Randomization–selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome

Hiten D. Madhani and Christine Guthrie

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143-0448 USA

Putative components of the spliceosomal active site include a bulged helix between U2 and U6 snRNAs (U2–U6 helix I) and the adjacent ACAGAG hexanucleotide in U6. We have developed an in vivo, bimolecular randomization–selection method to functionally dissect these elements. Although a portion of U2–U6 helix I resembles the G-binding site of group I introns, the data are inconsistent with an analogous functional role for this structure in the spliceosome. Instead, analysis of several novel covariants supports the existence of a structure in which the helix I bulge engages in a tertiary interaction with the terminal residue of the U6 hexanucleotide (ACAGAG). Such a higher order structure, together with other known interactions, would juxtapose the two clusters of residues of the U2–U6 complex that are specifically required for the second chemical step of pre-mRNA splicing with the 3' splice site. Indeed, mutations in the residues that participate in the tertiary interaction affect both the efficiency and fidelity of 3' splice site usage.

[Key Words: Splicing, spliceosome; snRNA; U2; U6; yeast]

Received January 20, 1994; revised version accepted March 22, 1994.

The splicing of messenger RNA precursors (pre-mRNAs) occurs in the spliceosome, a large and dynamic ribonucleoprotein complex containing five small nuclear RNAs (U1, U2, U4, U5, and U6 snRNAs) and at least 50 proteins (for review, see Green 1991; Guthrie 1991; Rymond and Rosbash 1992; Moore et al. 1993). Similarities between the two-step chemical pathway of pre-mRNA splicing and that of group II autocatalytic introns have led to the proposal that the former is also fundamentally an RNA-catalyzed process mediated by the snRNA components of the spliceosome (Sharp 1985 1991; Cech 1986; Guthrie 1991). A number of direct RNA–RNA interactions involving the pre-mRNA substrate and the spliceosomal snRNAs have been demonstrated. It is well-established that recognition of the 5’ splice site and intron branchpoint involves Watson–Crick base-pairing with U1 and U2 snRNAs, respectively (for review, see Green 1991; Guthrie 1991; Rymond and Rosbash 1992; Moore et al. 1993). More recent studies indicate that U5 also interacts with the pre-mRNA, in this case with the 5’ and 3’ exons (Newman and Norman 1991, 1992; Wyatt et al. 1992; Sontheimer and Steitz 1993). A non-Watson–Crick interaction between the guanosines at the 5’ and 3’ splice sites has been proposed to function in the second chemical step of splicing (Parker and Siliciano 1993). Several of these interactions have potential counterparts in group II self-splicing introns (for review, see Weiner 1993; Chanfreau and Jacquier 1994).

Our studies have focused on the U6 spliceosomal snRNA (Madhani et al. 1990; Madhani and Guthrie 1992). This highly conserved molecule is associated with U4 snRNA through an extensive base-pairing interaction, which is disrupted prior to the first chemical step of splicing (for review, see Green 1991; Guthrie 1991; Rymond and Rosbash 1992; Moore et al. 1993). We have previously used a mutational approach in yeast to demonstrate the existence of a conserved base-pairing interaction between U6 and U2 snRNAs that is mutually exclusive with the U4–U6 interaction (Madhani and Guthrie 1992). In this pairing, termed U2–U6 helix I, a conserved sequence in U6 snRNA interacts with sequences in U2 that are just upstream of the branchpoint recognition region of U2. As a result, highly conserved and functionally important residues in U6 can be juxtaposed with the intron branchpoint. Residues that participate in this structure have been shown to be required for cell viability (Madhani et al. 1990), as well as for both chemical steps of splicing in vitro and in vivo (Fabrizio and Abelson 1990; Madhani and Guthrie 1992; McPheeters and Abelson 1992). These properties led us to propose a model for the active site of the spliceosome (Fig. 1) in which U2–U6 helix I might participate directly in chemical steps of splicing (Madhani and Guthrie 1992). Because this helix is mutually exclusive with stem I of the U4–U6 interaction, its formation offers a mechanistic rationale for the destabilization of U4–U6 prior to the chemical steps of splicing (Madhani and Guthrie 1992). Further evidence consistent with a direct
Figure 1. RNA–RNA interactions between the intron and U2 and U6 snRNAs. *Saccharomyces cerevisiae* snRNA sequences are shown. Intron sequences are the *S. cerevisiae* consensus. Depicted are the U2–branchpoint region helix [Parker et al. 1987; Wu and Manley 1989; Zhuang and Weiner 1989], U2–U6 helix I [Madhani and Guthrie 1992], and the U6–5′-splice-site helix (Sawa and Abelson 1992, Wassarman and Steitz 1992, Kandels-Lewis and Séraphin 1993; Lesser and Guthrie 1993b). Residues in U2 and U6 required at or prior to the first chemical step of splicing are underlined; those involved in step 2 are circled [Fabrizio and Abelson 1990, Madhani and Guthrie 1992, McPheeters and Abelson 1992].

role for this structure in catalysis is suggested by in vitro cross-linking experiments (Sawa and Shimura 1992, Sawa and Abelson 1992; Wassarman and Steitz 1992, Sontheimer and Steitz 1993) and, more recently, genetic suppression experiments [Kandels-Lewis and Séraphin 1993; Lesser and Guthrie 1993b] that demonstrate base-pairing between the highly conserved ACA sequence upstream of U2–U6 helix I and the 3′ portion of the 5′ splice site consensus (Fig. 1A). Notably, the three RNA–RNA interactions depicted in Figure 1 provide a structural mechanism for juxtaposing the branchpoint nucleophile with the 5′ splice site.

U2–U6 helix I exhibits intriguing similarities to structures in self-splicing introns [for review, see Moore et al. 1993; Weiner 1993]. We have noted previously a resemblance to domain 5 of group II introns: Both structures are highly conserved, lie just upstream of helices involved in branchpoint recognition, and contain 2-nucleotide bulges that interrupt the 3′ side of their respective helices [Michel et al. 1989b; Madhani and Guthrie 1992]. A different analogy was suggested by an in vitro mutational analysis of the AGA sequence in yeast U2 snRNA [McPheeters and Abelson 1992; nucleotides 25–27 in Fig. 1, note that A25 is bulged]. On the basis of structural and phenotypic similarities to the AGA sequence in the P7 helix of group I introns, which contains the binding site for the guanosine [G] cofactor (Michel et al. 1989a), it was proposed that the essential role of the AGA sequence in U2 snRNA in the second chemical step of pre-mRNA splicing is to bind to the conserved G found at the 3′ splice site of nuclear introns [McPheeters and Abelson 1992]. This putative spliceosomal G-binding motif has been cited as a possible case of convergent evolution [Weiner 1993]. Moreover, it has recently been proposed based on stereochemical studies that the spliceosome generates a group I-like catalytic site to execute the second step of splicing [Moore et al. 1993].

Evaluation of the possible roles of snRNAs in catalysis and their relationship to components in self-splicing introns will ultimately require knowledge of the higher order foldings of the structures in which they participate.

In view of the large and dynamic nature of the spliceosome, a potentially powerful strategy for deriving such structural constraints is phylogenetic covariation analysis; this approach has already been applied successfully to diverse RNA molecules [Levitt 1969; Noller and Woese 1981; Woese et al. 1983; Michel et al. 1989b, Michel and Westhof 1990, Romero and Blackburn 1991]. In the case of group I introns [Michel and Westhof 1990], sufficient long-range constraints have been obtained in the form of tertiary interactions to allow the determination of a robust three-dimensional structural model, key aspects of which have been confirmed experimentally [for review, see Cech 1993]. Comparably detailed models are not yet available for group II introns.

Unfortunately, almost all of the U2 and U6 residues involved in the proposed active site model of the spliceosome [Fig. 1] are phylogenetically invariant, precluding a strictly analogous approach to the identification of tertiary structure [Guthrie and Patterson 1988; C. Guthrie, S. Mian, and H. Roiha, unpubl.]. In other systems this limitation has been overcome through the generation of “artificial phylogenies”: novel functional variants selected in vitro from pools of randomized sequences [for review, see Gold et al. 1993; Szostak and Ellington 1993]. This approach has been exploited to identify functionally important tertiary RNA interactions in the Rev-responsive element of human immunodeficiency virus type 1 [Bartel et al. 1991] and in the *sunY* group I ribozyme (Green and Szostak 1994). However, there are significant impediments to the application of such methods to the spliceosome. First, the region of interest (Fig. 1) is derived from several different RNA molecules, yet current in vitro methods do not allow for the selection of functional pairs of variants (intermolecular covariants). Second, the efficient reconstitution of splicing in vitro with multiple synthetic RNAs and the iterative selection of functional variants are likely to be difficult and laborious tasks.

To circumvent these limitations, we have employed a method that allows for the single-step selection in vivo of functional covariants of two interacting molecules.
from gene pools containing synthetically randomized residues. We have applied this procedure to functionally dissect residues in the U2–U6 helix I region that are involved in the second chemical step of splicing [Fig. 1]. The properties of the selected variants do not conform to those predicted by the G-binding site model for the AGA sequence in U2. Instead, analysis of several novel variants supports the existence of a structure in which the conserved A residue of the helix I bulge [AGA; U2–A25] forms a tertiary interaction with the terminal residue of the U6 hexanucleotide (ACAGAG; U6–G52). To our knowledge this represents the first evidence for higher order RNA structure involving spliceosomal snRNAs. We discuss the implications for the active site involved in the second chemical step of splicing.

Results
Randomization–selection strategy

Our strategy is summarized in Figure 2 [for details, see Materials and methods]. We first constructed libraries of yeast U6 and U2 genes in which particular residues had been randomized through chemical synthesis. These were used to cotransform a yeast strain, YHM118, in which both chromosomal copies of U2 and U6 are deleted and complemented by a single URA3-marked centromere plasmid that contains both wild-type genes. The cotransformants were then replica-plated to media containing 5-fluoro-orotic acid (5-FOA), which selects for transformants and complemented by a single yeast U6 and U2 genes in which particular residues had been randomized through chemical synthesis. These were used to cotransform a yeast strain, YHM118, in which both chromosomal copies of U2 and U6 are deleted and complemented by a single URA3-marked centromere plasmid that contains both wild-type genes. The cotransformants were then replica-plated to media containing 5-fluoro-orotic acid (5-FOA), which selects for the terminal residue of the U6 hexanucleotide (ACAGAG; U6–G52). To our knowledge this represents the first evidence for higher order RNA structure involving spliceosomal snRNAs. We discuss the implications for the active site involved in the second chemical step of splicing.

Results
Randomization–selection strategy

Our strategy is summarized in Figure 2 [for details, see Materials and methods]. We first constructed libraries of yeast U6 and U2 genes in which particular residues had been randomized through chemical synthesis. These were used to cotransform a yeast strain, YHM118, in which both chromosomal copies of U2 and U6 are deleted and complemented by a single URA3-marked centromere plasmid that contains both wild-type genes. The cotransformants were then replica-plated to media containing 5-fluoro-orotic acid (5-FOA), which selects against the URA3-marked wild-type U2–U6 plasmid (Boeke et al. 1987). Colonies that grow on 5-FOA contain U2 and U6 functional covariants that can then be isolated and analyzed by DNA sequencing. In deciding what residues to randomize in the libraries, we reasoned that residues involved in the same step of splicing would be the most likely to interact with each other. In the U2 and U6 libraries described here, regions that encompass residues required for the second chemical step of splicing were randomized [Fig. 1, nucleotides 51, 52, 58, 59 in U6 and nucleotides 23–26 in U2]. In addition, because two of the U2 residues randomized are thought to participate in an alternative intramolecular structure, U2 stem I (Guthrie and Patterson 1988; McPheeters and Abelson 1992), their putative base-pairing partners were also randomized [nucleotides 9 and 10]. The total number of residues randomized (in U2 and U6) was limited to 10. If only one combination of these nucleotides is functional, we expect it to be represented once in ~10⁴ variant pairs (4¹⁰= 10⁴), which approaches the upper limit of what is practical to screen in yeast [see Materials and methods].

We selected for functional cotransformants on 5-FOA under two different conditions, which represent arbitrary (but experimentally convenient) thresholds. First, we isolated non-temperature-sensitive variants by demanding growth after 2 days at 30°C and screening replica plates for temperature sensitivity at 37°C. Second, we reduced the stringency of selection by isolating temperature-sensitive variants. This was accomplished by selecting positives after 4 days, followed by screening of replica plates at 37°C. The 60 individual variant pairs (25 non-temperature-sensitive and 35 temperature-sensitive) obtained are listed in Figure 3. Diagrams showing the nucleotides recovered at each randomized position are shown in Figure 4. These sequences were visually inspected for nucleotide covariations.

Analysis of the selected variants

In the conserved hexanucleotide in U6, the last two positions were randomized [ACAGAG; positions 51 and 52]. In all variants, the wild-type A residue was observed at position 51 [Fig. 4]. In contrast, all three possible nucleotide changes were observed at position 52. The isolation of U6–G52U was not surprising, because this allele is viable at 30°C as a single mutation [Madhani et al. 1990]. However, the isolation of the other changes, U6–G52A and U6–G52C, was unexpected as these are unconditionally lethal as single mutations [Madhani et al. 1990]. We will return to the basis for this result in the next section.

In helix I, the residues in U2 and U6 flanking and including the 2-nucleotide bulge that separates helix Ia and helix Ib were randomized. We have previously reported evidence for base-pairing between U6–C58 and U2–G26, as well as between U6–A59 and U2–U23 [Fig 1; Madhani and Guthrie 1992]. The role of U2–G26 is of particular interest because this residue is the most critical in the proposed G-binding site described in the introductory section [McPheeters and Abelson 1992]. In the case of the G-binding site of group I introns, the

**Figure 2.** Strategy for in vivo selection of functional covariants of two interacting molecules. See Materials and methods for details.
A

| Variant | U6 (5'→3') | U2 (5'→3') |
|---------|------------|------------|
| WT      | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 030     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 031     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 032     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 033     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 034     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 035     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 036     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 037     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 038     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 039     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 040     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 041     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 042     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 043     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 044     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 045     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 046     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 047     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 048     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 049     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 050     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 051     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 052     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 053     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 054     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 055     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 056     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 057     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 058     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |

B

| Variant | U6 (5'→3') | U2 (5'→3') |
|---------|------------|------------|
| WT      | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 030     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 031     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 032     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 033     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 034     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 035     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 036     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 037     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 038     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 039     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 040     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 041     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 042     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 043     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 044     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 045     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 046     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 047     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 048     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |
| 049     | ACAGAAGAUGAGCA | UCUUUGCCUUUUGCCUGAAGUC |

**Figure 3.** Selected variant pairs. [A] Sequences of non-temperature-sensitive variant pairs selected after 2 days on 5-FOA at 30°C. [B] Sequences of temperature-sensitive variant pairs selected from the U2 and U6 libraries after 4 days on 5-FOA at 30°C. Boldface type indicates changes from wild type. Asterisks (*) indicate variants containing otherwise lethal mutations at U6 nucleotide 52. Prime (') entries are temperature-sensitive variants selected originally after 2 days on 5-FOA at 30°C.

**Figure 4.** Nucleotides recovered in the variant pairs. [A] Shown in the context of the U2-U6 structure are the nucleotides recovered at each randomized position in the variants shown in Fig. 3A. Large type indicates that only the wild-type nucleotide was recovered at that position. [B] Nucleotides recovered at each randomized position in the variants shown in Fig. 3B.

**Table 1.** Base appositions seen between U6 nucleotide 58 and U2 nucleotide 26 in functional variants shown in Fig. 3

| Variant class | Watson–Crick | G-U wobble | Other* |
|---------------|--------------|------------|-------|
| Non-temperature-sensitive variants | 25 | 0 | 0 |
| Temperature-sensitive variants | 26 | 5 | 4 |

Entries refer to the number of covariants of each type. *All are U2–G26A/U6–C58
Evidence for a tertiary interaction in the spliceosome

The phenotypes of these strains are summarized in Table 2 and shown in Figure 5A, for plates incubated at 30°C. As can be seen by examining the diagonals in both Figure 5A and Table 2, all Watson–Crick appositions support cell growth under the conditions tested. Both G-U wobble pairs are functional, though one (U6-C58/U2-G26) is temperature sensitive at 37°C (Fig. 5A, Table 2). One of the two A-C appositions confers slow, temperature-sensitive growth (Fig. 5A, Table 2); interestingly, rare A-C pairs that replace G-U pairs in rRNA phylogeny are always found at helical junctions (Gutell 1993). No other nucleotide apposition is functional at this position under any of the conditions tested (Fig. 5A, Table 2).

Previously, it was shown that double mutants that affect both G26 and A27 in U2 result in a more severe phenotype in vitro than single mutations at either position; a complete inhibition of the second step of splicing was observed (McPheeters and Abelson 1992). This observation was rationalized in terms of a recent “axial” G-binding site proposal for group I introns in which both the G and 3' A of the AGA sequence interact with the G substrate (Yarus et al. 1991a, b, McPheeters and Abelson 1992). However, given the results described above, an alternative possibility is that the severe defect of double mutants at nucleotides 26 and 27 in U2 instead might result from a defect in base-pairing with U6. To test this possibility, we constructed and analyzed a double transversion at these positions (U2-G26C, A27U). As expected, in the presence of wild-type U6, this mutant is lethal at 25°C, 30°C, and 37°C (Fig. 5B shows plates incubated at 30°C). In contrast, in the presence of a U6 allele that restores base-pairing (U6-U57A, C58G), wild-type growth is restored under all conditions tested (Fig. 5B). Moreover, suppression between the U2 and U6 alleles is mutual because the U6 allele is also lethal in the presence of wild-type U2 (Fig. 5B). Taken together with the results shown in Figure 5A and summarized in Table 2, the data indicate that base-pairing at these positions, rather than specific sequence, is crucial for cell viability and, presumably, the second step of splicing.

The other base pair randomized consists of U6 nucleotide 59 and U2 nucleotide 23. We demonstrated previously that the deleterious growth phenotypes of U6-A59G, a lethal allele, and U6-A59C, a temperature-sensitive allele, can be specifically suppressed by U2-U23C and U2-U23G, respectively. However, in the presence of wild-type U6, the U2 mutations have no effect on cell growth, although they exhibit a measurable defect in the second step of splicing in vivo (Madhani and Guthrie 1992). Thus, base-pairing at this position, although demonstrable, is not essential for cell growth. We expected, therefore, that the selected variants containing the wild-type nucleotide at U6 position 59 would contain all possible nucleotides at position 23 in U2, whereas those that contained U6-A59G would only be found in combination with the compensating U2-U23C mutation.

Table 2. Phenotypes of base appositions between U6 nucleotide 58 and U2 nucleotide 26

| U6 nucleotide 58 | U6 nucleotide 26 |
|------------------|------------------|
| **U2 nucleotide** | **G** | **C** | **U** | **A** | **G** |
| **U2 nucleotide** | **G** | + | ts<sup>a</sup> | - | - |
| 26                | A    | ts<sup>b</sup> | + | - | - |
| U                 | -    | - | + | + |
| C                 | -    | - | + | + |

YHM118 derivatives containing the indicated U2 and U6 allele combinations assayed on 5-FOA after 3 days at 25, 30, and 37°C. [+ ] Wild-type growth at 25, 30, and 37°C; [- ] no growth at 25, 30, and 37°C.

<sup>a</sup>Wild-type growth at 25 and 30°C and no growth at 37°C.

<sup>b</sup>Slow growth at 25 and 30°C and no growth at 37°C. The wild-type nucleotides are indicated in bold.

Figure 5. (A) Analysis of the U6-C59/U2-G26 base pair. Growth on 5-FOA of YHM118 derivatives harboring the indicated U2 and U6 alleles at nucleotides 26 and 58, respectively. Plates were incubated at 30°C for 3 days. (B) Analysis of double mutants in U2–U6 helix Ia. Growth on 5-FOA of YHM118 derivatives harboring the indicated U2 and U6 alleles at nucleotides 26–27 and 57–58, respectively. Plates were incubated at 30°C for 3 days.
This is what we observed (Figs. 3 and 4; Table 3). The U6-A59C mutation was not found among the non-temperature-sensitive variants even in combination with the predicted compensatory mutant U2–U23G (Table 3). It does appear in the temperature-sensitive variants but not always in combination with U2–U23G (Table 3). This latter result is expected because in the presence of wild-type U2, U6–A59C exhibits a deleterious effect on growth only above the 30°C temperature used in our selection experiments [Madhani and Guthrie 1992].

In an effort to better understand the structural requirements at this base pair, we used site-directed mutants to construct YHM118 strains containing all possible appositions between U6 nucleotide 59 and U2 nucleotide 23. The growth phenotypes of these strains on 5-FOA are shown in Table 4. Specific suppression of U6–A59C by U2–U23G and of U6–A59G by U2–U23C was observed (Table 4), confirming the occurrence of base-pairing at this position. Because the U6–A59C/U2–U23G combination supports growth at all temperatures tested, the reason for its absence in the non-temperature-sensitive selected variants is unknown but could reflect a defect that becomes apparent only in combination with other changes found in the variants. The U6–A59U site-directed mutation, which was also absent from the selected variants, was found to be inviable, even in combination with the compensatory mutation U2–U23A. The observation that base-pairing is not sufficient to confer growth is consistent with our previous suggestion for an additional role for U6–A59 [Madhani and Guthrie 1992].

Analysis of the variants revealed that the 2-nucleotide bulge [U2 nucleotides 24–25] that separates helix Ia and helix Ib can be replaced by a wide variety of dinucleotides (15/16 of the possible dinucleotides are found; Figs. 3 and 4). To further delineate the minimal features of the bulge required for growth, we constructed single and double nucleotide deletions of the bulge in addition to all possible single nucleotide substitutions. As expected from the selection data, all substitutions in U2 nucleotides 24 and 25 were fully viable (Table 5). Single nucleotide deletions are tolerated, whereas deletion of both nucleotides 24 and 25 is lethal (Table 5). The results at position 25 are consistent with the mild in vitro phenotypes of nucleotide substitutions and a point deletion of U2–A25 [McPheeters and Abelson 1992]. Finally, we engineered single nucleotide insertions in U6 at the kink (between nucleotides 58 and 59) opposite the bulge; all four insertions are lethal (Table 5). Thus, the bulge appears to be moderately tolerant to mutational insult.

| Table 3. Base appositions seen between U6 nucleotide 59 and U2 nucleotide 23 in functional variants shown in Fig. 3 |

| Identity of U6 nucleotide 59 | W-C | Other |
|-----------------------------|-----|-------|
| A (wild type)                | 21  | 24    |
| G                            | 6   | 0     |
| C                            | 1   | 8     |
| U                            | 0   | 0     |

In addition to the regions described above, U2 nucleotides 9 and 10 were randomized in the U2 library to test the importance of base-pairing in the lower portion of U2 stem I. However, because many functional variants are unpaired at either or both randomized positions (Fig. 3), it appears that neither the U2–C9/G26 pair nor the U2–U10/A25 pair is essential for growth. One variant (TS 109) contains a deletion that encompasses this region. These observations are in agreement with previous in vitro results in which no requirement for base-pairing in the lower part of the stem was observed [McPheeters and Abelson 1992].

| Table 4. Phenotypes of base appositions between U6 nucleotide 59 and U2 nucleotide 23 |

| U6 nucleotide 59 | U2 nucleotide | G | C | A | U |
|------------------|---------------|---|---|---|---|
| nucleotide 23    | G             | + | - | + | - |
|                  | A             | - | - | + | - |
|                  | U             | ts | - | + | - |
|                  | C             | - | - | + | + |

YHM118 derivatives containing the indicated U2 and U6 allele combinations assayed on 5-FOA after 3 days at 25, 30, and 37°C. (+) Wild-type growth at 25, 30, and 37°C; (−) no growth at 25, 30, and 37°C.

Unexpected variants containing lethal mutations at U6 position 52

As mentioned above, several unexpected temperature-sensitive variants were isolated from the libraries. These contain either an A or a C at nucleotide 52 in U6, which

| Table 5. Phenotypes of mutants in the bulge region of U2–U6 helix I |

| Mutant         | Growth |
|----------------|--------|
| U2–U24C        | +      |
| U2–U24A        | +      |
| U2–U24G        | +      |
| U2–A25G        | +      |
| U2–A25C        | +      |
| U2–A25U        | +      |
| U2–AU24        | +      |
| U2–AA25        | +      |
| U2–AU24, AA25  | −      |
| U6–i58G        | −      |
| U6–i58A        | −      |
| U6–i58U        | −      |
| U6–i58C        | −      |

YHM118 derivatives containing the indicated U2 or U6 allele assayed on 5-FOA after 3 days at 25, 30, and 37°C. U2 mutants assayed in combination with wild-type U6. U6 mutants assayed in combination with wild-type U2. (+) Wild-type growth at 25, 30, and 37°C; (−) no growth at 25, 30, and 37°C. i58 represents the insertion of the indicated nucleotide after U6 nucleotide 58.
we have shown previously to be lethal as single mutations (Madhani et al. 1990). This discrepancy is not the result of a difference in the yeast strains employed in the two studies because, as single changes, G52A and G52C are lethal at all temperatures tested in YHM118 in the presence of wild-type U2 [data not shown]. Therefore, in the viable variants, the lethality of G52A and G52C is being suppressed by additional mutations. Because U6–A51 does not deviate from wild type in any of the variants, mutations in the bulge region or in U2 nucleotides 9–10 must be responsible for suppression, because they are the only other randomized residues in the libraries. A single variant was isolated that contains U6–G52A, nine isolates contain U6–G52C. We describe an analysis of the U6–G52A-containing variant first.

The sequences of U2 and U6 in the randomized region of the single U6–G52A variant is shown in Figure 6A. In addition to G52A, this variant contains a transversion of the U6–C58/U2–G26 base pair (a flip to U6–C58G/U2–G26C) and mutations in the helix I bulge region as well as mutations in nucleotides 9 and 10 in the 5′ region of U2 stem I. To determine which of the additional mutations were responsible for suppression, we performed two different selection experiments (secondary selections). We first cotransformed YHM118 with plasmid DNA encoding U6–G52A together with the U2 library and selected for functional variants at 30°C for 4 days on 5-FOA media, as in the initial selection experiment (for details, see Materials and methods). However, no functional variants were obtained. We reasoned that this failure might be the result of a requirement for the flip in the U6–C58/U2–G26 base pair. When we attempted to select variants by cotransforming YHM118 with the double U6 mutant [U6–G52A, C58G] and the U2 library, positives were easily obtained. All are temperature sensitive at 37°C. Eighteen were analyzed by DNA sequencing [Fig. 6B]. Strikingly, all 18 contain a G residue at nucleotide 25 in the helix I bulge but no consistent changes at nucleotide 9, 10, or 24. As expected, all contain a G26C change in U2 that maintains base-pairing with U6–C58G. These results demonstrate that three changes are necessary for suppression of U6–G52A: U2–A25G and the flip of the U6–C58/U2–G26 base pair. To confirm these requirements using site-directed mutations, we constructed YHM118 strains containing U2 and U6 mutants in which U6–G52A was combined with U2–A25G, the transversion, or both. Although neither U2–A25G nor the flip alone is able to support growth on 5-FOA of U6–G52A under any condition tested (25–37°C), the combination of both results in slow growth at 25°C, 30°C, and 33°C [data not shown]. To test whether or not gross alterations of the bulge could suppress U6–G52A, we combined it with each of the viable bulge mutations listed in Table 5, including the two single-nucleotide bulge deletions. In none of the eight cases was suppression observed [data not shown].

In our initial screen, we identified nine variants containing U6–G52C. In this case, two classes of variants can be discerned on the basis of shared sequence motifs [Fig. 7A]. Class I can be divided into three subsets (IA, IB, and IC), all of which are characterized by the occurrence of one or more G residues in the helix I bulge. The four class IA variants all contain U2–U23C and U2–A25G, the former change disrupts the first base pair in U2–U6 helix I b. The two class IB variants also contain U2–U23C, but instead of a G at nucleotide 25 in U2, they contain a G at nucleotide 24 and a U at nucleotide 25 [Fig. 7A]. The single class IC variant contains, among other changes, U2–A25G and a flip of the U6–C58/U2–G26 base pair. In these respects it mimics the U6–G52A suppressors. The two members of the class II G52C-containing variants are grouped together because they share a subset of changes: Both contain U2–A25C, a flip of the U6–C58/U2–G26 base pair, as well as either a transition or a transversion in the U6–A59/U2–U23 base pair [Fig. 7A].

We performed a secondary selection, starting with U6–G52C. YHM118 was cotransformed with this mutant and the U2 library, and functional variants were selected on 5-FOA at 30°C. On the basis of our initial classification, we expected to obtain representatives of only class IA and class IB, as the other classes contain a flip of the U6–C58/U2–G26 base pair. Twelve functional variants, all of which are temperature sensitive at 37°C, were analyzed [Fig. 7B]. As predicted, 11 of the 12 match the consensus from class IA [U23C and A25G], the single exception contains U23C and U24G and, therefore, mimics class IB. Three of the class IA-like suppressors are also labeled class IB because these variants contain a G residue at the first position of the bulge. Further analysis of the class IC variant is described below.

Figure 6. Analysis of G52A-containing variants. (A) The single variant pair isolated from the U2 and U6 libraries that contains the lethal G52A mutation. (B) U2 alleles isolated from the U2 library after cotransformation of YHM118 with U6–G52A, C58G and selection at 30°C for 4 days on 5-FOA plates.
In principle, the mutations in the bulge region of helix I could function in suppression directly by affecting a physical interaction between this region and nucleotide 52 in U6 or, more indirectly, by forming or disrupting a different interaction. To address this issue, we asked whether or not bulge-region variants that suppress U6–G52C could also suppress U6–G52A and vice versa. Because the G52A-containing variants mimic the structure of the class IC U6–G52C-containing variants, we expected that the former class of bulge-region variants might at least weakly support growth in combination with U6–G52C (and vice versa). In contrast, we expected that the class IA, class IB, and class II U6–G52C-containing variants would not support growth when combined with U6–G52A, because these did not contain the required flip of the U6–G58/U2–G26 base pair found in the G52A-containing variants.

To attempt to dissect the class II G52C-containing variants, we constructed a pair of U2 and U6 libraries that contain the lethal G52C mutation. [A] Four classes of variants isolated from the U2 and U6 libraries that contain the lethal G52C mutation. [B] U2 alleles isolated from the U2 library after cotransformation of YHM118 with U6–G52C and selection at 30°C for 4 days on 5-FOA plates.

### Specificity of suppression

In principle, the mutations in the bulge region of helix I could function in suppression directly by affecting a physical interaction between this region and nucleotide...
Evidence for a tertiary interaction in the spliceosome

Table 6. Specificity of suppression of G52A and G52C by bulge-region variants

| Bulge-region sequencea | U6 nucleotide 52 |
|------------------------|------------------|
|                        |                  |
| U6                     | U2              | isolated as a suppressor of |
|                        |                  |
| WT                     | WT              | N.A.                      |
| C58G                   | TS100           | G52A                      |
| C58G                   | A52SUP1         | G52A                      |
| C58G                   | A52SUP17        | G52A                      |
| C58G                   | A52SUP4         | G52A                      |
| C58G                   | A52SUP9         | G52A                      |
| WT                     | TS086           | G52C (IA)                 |
| WT                     | C52SUP10        | G52C (IA)                 |
| WT                     | C52SUP15        | G52C (IA)                 |
| WT                     | C52SUP17        | G52C (IB)                 |
| WT                     | TS088           | G52C (IC)                 |
| C58G                   | TS092           | G52C (IC)                 |
| C58G                   | TS104           | G52C (II)                 |

|                      |                  |
| G                  | A     | C     | U     |
| wt                  | ++    | +     | +     | +     |
| wt                  | 30+   | +     | 25+   | +/-   |
| wt                  | 33++  | +     | 30+   | +/-   |
| wt                  | 30++  | +     | 30+   | +/-   |
| wt                  | 30+   | +     | 25+   |        |
| wt                  | 35+/- | 30+   | 35+   |        |
| wt                  | --    | 30+   | --    |        |
| wt                  | --    | 33+/- | --    |        |
| wt                  | --    | 35+/- | --    |        |
| wt                  | --    | --    | --    | --    |
| wt                  | 30+   | 30+   | 33+   | ++    |
| wt                  | --    | 35+   | --    | --    |
| wt                  | --    | --    | --    | --    |
| wt                  | 30+   | 30+   | 30+   | +/-   |
| wt                  | --    | 35+   | --    | --    |
| wt                  | --    | --    | --    | --    |

The highest temperature and degree of growth of the indicated allele combination that was observed is shown (relative to wild type). Boldface entries refer to cognate combinations.

The U6 entries refer to the sequence of the bulge region only (U6 nucleotide 52 varies depending on the column on the right side of the table). The U2 entries refer to the actual variant used in the experiment [see Figs. 6 and 7 for sequences].

Mutation of U2–A25 affects the second step of splicing in vivo

The specificity of suppression demonstrated above is consistent with the notion that U6–G52 and U2–A25 interact directly [see Discussion]. Previous work has demonstrated a role for U6–G52 in the second step of splicing: [1] the U6–G52A and G52C (but not G52U) mutations partially inhibit the second step of splicing in vitro [Fabrizio and Abelson 1990], and [2] the U6–G52U mutation can suppress multiple mutations in the AG dinucleotide at the 3' splice site (Lesser and Guthrie 1993b). Considering a direct interaction between U6–G52 and U2–A25, one predicts that U2–A25 might also influence the efficiency and specificity of the second step of splicing. To test these hypotheses, we constructed yeast strains containing either a wild-type U2 gene or one of the three possible nucleotide substitution mutants as the sole copy. These strains were transformed with intron-containing ACT1–CUP1 fusion constructs containing a wild-type 3' splice site or one of the six possible nucleotide substitutions in the AG dinucleotide at the 3' splice site [Lesser and Guthrie 1993a; Parker and Siliciano 1993]. Total RNA was prepared from these strains and analyzed by a primer extension method.

We first examined the splicing of wild-type pre-mRNAs in the U2 mutants. Shown in Figure 9A, lanes 1–4, are the effects of mutations in U2–A25 on the levels of splicing intermediates and products produced by wild-type ACT1–CUP1. Quantitation using a PhosphorImager revealed that the efficiency of the second step of splicing in the mutants is decreased to approximately one-half of wild-type as measured by the ratio of the levels of mature mRNA to lariat-intermediate (M/L ratio; data not shown). To confirm this result using an endogenous yeast pre-mRNA, we examined the levels of lariat–intermediate and mRNA produced from the chromosomal RP51A gene [Fig. 9B]. In this case, we observed a reduction in the M/L ratio in all three point mutations at U2–A25 to approximately one-third of wild type [Fig. 9B; data not shown]. These data demonstrate that alterations of this nucleotide decrease the efficiency of the second step of splicing.

We then examined the effects of the U2 mutations on the splicing of ACT1–CUP1 fusions that contain mutant 3' splice sites. Strikingly, the U2–A25G mutation partially suppresses the splicing defects produced by all six substitutions in the AG dinucleotide at the 3' splice site [Fig. 9A; lanes 6,10,14,18,22,26]. Quantitation revealed a two- to fivefold increase in the M/L ratio in the instances of suppression (data not shown). U2–A25C and A25U have no detectable effect. Thus, like the U6–G52U mutation, U2–A25G relaxes the specificity of the second step of splicing.

Discussion

Creating an artificial phylogeny of spliceosomal snRNAs in yeast

The constraints of higher order folding are essential to the development of robust models for three-dimensional structure. To seek tertiary interactions among putative active site components of the spliceosome [Fig. 1], we used a method that allowed us to select in vivo for functional covariants of two interacting molecules from gene libraries that contain randomized residues [Fig. 2]. This method allowed us to create an artificial phylogeny of the U2–U6 helix-I region [Figs. 3 and 4]. This method differs from in vitro randomization–selection methods [for review, see Gold et al. 1993; Szostak and Ellington 1993] in several important ways that affect its applica-
bility. First, unlike current in vitro methods, in vivo selection permits the isolation of functional pairs of covariants versus variants of a single molecule. Because each covariant is isolated within a living cell, specific pairs are sequestered from other covariants and are propagated clonally as cells divide. Second, these variants can be isolated in a single selection step; in vitro methods are generally iterative because they are limited by the enrichment obtainable in a single cycle of biochemical selection. Moreover, because our selection method is based on a genetic phenotype, its potential usefulness can be extended to molecules other than RNA and to situations where biochemical selections are impractical. Finally, a limitation of our method is that many fewer variants can be introduced into cells \(10^6\) than can be manipulated in vitro \(10^{15}\); Bartel and Szostak 1993), thereby limiting the number of residues that can be analyzed in a single experiment. If one desires to analyze all possible variants, selection in vitro raises the number of residues that can be randomized from \(\sim 10\) to \(25\) \(4^{10} = 10^6, 4^{25} = 10^{15}\). Of course, in both cases, decreasing the level of degeneracy would permit the analysis of larger sets of residues [e.g., Green and Szostak 1992].

**Does the AGA sequence of U2 function as a group I intron-like G-binding site?**

As expected, our analysis of the selected variants revealed Watson–Crick covariation at two positions previously proven to engage in base-pairing [Tables 1 and 3; Madhani and Guthrie 1992]. Of particular interest are the data that relate to the U6--C58/U2-G26 base pair. As discussed in the introductory section, U2-G26 has been proposed to function as part of a group I intron-like G-binding site during the second chemical step of splicing (McPheeters and Abelson 1992). Such a model would be consistent with the recent proposal, based on stereochemical studies, that the spliceosome generates a group I-like catalytic site to execute the second chemical step of splicing (Moore and Sharp 1993). In group I introns, analysis of the putatively homologous base pair (G264-C311 in the *Tetrahymena* rDNA intron) indicates that the identity of the G at position 264 is more important

---

**Figure 8. Specificity of suppression.**

Growth on 5-FOA of YHM118 containing all possible nucleotides at U6 position 52 and the indicated bulge variants in U2–U6 helix I. Plates were incubated at 30°C for 4 days.
Evidence for a tertiary interaction

In addition to providing further insight into the constraints on established RNA–RNA interactions, our approach permits the identification of nucleotide covariations indicative of previously unknown secondary and/or tertiary interactions. Below we summarize evidence, based on the analysis of several novel variants, for a tertiary interaction between U6-G52 and the helix I bulge.

Novel variants containing lethal substitutions in the ACAGAG hexanucleotide  We isolated temperature-sensitive variants from the randomized libraries that contained otherwise lethal mutations, G52A and G52C, in the terminal residue of the ACAGAG hexanucleotide in U6. We demonstrated that specific changes in the bulge region of U2-U6 helix I suppress the lethality of these substitutions. In the case of U6-G52A, an A → G substitution in the second nucleotide of the bulge together with a flip of the adjacent base pair of helix Ia are necessary and sufficient to suppress its lethal phenotype. In the case of U6-G52C, two classes of suppressors could be discerned on the basis of common sequence motifs. Class I contains three subsets [IA, IB, IC], all of which contain a G residue in one of the two residues of the bulge. Class IA and class IB suppressors contain, in addition, a U2-U23C mutation that disrupts the first base pair of helix Ib. Class IC, instead, contains a flip of the U6-C58/U2-G26 base pair and therefore mimics the structure of the G52A-containing variants. The class II bulge variants found in association with G52C are more complex, containing changes in both base pairs that flank the bulge as well as changes in the bulge itself.

Specificity of suppression suggests a direct interaction  An important question is whether suppression reflects the formation of a direct tertiary interaction be-
between U6 nucleotide 52 and the helix I bulge or a more indirect interaction. To address this issue, we examined the specificity of suppression. We found that variants isolated initially in combination with G52A preferentially suppress this U6 mutation. However, as expected from their similarity to the class IC G52C-containing variants, weak suppression of G52C was also observed. In contrast, three of the four classes of G52C-containing bulge-region variants exhibit dramatic specificity: no detectable suppression of G52A was observed; moreover, these variants exacerbate the temperature-sensitive phenotype of G52U. The other class [IC] of G52C-containing variants structurally mimics the G52A-containing variants and, as expected, can support growth of G52A. The substantial specificity of suppression can most simply be explained by proposing a direct interaction between U6 nucleotide 52 and the helix I bulge. In this scenario, the examples of incomplete allele specificity [i.e., the G52A and G52C class IC variants] could be accounted for by the existence of a common element to the physical interaction in the different cases of suppression. Concordant with the latter notion is the observation that several classes of suppressors are characterized by the occurrence of one or two G residues in the helix I bulge; the exception is the more complex G52C-containing class II suppressors.

It should be emphasized that the specific interaction between U6–G52 and the bulge is not absolutely essential for cell growth because when nucleotide 52 is wild type or a U, the sequence requirements in the bulge are flexible [Fig. 3]. However, such an asymmetric requirement for nucleotides involved in an RNA–RNA interaction is not unusual. For instance, this pattern is seen in the U6–A59/U2–U23 base pair of U2–U6 helix Ib where a direct Watson–Crick interaction can be demonstrated only when U6–A59 is mutated to a G or a C [Table 4]. A second example is seen in U4–U6 stem I, where mutations in U6 have much more dramatic effects on growth that mutations in U4 [Madhani et al. 1990]. As with the previous examples, the more stringent requirement for U6–G52 suggests a function for this nucleotide in addition to its interaction with the helix I bulge. This additional function could either be required at the same time as the tertiary interaction or at some other point during the spliceosome cycle.

Model for suppression Any model that accounts for the novel variants described above must rationalize the following observations: [1] the ability of highly specific changes in the bulge region to suppress otherwise lethal mutations in U6 nucleotide 52, [2] a requisite G residue in the 2-nucleotide bulge in the majority of both G52C- and G52A-containing variants; [3] the requirement for alterations in one of the base pairs flanking the bulge, and [4] substantial but partially overlapping specificity of suppression. Although the complexity of the suppressors makes it difficult to derive a single structural model that accounts for all of the data, the following model provides a simple explanation that can account for the four observations summarized above.

We propose that U6–G52 forms a symmetrical heteropurine base pair with U2–A25 [Fig. 10A]. Such a base pair is found at the junction of the anticodon and D stems in yeast tRNAPhe [for review, see Saenger 1984]. We further propose that the identity of flanking residues influences the stability of this interaction. In this scenario, the G52A suppressors result in an isosteric flip of the U6–G52/U2–A25 heteropurine base pair [Fig. 10B]. The required flip of the adjacent Watson–Crick base pair could reflect a necessity for optimal stacking between the proposed U6–G52/U2–A26 base pair and the adjacent base pair, additional hydrogen bonding interactions, or a steric incompatibility between the U6–G52A/U2–A25G apposition with the wild-type base pair in helix Ia. Several studies have documented the influence of flanking residue sequence context on the thermodynamic stability of hydrogen bonding interactions in RNA, including tertiary interactions [for review, see Turner and Bevilacqua 1993]. Michel et al. [1990] have observed cooperative interactions between adjacent tertiary interactions [base-triples] in the catalytic core of group I introns.

The class IA-, class IB-, and class IC G52C-containing variants can be rationalized by the creation of a Watson–Crick U6–G52C/U2–A25G base pair that would be a similar but not isomorphous replacement of the original heteropurine apposition [Fig. 10C]. Such substitutions have been observed in rRNA phylogeny and are consistent with the similarities in symmetry and backbone geometry between symmetric heteropurine pairs and Watson–Crick pairs [Saenger 1984; Gutell 1993]. Presumably, these shared characteristics outweigh in im-

Figure 10. Model for suppression. Comparison of symmetric G-A [A], A-G [B], and C-G [C] base pairs [Saenger 1984]. [D] Graphic depiction showing the steric feasibility of the proposed heteropurine pair in the context of U2–U6 helix I.
Evidence for a tertiary interaction in the spliceosome

Figure 11. Network of RNA–RNA interactions involving the 3' splice site. The intron is shown in lariat–intermediate form. In addition to the interactions shown in Fig. 1, the non-Watson–Crick between the guanosines at the 5' and 3' splice sites is depicted here (Parker and Siliciano 1993), as well as a site-specific cross-link between the second residue of the 5' splice site and U6–A51 (Sontheimer and Steitz 1993). U5–exon interactions [Newman and Norman 1992, Wyatt et al. 1992, Sontheimer and Steitz 1993], and the proposed tertiary interaction between U6–G52 and the bulge of U2–U6 helix I [this paper].

Functional implications

The model for a direct interaction between U6–G52 and U2–A25 explains the bulk of the experimental data and is structurally reasonable. Given the complexity of the suppression data, however, it seems likely that this depiction of the interaction is oversimplified. As with other examples of tertiary interactions proposed on the basis of functional covariation data [Michel and Westhof 1990; Bartel et al. 1991; Green and Szostak 1994], high-resolution structural studies will be necessary to understand its details. Nonetheless, it is interesting to consider a direct interaction between U6–G52 and the helix I bulge in the context of other recent observations. As described in the introductory section, two clusters of residues in U2 and U6 are specifically required for the second chemical step of splicing [circled in Fig. 1A; see introductory section]. One cluster involves nucleotides 51 and 52 in U6, the other involves the residues that immediately flank the helix I bulge. Notably, a direct interaction between U6–G52 and U2–A25 would closely juxtapose these two groups of residues. Especially in light of other recent results, it is provocative to consider the possibility that this long-range interaction reflects the formation of an active site involved in the second chemical step. First, cross-linking studies performed in mammalian extracts demonstrate proximity between the second nucleotide of the intron and the equivalent of A51 in U6 prior to and/or during the second step of splicing [Sontheimer and Steitz 1993]. Second, a non-Watson–Crick interaction between the conserved Gs at the 5' and 3' ends of the intron has been shown to be important for the second step of splicing [Parker and Siliciano 1993]. If these interactions occur simultaneously such that the 5' and 3' splice sites are juxtaposed and placed in proximity to A51 in U6 [Fig. 11], then an interaction between U6–G52 and U2–A25 would facilitate the recruitment of the second group of residues critical for the second step of splicing to this region. Thus, we suggest that the tertiary interaction proposed here might function in a network of interactions.
RNA–RNA interactions that together form an active site involved in 3′ splice site cleavage and exon ligation.

Intriguingly, mutation of the residues that participate in this tertiary structure affects not only the efficiency of this step of splicing but its fidelity as well (Fabrizio and Abelson 1990; Lesser and Guthrie 1993b; Fig. 9). Specifically, both the U6–G52U and U2–A25G mutations can partially reverse the splicing defect produced by several different mutations in the AG dinucleotide at the 3′ splice site. In the context of the model presented in Figure 11, this apparent relaxation of specificity might be the result of alterations in the proposed tertiary interaction in a manner that can in some way loosen or change the fit of the 3′ splice site into the spliceosomal active site such that noncanonical 3′ splice sites can now be accommodated. The precise structural mechanism by which this might occur is unclear, one possibility is that the U2 and U6 mutations cause increased structural flexibility in the active site. A precedent for this notion exists in proteins: crystallographic analyses of mutations that relax the specificity of a bacterial serine protease suggests that increased active site flexibility underlies effects of mutations that decrease specificity (Bone et al. 1991). In any case, the present data demonstrate that both of the residues that participate in the proposed tertiary interaction can influence events at the 3′ splice site and are consistent with the notion that this higher order interaction functions intimately in the second step of splicing.

Materials and methods

Strain construction

Yeast procedures used in this study have been described previously (Guthrie and Fink 1991). YHM118 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 snr20::LYS2 pU2U6]/SNR6 [SNR20 URA3 CEN]) was constructed as follows: Two parental strains were employed. One contained a disruption of the chromosomal U6 gene, YHM2 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 pSX6U(SNR6 URA3 CEN)] and the other, YHM111[a lys2 trp1 his3 ura3 ade2 snr20::LYS2 pU2U]/SNR20 URA3 CEN], contains a disruption of the chromosomal U2 gene. Each contains a complementing U3A3-marked plasmid containing the wild-type snRNA gene. YHM2 and YHM111 were mated, and diploids were selected on 5-FOA which also cures them of both the U2 and U6 mutations cause increased structural flexibility in the active site. A precedent for this notion exists in proteins: crystallographic analyses of mutations that relax the specificity of a bacterial serine protease suggests that increased active site flexibility underlies effects of mutations that decrease specificity (Bone et al. 1991). In any case, the present data demonstrate that both of the residues that participate in the proposed tertiary interaction can influence events at the 3′ splice site and are consistent with the notion that this higher order interaction functions intimately in the second step of splicing.

Materials and methods

Strain construction

Yeast procedures used in this study have been described previously (Guthrie and Fink 1991). YHM118 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 snr20::LYS2 pU2U6]/SNR6 [SNR20 URA3 CEN]) was constructed as follows: Two parental strains were employed. One contained a disruption of the chromosomal U6 gene, YHM2 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 pSX6U(SNR6 URA3 CEN)] and the other, YHM111[a lys2 trp1 his3 ura3 ade2 snr20::LYS2 pU2U]/SNR20 URA3 CEN], contains a disruption of the chromosomal U2 gene. Each contains a complementing U3A3-marked plasmid containing the wild-type snRNA gene. YHM2 and YHM111 were mated, and diploids were selected on 5-FOA which also cures them of both the U2 and U6 mutations cause increased structural flexibility in the active site. A precedent for this notion exists in proteins: crystallographic analyses of mutations that relax the specificity of a bacterial serine protease suggests that increased active site flexibility underlies effects of mutations that decrease specificity (Bone et al. 1991). In any case, the present data demonstrate that both of the residues that participate in the proposed tertiary interaction can influence events at the 3′ splice site and are consistent with the notion that this higher order interaction functions intimately in the second step of splicing.

Materials and methods

Strain construction

Yeast procedures used in this study have been described previously (Guthrie and Fink 1991). YHM118 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 snr20::LYS2 pU2U6]/SNR6 [SNR20 URA3 CEN]) was constructed as follows: Two parental strains were employed. One contained a disruption of the chromosomal U6 gene, YHM2 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 pSX6U(SNR6 URA3 CEN)] and the other, YHM111[a lys2 trp1 his3 ura3 ade2 snr20::LYS2 pU2U]/SNR20 URA3 CEN], contains a disruption of the chromosomal U2 gene. Each contains a complementing U3A3-marked plasmid containing the wild-type snRNA gene. YHM2 and YHM111 were mated, and diploids were selected on 5-FOA which also cures them of both the U2 and U6 mutations cause increased structural flexibility in the active site. A precedent for this notion exists in proteins: crystallographic analyses of mutations that relax the specificity of a bacterial serine protease suggests that increased active site flexibility underlies effects of mutations that decrease specificity (Bone et al. 1991). In any case, the present data demonstrate that both of the residues that participate in the proposed tertiary interaction can influence events at the 3′ splice site and are consistent with the notion that this higher order interaction functions intimately in the second step of splicing.

Materials and methods

Strain construction

Yeast procedures used in this study have been described previously (Guthrie and Fink 1991). YHM118 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 snr20::LYS2 pU2U6]/SNR6 [SNR20 URA3 CEN]) was constructed as follows: Two parental strains were employed. One contained a disruption of the chromosomal U6 gene, YHM2 [a leu2 lys2 trp1 his3 ura3 ade2 snr6::LEU2 pSX6U(SNR6 URA3 CEN)] and the other, YHM111[a lys2 trp1 his3 ura3 ade2 snr20::LYS2 pU2U]/SNR20 URA3 CEN], contains a disruption of the chromosomal U2 gene. Each contains a complementing U3A3-marked plasmid containing the wild-type snRNA gene. YHM2 and YHM111 were mated, and diploids were selected on 5-FOA which also cures them of both the U2 and U6 mutations cause increased structural flexibility in the active site. A precedent for this notion exists in proteins: crystallographic analyses of mutations that relax the specificity of a bacterial serine protease suggests that increased active site flexibility underlies effects of mutations that decrease specificity (Bone et al. 1991). In any case, the present data demonstrate that both of the residues that participate in the proposed tertiary interaction can influence events at the 3′ splice site and are consistent with the notion that this higher order interaction functions intimately in the second step of splicing.
clones (each variant represented $1 \times 10^2$ times). Control ligations lacking insert indicated negligible background. Unlike pΔSX, pΔES cannot replicate in yeast. Therefore, the U2 library was subcloned en masse into the yeast vector pSE362 [HIS3, CEN] using SalI and SnaI sites that flank the U2 gene. Little, if any, complexity was lost during this subcloning step because $>10^6$ transformants were recovered. DNA sequencing of the pool DNA revealed that the appropriate residues had been randomized. Moreover, at these positions, we observed an even distribution of nucleotides except for G, which was underrepresented approximately twofold.

Selection experiments

For transformation, competent YHM118 was prepared from 1 liter of an exponentially growing culture using a modification of the lithium acetate method [Schiestl and Gietz 1989]. Cells were pelleted by centrifugation, washed once with 20 ml of 0.1 M lithium acetate in TE, and, finally, resuspended in 1 ml of 0.1 M lithium acetate in TE to form a concentrated paste. Prior to use, 0.5 ml of 50% glycerol was added and the cells were frozen by an overnight incubation at $-70^\circ$C. The use of highly concentrated, frozen cells was found to increase cotransformation efficiency 10-fold. In the initial selection experiments, 1 ml of thawed competent cells was transformed with 5 $\mu$g of each library. Transformants ($10^6$) were selected on 10 SD–HIS–TRP plates at 30°C for 3 days. These were replica-plated to prewarmed 5-FOA plates that were then incubated for either 2 or 4 days at 30°C. Approximately 0.5% and 2% of the transformants, respectively, grew on 5-FOA. These were replica-plated to prewarmed YEPD plates that were then incubated at 37°C for 2 days to test for temperature sensitivity. Non-temperature-sensitive variants that grew initially after 2 days on 5-FOA and temperature-sensitive variants that grew after 4 days on 5-FOA were selected for further analysis. Following colony purification, the U2 and U6 plasmids were recovered from the variants using standard methods [Guthrie and Fink 1991] and were used to retransform YHM118 to verify the growth phenotype. The DNA sequences of the randomized regions of a total of 60 U2 functional variants and their 60 U6 partners were determined.

Secondary selections in which specific lethal U6 mutants were used in combination with the U2 library were accomplished by cotransformation of YHM118 followed by selection of the resultant transformants on 5-FOA for 4 days at 30°C. In these cases, $\sim3\%$ of the transformants grew on 5-FOA. These were analyzed as above.

RNA analysis

The ACT1–CUP1 fusions used in this work have been described elsewhere [Lesser and Guthrie 1993a; Parker and Siliciano 1993]. Primer extension analysis was performed as described previously [Madhani and Guthrie 1992].

Model construction

Parts for the RNA physical model were purchased from Maruzen, Inc. (Tokyo). Helices were constrained into A-form using spacers and supports supplied by the manufacturer. A model was built that corresponds to the sequences shown in Figure 1. To examine the tertiary interaction suggested by the results, we assumed the symmetric G-A pair between U6 nucleotide 52 and U2 nucleotide 25 discussed above. This was found to be sterically feasible in many different conformations, including one in which the base pair is stacked on the end of helix Ia. Computer modeling was done using the interactive modeling program SYBYL (Tripos Associates). In this case, only helix I and the upstream AGA sequence were modeled.

Acknowledgments

We thank D. Bartel, A. Frankel, R. Green, R. Gutell, and members of our laboratory for helpful discussions and U. Schmitz for help with computer modeling. ACT1–CUP1 fusion constructs containing 3' splice site mutations were kindly provided by R. Parker and P. Siliciano. We are grateful to S. Noble, J. Umen, A. Murray, I. Herskowitz, S. Burgess, D. McPheeters, and J. Abelson for comments on the manuscript. We also thank L. Esperas, C. Pudlow, and H. Roitha for technical assistance. H.D.M. is a trainee of the University of California Medical Scientist Training Program. C.G. is an American Cancer Society Research Professor of Molecular Genetics. This work was supported by a grant from the National Institutes of Health.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bartel, D.P. and J.W. Szostak. 1993. Isolation of new ribozymes from a large pool of random sequences. Science 261: 1411–1418.

Bartel, D.P., M.L. Zapp, M.R. Green, and J.W. Szostak. 1991. HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. Cell 67: 529–536.

Bocke, J.D., J. Trueheart, G. Natsoulis, and G.R. Fink. 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol. 154: 164–175.

Bone, R., A. Fujishige, C.A. Kettnner, and D.A. Agard. 1991. Structural basis for broad specificity in alpha-lytic protease mutants. Biochemistry 30: 10388–10398.

Cech, T.R. 1986. The generality of self-splicing RNA: Relationship to nuclear mRNA splicing. Cell 44: 207–210.

Cech, T.R. 1993. Structure and mechanism of the large catalytic RNAs: Group I and Group II introns and ribonuclease P. In The RNA world (ed. R.F. Gesteland and J.F. Atkins), pp. 239–270. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Chanfreau, G. and A. Jacquier. 1993. Interaction of intronic boundaries required for the second step splicing efficiency of a group II intron. EMBO J. 12: 5173–5180.

Dower, W.J., J.F. Miller, and C.W. Ragsdale. 1988. High efficiency transformation of E. coli by high voltage electroporation. Nucleic Acids Res. 16: 6127–6145.

Fabrizio, P. and J. Abelson. 1990. Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precur- sor messenger RNA splicing. Science 250: 404–409.

Gold, L., C. Tuerk, P. Allen, J. Binkley, D. Brown, L. Green, S. MacDougall, D. Schneider, D. Tasset, and S.R. Eddy. 1993. RNA: The shape of things to come. In The RNA world (ed. R.F. Gesteland and J.F. Atkins), pp. 497–510. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Green, M.R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. 7: 599–599.

Green, R. and J.W. Szostak. 1992. Selection of a ribozyme that functions as a superior template in a self-copying reaction. Science 258: 1910–1915.

——. 1994. In vitro genetic analysis of the hinge region between helical elements P5-P4-P6 and P7-P3-P8 in the sunY group I self-splicing intron. J. Mol. Biol. 235: 140–155.

Gutell, R. 1993. Comparative studies of RNA: Inferring higher-order structure from patterns of sequence variation. Currr.
Madhani and Guthrie

Opin. Struct. Biol. 3: 313–322.

Guthrie, C. 1991. Messenger RNA splicing in yeast: Clues to why the spliceosome is a ribonucleoprotein. Science 253: 157–163.

Guthrie, C. and G.R. Fink. 1991. Guide to yeast genetics and molecular biology. Academic Press, San Diego, CA.

Guthrie, C. and B. Patterson. 1988. Spliceosomal snRNAs. Annu. Rev. Genet. 22: 387–419.

Hou, Y.-M., E. Westhof, and R. Giege. 1993. An unusual RNA tertiary interaction has a role for the specific aminoacylation of a transfer RNA. Proc. Natl. Acad. Sci. 90: 6776–6870.

Kandels-Lewis, S. and B. S6raphin. 1993. Role of U6 snRNA in the yeast U5 snRNA interacts with exon sequences at 5’.

Lesser, C.F. and C. Guthrie. 1993a. Mutational analysis of pre-
mRNA splicing in Saccharomyces cerevisiae using a sensitive new reporter gene, CUP1. Genetics 133: 851–863.

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

Levitt, M. 1969. Detailed molecular model for transfer ribonucleic acid. Nature 224: 759–763.

Madhani, H.D. and C. Guthrie. 1992. A novel base-pairing in-
teraction between U2 and U6 snRNAs suggests a mecha-
nism for the catalytic activation of the spliceosome. Cell 71: 803–817.

Madhani, H.D., R. Bordonne, and C. Guthrie. 1990. Multiple roles for US snRNA in the splicing pathway. Genes & Dev. 4: 2264–2277.

McPheters, D.S. and J. Abelson. 1992. Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. Cell 71: 819–831.

Michel, F. and E. Westhof. 1990. Modelling of the three-dimen-
sional architecture of group I catalytic introns based on comparative sequence analysis. J. Mol. Biol. 216: 585–610.

Michel, F., M. Hanna, R. Green, D.P. Bartel, and J.W. Szostak. 1989a. The guanosine binding site of the Tetrahymena ribozyme. Nature 342: 391–395.

Michel, F., K. Umesono, and H. Ozeki. 1989b. Comparative and functional anatomy of group II catalytic introns—A review. Gene 8222: 5–30.

Michel, F., A.D. Ellington, S. Couture, and J.W. Szostak. 1990. Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns. Nature 347: 578–580.

Moore, M.J. and P.A. Sharp. 1993. Evidence for two active sites in the splicing reaction provided by stereochemistry of pre-
mRNA splicing. Nature 365: 364–368.

Moore, M.J., C.C. Querry, and P.A. Sharp. 1993. Splicing of pre-
cursors to mRNA by the spliceosome. In The RNA World (ed. R.F. Gesteland and J.F. Atkins), pp. 303–358. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Newman, A. and C. Norman. 1991. Mutations in yeast US sn-
RNA alter the specificity of 5’ splice-site cleavage. Cell 65: 115–123.

—. 1992. US snRNA interacts with exon sequences at 5’ and 3’ splice sites. Cell 68: 743–754.

Noller, H.F. and C.R. Woese. 1981. Secondary structure of 16S ribosomal RNA. Science 212: 403–411.

Parker, R. and P.G. Siliciano. 1993. Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. Nature 361: 660–662.

Parker, R., P.G. Siliciano, and C. Guthrie. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2–like snRNA. Cell 49: 229–239.

Peterson, E.T., J. Blank, M. Sprinzl, and O.C. Uhlenbeck. 1993. Selection for active E. coli tRNA[Phe] variants from a random library using two proteins. EMBO J. 12: 2959–2967.

Puglisi, J.D., J.R. Wyatt, and I.J. Tinoco. 1990. Conformation of an RNA pseudoknot. J. Mol. Biol. 214: 437–453.

Romero, D.P. and E.H. Blackburn. 1991. A conserved secondary structure for telomerase RNA. Cell 67: 343–353.

Rymond, B. and M. Rosbash. 1992. Yeast pre-mRNA splicing. In The molecular and cellular biology of the yeast Saccharomyces (ed. E.W. Jones, J.R. Pringle, and J.R. Broach), pp. 143–192. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Saenger, W. 1984. Principles of nucleic acid structure. Springer-Verlag, New York.

Sawa, H. and Y. Shimura. 1992. Association of U6 snRNA with the 5’ splice site region of pre-mRNA in the spliceosome. Genes & Dev. 6: 244–254.

Schiestl, R.H. and R.D. Gietz. 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. 16: 339–346.

Sharp, P.A. 1985. On the origin of RNA splicing and introns. Cell 42: 397–400.

—. 1991. Five easy pieces. Science 254: 663.

Sontheimer, E.J. and J.A. Steitz. 1993. The US and U6 small nuclear RNAs as active site components of the spliceosome. Science 262: 1989–1996.

Szostak, J.W. and A.D. Ellington. 1993. In vitro selection of functional RNA sequences. In The RNA World (ed. R.F. Gesteland and J.F. Atkins), pp. 511–534. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Wasserman, D.A. and J.A. Steitz. 1992. Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. Science 257: 1918–1925.

Weiner, A.M. 1993. mRNA splicing and autocatalytic introns: Distant cousins or the products of chemical determination? Cell 72: 161–164.

Woese, C.R., R. Gutell, R. Gupta, and H.F. Noller. 1983. Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. Microbiol. Rev. 47: 621–669.

Wu, J. and J.L. Manley. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes & Dev. 3: 1553–1561.

Wyatt, J.R., J.D. Puglisi, and I.J. Tinoco. 1990. RNA pseudoknots. Stability and loop size requirements. J. Mol. Biol. 214: 455–470.

Wyatt, J.R., E.J. Sontheimer, and J.A. Steitz. 1992. Site-specific cross-linking of mammalian US snRNP to the 5’ splice site before the first step of pre-mRNA splicing. Genes & Dev. 6: 2542–2553.

Yarus, M., M. Illangeskare, and E. Christian. 1991a. An axial binding site in the Tetrahymena precursor RNA. J. Mol. Biol. 222: 995–1012.

—. 1991b. Selection of small molecules by the Tetrahymena catalytic center. Nucleic Acids Res. 19: 1297–1304.

Zhuang, Y. and A.M. Weiner. 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. Genes & Dev. 3: 1545–1552.
Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome.

H D Madhani and C Guthrie

*Genes Dev.* 1994, 8:
Access the most recent version at doi:10.1101/gad.8.9.1071