994) Architecture of the U5 small nuclear RNA. *Mol. Cell. Biol.* **14**, 2180-2190
36. Porter, G., Brennwald, P., and Wise, J.A. (1990) U1 small nuclear RNA from *Schizosaccharomyces pombe* has unique and conserved features and is encoded by an essential single-copy gene. *Mol. Cell. Biol.* **10**, 2874-2881
37. Patterson, B. and Guthrie, C. (1987) An essential yeast snRNA with a U5-like domain is required for splicing *in vivo*. *Cell* **49**, 613-624
38. Woodson, S.A. and Cech, T.R. (1989) Reverse self-splicing of the *Tetrahymena* group I intron: Implication for the directionality of splicing and for intron transposition. *Cell* **57**, 335-345
39. Augustin, S., Müller, M.W., and Schweyen, R.J. (1990) Reverse self-splicing of group II intron RNAs *in vitro*. *Nature* **343**, 383-386
40. Mörl, M. and Schmelzer, C. (1990) Integration of group II intron bl1 into a foreign RNA by reversal of the self-splicing reaction *in vitro*. *Cell* **60**, 629-636
41. Myslinski, E., Ségault, V., and Branlant, C. (1990) An intron in the genes for U3 small nucleolar RNAs of the yeast *Saccharomyces cerevisiae*. *Science* **247**, 1213-1216
42. Kretzner, L., Krol, A., and Rosbash, M. (1990) *Saccharomyces cerevisiae* U1 small nuclear RNA secondary structure contains both universal and yeast-specific domains. *Proc. Natl. Acad. Sci. USA* **87**, 851-855
43. Dahlberg, J.E. and Lund, E. (1988) The genes and transcription of the major small nuclear RNAs in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles* (Birnstiel, M.L., ed.) pp. 38-70, Springer-Verlag, Heidelberg
44. Steitz, J.A. (1992) Splicing takes a Holliday. *Science* **257**, 888-889