U2 RNA from Yeast Is Unexpectedly Large and Contains Homology to Vertebrate U4, U5, and U6 Small Nuclear RNAs

Manuel Ares, Jr.
Department of Molecular Biophysics and Biochemistry
Yale University School of Medicine
New Haven, Connecticut 06510

Summary

I have determined the structure of the gene from Saccharomyces cerevisiae coding for the yeast homolog of vertebrate U2 snRNA. Surprisingly, the RNA is 1175 nucleotides long, six times larger than U2 RNAs from other organisms, including Schizosaccharomyces pombe. Nearly 100 nucleotides of the large RNA share sequence homology and potential secondary structure with metazoan U2. The large RNA also contains homology to vertebrate U4, U6, and U6 snRNAs, implying a "poly-snRNP" structure for the RNP containing the large RNA. The gene LS71, encoding the large RNA, is essential for growth, suggesting that the yeast spliceosome can be dissected using genetic approaches. The different organization of spliceosomal RNA may underlie differences in splicing between yeast and metazoans.

Introduction

Small nuclear RNAs U1, U2, U4, U5, and U6 (U-snRNAs) are metabolically stable structural RNAs of the cell nucleus present at about 10^5-10^6 copies per cell (Weinberg and Ponman, 1988). Individual snRNAs are preserved in sequence among vertebrates and other metazoans, contain a trimethylguanosine cap structure at the 5' end (except U6, which contains an unidentified cap), and usually contain modified nucleotides at specific internal positions (see Reddy, 1986). The U-snRNAs are integral components of small nuclear ribonucleoprotein particles (U-snRNPs) that share a common core of polypeptides, but which differ with respect to the snRNA and accessory polypeptides they contain (Steitz et al., 1983; Brüngmann et al., 1983). The yeast Saccharomyces cerevisiae contains trimethylguanosine-capped small nuclear RNAs (Wise et al., 1983; Tollervey and Guthrie, 1985) and, although these RNAs are less abundant, as a family they seem more heterogeneous in size (up to about 1 kb) than the vertebrate U-snRNAs (Riadel et al., 1986).

The abundant family of metazoan U-snRNPs is required for nuclear pre-messenger RNA (pre-mRNA) splicing in vitro. Destruction of individual U-snRNPs using RNAaseH and oligonucleotides complementary to the snRNA shows that intact U1 and U2 snRNPs are individually required for splicing in vitro (Kramer et al., 1984; Krainer and Maniatis, 1985; Black et al., 1985). U4 and U6 are found in the same snRNP (Brüngmann et al., 1984; Hashimoto and Steitz, 1984; Rinko et al., 1985), and are also required for splicing (Black and Steitz, 1985). Antibodies directed against the U1 snRNP or against the snRNA trimethylguanosine cap (anti-TMG) or the Sm protein determinant present on the entire family of snRNPs also inhibit splicing in vitro (Yang et al., 1981; Padgett et al., 1983; Kramer et al., 1984; Fren dewey and Keller, 1985; DiMarino et al., 1985), although oddly an antibody against the U2 snRNP does not (Padgett et al., 1983).

The interactions of individual snRNPs with specific regions of the pre-mRNA have been mapped using the U-snRNP specific antibodies. The U1 snRNP interacts with the 5' splice site (Mount et al., 1983), the U2 snRNP with the branchpoint (Black et al., 1985; Chabot et al., 1985), and the U6 snRNP may interact with the 3' splice site (Chabot et al., 1985). In addition, there is genetic evidence for the interaction between U1 and the 5' splice site because a compensatory mutation in U1 suppresses a 5' splice site mutation (Zhuang and Weiner, 1986). In Hela cell nuclear extracts, splicing occurs in a 50-60S complex called the spliceosome, which requires U-snRNPs for its formation (Grabowski et al., 1985; Fren dewey and Keller, 1985; Bindslev and Green, 1986). Taken together, these studies imply that members of the abundant family of U-snRNPs recognize different regions of the pre-mRNA transcript and assemble as structural subunits of the spliceosome.

The yeast spliceosome (Brody and Abelson, 1985) also seems to contain essential RNA components. Yeast splicing extracts (Lin et al., 1985) contain a micrococcal nuclease-sensitive component and can be inactivated by incubation with the anti-TMG antibody (Cheng and Abelson, 1986). Spliceosome purification schemes enrich for the presence of three small anti-TMG precipitable RNAs (Pikielny and Rosbash, 1986; J. Abelson and C. Guthrie, personal communication) and in some cases at least one anti-TMG precipitable RNA greater than 1 kb in length (Pikielny and Rosbash, 1986; J. Abelson and C. Guthrie, personal communication; M. Rosbash, personal communication).

In this paper, I show that both Saccharomyces cerevisiae and the fission yeast Schizosaccharomyces pombe contain an RNA homologous to vertebrate U2 snRNA, a component of the vertebrate spliceosome (Konarska and Sharp, 1986; P. Grabowski and P. Sharp, personal communication; B. Chabot and J. Steitz, personal communication). The U2 homolog from S. pombe is the same size as vertebrate U2, while that from S. cerevisiae is unexpectedly large, and contains additional sequence homology to vertebrate U4, U5, and U6 snRNAs. These observations suggest that there may be substantial evolutionary variation in certain aspects of spliceosome structure and strict conservation of others, and may explain some of the differences in splicing between yeast and metazoans.

Results

Yeast Contains RNA with a 5' End Similar to Human U2 RNA

During a pilot experiment to determine the usefulness of an oligonucleotide complementary to human U2 ("L155"; Black et al., 1985) for transcription studies (Ares et al.,
1985; Mangin et al., 1986), I found that a commercial preparation of yeast tRNA, intended as a negative control, gave rise to oligonucleotide primed cDNA one nucleotide longer than that derived from vertebrate U2 RNA. To determine whether the yeast RNA was similar to U2, the RNA was sequenced using the L15 oligomer as a primer. The results are shown in Figure 1. Both Saccharomyces cerevisiae (Sc) and Schizosaccharomyces pombe (Sp) contain RNA that is complementary to the oligonucleotide and similar in sequence to human U2 (H). Moreover, the distance from the 5′ end of the RNA to the region complementary to the primer is very similar. Therefore, U2 RNA sequence is conserved in organisms as divergent as vertebrates and yeasts.

The U2 Domain of S. cerevisiae, but Not S. pombe, Is Carried on a Large RNA

To determine the size of the yeast U2 homologs, RNA from Hela cells, S. cerevisiae, S. pombe, and E. coli were separated on a denaturing gel, transferred to nylon membrane, and probed with labeled L15 oligonucleotide. The results are shown in Figure 2. As expected, L15 hybridizes to U2 RNA (189 nucleotides) from human cells (lanes 2 and 3). By comparison, the L15 complementary RNA from S. cerevisiae is huge (large arrow, lanes 5 and 6), migrating much more slowly than the largest marker (622 nucleotides; lanes 1 and 8). In contrast, the probe recognizes an RNA from S. pombe one to three nucleotides shorter than human U2 (based on shorter exposures of this blot; compare lanes 3 and 4). The bacterium E. coli does not contain detectable L15 complementary RNA (lane 7). Thus, while the primary sequence near the 5′ end of the U2 equivalent RNAs is conserved between man and yeast, the size of the RNA that carries the U2 domain is not necessarily conserved, even among yeasts.
Table 1. Oligonucleotides Used in This Study

<table>
<thead>
<tr>
<th>Oligomer</th>
<th>Length</th>
<th>Sequence (5'-3')</th>
<th>Complementary Positions (5'-3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>L15</td>
<td>15</td>
<td>CAGATACTACACTTG</td>
<td>43 to 29</td>
</tr>
<tr>
<td>15A</td>
<td>15</td>
<td>GAATCTCTTGGCCTT</td>
<td>3 to 17</td>
</tr>
<tr>
<td>21A</td>
<td>21</td>
<td>AGCCAAAAAGGCAAGAGATTC</td>
<td>23 to 3</td>
</tr>
<tr>
<td>17A</td>
<td>17</td>
<td>TTTTGTGAAATAGGAT</td>
<td>620 to 638</td>
</tr>
<tr>
<td>20B</td>
<td>20</td>
<td>GGATATATAATGGAGGTTCG</td>
<td>834 to 853</td>
</tr>
<tr>
<td>20A</td>
<td>20</td>
<td>GTAAAGTCAGAAGACGACTC</td>
<td>1134 to 1115</td>
</tr>
<tr>
<td>18A</td>
<td>18</td>
<td>GGGCCTTAGCTGTTCCC</td>
<td>1339 to 1356</td>
</tr>
<tr>
<td>19A</td>
<td>19</td>
<td>TGAGTACGGGCACTGC</td>
<td>1409 to 1391</td>
</tr>
<tr>
<td>M13</td>
<td>17</td>
<td>GTAACGACGGCCAGT</td>
<td>vector flanking polylinker</td>
</tr>
</tbody>
</table>

The Structure of the Gene Encoding the Large RNA

By analogy to the situation in vertebrates, where the U2 snRNP plays an essential role in splicing (Krainer and Maniatis, 1985; Black et al., 1985; Chabot et al., 1985) and has been identified as a component of the vertebrate spliceosome (Konarska and Sharp, 1986; P. Grabowski and P. Sharp, personal communication; B. Chabot and J. Steitz, personal communication), the yeast U2 homolog would be expected to be a component of the yeast spliceosome, possibly the same large RNA independently discovered by others (Pikis and Rosbash, 1986; J. Abelson and C. Guthrie, personal communication). Therefore, although this paper presents no evidence showing that the yeast U2 homolog actually is a component of the yeast spliceosome, the identity of the U2 domain (see below) makes the strong prediction that it is a part of the spliceosome. Hence the gene coding for the yeast U2 homolog should be called LSR1, for large spliceosomal RNA.

To isolate the gene LSR1 coding for the U2 homologous RNA of S. cerevisiae, the sequence obtained in Figure 1 was used to design two additional oligonucleotides (15A and 21A, see Table 1) to screen a library of S. cerevisiae genomic DNA fragments carried in phage lambda. Three independent overlapping clones were isolated and mapped using the oligonucleotides as probes. A 2.4 kb Sall-Smal fragment was subcloned from one of the phages into pUC13. Part of this fragment was sequenced according to the strategy outlined in Figure 3A, with additional oligonucleotides as needed (see Experimental Procedures; Table 1). The sequence is shown in Figure 3B.

The 5' end of the RNA was mapped to the DNA sequence of the template strand by comparing reverse transcripts of yeast RNA primed by labeled L15 oligomer (Figure 4A, lanes 5 and 6) to a sequence pattern generated by the same primer (lanes 1 through 4). The major cDNA product derived from yeast RNA comigrates with a product derived from dIddTTP termination in the sequencing reaction (large arrow; compare lane 4 with lanes 5 and 6). This position would code for an adenosine as the first nucleotide in the RNA, and is designated position 1. A minor product (small arrow), probably representing initiation at position -15, represents about 5% of the total.

The 3' end of the RNA was mapped using an antisense RNA probe (Melton et al., 1984). The probe was complementary to the RNA-like (nontemplate) strand of the gene between positions 1658 and 1082 (see Figure 5). After incubation in the presence of S. cerevisiae (lanes 2–5) or E. coli (lanes 6–9) RNA in mock reactions, hybrids were detected by protection from RNAases (Figure 4B). The probe is protected by hybridization to yeast RNA over a region of 95–98 or 90–93 nucleotides depending on the distribution of nucleosensitive sites. Therefore, the 3' end of the coding region is 90–98 nucleotides past position 1062, and the LSR1 gene codes for an RNA about 1175 nucleotides long. The RNA is not retained on oligo(dI) cellulose (data not shown), and therefore is not extensively polyadenylated. However, other posttranscriptional modifications at the 3' end of the RNA are not excluded.

The LSR1 gene contains sequence elements similar to those found in yeast snRNA genes (Figure 3B). A TATA box at position -90 (TATAAAAGG) is similar to those found in the same position of SnR3 and SnR10 (Tollervey et al., 1983; Tollervey and Guthrie, 1985). The sequence just past the 3' end of the coding region is a run of thymidine residues (24/28) similar to SnR3 (Tollervey et al., 1993). Past this is a sequence with striking dyad symmetry (CATGAATGAGCTTGATTG) which extends imperfectly an additional ten nucleotides to either side (Figure 3B).

The Yeast and Metazoan U2 Domains Are Similar in Primary Sequence and Potential Secondary Structure

Comparison of the U2 homologous sequence of the yeast RNA with the sequence of vertebrate U2 reveals several conserved features, and argues strongly that the large RNA is the yeast equivalent of vertebrate U2 snRNA. Figure 5A represents the sequence of the yeast, human, and Drosophila U2 domains drawn in a potential secondary structure similar to that proposed for Drosophila U2 by Keller and Noon (1985). The first seven nucleotides of the yeast U2 domain are not homologous to the first six of human or fly U2; thereafter the yeast RNA is identical to human U2 in 43 of the next 47 residues. Primary sequence homology is reduced beyond position 53 in all the U2 sequences, but all contain sequences that can be folded into a large hairpin with an internal loop (Figure 5A). Alternative structures for the 5' end of vertebrate U2 proposed originally (Reddy et al., 1981; Mattaj and Zoller, 1983) are not supported by the yeast sequence because two adenosines (positions 24 and 49) required to form a stem between nucleotides 18–26 and 44–51 in the vertebrate sequence are uridines in the corresponding positions (25 and 50) of the yeast RNA, and would greatly destabilize the proposed stem. The sequence 5'-CUCA-3' is conserved in the internal loop of the large stem, as is the local-
tion of the potential Sm binding site (see below) relative to the rest of the structure.

Sequences Found in Vertebrate U4, U5, and U6 Are Also Present in the Large RNA

To determine whether the large RNA contains sequences found in other vertebrate snRNAs, the sequence of the large yeast RNA predicted by the LSR1 gene was compared by computer to the sequences of vertebrate U1, U4, U5, U6, and U8 RNAs (Reddy, 1986) using the University of Wisconsin COMPARE and FIND programs. Figure 5B shows primary sequence homology between the large yeast RNA and vertebrate U4, U5, and U6 RNAs. Sequence homology between the yeast RNA and vertebrate U1 or U8 was not detected. There are no sequences perfectly complementary to either the strictly conserved 5′ splice site GUAGAG or the branchpoint consensus UAUAC of yeast pre-mRNAs (Teem et al., 1984).

A conserved sequence of the abundant snRNAs of metazoans is the domain A (Branlant et al., 1982) or “Sm binding site” with the consensus AU₄₆G. This sequence has been shown to be necessary for assembly of RNA into anti-Sm antibody precipitable RNP particles (Mattaj and De Robertis, 1985). Three perfect matches to the Sm bind-
The Gene Encoding the Large RNA Is Essential for Growth in Yeast

To determine the functional importance of the large yeast RNA, selectable null alleles of the \( LSR1 \) gene were constructed, introduced into diploid yeast by gene replacement, and tested for the ability to produce viable haploid spores after meiosis. Sequences from position −66 to 433 in the \( LSR1 \) gene, a region that includes the highly conserved U2 domain, were deleted and substituted with a 1.1 kb fragment containing the selectable \( URA3 \) gene. Homozygous \( ura3 \) diploid yeast was transformed with a linear DNA fragment derived from the \( LSR1 \) deletion-\( URA3 \) substitution clone, homologous at its ends to the \( LSR1 \) locus, to produce \( URA3 \) diploids containing one wild-type \( LSR1 \) allele and one deletion–substitution allele (transplacement heterozygotes) by the one-step gene disruption technique (Rotstein, 1983). Figure 6 (center) shows a blot of genomic DNA from the parent (lane A) and one of the transplacement heterozygotes (lane B). The structure of each strain at the \( LSR1 \) locus is shown in the diagram (left, wild-type; right, transplacement heterozygote) and is derived from additional genomic blots (not shown). The panels at the bottom show viability of spores dissected after meiosis and sporulation of the parent homozygous for the wild-type allele \( LSR1/LSR1 \) (left panel), and the transplacement heterozygote \( LSR1/lsr1\textsuperscript{Δ}URA \) (right panel). The heterozygote did not give rise to more than two viable spores, whereas the parent nearly always produced four viable spores. All viable spores from the heterozygote must contain the wild-type \( LSR1 \) allele, because none are \( URA3 \), the marker associated with the null mutation (not shown). This indicates that cells lacking a wild-type copy of the \( LSR1 \) gene are unable to survive. Microscopic examination of the dissected spores reveals that the inviable spores germinate and divide two to four times before dying. Diploids containing the disrupted allele with the \( URA3 \) gene inserted in the other orientation behave the same way (not shown). Therefore, the region of the genome containing the \( LSR1 \) gene is single copy and essential for spore germination or viability in yeast.

Discussion

I have described the structure of a gene from the yeast Saccharomyces cerevisiae that codes for an RNA homologous to vertebrate U2 snRNA. U2 RNAs from human, rat, mouse, frog, chicken, fruit fly (Reddy, 1986), S. pombe (Figures 1 and 2), and trypanosome (C. Tschudi, F. Richards, and E. Ullu, personal communication) are less than 200 nucleotides long. The U2 homologous RNA from S. cerevisiae is 1175 nucleotides long (Figures 3 and 4). The consensus U2 domain is about 100 nucleotides long and resides six to seven nucleotides from the 5' end of the RNA (Figure 5A). Sequence homology to U4, U5, and U6 snRNAs are also found in the large yeast RNA.
Structure of the U2 Domain

The highly conserved U2 domain could potentially be folded into a tripartite structure containing: one, a small stem and loop with an internal loop near the 5' end; two, an unpaired region conserved in primary sequence; and three, a large stem and loop with an internal loop (see Figure 5A; Keller and Noon, 1985). Although there are many compensatory base changes in the large stem, lending phylogenetic support for this part of the structure, there are few base changes in other parts of the U2 domain, and none of these are in paired nucleotides of the small stem. U2 sequences from more divergent organisms (Paco et al., 1986) will be necessary to confirm this part of the structure phylogenetically, but the striking degree of con-
Does a “Poly-snRNP” Constitute a Core Subunit of the Yeast Spliceosome?

Although evidence to address directly the function of the large RNA of S. cerevisiae is not yet available, the structural identity of the U2 domain (Figure 5A), the demonstrated role of the U2 snRNP in vertebrate splicing (Krainer and Maniatis, 1985; Black et al., 1985; Chabot et al., 1985), and the localization of U2 snRNA in the vertebrate spliceosome (Konarska and Sharp, 1986; P. Grabowski and P. Sharp, personal communication) together with the finding that a large RNA is enriched in yeast spliceosome fractions (Pikielny and Rosbash, 1986; J. Abelson and C. Guthrie, personal communication) predict that it is a component of the yeast spliceosome. As part of the spliceosome, the large RNA would be expected to exhibit additional functions not represented by RNA in the U2 snRNP of S. pombe or metazoans since the U2 domain from these organisms is carried on an RNA less than 190 nucleotides long (Figure 2; Reddy, 1986). The U4, U5, and U6 sequence homologies, though perhaps statistically less significant than that for U2, are found in a convincing context, raising the possibility that the large RNA is a part of an RNP equivalent to a complex of the U2, U5, and U4/U6 snRNPs in metazoans. If this is the case, some of the functional constraints on small RNAs, which must be individually transcribed and packaged in separate particles, may be relieved when they are cotranscribed, packaged together, and coexist in the same poly-snRNP particle, and direct structural homology would not necessarily be expected.

The large RNA may contain secondary interactions homologous to those known to occur within and between vertebrate snRNPs. Secondary structure models proposed for U4 and U6 account for the fact that U4 and U6 are in the same particle and can be dissociated under near physiological conditions (Bringmann et al., 1984; Hashimoto and Steitz, 1984). A site of interaction between vertebrate U4 and U6 RNAs within the U4/U6 snRNP has been mapped by psoralen crosslinking (Rinke et al., 1985), and there is potential to form a similar interaction between the U4 region (positions 513–519) and the U6 region (positions 156–163) of the yeast RNA (Figure 5B). There is also an 18 bp complementarity between part of the U4 homology (positions 499–516) and positions 727–735 of the yeast RNA (Figure 3B). The large RNA shares 10 of 11 contiguous nucleotides with vertebrate and pea U5 (Figure 5B). This 11 nucleotide stretch of U5 is the center of the only primary sequence homology shared among U5 RNAs from different organisms (except the Sm binding site). The U5 homology is in the major loop of the conserved U5 secondary structure and is the site of most of the modified nucleotides in U5 (Krol et al., 1981). The U5 homologous region of the yeast RNA can be drawn in a 22 nucleotide loop bounded by a 5 bp stem (positions 754–758 with 782–786), but not into an 11 nucleotide loop bounded by either a 5 bp stem (as in pea cells) (Krol et al., 1983) or an 8 bp stem (as in animal cells) (Branlant et al., 1983).

The large RNA contains much additional sequence not accounted for by the snRNA homologies. The function of these regions could include “linker RNA” for connecting different domains of the poly-snRNP (for example, note the low complexity of the sequences between positions 290 and 370 of the LSR1 gene, Figure 3B) or they could be regions represented in metazoans by as yet uncharacterized spliceosomal RNAs (P. Grabowski and P. Sharp, personal communication). Alternatively, the RNA 3' of the U2 domain may have a function performed by proteins in metazoan spliceosomes, or functions not required in splicing of metazoan mRNAs, or may even have no function at all.

A detailed higher order structure of the RNA in native form (for example see Moazed et al., 1986) would be useful in understanding the function of the different regions. The RNA almost certainly binds proteins in vivo, possibly modifying the configuration of the RNA. The small RNAs of the yeast spliceosome (Pikielny and Rosbash, 1986; J. Abelson and C. Guthrie, personal communication) could potentially interact with the large RNA and affect its conformation. The LSR1 gene appears to be essential (Figure 6), hence the functions of the various structural features of the large RNA may be evaluated by making and testing mutations in the LSR1 gene.
Does U2 Recognize the Branchpoint by Base Pairing?
By analogy to the role of the vertebrate U2 snRNP, the U2 domain of the yeast large spliceosomal RNA would be expected to associate with the branchpoint region of the pre-mRNA (Black et al., 1985; Chabot et al., 1985). Models involving recognition of the branchpoint region of pre-mRNA by base pairing to sequences in the first 55 nucleotides of U2 RNA within the U2 snRNP (Keller and Noon, 1984, 1985; Black et al., 1985) have yet to be tested. Although there are no perfect complements to the branchpoint consensus sequence UACUAAC anywhere in the large RNA, residues 32–39 of the yeast U2 domain (or 31–38 in vertebrate U2) could potentially interact with UACUAAC in a number of ways. One would include all of UACUAAC in base pairing interactions and would bulge the guanosine at position 34 of U2 (Black et al., 1985). Another would pair UACUAAC with a bulged second or third adenosine to positions 34–39 of U2. These possibilities are intriguing because they might identify or position the branchpoint 2'-OH for attack on the 5'-phosphate of the intron during the first step of splicing (see Keller, 1984).

If the conserved U2 domain is to pair with the branchpoint in yeast and metazoa, why is the branchpoint UACUAAC signal strictly conserved in yeast but not in metazoa? Perhaps the polypyrimidine stretch, an element of the metazoan 3' splice site not obviously conserved in yeast pre-mRNAs, is recognized by a component that potentiates the U2 snRNP interaction in metazoan pre-mRNAs lacking good complementarity to U2 RNA at the branchpoint. If base pairing between U2 and the branchpoint is important, it need not be extensive or perfect. Mutations in a number of positions of UACUAAC in yeast introns have surprisingly mild effects on the efficiency of splicing (Langford et al., 1984; Jaquier et al., 1985; Fouser and Friesen, 1986). The U2 RNA domain might also be involved in functions other than substrate recognition. For example, U2 might be a catalytic domain participating directly in the splicing reactions, or the U2 domain might bind factors requiring conserved sequences as binding sites.

Spliceosomes in Yeast and Metazoa
The dramatically different size of the U2 homologs (Figure 2) may reflect a difference in the organization of the spliceosome in yeast and metazoa, and offers a tempting explanation for the well-documented differences in pre-mRNA splicing in yeast and metazoa. Such differences include, but are not limited to, the requirement during in vitro splicing for the conserved branchpoint sequence UACUAAC (Newman et al., 1985; Vijayraghavan et al., 1986) and lack of an apparent polypyrimidine stretch in yeast (Rymond and Rosbash, 1985; Cellini et al., 1986), as opposed to the requirement for a polypyrimidine tract (Reed and Maniatis, 1985; Fredewey and Keller, 1985; Ruskin and Green, 1985), and promiscuous use of cryptic branchpoints defined less by sequence than by position of adenosine residues relative to the 3' splice site region in higher cells (Reed and Maniatis, 1985; Ruskin et al., 1985). Whatever the meaning of the different organization of spliceosomal RNA, the results suggest that spliceosomes share certain critical features such as the fine structure of the U2 RNA domain (Figure 5A), but retain differences in mechanism of substrate recognition (Newman et al., 1985; Rymond and Rosbash, 1985; Ruskin and Green, 1985; Cellini et al., 1986; Reed and Maniatis, 1985; Fredewey and Keller, 1985; Ruskin et al., 1985) and spliceosome assembly (Brody and Abelson, 1985; Grabowski et al., 1985; Fredewey and Keller, 1985; Bindereif and Green, 1986; Vijayraghavan et al., 1986). Since substrate requirements for stable spliceosome formation in vitro (Brody and Abelson, 1985; Vijayraghavan et al., 1986; Fredewey and Keller, 1985; Grabowski et al., 1985; Bindereif and Green, 1986), parallel the substrate requirements for the first splicing reaction in vitro (Newman et al., 1985; Rymond and Rosbash, 1985; Ruskin and Green, 1985; Cellini et al., 1986; Reed and Maniatis, 1985; Fredewey and Keller, 1985; Ruskin et al., 1985), the species-specific splicing differences must be manifest at an early stage of the splicing reaction, probably during spliceosome assembly (see also Ruskin et al., 1986).

Pre-mRNA structure (Toda et al., 1984; Hiraoka et al., 1984; Fukui and Kaziro, 1985) and U2 organization (Figures 1 and 2) of S. pombe are more like the metazoa than S. cerevisiae. This concurs nicely with the observation that S. pombe is able to splice at least one higher cell pre-mRNA correctly (Kaufner et al., 1985), and suggests that pre-mRNA splicing in S. pombe may be more similar to splicing in higher cells than in S. cerevisiae.

All spliceosomes must arrange the RNA precursor and the splicing enzymes so that splicing can occur. Because the reaction mechanism of nuclear pre-mRNA splicing seems to be conserved from yeast to metazoa (see Keller, 1984), the reaction center of the assembled spliceosome should contain conserved structures. The U2 RNA domain and the conserved U4, U5, and U6 sequences may be the first examples of such structures.

Experimental Procedures
Screening Lambda Plaques
A lambda library of yeast genomic fragments was constructed by M. Graham (Maynard Olson's lab, Washington University) using the left arm of lambda/009, the right arm of Charon 30, and a Sau3A partial digest of yeast DNA (strain AB1772, a derivative related to S288c and X2180-1B) and was kindly provided by Charlotte Hammond (Wesleyan University). Phages were plated on E. coli Q35B and duplicate filters were lifted. Filters were prehybridized 1 hr in and hybridized 10 hr in 5X SSC, 5X Denhardt's solution, 5 mM EDTA, and 50 μg/ml E. coli tRNA (Sigma) at 37°C with 106 cpm/ml kinase labeled oligonucleotide, and washed in 5X SSC at room temperature. One set of filters was probed with L15, the other with 21A. Areas of the lawn giving rise to a signal with both probes were eluted, plated, and rescreened. Of about 5 × 108 phages screened, three were detected and purified.

Oligonucleotides
Oligonucleotides used in this study are listed in Table 1 and were synthesized by Aino Riussala (Grothers Lab, Yale) except for the M13 primer, which was purchased from Collaborative Research. Oligonucleotides were purified by electrophoresis on 20% polyacrylamide-7 M urea gels, detected by UV shadowing, eluted in 10 mM Tris, 2.5 mM EDTA, and desalted on columns of P-4 in dH2O.
DNA Cloning and Sequencing

Construction of various LSR1 subclones in M13 phage and plasmid vectors (Messing, 1983) was performed using standard techniques (Maniatis et al., 1982). Sequencing was done according to Sanger et al. (1977) using the engineered Klenow polymerase ("Joyce" polymerase; Joyce and Grindley, 1983) according to the strategy shown in Figure 3A.

Analysis of RNA

RNA was extracted from S. cerevisiae strain YH-05 (Hottinger et al., 1982) and from S. pombe strain h' L927 by homogenization in the presence of phenol and glass beads essentially as described by Jensen et al. (1983). Hela RNA was isolated as described by Hernandez (1985). RNA from Hela cell Sm RNPs was isolated by immunoprecipitation of Hela cell nuclear extracts with monoclonal Sm antibody (Steitz et al., 1983). Primer extension using the L15 oligonucleotide was performed as described previously (Ares et al., 1985), except that MgCl2 was omitted from the annealing reaction and added later, and reverse transcription was done in the presence of 50 ^M ActinomycinD. Dideoxy sequencing with reverse transcriptase (AMV enzyme, Life Sciences) was done according to Wagner et al. (1981), except that ActinomycinD (40 ^g/ml) was included and a twentieth volume of 10 mM each dNTP was used for the chase. Reactions were performed for 10 min at room temperature, 20 min at 42°C, the chase mix was added and incubation at 42°C continued for 10 more min, and then for 20 min at 25°C in Northern blots. RNA was fractionated on 0.4 mm thick polyacrylamide-agar gels, and transferred electrophoretically to nylon membrane (Biobra Zeta-Probe) according to the recommendations of the manufacturer. The blots were baked and hybridized as described above. Synthesis of SP6 transcripts and ribonuclease protection mapping was done according to Melton et al. (1984) using ^P-CTP for uniform labeling of transcripts. The 3' end mapping probe (Figure 4B) was synthesized with Drai-digested DNA from a clone containing the Smal-HindIII fragment of the LSR1 gene cloned in the Smal-HindIII sites of pSP65.

Yeast Genetics

Selectable null alleles of the LSR1 gene were created in vitro and introduced into yeast. The 964 bp HindIII fragment (position—326 to 630) was cloned into the HindIII site of pUC13. The DNA was digested with Clal, and mixed with a purified 1.1 kb fragment carrying the URA3 gene. Ends were filled, ligated, and redigested with Clal and NruI to select against clones that had not taken an insert. Clones were isolated that substituted the 1.1 kb URA3 gene fragment in either orientation for LSR1 sequences between —86 and 433. Diploid NY836 was an H O/hs-fS-19(h·u·c·2, 112LEU2, ura-3-SU2-12·3 from P·Novick was transformed (Itt et al., 1983) with HindIII-cut DNA from the deletion-substitution plasmids and plated on minimal plates. Southern blot hybridization of DNA from uracil-independent strains was done using an LSR1 SP6 probe synthesized from Clal-digested DNA of an SP6 clone containing the HindIII·AsuI fragment in the HindIII·AccI sites. In most cases correct replacement of one LSR1 allele with the appropriate lpsCARNA allele generated in vitro occurred. Diploids containing the correctly replaced allele or the wild-type recipient diploid were sporulated. Tetrad were dissected on YPD plates and spores were germinated at 25°C or 30°C.

Acknowledgments

I am indebted to Alan Weiner for laboratory space and support, advice, and encouragement. Thanks to Ian Willis and Guido Kripp for generously providing yeast RNA samples, and Peter Novick for teaching me tetrat dissection. Thanks also to Kim Mowry for Sm RNA and Cathy Joyce for DNA polymerase. Thanks to Chris Tischord, Magda Konsa, Paula Grabowski, and members of the Abelson, Guthrie, Rostash, and Steitz labs for communicating results prior to publication. Special thanks to Joan Steitz and Nouria Hernandez for advice. Thanks to Doug Black, Benoit Chabot, Kim Mowry, Ellen Gottlieb, David Solnick, Joan Steitz, and Alan Weiner for critical comments on the manuscript. I was supported by a postdoctoral fellowship from the National Institutes of Health (GM 09418), and by NIH grants (GM 31073 and GM 31335) to Alan Weiner.

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

Received July 21, 1986; revised August 13, 1986.

References


Note in Proof

Using the L16 oligomer, it has been shown that an RNA that has been identified as a trimethylguanosine-capped nuclear RNA called snR90 is the same as that described here (C. Guthrie, personal communication; see Riedel et al., 1986). The work cited as Grabowski and Sharp, personal communication, has been published: Grabowski, P. J., and Sharp, P. A. (1986). Affinity chromatography of splicing complexes: U2, U5, and U4 – U6 small nuclear ribonucleoprotein particles in the spliceosome. Science 233, 1294–1299.