Structural Basis for a Lethal Mutation in U6 RNA[†,‡]

Dipali G. Sashital,[§] Anne M. Allmann,[§] Steven R. Van Doren,[||] and Samuel E. Butcher*[,§]

Department of Biochemistry, University of Wisconsin−Madison, Madison, Wisconsin 53706, and Department of Biochemistry, 117 Schweitzer Hall, University of Missouri−Columbia, Columbia, Missouri 65211

Received November 8, 2002; Revised Manuscript Received December 5, 2002

ABSTRACT: U6 RNA is essential for nuclear pre-mRNA splicing and has been implicated directly in catalysis of intron removal. The U80G mutation at the essential magnesium binding site of the U6 3′ intramolecular stem−loop region (ISL) is lethal in yeast. To further understand the structure and function of the U6 ISL, we have investigated the structural basis for the lethal U80G mutation by NMR and optical spectroscopy. The NMR structure reveals that the U80G mutation causes a structural rearrangement within the ISL resulting in the formation of a new Watson−Crick base pair (C67·G80), and disrupts a protonated C67·A79 wobble pair that forms in the wild-type structure. Despite the structural change, the accessibility of the metal binding site is unperturbed, and cadmium titration produces similar phosphorus chemical shift changes for both the U80G mutant and wild-type RNAs. The thermodynamic stability of the U80G mutant is significantly increased ($\Delta\Delta G_{fold} = -3.6 \pm 1.9$ kcal/mol), consistent with formation of the Watson−Crick pair. Our structural and thermodynamic data, in combination with previous genetic data, suggest that the lethal basis for the U80G mutation is stem−loop hyperstabilization. This hyperstabilization may prevent the U6 ISL melting and rearrangement necessary for association with U4.

Proteomic diversity in eukaryotes is generated by alternative splicing of exons from nuclear premessenger RNA (pre-mRNA) by the spliceosome. The spliceosome is a large ribonucleoprotein complex, made up of five small nuclear RNAs (snRNAs), U1, U2, and U4, U5, U6, and more than 70 proteins (*1*−*4*). The spliceosome catalyzes a two-step transesterification reaction, speculated to be RNA-catalyzed by analogy to mitochondrial group II self-splicing ribozymes (*2*, *3*, *5*). Two snRNAs (U2 and U6) likely comprise part of the spliceosome active site, and U6 is essential for catalysis (*2*, *3*, *5*). In the active spliceosome, U2 and U6 form a complex by base pairing to each other, and have also been shown to base pair to pre-mRNA at the first step of splicing. Moreover, mutagenesis data have shown that certain regions of U6 must be intact for the assembled spliceosome to catalyze the first or second step of splicing (*6*). Other atomic substitution studies of U6 RNA have revealed several phosphate oxygens that are essential for splicing (*7*−*9*), which are potential sites of magnesium coordination required by the spliceosome. Recently, the U2−U6 complex was shown to catalyze a reaction similar to the first step of splicing in the absence of protein, lending more evidence to the hypothesis of an RNA active site in the spliceosome (*10*).

The formation of the U2−U6 complex is initiated by a large conformational change in U6 RNA (*1*, *11*). During spliceosome assembly, U6 extensively base pairs with U4 RNA, forming a helical secondary structure. After the spliceosome is assembled completely, base pairing between U4 and U6 is disrupted and U6 undergoes a large conformational change, in which an intramolecular stem−loop (ISL)[1] region forms near the 3′ end of U6 (*1*, *11*). This ISL is mutually exclusive with the U4−U6 complex, and allows the U2−U6 complex to form.

The structurally significant U6 ISL has also been implicated in catalytic steps of the spliceosome. U6 RNA stereospecifically binds a metal ion at the S_p phosphate oxygen 5′ of the highly conserved residue U80 of the ISL (*9*, *12*). The S_p phosphate oxygen has been shown to be essential for the first step of splicing (*7*−*9*); a sulfur substitution at this position in *Saccharomyces cerevisiae* U6 halts splicing after the assembly of the spliceosome (*9*). Splicing activity is rescued with the addition of the thiophilic metal ion cadmium, demonstrating the importance of metal ion coordination at U6 ISL residue U80 for spliceosome function (*9*). Recently, structural studies have shown that protonation of a C67·A79 wobble pair modulates metal ion binding at U80, suggesting a possible method for splicing regulation (*12*). Interestingly, mutation of the U80 residue to G is lethal in yeast (*13*, *14*). Despite the recent evidence

[†] This work was supported by NIH Grant GM65166. S.R.V.D. acknowledges support from the Missouri F21C program and Agriculture Experiment Station. This work was supported by an NIH Grant to S.E.B. and a Molecular Biosciences Training Grant to D.G.S.

[‡] Structure coordinates have been deposited in the Protein Data Bank as entry 1NC0. Chemical shift data have been deposited in BioMagResBank (BMRB), entry code 5655.

* To whom correspondence should be addressed: 433 Babcock Dr., Madison, WI 53706. E-mail: butcher@nmrfam.wisc.edu.

[§] University of Wisconsin−Madison.

[||] University of Missouri−Columbia.

[1] Abbreviations: CT, constant time; COSY, correlation spectroscopy; HCCH, proton, carbon, carbon, proton; HCN, proton, carbon, nitrogen; HNN, proton, nitrogen, nitrogen; HMQC, heteronuclear multiple-quantum correlation spectroscopy; HSQC, heteronuclear single-quantum correlation spectroscopy; ISL, U6 3′ intramolecular stem−loop region; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; NOE, nuclear Overhauser effect; RDC, residual dipolar coupling; rmsd, root-mean-square deviation; TOCSY, total correlation spectroscopy.

10.1021/bi027137h CCC: $25.00
Published on Web 01/18/2003

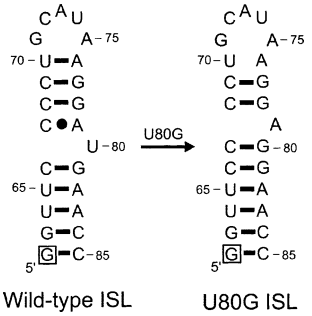

FIGURE 1: Secondary structures of the wild-type and U80G mutant U6 ISL RNAs, representing nucleotides 62−85 from *S. cerevisiae*. The A62G substitution is boxed.

of the functional importance of residue U80, as well as the structural significance of the ISL, the basis of U80G lethality is unknown. To gain more insight into the structural requirements of U6 RNA for splicing, we sought to determine the solution structure of the U80G mutant U6 ISL.

Recent advances in NMR structure determination methods have improved the precision of solution structures through the use of residual dipolar couplings (*15*), and the direct detection of hydrogen bonds via scalar couplings (*16*). Using these methods, we have determined the structure of the U6 ISL harboring the lethal U80G mutation. The mutation leads to the formation of a new Watson−Crick base pair in the internal loop, between G80 and C67, which in turn disrupts the phylogenetically conserved C67·A79 wobble pair that forms in the wild-type structure (*12*) (Figure 1). Thermodynamic studies reveal that this altered secondary structure causes an increase in stability ($\Delta\Delta G_{fold} = -3.6 \pm 1.9$ kcal/mol) over that of the wild-type ISL. Surprisingly, we also find that the U80G mutant and wild-type ISL RNAs produce similar phosphorus chemical shift changes upon cadmium binding, which suggests that the mutant is still capable of binding metal ion. We propose that the basis for lethality of the U80G mutation is likely due to a structural rearrangement within the conserved internal loop of the U6 ISL, resulting in hyperstabilization of the structure, which normally must function as a marginally stable stem−loop structure that can rearrange.

MATERIALS AND METHODS

RNA Synthesis and Purification. The U6 ISL RNA with the A62G and U80G substitutions was transcribed in vitro using purified His_6-tagged T7 RNA polymerase and synthetic DNA oligonucleotides (Integrated DNA Technologies, Inc.). RNA was purified by denaturing 20% polyacrylamide gel electrophoresis, identified by UV absorbance, and excised from the gel. RNA was recovered by soaking gel pieces in 0.3 M sodium acetate, ethanol precipitated, purified on a 6 mL DEAE anion exchange column, ethanol precipitated again, and desalted on a 15 mL G-25 gel filtration column. The purified RNA was lyophilized, resuspended in water, and brought to pH 6.5 by the addition of 1 M NaOH. ^{13}C- and ^{15}N-labeled RNA was prepared using ^{13}C- and ^{15}N-labeled rNTPs (Silantes GmbH, München, Germany).

NMR Spectroscopy. All NMR spectra were obtained on Bruker DMX spectrometers at the National Magnetic Resonance Facility at Madison (NMRFAM, Madison, WI). All spectrometers used for data collection were equipped with either conventional HCN triple-resonance, triple-axis pulsed-field gradient probes, a cold (cryoprobe, Bruker) single *Z*-axis gradient HCN probe, or a Bruker QNP (quadruple-nucleus probe) tuned to phosphorus with a *Z*-axis gradient. Exchangeable resonances were assigned by reference to two-dimensional (2D) NOESY (mixing time of 150 ms) and 2D $^1H-^{15}N$ HMQC spectra of the RNA in a 90% H_2O/10% D_2O mixture at 283 K. Nonexchangeable resonances were assigned by reference to 2D NOESY spectra (mixing times of 75, 150, and 250 ms) and 2D TOCSY, 2D $^1H-^{13}C$ HSQC, three-dimensional (3D) $^1H-^{13}C-^1H$ TOCSY, 3D $^1H-^{13}C-^1H$ HCCH COSY, and 3D $^1H-^{13}C$ NOESY-HMQC spectra of the RNA in 99.99% D_2O at 303 K, as described previously (*12*). Water suppression for samples in a 90% H_2O/10% D_2O mixture was achieved with a 1-1 spin-echo pulse sequence. For experiments in 99.99% D_2O, the residual HDO resonance was suppressed with a low-power presaturation pulse.

Hydrogen bonds were detected on the fully labeled ^{13}C and ^{15}N sample by 2D J_{NN} HNN-COSY in a 90% H_2O/10% D_2O mixture and in 99.99% D_2O at pH 6.3 and 298 K (*17*−*19*). Water suppression for the 90% H_2O/10% D_2O sample was achieved using Watergate suppression methods.

Partial alignment of RNA for residual dipolar coupling measurements was achieved by adding 17 mg/mL Pf1 filamentous bacteriophage (ASLA Ltd., Riga, Latvia) to the ^{13}C- and ^{15}N-labeled samples (*15*, *20*). $^1H-^{15}N$ residual dipolar couplings (RDCs) were measured for exchangeable protons using ^{15}N-coupled one-dimensional (1D) ^1H spectra, and ^1H-^{13}C RDCs were measured in the carbon dimension of 2D $^1H-^{13}C$ CT-HSQC spectra for a uniformly ^{13}C- and ^{15}N-labeled sample and a G only-labeled sample in a 90% H_2O/10% D_2O mixture at 303 K, for both isotropic and oriented samples. Uncertainty in RDC measurements was estimated to be ±1.2 Hz.

All data were processed using XWINNMR software from Bruker. Resonance assignments were completed using Felix 98 (MSI) and Sparky (http://www.cgl.ucsf.edu/ home/sparky/).

Interproton Distances. NOE distances were obtained using information from nonexchangeable 2D NOESYs at varying mixing times (75, 150, and 250 ms). NOE peak volumes were integrated using the Gaussian peak fitting function in Sparky. Distances were calibrated by setting the peak volumes of pyrimidine H5−H6 NOEs to the fixed distance of 2.4 Å, using the CALIBA macro in DYANA (*21*). These distances were used to group NOEs into three classes: strong (1.8−3.0 Å), medium (1.8−4.5 Å), and weak (3.0−6.0 Å). NOE distances for exchangeable protons were qualitatively assigned to one of these three classes.

Torsion Angle Constraints. All ^{31}P chemical shifts fall between −4 and −5 ppm, suggesting that no α and ζ in the structure are in the *trans* range. Therefore, all α and ζ values were set to exclude the *trans* range (0 ± 120°) (*22*). Nucleotides with strong H1′−H2′ couplings (C72 and U74), as observed in a 45 ms mixing time 2D $^1H-^1H$ TOCSY experiment, were constrained as C2′-*endo*, while all other nucleotides were constrained as C3′-*endo*. Analysis of intranucleotide H1′−aromatic NOEs from the 75 ms 2D NOESY experiment indicated that all nucleotides fell into the *anti* range, and were thus constrained with a χ value of −160 ± 15°. Other backbone torsion angles (β, γ, and ε) were set to standard A-form values (±15°) only in the helical

regions of the structure known to be A-form Watson−Crick helices from NOE, dihedral, and through hydrogen-bond J_{NN} coupling data.

Residual Dipolar Coupling Analysis. Residual dipolar couplings (RDCs) were measured using XWINNMR (Bruker) by determining the difference between ^{1}H and ^{13}C coupling for isotropic and partially aligned samples. Values for the axial and rhombic components (D_a and R) of the alignment tensor were estimated by analysis of a powder-pattern distribution of RDC values (*23*), yielding an axial component (D_a) of −17 Hz and a rhombic component (R) of 0.14. PALES (*24*) (http://spin.niddk.nih.gov/bax/software/PALES) was used to predict the alignment tensor from converged, low-energy structures calculated in the absence of RDCs, having an rmsd from the mean of 1.2 Å. However, these predictions appear to be less reliable than the powder-pattern analysis, both because the structures were underdetermined with respect to the overall helical axis prior to incorporation of the RDC measurements and because significant uncertainty in D_a and R results from even moderate structural uncertainty (*25*). Finally, CNS structure calculations with a grid search for optimal D_a and R values (*23*) revealed that the best agreement between structure and RDCs, as evidenced by very low energies and no violations, is obtained with a D_a of −16 Hz and an R of 0.15, which are in close agreement with the values obtained from the powder-pattern analysis. Therefore, a D_a of −16 Hz and an R of 0.15 were used in the final calculations. The final converged structures were used to back-calculate predicted RDC values, and the calculated RDCs were plotted against the experimentally measured values and were found to be in excellent agreement.

Structure Calculation. CNS 1.1 (*26*) (http://cns.csb.yale.edu/v1.1) was used to calculate structures using NOE distance and dihedral restraints, and residual dipolar couplings. CNS was recompiled with an improved version of the harmonic potential for RDCs, which corrects the susceptibility anisotropy refinement protocol (*27*, *28*). The CNS structure calculations closely followed the default values for NMR structure determination of nucleic acids with CNS, with minor changes. First, an extended (completely unfolded) structure was generated, from which 100 starting structures were calculated from random initial velocities. The 100 starting structures were subjected to 60 ps of restrained molecular dynamics in torsion angle space, using 15 fs time steps, followed by 90 ps of slow cooling. Finally, 30 ps (5 fs time steps) of restrained molecular dynamics in Cartesian coordinate space was performed. Planarity for all base pairs detected using exchangeable 2D NOESY and trans-hydrogen bond experiments was enforced during the calculations, and hydrogen bonds were maintained by distance restraints for Watson−Crick base pairs as well as one G·A pair. Structures were also calculated without RDCs for comparison. After Cartesian space refinement, the structures were evaluated for convergence. Acceptance criteria of converged structures were low overall energies and no significant NOE (>0.5 Å) or dihedral (>5°) violations. Finally, the 20 lowest-energy structures were subjected to a gentle refinement in XPLOR 3.843 (*29*) for 4 ps (0.2 fs time steps) at 300 K followed by 200 steps of energy minimization. This latter step does not significantly change the structures (same overall rmsd), but includes a script for maintaining the proper chirality of phosphate oxygens and amino protons which can be reversed in Cartesian space refinement protocols (*30*). The final structures were viewed and analyzed using MOLMOL (*31*).

Metal Binding Studies. A62G, U80G U6 ISL RNA substituted with a phosphorothioate linkage between A79 and G80 was purchased from Dharmacon, Inc., and deprotected according to recommended procedures. The phosphorothioate-substituted RNA was purified by denaturing PAGE, anion exchange, and gel filtration chromatography as described above. Cadmium binding was monitored by titrating cadmium chloride into the sample and by following the chemical shift changes in the S_p and R_p phosphorothioate signals by ^{31}P 1D NMR at 202 MHz (phosphorus frequency), as previously described (*12*).

Optical Spectroscopy. Thermal stability studies were conducted on purified U6 ISL and U80G mutant RNAs using a Cary model 1 Bio UV−visible spectrophotometer equipped with a Peltier heating accessory and temperature probe. All samples contained 10 mM sodium phosphate buffer (pH 7.0 or 5.5), 200 mM KCl, and approximately 1 μM RNA. Samples were heated to 90 °C and cooled to 20 °C at a rate of 2 °C/min, while absorbance data were collected at 260 nm in 1 °C increments. Two scans were taken for each RNA. Thermodynamic values and transition temperatures (T_r) were calculated from normalized data using SigmaPlot version 8.0 (SPSS Science). For the process of unfolding a stem−loop structure composed of a single RNA molecule (F → U) where $\Delta C_p = 0$

$$K = U/F = \theta/(1 - \theta) = \exp[-\Delta H/(RT) + \Delta S/R]$$

and

$$\Delta S = \Delta H/T_r$$

where θ is the fraction of RNA in the unfolded state, R is the gas constant, and T is the temperature in kelvin (*32*). To fit a curve to the absorbance data for the U6 ISL and U80G RNAs, the pretransitional and post-transitional baselines of each thermal stability curve were first determined by a linear curve fit (*32*). These baselines were used to find the fraction of unfolded RNA as a function of temperature and subsequently the values of ΔH, ΔS, and T_r (*32*). The difference in folding free energies (ΔG_{fold}) of the two RNAs was calculated from their ΔH and ΔS values.

RESULTS

NMR Analysis. (A) Watson−Crick Pairing As Determined by Exchangeable Proton Spectra and a Trans-Hydrogen Bond Experiment. The sequence of the U6 ISL, corresponding to nucleotides 62−85 of *S. cerevisiae* U6 RNA and the U80G mutant, is shown in Figure 1. The "wild-type" and U80G U6 ISL RNAs incorporate an additional A62G substitution, which has no effect on growth rates at 30 °C (*11*) and no effect on the overall structure of the U6 ISL (*12*), but yields higher-quality NMR spectra and allows for optimal in vitro transcription using T7 RNA polymerase.

Comparison of wild-type and mutant ISL 1D ^{1}H NMR imino spectra at pH 6.5 reveals that seven of eight imino peaks present in the wild type are readily observed in the U80G mutant (Figure 2A). The similarity in the data indicates that the mutant ISL contains two helices separated by an

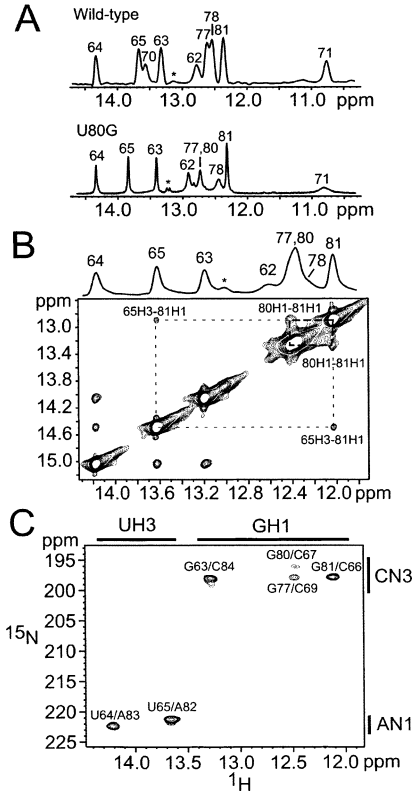

FIGURE 2: NMR evidence for Watson−Crick base pairs in the yeast U80G U6 ISL. (A) 1D ^{1}H NMR spectra for wild-type and U80G imino peaks taken at pH 7.0 and 283 K. Peak assignments are indicated. An asterisk indicates a second resonance for G62 due to a small amount of nontemplated addition at the 3′ end of the RNA (n + 1). (B) 1D ^{1}H imino spectrum and 2D ^{1}H−^{1}H NOESY spectrum for the U80G ISL at pH 5.4 and 283 K. A portion of the imino−imino region of the NOESY spectrum is shown. The 65 H3−81 H1 and 80 H1−81 H1 NOEs are labeled. (C) Direct detection of Watson−Crick base pairs by heteronuclear J_{NN} COSY correlation through the N−H···H hydrogen bond in U·A and G·C base pairs at pH 6.3 and 298 K.

internal loop, as well as a G71·A75 pair at the base of a pentaloop, as in the wild-type structure.

A key feature of the U80G ISL structure is an additional Watson−Crick base pair between G80 and C67. Identification of this base pair was complicated by exchange of the G80 imino proton and complete spectral overlap with the imino proton and nitrogen resonances of G77. However, a 2D ^{1}H−^{1}H NOESY spectrum at pH 5.9 revealed a sequential NOE from G81 to a peak which was tentatively assigned to the imino proton of G80 (Figure 2B). To confirm this assignment and the possible existence of a new base pair, we performed a $^{2h}J_{N1N3}$ HNN-COSY experiment that directly detects scalar couplings across hydrogen bonds. This experiment identifies two U·A (H1−N1 to N3) and four G·C (H3−N3 to N1) base pairs, including a G80 N3−C67 N1 correlation, indicating that a new base pair involving G80 does indeed form in the U80G ISL (Figure 2C). Correlations for G62 and G78 are not observed due to fast exchange of these imino peaks with water.

(B) Assignment and Analysis of Nonexchangeable Spectra. All of the nonexchangeable proton distance restraint information used in structure calculations was taken from the 2D ^{1}H−^{1}H NOESY spectra and 3D ^{1}H−^{13}C NOESY-HMQC spectrum. The 2D NOESY revealed several similarities and differences between the structure of the wild-type and U80G ISL structures. Chemical shifts for protons of the lower helix (residues 62−66 and 81−85) are nearly identical for the two RNAs, indicating that the first segment of the helical stem remains the same with the U80G mutation. These results agree with the Watson−Crick pairs expected from the exchangeable proton spectra and trans-hydrogen bond experiments. Though the U70 imino peak is absent from U80G spectra, unlike that of the wild type, the chemical shifts of the mutant ISL pentaloop (residues 71−75) are also similar to those of the wild type, suggesting a structure with U74 bulged out of the loop as indicated by NOEs between A73 and A75 (*12*). As expected, the major chemical shift changes for the U80G ISL occur in the internal loop region where the mutation is located. In particular, two NOEs of nearly equal intensity arise between the H2 proton of A79 and H1′ protons of C68 and C69, the latter of which is not observed in wild-type spectra. Additionally, the protonation state of residue A79 was investigated by ^{1}H−^{13}C 2D HSQC at pH 7.4, 6.4, and 5.3. In wild-type studies, the chemical shift of A79 C2 changes significantly as a function of pH, indicating that the adjacent N1 atom is protonated with a pK_a of 6.5 (*12*). The U80G spectra show no significant change in A79 C2 chemical shift as a function of pH; we therefore conclude that the unpaired A79 does not undergo protonation until the pH is below 5.3.

Table 1: Structural Statistics for the U80G U6 ISL

	with dipolar coupling data	without dipolar coupling data
no. of structures		
accepted	18	20
calculated	100	100
no. of NOE-derived distance restraints	474	474
intranucleotide	178	178
internucleotide	294	294
no. of dihedral restraints	179	179
no. of hydrogen bond restraints	24	24
no. of dipolar coupling restraints	31	0
rmsd (for all heavy atoms to mean structure, residues 62−71, 73, and 75−85) (Å)	0.95	1.15
no. of NOE violations of >0.5 Å	0	0
no. of dihedral violations of >5°	0	0
average NOE rmsd (Å)	0.028	0.029
average dihedral rmsd (deg)	1.27	0.78
average RDC rmsd (Hz)	2.3	−

Structure Calculations with Residual Dipolar Couplings. The solution structure of the U80G U6 ISL was calculated with 474 NOE restraints (an average of 19.8 NOEs per residue) and torsion angle restraints for the A-form helical regions (Table 1). Additional Watson−Crick hydrogen bond restraints were used for base pairs established through imino spectra and the trans-hydrogen bond HNN-COSY experiment (Figure 2). Thirty-three residual dipolar couplings were measured and used to improve the long-range order of the structure. RDCs were not included for regions of the pentaloop structure that are suspected to be dynamic from previous structure calculations (*12*). The axial component for the principal alignment tensor (D_a) and rhombicity (R) was determined from a powder-pattern analysis and adjusted by a grid search (*23*). Back-calculated RDCs from the final structures agree well with the measured values (Figure 3). Comparison of structures with and without RDCs reveals a significant improvement of the long-range order in the ensemble of structures (Figure 4). The 18 lowest-energy

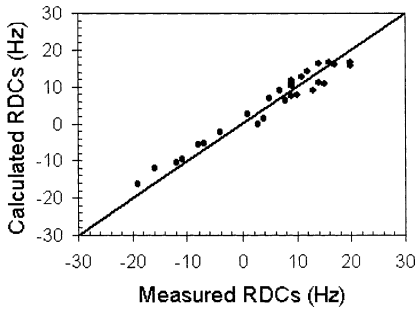

FIGURE 3: Comparison of measured and calculated residual dipolar couplings for the yeast U6 U80G ISL ($R^2 = 0.95$).

structures calculated with RDCs have a pairwise root-mean-square deviation (rmsd) of 0.95 Å (Table 1).

The inner loop region adopts a conformation different from that of the wild type, as illustrated in Figure 5. The C67·G80 pair makes it impossible to form the protonated C67·A79 wobble pair, leaving A79 unpaired and stacked in the helix instead of U80. Nevertheless, the location of the S_p phosphate oxygen is unchanged, with U80 and G80 occupying the same space in the helix in the two structures (Figure 5). The wild-type and U80G structures have an rmsd of 0.7 Å over all common heavy atoms in the internal loop region (residues 66−68 and 79−81). The resulting conformation is a rearrangement of base pairings that do not significantly change the position or accessibility of the S_p phosphate oxygen at position 80.

Metal Binding Studies. The structure of the U80G mutant U6 ISL suggests that metal binding should still be able to occur at the S_p phosphate oxygen of position 80 as in the wild-type U6 ISL (*9*, *12*). We assayed for metal binding at the U80 phosphate by introducing S_p and R_p phosphorothioate substitutions into the ISL RNAs. The phosphorothioate substitution shifts the U80 phosphate by >50 ppm from unmodified phosphorus signals and creates a cadmium-specific metal binding site in both the assembled spliceosome (*9*) and the isolated ISL (*12*). We find that the U80G ISL displays stereospecific chemical shift changes at G80 upon cadmium addition that are qualitatively similar to those of the wild-type ISL (Figure 6A). Cadmium binding to phosphorothioates typically produces a diagnostic upfield change in the ^{31}P chemical shift, as observed for the S_p resonance 5′ to nucleotide 80 for both the wild-type and mutant ISL RNAs (Figure 6A) (*12*, *33*, *34*). As previously observed for the wild-type U6 ISL, these data cannot be fit to a single-binding site model, because cadmium ions also bind nonspecifically to the electronegative major groove of the ISL, and sample aggregation and line broadening occur with higher molar equivalents of cadmium (*12*).

Optical Spectroscopy and Thermodynamics. The structure of the U80G mutant U6 ISL also suggests that the mutation should confer an increase in thermodynamic stability. Therefore, we used optical spectroscopy at 260 nm as a function of temperature to monitor the thermal stability of the wild-type and U80G U6 ISL RNAs. We find that the U80G mutant is significantly more stable than the wild type at pH 7.0, with a 15 °C increase in melting temperature (T_m), corresponding to a difference in the free energy of folding ($\Delta\Delta G_{fold}$) of -3.6 ± 1.9 kcal/mol. The observed difference in the folding free energies is in close agreement with calculated values based on nearest-neighbor free energies, which predict that the U80G U6 ISL should be −5 kcal/mol more stable than the wild type (*35*). The stability of the wild-type ISL also increases significantly when the pH is decreased from 7.0 to 5.5, resulting in an observed 3.9 °C increase in T_m (Figure 6B). This increase in stability can be attributed to the formation of an extra hydrogen bond due to the protonation of the C67·A79 base pair, which contributes an additional -0.9 ± 0.4 kcal/mol of folding free energy. In contrast, a low pH has little effect on the stability of the U80G ISL due to the disruption of the C·A base pair.

DISCUSSION

Structure and Stability of the U80G ISL versus the Wild Type. The structure of the mutant U80G U6 ISL retains many of the features of the wild type. Both consist of two helices, separated by an internal loop composed of one unpaired residue, and a GCA(U)A pentaloop that adopts a GNRA tetraloop conformation, with the U73 residue bulged out of the loop (*12*). Given the similarities between the mutant and wild-type structures, and the fact that both bind metal stereospecifically, the lethal phenotype is surprising and highlights the importance of an unpaired residue at position 80 for splicing. The main difference between the two structures arises from the formation of a new Watson−Crick pair between G80 and C67, and the resulting disruption of a C67·A79 wobble pair that is present in the wild-type structure (Figures 1 and 5). Trans-hydrogen bond scalar couplings and residual dipolar coupling data provide strong evidence for the validity of this structure.

The formation of the G80·C67 base pair at the expense of the C67·A79 wobble pair causes a dramatic increase in the stability of the ISL when compared to the wild-type ISL. Thermodynamic studies show that the melting temperature of the mutant ISL is 15 °C higher than that of the wild type at pH 7.0 (Figure 6B). During spliceosome assembly and turnover, the region of U6 RNA that forms the ISL must unfold to base pair with U4 RNA (*1*). Improved stabilization of the ISL may inhibit formation of the U4−U6 complex, and thus slow the turnover rate of the spliceosome, as suggested by Madhani and Guthrie (*36*). Genetic studies in yeast have shown that other mutations that potentially hyperstabilize the ISL by forming new Watson−Crick pairs (such as A62G and A79G) cause slower growth rates at 18 °C, and that the U4−U6 complex concentration decreases under these conditions (*11*). Additional destabilizing mutations at other points in the ISL suppress the effect of hyperstabilizing mutations (*11*, *37*), although this type of study has never been performed for the U80G mutation.

The U80G mutation in the U6 ISL lengthens the first helix by one base pair, and actually shortens the upper helix by one base pair (Figure 1). Overall, this is a more stable base pairing arrangement even though the upper helix may be less stable than the wild type. We observe that the G71 imino proton in the pentaloop G·A pair exchanges more rapidly in U80G than in the wild type (Figure 2A), and that the loss of the U70 imino peak in U80G suggests the U70·A76 base pair proximal to the pentaloop is destabilized in the U80G ISL (Figures 1 and 2A). These data suggest that the cooperative melting of the U80G ISL is influenced more by the stability of the first helix than by the pentaloop-proximal helix. Although the U70·A76 base pair is destabilized in the

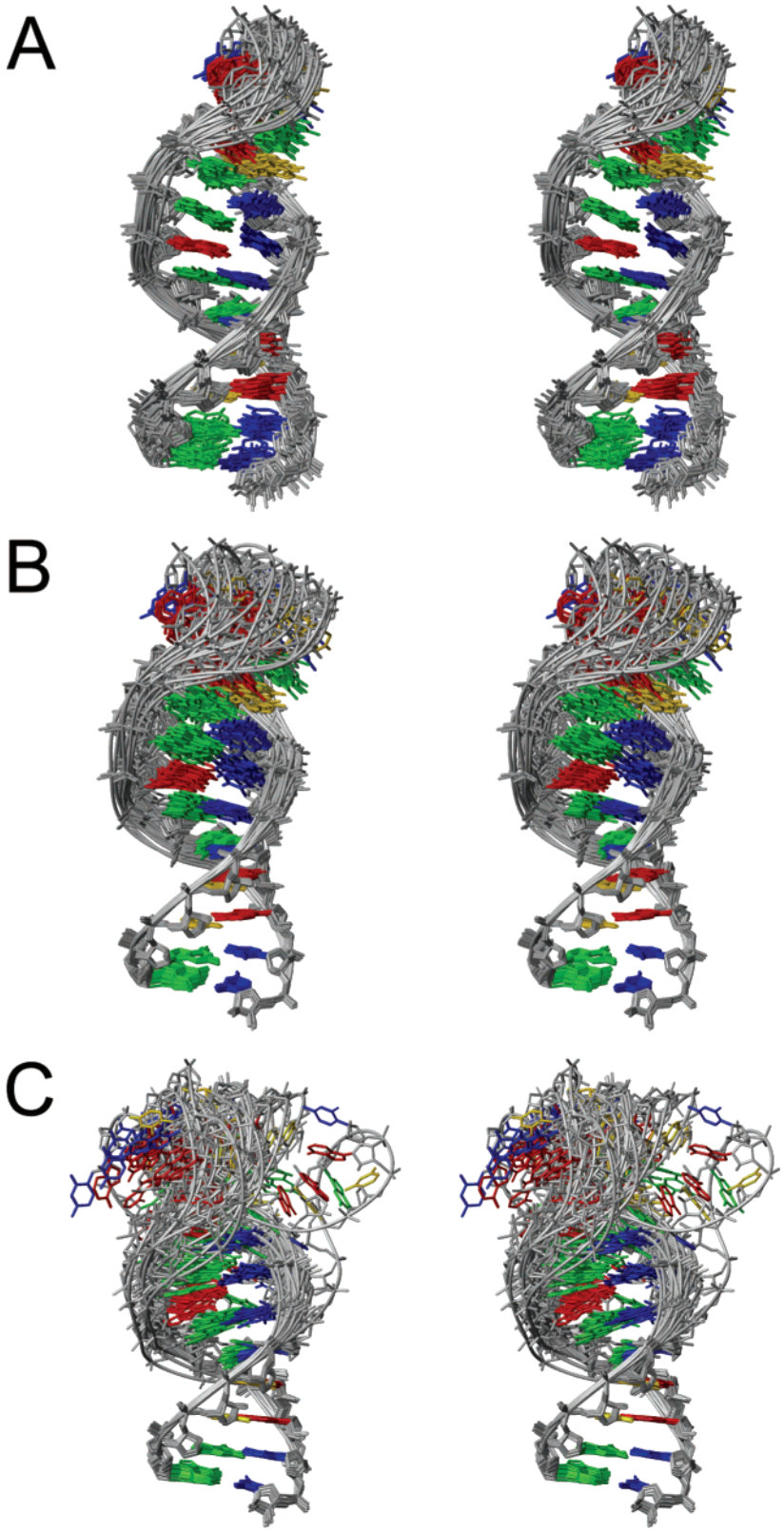

FIGURE 4: Stereoview of the yeast U6 U80G ISL. The phosphoribose backbone is gray, and guanines are green, uracils gold, cytosines blue, and adenines red. (A) The 18 structures calculated with RDCs superimposed over all heavy atoms. (B) The 18 structures calculated with RDCs superimposed over the first five base pairs. (C) The 20 structures calculated without RDCs superimposed over the first five base pairs.

U80G ISL structure, it is well stacked below the G71·A75 base pair and may form one or more water-mediated hydrogen bonds, as suggested by increased distances between hydrogen bond acceptors and donors. However, such water molecules have yet to be identified.

Metal Binding in the U80G ISL. The U80G ISL displays phosphorus chemical shift changes upon cadmium addition that are similar to those of the wild type (Figure 6A), which suggests that U80G lethality is not due to the loss of the essential phosphate metal binding site. Additionally, the

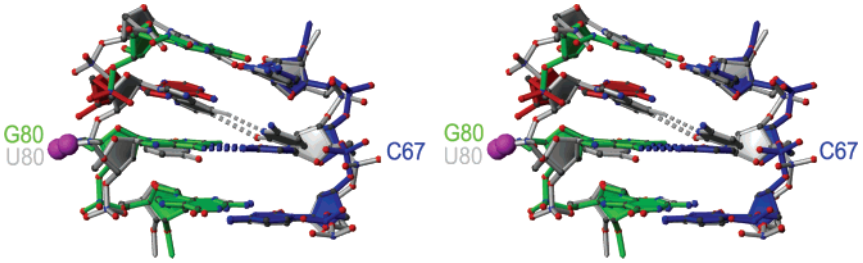

FIGURE 5: Stereoview of the internal loop of the U80G ISL and the wild-type ISL superimposed over heavy atoms. The U80G ISL is colored as in Figure 4, and the wild-type ISL is gray. The S_p phosphate oxygen of residue 80 is shown as a magenta sphere in both structures. Hydrogen bonds are denoted with dashed lines. The rmsd over common heavy atoms in this region for U80G and the wild type is 0.7 Å.

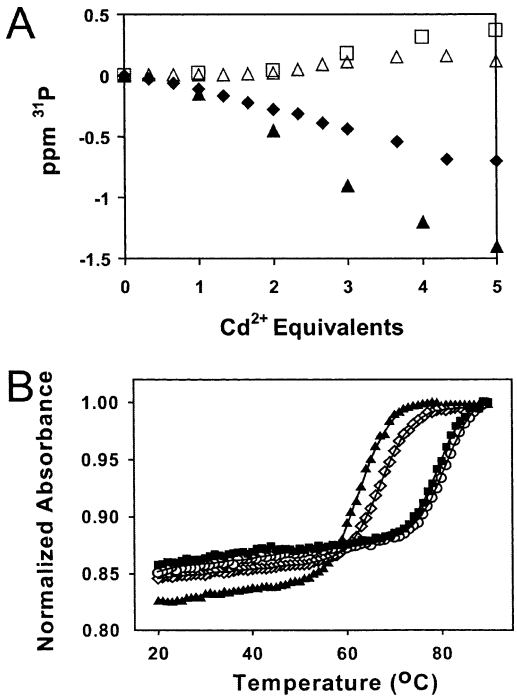

FIGURE 6: Thermal stability and metal binding data for the yeast U6 U80G and wild-type ISL. (A) ^{31}P 1D NMR chemical shift data for cadmium binding to phosphorothioate-substituted residue 80 in the wild-type and U80G ISL. Both the wild-type S_p (◆) and the U80G S_p (▲) display upfield shifts in the ^{31}P signal upon Cd^{2+} binding. Wild-type R_p (△) and U80G R_p (□) substitutions shift downfield and to a lesser degree. (B) Ultraviolet (260 nm) melting curve data for the yeast U6 U80G and wild-type ISL in 200 mM KCl at pH 7.0 [wild type (▲) and U80G (■)] and pH 5.5 [wild type (◇) and U80G (○)].

position and accessibility of the metal-binding *pro*-S_p phosphate oxygen are unperturbed in the U80G mutant structure. Previous wild-type studies in our laboratory have shown that metal binding and protonation of the N1 atom of A79 are mutually antagonistic, suggesting a possible mechanism for RNA-mediated regulation of splicing (*12*). The pK_a of A79 is near physiological pH when it is wobble paired with C67, making protonation possible in vivo. In the mutant ISL, the C67·A79 wobble pair is disrupted due to the formation of the G80·C67 base pair. Due to this change in conformation, protonation of A79 is not favorable in the mutant ISL, and on the basis of our measurements must be significantly lower than 5.3. Other mutations that likely disrupt the C67·A79 wobble pair, such as C67A and A79G, are also lethal or cause slow growth at low temperatures (*11*, *14*). Thus, it is possible that loss of the C·A wobble pair could contribute to the lethality of U80G, although it is unlikely that its disruption is a major factor since it does not cause lethality in all mutant ISLs (*14*, *38*). It is interesting to note that larger phosphorus chemical shift changes for residue 80 are observed with cadmium binding to the U80G mutation than to the wild-type ISL. Guanine is more electronegative than uracil and may provide additional metal binding affinity via unpaired electrons on its N7 and O6 functional groups.

Importance of Marginal Stability in the U6 ISL Structure. The U80G U6 ISL structure presented here, in combination with previous genetic studies, suggests that U80G lethality is due to the loss of a marginal stability required for the U6 ISL to rearrange. It is unlikely that the lethal phenotype of U80G is due to the loss of important catalytic functional groups, because U80A and U80C mutations have been shown to have no effect on growth in yeast (*14*). One obvious need for marginal stability is that U6 must undergo large conformational changes during spliceosome assembly and activation (*1*, *11*), and hyperstabilization of the U6 ISL may inhibit formation of the U4−U6 complex. Indeed, some mutations that are capable of conferring additional base pairings within the U6 ISL are cold sensitive and result in decreased concentrations of the U4−U6 complex in vivo, a phenotype that can be partially suppressed by overexpression of U4 (*11*). The cold sensitive mutations have normal growth rates at 30 °C. However, the U80G mutation is lethal at all temperatures, which suggests it is more stable than other cold sensitive mutations or the existence of an additional functional role for U80. Phylogenetic data indicate that an unpaired uracil is present in the U6 ISL internal loop of all genetically characterized eukaryotes, suggesting that base identity may contribute to tertiary structure. Therefore, the simplest explanation for all the genetic and structural data to date is a model in which the U6 ISL requires the ability to form tertiary contacts and undergo structural rearrangement. Such a tertiary interaction would be required to position the U6 ISL-bound metal ion close to the active site (*9*, *39*). The ability to undergo structural rearrangement is limited in the U80G mutant ISL, because of hyperstabilization due to the G80·C67 Watson−Crick pair.

CONCLUSION

The U6 RNA is evolutionarily the most highly conserved sequence in the spliceosome (*40*). The U6 ISL structure is conserved in all eukaryotes and functions to bind a catalytically essential metal ion at position 80 (*9*). The U80G U6 ISL structure presented here suggests a structural basis for the lethal mutation within U6 RNA: a base pairing rearrangement that leads to hyperstabilization of the structure. The structural, thermodynamic, and genetic evidence to date

indicates a required degree of marginal stability within the U6 ISL structure. We speculate that flexibility within the ISL is required for U4 association, as well as tertiary structure formation within the U2−U6 complex. We note that the U2−U6 secondary structure contains an imperfect four-way junction, and four-way junctions have known magnesium-dependent folding patterns that bring coaxially stacked helices into parallel alignment (*41*, *42*). Interestingly, the folding topology rules that can be derived from studies of four-way junctions suggest that the ISL would fold down onto helix 1 of the U2−U6 complex, bringing the metal binding site into the proximity of the ACAGAGA loop, as previously suggested (*39*).

ACKNOWLEDGMENT

We thank members of the Butcher lab and Gabriel Cornilescu who provided helpful discussion, Professor David Brow for comments on the manuscript, and the NMRFAM staff for technical support. We also thank Professor Thomas Record for spectrophotometer measurement time and Jonathan Cannon and Daniel Felitsky for technical assistance. NMR studies were carried out at NMRFAM with support from the NIH Biomedical Technology Program and additional equipment funding from the University of Wisconsin, NSF Academic Infrastructure Program, NIH Shared Instrumentation Program, NSF Biological Instrumentation Program, and the U.S. Department of Agriculture.

REFERENCES

1. Staley, J. P., and Guthrie, C. (1998) *Cell 92*, 315−326.
2. Burge, C. B., Tuschl, T. A., and Sharp, P. A. (1999) in *RNA World* (Atkins, J. F., Ed.) pp 525−560, Cold Spring Harbor Laboratory Press, Plainview, NY.
3. Nilsen, T. W. (1998) *RNA-RNA interactions in nuclear pre-mRNA splicing*, Cold Spring Harbor Laboratory Press, Plainview, NY.
4. Brow, D. A. (2002) *Annu. Rev. Genet. 36*, 333−360.
5. Collins, C. A., and Guthrie, C. (2000) *Nat. Struct. Biol. 7*, 850−854.
6. Fabrizio, P., and Abelson, J. (1990) *Mol. Biol. Rep. 14*, 135.
7. Fabrizio, P., and Abelson, J. (1992) *Nucleic Acids Res. 20*, 3659−3664.
8. Yu, Y. T., Maroney, P. A., Darzynkiewicz, E., and Nilsen, T. W. (1995) *RNA 1*, 46−54.
9. Yean, S. L., Wuenschell, G., Termini, J., and Lin, R. J. (2000) *Nature 408*, 881−884.
10. Valadkhan, S., and Manley, J. L. (2001) *Nature 413*, 701−707.
11. Fortner, D. M., Troy, R. G., and Brow, D. A. (1994) *Genes Dev. 8*, 221−233.
12. Huppler, A., Nikstad, L. J., Allmann, A. M., Brow, D. A., and Butcher, S. E. (2002) *Nat. Struct. Biol. 9*, 431−435.
13. Madhani, H. D., Bordonne, R., and Guthrie, C. (1990) *Genes Dev. 4*, 2264−2277.
14. McPheeters, D. S. (1996) *RNA 2*, 1110−1123.
15. Zhou, H., Vermeulen, A., Jucker, F. M., and Pardi, A. (1999) *Biopolymers 52*, 168−180.
16. Grzesiek, S., Cordier, F., and Dingley, A. J. (2001) *Methods Enzymol. 338*, 111−133.
17. Dingley, A. J., Masse, J. E., Feigon, J., and Grzesiek, S. (2000) *J. Biomol. NMR 16*, 279−289.
18. Luy, B., and Marino, J. P. (2000) *J. Am. Chem. Soc. 122*, 8095−8096.
19. Hennig, M., and Williamson, J. R. (2000) *Nucleic Acids Res. 28*, 1585−1593.
20. Hansen, M. R., Hanson, P., and Pardi, A. (2000) *Methods Enzymol. 317*, 220−240.
21. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
22. Allain, F. H., and Varani, G. (1995) *Nucleic Acids Res. 23*, 341−350.
23. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) *J. Magn. Reson. 131*, 159−162.
24. Zweckstetter, M., and Bax, A. (2000) *J. Am. Chem. Soc. 122*, 3791−3792.
25. Zweckstetter, M., and Bax, A. (2002) *J. Biomol. NMR 23*, 127−137.
26. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grossekunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
27. Leeper, T. C., Martin, M. B., Kim, H., Cox, S., Semenchenko, V., Schmidt, F. J., and Van Doren, S. R. (2002) *Nat. Struct. Biol. 9*, 397−403.
28. Warren, J. J., and Moore, P. B. (2001) *J. Magn. Reson. 149*, 271−275.
29. Brunger, A. T. (1992) *X-PLOR (Version 3.1) Manual*, Yale University Press, New Haven, CT.
30. Schultze, P., and Feigon, J. (1997) *Nature 387*, 668.
31. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55, 29−32.
32. Breslauer, K. J. (1995) *Methods Enzymol. 259*, 221−242.
33. Maderia, M., Horton, T. E., and DeRose, V. J. (2000) *Biochemistry 39*, 8193−8200.
34. Maderia, M., Hunsicker, L. M., and DeRose, V. J. (2000) *Biochemistry 39*, 12113−12120.
35. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
36. Madhani, H. D., and Guthrie, C. (1992) *Cell 71*, 803−817.
37. Wolff, T., and Bindereif, A. (1995) *Biochim. Biophys. Acta 1263*, 39−44.
38. Sun, J. S., and Manley, J. L. (1997) *RNA 3*, 514−526.
39. Valadkhan, S., and Manley, J. L. (2002) *Nat. Struct. Biol. 9*, 498−499.
40. Brow, D. A., and Guthrie, C. (1988) *Nature 334*, 213−218.
41. Zhao, Z. Y., Wilson, T. J., Maxwell, K., and Lilley, D. M. (2000) *RNA 6*, 1833−1846.
42. Rupert, P. B., and Ferre-D'Amare, A. R. (2001) *Nature 410*, 780−786.

BI027137H