Structure of the Polyadenylation Regulatory Element of the Human U1A Pre-mRNA 3′-Untranslated Region and Interaction with the U1A Protein

Charles C. Gubser and Gabriele Varani*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH England

Received September 28, 1995; Revised Manuscript Received December 6, 1995[⊗]

ABSTRACT: The N-terminal RNP domain of U1A binds two different RNA substrates with high affinity and specificity: stem−loop II of the U1 snRNA and a complex secondary structure in the 3′-untranslated region (3′-UTR) of the U1A pre-mRNA. Both RNAs contain a single-stranded sequence which is the main site of interaction with the protein, but in completely different structural contexts. Here we describe the solution structure of the free 3′-UTR RNA molecule and the NMR characterization of its complex with the U1A protein N-terminal domain. The structure of the free RNA indicates that the stems are nearly canonical A-form helices and that the single-stranded region contains local stacking interactions in the context of a generally flexible structure. Upon protein binding, the internal loop region folds into an ordered structure containing significant changes in the local stacking interactions. These results demonstrate the role of RNA structure and folding in specific RNA−protein recognition.

The U1A protein is an essential component of U1 snRNP,[1] one of five RNA−protein complexes which carry out pre-mRNA splicing in the eukaryotic nucleus. U1A binds stem−loop II of the U1 snRNA with very high affinity ($K_d \approx 10^{-11}$ M) (Hall & Stump, 1992; Jessen et al., 1991; Scherly et al., 1990). Although U1A is found in several species, its function varies. In yeast, the U1A protein can be deleted without interfering with splicing, whereas in *Drosophila* the *Snf* gene encodes a protein highly homologous to U1A which participates in sex determination and is embryonic lethal if deleted (Flickinger & Salz, 1994). *In vitro* experiments have suggested that U1A may participate in coupling splicing and polyadenylation (Lutz & Alwine 1994). The human U1A protein consists of 283 amino acids and contains two RNP (or RRM) domains, a ubiquitous RNA binding motif (Burd & Dreyfuss, 1994; Mattaj, 1993), separated by a diffuse nuclear localization signal (Kambach & Mattaj, 1992). The N-terminal RNP domain and its proximal C-terminal region are sufficient to confer full RNA binding affinity and specificity to the U1A protein (Scherly et al., 1990, 1991). The N-terminal RNP domain of U1A has been extensively investigated biochemically (Hall, 1994; Hall & Stump, 1992; Jessen et al., 1991; Scherly et al., 1990; Stump & Hall, 1995) and also structurally by crystallography (Jessen et al., 1991; Nagai et al., 1990, 1995; Oubridge et al., 1994) and NMR (Avis et al., 1996; Hoffman et al., 1991; Howe et al., 1994).

In addition to its role in splicing, U1A regulates its own production by binding to the 3′-untranslated region (3′-UTR) of its own pre-mRNA and preventing polyadenylation (Boelens et al., 1993; Gunderson et al., 1994; van Gelder et al., 1993). The 3′-UTR of the U1A pre-mRNA contains an extended double-helical structure interrupted by two internal loops (Figure 1A) (van Gelder et al., 1993). The two internal loops contain a seven nucleotide single-stranded region very similar in sequence to the U1A binding site in stem−loop II of the U1 snRNA. In addition to this conserved sequence, the integrity of the helical stems flanking the internal loops is also necessary for binding (van Gelder et al., 1993). When two U1A proteins bind to the 3′-UTR, polyadenylation is inhibited, apparently by a direct interaction between U1A and poly(A) polymerase: the first ≈138 amino acids of U1A are both necessary and sufficient for this function (I. Mattaj, personal communication).

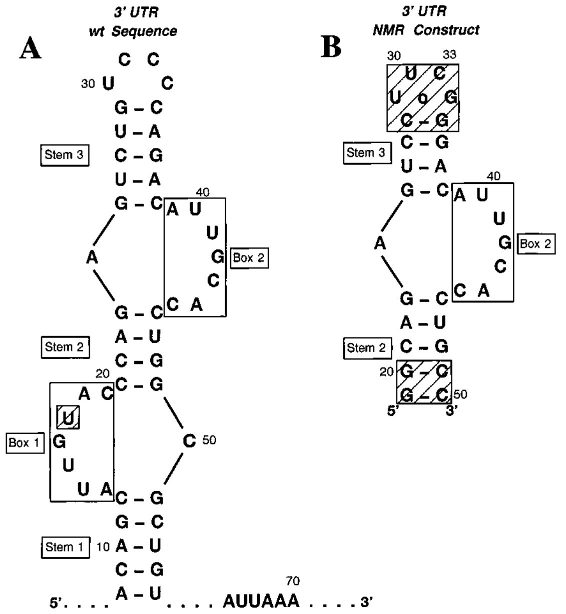

FIGURE 1: (A) Sequence and secondary structure of part of the U1A 3′-UTR. Box 1 and box 2 are the sites of interaction with the U1A N-terminal RNP domain. The nucleotide in the cross-hatched box, U17, is the only deviation from the consensus binding site for the protein: AUUGCAC. The polyadenlyation signal AUUAAA is just downstream of the extended helical structure. (B) The sequence and secondary structure of the 3′-UTR NMR construct, with the box 2 sequence highlighted. The cross-hatched boxes indicate the nucleotides which differ from the wild type sequence.

The N-terminal RNP domain of U1A binds to both of its RNA cognates with very high affinity: the K_d for the U1

[⊗] Abstract published in *Advance ACS Abstracts,* February 1, 1996.

[1] Abbreviations: 3′-UTR, 3′-untranslated region; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy; 2D, two dimensional; 3D, three dimensional; RNP, ribonucleoprotein; RRM, RNA recognition motif; snRNA, small nuclear RNA.

0006-2960/96/0435-2253$12.00/0

snRNA stem−loop II is ≈10^{-11} M (Hall & Stump, 1992), whereas the K_ds for the internal loops in the 3′-UTR of the U1A pre-mRNA are ≈10^{-10} M (box 2) and 10^{-9} M (box 1) (van Gelder et al., 1993). The sequence homology between stem−loop II and the internal loops of the 3′-UTR suggests that at least some of the contacts between the protein and the single-stranded stretches of RNA are conserved (Nagai et al., 1995; Oubridge et al., 1994). However, the secondary structure context of these single-stranded regions is completely different and yet critical, since U1A binds very weakly to single-stranded RNAs containing the consensus binding site (Hall, 1994; Hall & Stump, 1992; Tsai et al., 1991; van Gelder et al., 1993). Not surprisingly, disrupting the secondary structure of the 3′-UTR interferes with the biological activity of the protein as assayed in nuclear extracts (van Gelder et al., 1993). It is puzzling how U1A can bind two different structures, the 3′-UTR and the stem−loop II, with such high affinity, and we have therefore used NMR to analyze the interaction between U1A and the 3′-UTR of its own pre-mRNA.

The U1A binding site from the 3′-UTR of the U1A pre-mRNA is 45 nucleotides long and contains two binding sites for the protein. A complex between this RNA and two RNP domains would be very large for NMR investigation (45 kDa). We have thus prepared a shortened RNA construct containing a single protein binding site. All of the resonances in the molecule have been assigned using heteronuclear multidimensional NMR, and the free RNA structure has been determined from 971 experimentally derived distance and dihedral angle constraints. The structural analysis shows that the stems flanking the internal loop are formed in the absence of the protein, and that the internal loop is single stranded. No base pairing interactions were detected in the loop, but extensive stacking interactions were formed in the context of a flexible loop structure. Nearly complete assignments have also been obtained for the 30-nucleotide RNA in complex with the 102 amino acid N-terminal domain of U1A. This analysis reveals the sites of interactions between the U1A protein and the 3′-UTR RNA structure and demonstrates that folding of the RNA single-stranded region accompanies protein binding. The assignments and the structural investigation reported here form the basis for the formal determination of the structure of the protein−RNA complex.

EXPERIMENTAL PROCEDURES

Isotopically Labeled Nucleotides. ^{13}C- and/or ^{15}N-labeled nucleotides were isolated from *Escherichia coli* as described (Batey et al., 1992; Hines et al., 1994; Nickonowicz et al., 1992). Briefly, strain JM101 cells were grown on minimal medium with 99% ^{13}C-labeled glucose and/or 99% ^{15}N-labeled NH_4Cl. Cells were lysed, and ribosomes were harvested by centrifugation and then phenol/chloroform extracted before ethanol precipitation. One liter of bacterial growth ordinarily yielded 60−80 mg of labeled RNA. The RNA was hydrolyzed to nucleotide monophosphates by treatment with RNase P1. After P1 digestion, the nucleotides were enzymatically rephosphorylated as described. NTPs were purified by affinity chromatography on a boronate−polyacrylamide column (Bio-Rad).

RNA Synthesis by Single-Stranded Runoff Transcription. RNA was synthesized from single-stranded DNA templates (Milligan et al., 1987). Two DNA oligonucleotides were chemically synthesized, a top strand (5′ TAA TAC GAC TCA CTA TAG 3′) and a template strand (5′ GGC AGG TGC AAT GTC CCG AAG GAC TCT GCC TAT AGT GAG TCG TAT TA 3′). The top and template strands form a promoter for the phage T7 RNA polymerase. The RNA oligonucleotide that results from T7 runoff transcription has the sequence 5′ pppGGCAGAGUCCUUCGGGACAUUGCACCUGCC 3′. Note the presence of the UUCG tetraloop. RNA was purified from transcription mixtures by electrophoresis on 7 M urea 20% polyacrylamide gels. After electroelution from the gel followed by ethanol precipitation, the RNA was dialyzed for 24 h against 0.5 M NaCl and then for 24 h against water. After dialysis, the RNA was lyophilized and resuspended in H_2O and then loaded on a G-15 Sephadex column equilibrated with 2 mM NaH_2PO_4 (pH 5.5) and eluted with the same buffer. Fractions containing RNA were lyophilized, resuspended in water, and then dialyzed against the final NMR buffer (5 mM NaH_2PO_4, pH 5.5) for 24 h.

T7 RNA Polymerase. T7 RNA polymerase was prepared according to the protocol of Dr. Kiyoshi Nagai (unpublished). *E. coli* cells strain BL21 expressing phage T7 RNA polymerase under the control of an inducible promoter were the gift of Dr. Jonathan Karn.

U1A Protein. The N-terminal RNA-binding domain of U1A, a fragment consisting of amino acids 2-102 after posttranslational cleavage of the N-terminal methionine, was expressed in pET13a *E. coli* (Gerchman et al., 1994) and purified as described (Nagai et al., 1990). Two amino acid substitutions have been introduced to increase the solubility of the RNA−protein complex without affecting RNA binding (Oubridge et al., 1994): Tyrosine 31 was mutated to histidine and glutamine 36 to arginine. The protein was ^{13}C- and/or ^{15}N-labeled by growing *E. coli* cells on M9 minimal medium supplemented with $^{15}NH_4Cl$ and/or [^{13}C]glucose.

Band Shift Experiments. Band shift experiments were performed with the A102 truncation of U1A and the 3′-UTR NMR construct. RNA was treated with alkaline phosphatase and then with T4 polynucleotide kinase in the presence of [γ-^{32}P]ATP. The ^{32}P-labeled RNA was repurified on a 20% acrylamide−7 M urea gel, eluted from the gel strips, and ethanol precipitated. Binding mixtures were assembled from a 20 nM stock of RNA, 10× TK (0.5 M Tris pH 7.1, 0.2 M KCl), 10% Triton X-100, 1 M DTT, and 100 mg/mL tRNA (Sigma). Ordinarily, 100 μL of RNA mix was prepared, containing the following: 20 μL of RNA stock, 20 μL of 10× TK, 2 μL of 10% Triton X-100, 20 μL of DTT, 1.4 μL of tRNA, and 36.6 μL of water. Each binding reaction contained 10 μL of this mixture and 10 μL of a protein stock of known concentration. Binding mixtures were incubated at room temperature for 20−30 min. The RNA concentration in the binding mixtures was 2 nM, and the protein concentration was varied between 4 and 100 nM. Binding mixtures were loaded on 10% acrylamide nondenaturing gels. Fifty milliliters of the gel stock contained 16.7 mL of 30% acrylamide (30% acrylamide, 0.8% bis acrylamide, w/v), 2.5 mL of 10× TB (0.9 M boric acid, 0.9 M Tris, pH 8.1), 0.5 mL of 10% Triton X-100, and 34.5 mL of water. The gels were run at 4 °C in 1/2× TB + 0.1% Triton X-100 for 1.5 h.

NMR Spectroscopy. NMR spectra were recorded on Bruker AMX-500 or DMX-600 spectrometers equipped with triple-resonance gradient probes. The AMX-500 operates at 500.14 MHz and the DMX-600 operates at 600.13 MHz for ^{1}H. The spectra were referenced to TSP (^{1}H and ^{13}C) and phosphoric acid (^{31}P), either directly or by comparison with published values, or to external NH_3 (^{15}N, D. Live, personal communication). Data were processed using FELIX (Biosym, San Diego). For two-dimensional spectra, 1024 complex points were acquired in t_2 and (typically) 512 real points in t_1, employing the TPPI scheme (Marion & Wüthrich, 1983). Spectra were zero-filled to a final 2K × 2K (real) data size after apodization with 40°−60° shifted sine bell functions.

One-dimensional spectra in 95% H_2O/5% D_2O of the free RNA were recorded with a 12 500 Hz spectral width at 1.5 mM RNA in 5 mM sodium phosphate buffer at pH 6 using the jump−return scheme for water suppression (Plateau & Gueron, 1982). The 2D NOESY spectra in H_2O were recorded at 250 ms mixing time (free RNA) or 100 ms mixing time (RNA−protein complex), under the same conditions as the 1D spectra and using the jump−return water suppression scheme on the last pulse of the sequence. The 2D NOESY (60, 100, 120, 200, and 300 ms mixing times), TOCSY (60 ms mixing time), and 2QF-COSY (^{31}P-decoupled) spectra were recorded in 99.99% D_2O at 1 mM RNA concentration, at a temperature of 293 K using very low power presaturation on the residual H_2O resonance. The spectral width was 8,012 Hz in each dimension, and the total acquisition time was ≈22 h per spectrum. A 2QF-COSY spectrum was also recorded with reduced spectral width (2500 Hz) to facilitate the analysis of the sugar proton region.

Assignments of the RNA spectrum were aided by a ^{15}N-labeled sample. HMQC spectra (Szewczak et al., 1993) were collected in 95% H_2O/5% D_2O and used to distinguish G and U imino protons and assign amino nitrogen resonances. HSQC spectra in D_2O were collected with long INEPT delays (60−80 ms) to optimize 2−3 bond transfer between base nitrogen and proton resonances (Sklénar et al., 1994). Spectra on the RNA−protein complex were acquired with both ^{15}N-labeled and unlabeled protein samples. HSQC and HSQC-NOESY spectra (mixing time 100 ms) were acquired using spin lock pulses (Messerle et al., 1989) to surpress the solvent while reducing bleaching of the exchangeable protein and RNA resonances observed in spectra acquired with solvent presaturation.

A 3D ^{13}C-edited NOESY-HMQC spectrum (Zuiderweg et al., 1990) was recorded in ≈66 h at 200 ms mixing time (free RNA) or 100 ms mixing time (RNA−protein complex) with ^{15}N−^{13}C GARP (Shaka et al., 1985) decoupling during acquisition. The spectral widths were 8012 Hz (each ^{1}H dimension) and 6944 Hz (^{13}C). A total of 256 points were recorded in the first ^{1}H dimension ($t_1^{max} = 15.9$ ms) and 64 in the ^{13}C dimension ($t_2^{max} = 5.3$ ms). The data were processed to a final size of 512 × 128 (^{13}C) × 1024 real points. Folded resonances (C8, C6, C2) had opposite phase to all other resonances (Archer et al., 1992). A three-dimensional HMQC-“clean” TOCSY spectrum was acquired with a 50 ms mixing time without ^{13}C-decoupling in either dimension (Hines et al., 1994). Processing and acquisition parameters were similar to those used for the 3D NOESY experiment, but the ^{1}H spectral widths were reduced to 2500 Hz to facilitate the analysis of the sugar proton region. Two- and three-dimensional HCCH-E.COSY (Griesinger & Eggenberger, 1992; Schwalbe et al., 1994) experiments were acquired at 500 MHz using spectral widths of 2500 Hz (^{1}H) and 6000 Hz (^{13}C) in 22 and 66 h, respectively. No deconvolution procedure was used to correct the peak to peak separations for relaxation artifacts (Schwalbe et al., 1994). An HCCH-COSY spectrum (Kay et al., 1990) was acquired in 70 hs for the RNA−protein complex with acquisition parameters similar to those employed for the 3D NOESY spectra. ω_1-^{13}C-half-filtered NOESY spectra (Gemmecker et al., 1992; Otting & Wüthrich, 1990) were recorded at 100 ms mixing time for the complex of labeled protein and unlabeled RNA. 3D NOESY spectra in water were acquired using gradient pulses for suppressing the solvent with [^{15}N-edited NOESY (Jahnke et al., 1995; Stonehouse et al., 1994)] or without [^{13}C-edited NOESY (Majumdar & Zuiderweg, 1993)] sensitivity enhancement and water flip-back to minimize sensitivity losses due to exchange with solvent.

A series of ^{31}P−^{1}H correlation spectra were acquired for both the free and bound RNA using the AMX-500 spectrometer equipped with an inverse probe. Two-dimensional Het-Cor (Sklénar et al., 1986), Hetero-Tocsy (Kellogg, 1992), HSQC, and HSQC-NOESY (Gronenborn et al., 1989) (INEPT delay 20 ms for both HSQC experiments, mixing time 200 ms for the NOESY portion of the experiment) spectra were acquired with spectral widths of 2500 Hz (^{1}H) and 1000 Hz (^{31}P). A total of 128 FIDs of 512 real points were collected ($t_2^{max} \approx 160$ ms, $t_1^{max} \approx 64$ ms) with total acquisition times of ≈20 h per experiment. Three-dimensional HCP (Marino et al., 1994) and ^{13}C-edited Het-cor (P,H-COSY-H,C-HMQC) (Varani et al., 1995) spectra were recorded in ≈22 h at 600 Mhz with spectral widths of 3004 Hz (^{1}H), 1000 Hz (^{31}P), and 4000 Hz (^{13}C). Sixty-four points were collected in each indirectly detected dimension ($t_1^{max} \approx 32$ ms, $t_2^{max} \approx 8$ ms), and the final size of the processed data was 64 (^{31}P) × 128 (^{13}C) × 512 real points.

Interproton Distance Constraints. NOE cross peaks from 2D NOESY spectra recorded in D_2O at mixing times of 60, 100, 120, and 200 ms and from the 3D NOESY with a mixing time of 200 ms were used to determine interproton distances. For the RNA−protein complex, mixing times of 40, 80, and 120 ms were used instead. NOE cross peaks were subdivided into five classes based on their intensities: strong (0−3 Å), medium (0−4 Å), weak (0−5 Å), very weak (0−6 Å). Peaks present only at long mixing time were given even looser upper bounds (0−7 Å). Pyrimidine H5−H6 cross peaks were used as reference since the distance between them is fixed by covalent geometry (2.41 Å). Many cross peaks which were overlapped in the 2D NOESY spectra could be resolved in the 3D ^{13}C-edited NOESY spectrum. Cross peak intensities from 3D experiments were used more conservatively, since the cross peak intensity is dependent not only on the interproton distance, but also the one-bond H−C couplings and the different relaxation rates of different ^{13}C resonances (this problem is particularly severe for the RNA−protein complex). Furthermore, only one long mixing time NOESY was used. Finally, the aromatic carbon resonances are ≈60 ppm away from the carrier: off-resonance pulse imperfections compromise the reliability of cross peak intensities.

NOE cross peaks involving exchangeable protons were translated into distance constraints with even greater attention, since exchange rates and the excitation profile of the jump−return water suppression affect the cross peak intensities. Most of the NOEs involving exchangeable protons were classified as weak, or very weak, with the exception of NOEs between uracil NH and adenine H2 protons and guanosine NH to cytosine NH$_2$ in Watson−Crick base pairs. These were included as medium constraints.

Hydrogen Bonding Constraints. Hydrogen bonds in Watson−Crick base pairs were included in the constraint list. Three constraints were used in G-C pairs and two constraints in A-U pairs. For N···H−N hydrogen bonds, the heavy atoms were constrained to 2.65−3.25 Å. For O···H−N hydrogen bonds, the heavy atoms were constrained to 2.64−3.17 Å. In order for a hydrogen bonding constraint to be included, it was necessary to observe a significant downfield ^{1}H shift and a slow rate of exchange with solvent for the proton involved in the hydrogen bond.

Dihedral Angle Constraints. The procedure used to derive dihedral angle constraints has recently been described in detail (Allain & Varani, 1995; Varani et al., 1995). Briefly, α (O3′−P−O5′−C5′) and ζ (C3′−O3′−P−O5′) were constrained only qualitatively from ^{31}P chemical shifts (Gorenstein, 1984). Whenever the ^{31}P chemical shift values were found in the −4 to −5 ppm range common to phosphates in regular A- or B-form structures, α and ζ were very loosely restrained to exclude the *trans* conformation (constraints of 0° ± 120° were used for both dihedrals). The β (P−O5′−C5′−C4′) dihedral angles were constrained using semiquantitative estimates of the $^3J_{\mathrm{PH5'}}$, $^3J_{\mathrm{PH5''}}$ and $^3J_{\mathrm{PC4'}}$ scalar couplings (Altona, 1982) obtained from 3D or, when possible, 2D Het-TOCSY and Het-COR experiments. In almost all cases, the PH5′ and PH5″ cross peaks were clearly absent or very weak, as expected for regular A-form helices with β in the *trans* conformation. Cross peaks involving the C3′ or C5′ resonances in HCP spectra (generated by 4−5 Hz two-bond active couplings), and the four-bond PH4′ couplings in Het-COR spectra, provided internal controls: in the flexible internal loop, analysis of cross peaks corresponding to these known couplings identified resonances which were broadened. Semiquantitative estimates of the scalar coupling are sufficient to constrain β to the *trans* domain (180 ± 30° or 50°). The ϵ (C4′−C3′−O3′−P) dihedral angles were constrained from values of $^3J_{\mathrm{C4'P}}$, $^3J_{\mathrm{C2'P}}$, and $^3J_{\mathrm{H3'P}}$ (Altona, 1982). Constraints of 210 ± 30° (ϵ *trans*) or 260 ± 30° (ϵ *gauche*$^-$) were typically applied. The γ (O5′−C5′−C4′−C3′) dihedral angles were constrained using estimates of $^3J_{\mathrm{H4'H5'}}$ and $^3J_{\mathrm{H4'H5''}}$ couplings obtained from the 3D ^{13}C-edited TOCSY experiment and (in a few cases) 2D and 3D HCCH-E. COSY experiments. In the *gauche*$^+$ conformers found in A-form helices, both couplings are very small, whereas either $^3J_{\mathrm{H4'H5'}}$ or $^3J_{\mathrm{H4'H5''}}$ are large (approximately 10 Hz) in the *trans* or *gauche*$^-$ conformations. γ was constrained to 50 ± 20° or 40° for the *gauche*$^+$ conformation. The observation of resolved four-bond PH4′ couplings in heteronuclear ^{1}H−^{31}P correlation spectra provided in several cases independent support to this conclusion. The sugar puckers, identifying the δ (C5′−C4′−C3′−O3′) dihedral angles, were constrained using a variety of ^{1}H−^{1}H and ^{1}H−^{13}C scalar couplings to C3′-*endo* ($\delta = 85 \pm 25°$) or C2′-endo (160 ± 30°) conformations or were left unconstrained in cases where significant conformational averaging was present.

Structure Calculation. Structure calculations were carried out with the program X-PLOR 3.1 (Brünger, 1990). As noted (Allain & Varani, 1995), the parameter file specifying sugar ring conformation was modified to ensure more realistic sugar puckers. A total of 971 NMR derived restraints were used in the structure calculation: 820 NOE derived distance constraints, 25 hydrogen bonding constraints, and 126 dihedral angle restraints. No constraints were added to keep the base pairs planar. The details of the calculation are identical to those just described in detail (Allain & Varani, 1995). Briefly, a set of 50 structures with randomized backbone torsion angles were used as starting structures for restrained molecular dynamics. The calculations were carried out in three steps: (1) a simulated annealing protocol used the distance restraints to fold the molecule in a 15 ps dynamics run at 1000 K, after which the system was cooled to 300 K, (2) a second dynamics protocol refined the structure in a 2 ps dynamics run at 1000 K in which the dihedral angle constraints were introduced gradually, followed by slowly cooling the system to 300 K, and (3) finally, the structures were energy minimized taking van der Waals interactions into account, but excluding electrostatics from the calculation. The details of convergence and agreement with the data are discussed below.

Structure Analysis. Root mean square deviations from average structures were calculated using the program Superpose (Diamond, 1992). Helical parameters were calculated with the program RNA (Babcock et al., 1993), kindly provided by Dr. Marla Babcock.

RESULTS

The RNA Construct. The sequence and secondary structure of the 3′-UTR of the U1A pre-mRNA are shown in Figure 1A. Box 1 and box 2 are the binding sites for the N-terminal RNP domains of U1A (van Gelder et al., 1993). The seven nucleotides of box 2 are identical in sequence to the first seven nucleotides of the loop of the U1 snRNA stem−loop II, AUUGCAC, but the sequence of box 1 differs from this consensus by the substitution of a U for the first C, AUUGUAC. U1A binds to box 2 slightly more tightly than to box 1 (van Gelder et al., 1993). We have prepared an RNA containing box 2 as well as three base pairs from each of the stems flanking the internal loop. Figure 1B shows the RNA construct used in the present investigation. Two features have been added: (1) the 5′ terminus has been modified to include two guanosine residues for the purpose of increasing the efficiency of transcription with T7 RNA polymerase, and (2) a UUCG tetraloop has been added to stem 3 to ensure the stability of that stem and to provide a starting point for NMR assignments.

Binding of U1A to this 30-nucleotide RNA was assayed by native gel electrophoresis. The results show that even in the presence of a high concentration of competitor tRNA (0.7 mg/mL, corresponding to >10 000-fold molar excess), the N-terminal RNP domain of U1A binds to the 3′-UTR NMR construct with $K_{\mathrm{d}} \leq 10^{-9}$ M in 20 mM KCl and 50 mM Tris (pH 7.1).

Spectral Assignments. The NMR spectrum of the free RNA molecule was assigned as described in the remainder of this section, and assignments are reported in Tables 1 and 2. Details of the UUCG assignments are in complete accord with previous work (Allain & Varani, 1995; Allain & Varani, 1995). The 2′-OH resonance of U29 was clearly visible in the present set of spectra as well and provides several important experimental constraints for the tetraloop structure.

Table 1: Proton Chemical Shifts for the Free and Bound 3′-UTR RNA at 300 K and 5 mM Sodium Phosphate Buffer, pH 5.5, Referenced to TSP

nt	**1′**	**2′**	**3′**	**4′**	**5′/5″**	**6/8**	**2/5**	**i**	**am**
G19	**5.83**	**5.00**	**4.80**	**4.60**	**4.50/4.34**	**8.21**		**13.46**	
G20	**5.95**	**4.61**	**4.66**	**4.46**	**4.62/4.28**	**7.67**		**13.43**	**8.20/7.98**
C21	**5.57**	**4.63**	**4.63**	**4.50**	**4.57/4.13**	**7.74**	**5.34**		**8.50/6.40**
	5.58	**4.63**	**4.59**	**4.48**	**4.58/4.11**	**7.74**	**5.31**		***8.74/7.00***
A22	**5.98**	**4.69**	**4.72**	**4.57**	**4.57/4.22**	**8.01**	**7.07**		**8.47/7.00**
	5.97	**4.69**	**4.69**	**4.56**		***8.20***	**7.11**		***8.21/6.51***
G23	**5.61**	**4.36**	**4.60**	**4.50**	**4.58/4.16**	**7.22**		**13.24**	
	5.65	**4.32**	***4.70***	**4.48**	**/4.19**	***7.56***		**13.23**	**8.35/6.97**
A24	**6.14**	**4.74**	**4.96**	**4.64**	**4.47/4.30**	**8.17**	**8.15**		
	6.28	***4.32***	***4.66***	***4.52***	**4.48/4.29**	**8.13**	***8.24***		
G25	**5.85**	**4.75**	**4.66**	**4.62**	**4.58/4.29**	**8.13**		**12.70**	**8.49/6.50**
	5.98	**4.70**	**4.61**	**4.61**		**8.20**		***12.82***	***8.79/6.88***
U26	**5.70**	**4.61**	**4.61**	**4.57**	**4.64/4.22**	**7.99**	**5.23**	**14.50**	
C27	**5.65**	**4.52**	**4.52**	**4.50**	**4.60/4.20**	**7.96**	**5.74**		**8.57/7.08**
C28	**5.62**	**4.48**	**4.71**	**4.51**	**4.60/4.22**	**7.72**	**5.55**		**8.60/7.00**
U29	**5.63**	**3.83**	**4.57**	**4.42**	**4.53/4.22**	**8.04**	**5.60**	**11.92**	
U30	**6.15**	**4.72**	**4.08**	**4.52**	**4.30/4.10**	**8.09**	**5.92**	**11.36**	
C33	**6.02**	**4.15**	**4.56**	**3.86**	**3.67/2.78**	**7.73**	**6.20**		**7.18/6.48**
G34	**6.02**	**4.92**	**5.70**	**4.47**	**4.47/4.26**	**7.91**		**9.90**	6.89
G35	**4.56**	**4.61**	**4.37**	**4.46**	**4.60/4.36**	**8.34**		**12.87**	**8.55/6.44**
G36	**5.83**	**4.61**	**4.66**	**4.48**	**4.54/4.10**	**7.32**		**12.44**	**8.13**
A37	**6.06**	**4.64**	**4.62**	**4.54**	**4.64/4.16**	**7.87**	**7.86**		**8.30/6.72**
C38	**5.39**	**4.53**	**4.33**	**4.45**	**4.50/4.09**	**7.27**	**5.23**		**8.24/7.16**
	5.39	**4.51**	**4.32**	**4.47**	**4.50/4.07**	**7.28**	**5.17**		***7.93/7.27***
A39	**5.96**	**4.40**	**4.45**	**4.47**	**4.50/4.11**	**7.92**	**7.35**		
	5.97	**4.35**	**4.51**	4.50		**7.93**	***7.09***		
U40	**5.31**	**3.96**	**4.34**	**4.18**	**4.29/3.98**	**7.36**	**5.25**	11.11	
	5.34	***3.75***	***4.24***	4.23	**4.25/3.93**	***7.07***	**5.22**	***13.50***	
U41	**5.76**	**4.23**	**4.47**	**4.17**	**3.92/3.92**	**7.61**	**5.68**	11.39	
	5.25	***3.84***	*4.37*	***3.93***		***7.33***	**5.62**	***11.95***	
G42	**5.73**	**4.77**	**4.79**	**4.47**	**4.20/4.14**	**7.87**		**10.64**	
	6.12	***5.26***	***5.11***			***7.69***		***9.69***	
C43	**5.79**	**4.43**	**4.49**	**4.37**		**8.06**	**6.04**		
	6.24	***4.05***	**4.43**	***5.24***		***7.27***	***5.30***		
A44	**5.89**	**4.75**	**4.62**	**4.48**	**4.39/4.23**	**8.24**	**7.95**		
	5.47	***4.54***	***4.74***	**4.43**		***8.50***	***7.50***		
C45	**6.02**	**4.44**	**4.67**	**4.54**	**4.22/4.42**	**7.82**	**5.87**		
	5.68	***4.27***	***4.43***	***4.32***		***7.77***	***5.30***		
C46	**5.50**	**4.55**	**4.35**	**4.46**	**4.57/4.13**	**7.71**	**5.65**		**8.62/6.98**
	5.67	**4.53**	**4.31**	**4.50**	**4.64/4.14**	***8.13***	***6.16***		***8.85/7.17***
U47	**5.60**	**4.65**	**4.62**	**4.41**	**4.66/4.21**	**7.84**	**5.82**	**13.79**	
G48	**5.87**	**4.56**	**4.66**	**4.55**	**4.63/4.22**	**7.84**		**12.57**	**8.17/6.51**
	5.84	**4.56**	**4.66**	**4.53**	4.63/4.19	**7.85**		***12.74***	
C49	**5.55**	**4.29**	**4.55**	**4.48**	**4.57/4.13**	**7.74**	**5.30**		**8.61/7.03**
C50	**5.81**	**4.06**	**4.19**	**4.22**	**4.54/4.11**	**7.72**	**5.55**		**8.52/6.91**

[a] The chemical shifts for the free RNA molecule are uppermost in each box, and the bound RNA chemical shifts are reported below. Only resonances at or near the internal loop region are significantly shifted in the presence of the U1A protein, and are included in this and the following table. Chemical shift values in lightface type indicate tentative assignments.

The assignments of stems 2 and 3 were straightforward. The two A H2s were identified using NOESY spectra acquired in H_2O (Figure 2A,B) and starting from the U imino resonances, which were assigned from their connections with one another and with known tetraloop resonances. In addition to this pivotal information, several tentative assignments could be made from these spectra for H1's and H5s in the stems. Using the A H2 assignments as starting points, the majority of the H8, H6, H5, H2, and H1′ resonances were assigned from a combination of NOESY (Figure 3A,B) and correlated experiments.

Although complete assignments for the internal loop required ^{13}C labeling, initial assignments were made on the basis of homonuclear spectra. The preservation of A-form stacking between G23 and A24 led to the identification of the A24 resonances. Connections between the G25 H1′ and A24 H2, and A24 H2 and A24 H1′ led to assignments of the A24 H2 and most sugar resonances. Similarly, C38 and A39 were stacked together, as were A39, U40, and U41, as revealed by characteristic NOE interactions: the first three loop nucleotides display regular features of A-form stacking. G42 and A44 were identified by a process of elimination (there were no other purines unassigned in the sequence), and A44 was distinguished from G42 on the basis of A44's connections with C45. C45 was stacked on C46, which allowed unambiguous assignment of C45 and consequently A44. C43 appears to be flexible: the H5−H6 cross peak in the 2QF-COSY is weak, indicating line broadening, and there are very few NOEs to the resonances at this position.

With ^{13}C- and ^{15}N-labeled samples, various heteronuclear experiments were used to identify resonances which were overlapped in the 2D spectra, thereby completing the spectral assignments. The most valuable spectrum was the 3D ^{13}C edited NOESY-HMQC. It proved possible to assign all of the protons in the molecule with the exception of C43 H5′ and H5″. The greatest difficulty in assigning RNA when ^{13}C-labeling is available is to distinguish H2′ and H3′ resonances within a given nucleotide. Although the C2′

Table 2: Carbon, Nitrogen, and Phosphorus Chemical Shifts for the Free and Bound 3′-UTR RNA at 300 K and 5 mM Sodium-Phosphate Buffer, pH 5.5[a]

nt	1′	2′	3′	4′	5′/5″	6/8	2/5	i	P
G19	**91.6**	**74.3**	**73.3**	**82.7**	**65.7**	**138.9**			
G20	**92.5**	**74.5**	**71.9**	**82.0**	**63.6**	**136.4**		**150.1**	**−4.04**
C21	**93.2**	**74.5**	**72.0**	**81.2**	**63.6**	**140.3**	**96.8**		**−4.41**
A22	**92.7**	**75.0**	**72.3**	**81.3**	**65.0**	**139.0**	**152.0**		**−4.06**
G23	**92.0**	**75.1**	**72.5**	**81.8**	**64.6**	**135.7**		**150.2**	**−3.99**
	92.4	**75.0**	**71.7**	**82.3**	**65.0**	**135.7**		***149.0***	−3.95
A24	**89.4**	**75.5**	**75.2**	**83.8**	**65.6**	**140.5**	**154.8**		**−4.11**
	91.7	***76.6***	***71.8***	***81.9***	**66.0**	**139.8**	**154.6**		*−3.84*
G25	**92.5**	**74.3**	**73.2**	**82.3**	**65.8**	**138.3**		**149.3**	**−3.40**
	92.5	**74.4**	***74.3***	**82.3**		**138.8**		**149.3**	***−3.87***
U26	**93.1**	**74.6**	**71.5**	**81.4**	**64.1**	**142.0**	**102.0**	**164.9**	**−4.46**
C27	**93.4**	**74.9**	**71.7**	**81.0**	**64.0**	**141.4**	**97.1**		**−4.13**
C28	**93.6**	**74.8**	**71.7**	**81.2**	**63.8**	**140.5**	**97.5**		**−4.13**
U29	**93.6**	**75.3**	**72.2**	**82.0**	**64.2**	**141.5**	**103.3**	**161.6**	**−4.40**
U30	**88.9**	**73.6**	**77.0**	**86.3**	**66.0**	**144.3**	**105.0**	**160.0**	**−3.48**
C33	**88.6**	**76.6**	**79.6**	**83.7**	**66.4**	**142.3**	**98.1**		**−5.05**
G34	**94.1**	**76.3**	**75.2**	**82.3**	**68.2**	**142.5**		**144.8**	**−4.97**
G35	**92.4**	**74.5**	**73.8**	**82.0**	**68.8**	**138.5**		**149.1**	**−2.44**
G36	**92.3**	**74.4**	**71.7**	**81.2**	**63.7**	**135.9**		**149.1**	**−3.92**
A37	**92.3**	**74.6**	**71.9**	**81.4**	**63.9**	**139.0**	**153.9**		**−4.38**
C38	**93.0**	**74.8**	**71.8**	**81.1**	**64.2**	**139.6**	**97.0**		**−4.29**
	93.3	**74.6**	**72.0**	**81.4**	**64.0**	**140.0**	**97.2**		**−4.27**
A39	**93.0**	**74.9**	**72.2**	**81.7**	**64.3**	**139.0**	**153.9**		**−3.94**
	92.6	**75.0**	**71.7**	81.7		***140.7***	**153.2**		−3.87
U40	**91.2**	**74.9**	**73.6**	**82.8**	**64.5**	**140.9**	**103.4**	159.9	**−4.40**
	93.0	**75.4**	***70.7***	*81.7*	***65.6***	***139.8***	**102.8**	**165.2**	*−4.12*
U41	**90.2**	**74.7**	**76.0**	**83.8**	**66.7**	**143.0**	**104.6**	160.0	**−4.17**
	87.1	**75.4**	***77.2***	**83.6**		**142.1**	**104.8**	**161.6**	
G42	**90.0**	**74.3**	**74.6**	**84.0**	**66.7**	**139.8**		**147.8**	**−3.86**
	89.6	***75.5***	**74.3**			***144.8***		**146.0**	***−3.42***
C43	90.3		**73.7**	**83.8**		**141.1**	**98.1**		**−3.76**
	87.1	**77.3**	***74.7***	***80.8***		**141.1**	***99.5***		***−2.97***
A44	**90.0**	**75.4**	**74.5**	**83.7**	**65.8**	**140.4**	**154.7**		**−3.86**
	93.3	***74.7***	***71.2***	***81.3***		**140.1**	***152.2***		*−3.97*
C45	**90.6**	**75.1**	**75.3**	**82.2**	**65.5**	**140.3**	**98.2**		**−3.23**
	92.4	***76.5***	***71.6***	**81.9**		***142.3***	***96.5***		***−3.79***
C46	**93.2**	**74.7**	**71.2**	**81.1**	**63.4**	**142.4**	**97.9**		**−3.55**
	93.3	**74.7**	**71.4**	**81.5**	***64.4***	***140.5***	**98.4**		***−4.26***
U47	**93.0**	**74.6**	**72.1**	**81.8**	**63.5**	**140.4**	**104.5**	**164.2**	**−4.24**
G48	**91.7**	**74.5**	**71.8**	**81.4**	**63.9**	**135.0**		**149.3**	**−4.14**
C49	**93.3**	**74.9**	**74.5**	**81.1**	**63.5**	**140.3**	**96.5**		**−4.22**
C50	**93.2**	**76.8**	**68.9**	**82.7**	**63.7**	**140.5**	**97.5**		**−4.32**

[a] The chemical shifts for the free RNA molecule are uppermost in each box, and the bound RNA chemical shifts are reported below. Chemical shift values in lightface type indicate tentative assignments.

carbons resonate slightly downfield of the C3′ resonances in helical regions, the difference of only 2−3 ppm invites some caution. The pattern of NOEs for these two resonances in helical regions is similar, and it is often not possible to use through-bond connectivities to distinguish them, because the spectra are very crowded. Whenever possible, different NOEs to sugar protons were used to complete the assignments: the H2′ shows a stronger NOE to the H1′, and the H3′ shows a stronger NOE to the H4′ resonances. A 3D HCCH-COSY spectrum was valuable for obtaining assignments of the bound RNA molecule and for distinguishing H2′ and H3′ resonances in other RNA molecules studied in this laboratory.

The phosphorus resonances were assigned using a pair of complementary triple-resonance experiments, HCP and P,H-COSY-H,C-HMQC, in conjunction with two-dimensional Het-COR and Het-TOCSY experiments (Varani et al., 1995). The HCP experiment provides connections between ^{1}H, ^{13}C, and ^{31}P nuclei within the backbone using $^{13}C-^{31}P$ scalar couplings for magnetization transfer, whereas the P,H-COSY-H,C,-HMQC relies on $^{1}H-^{31}P$ couplings (Varani et al., 1995). Since all proton and carbon resonances had been assigned, it was possible to connect phosphorus resonances with known sets of carbon and proton resonances. The greatest difficulty was the poor dispersion of the spectra: many of the strongest cross peaks in the HCP spectra involve C4′ and H4′ resonances, which are clustered in a narrow spectral region, as are the H3′−C3′ resonances that generate the strongest cross peaks in P,H-COSY-H,C-HMQC spectra. In addition, the phosphorus resonances themselves are all in a very narrow range. These difficulties, together with the unfavorable relaxation properties of phosphorus nuclei, make analysis of these spectra very challenging, and a few assignments in the stems are tentative.

NMR Constraints. Three classes of experimental constraints were used to determine the RNA structure: (1) interproton distance constraints derived from NOESY spectra, (2) hydrogen bonding constraints inferred from the pattern of exchange rate and chemical shift of imino and amino resonances, and (3) dihedral angle constraints inferred from scalar coupling constants. The number and position of the constraints is presented in Table 3. Altogether, 477 intranucleotide distances, 260 sequential distances, and 83 long- and medium-range distances were derived from

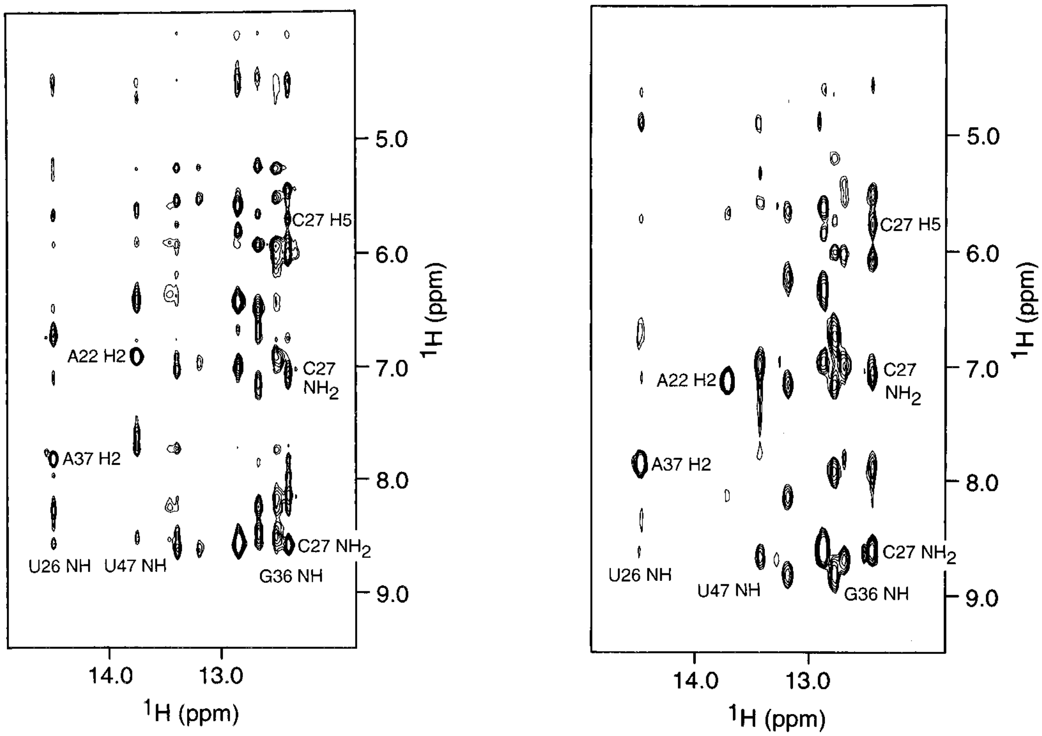

FIGURE 2: (A, left) Portion of a NOESY spectrum at 200 ms mixing time recorded in H_2O of the free RNA showing NOEs between imino protons (horizontal axis) and NH_2, H2, H6, H5, and H1′ protons. (B, right) The same spectral region for the RNA−protein complex (at 100 ms mixing time). Notice the remarkable high quality of the spectrum for this 22 kDa complex.

NOESY spectra. Twenty-five hydrogen bonding constraints and 126 dihedral angle constraints were also included. A total of 971 NMR derived constraints were obtained, corresponding to ≈5 constraints for each conformational degree of freedom. A comparable number of constraints (725) has also been obtained for the bound RNA, indicating that this very high density of constraints is obtainable at a molecular mass of 22 kDa. Some of these constraints, particularly NOE constraints between protons of a single sugar, are informationally redundant: the sugar ring is already constrained by the covalent structure. However, it is important to analyze all of the cross peaks in the spectrum: interresidual sugar−sugar contacts are often visible, and it is necessary to sift through all of the cross peaks in order to unambiguously identify conformationally important NOE interactions. Moreover, it is not clear whether the conformationally redundant NOE constraints are truly redundant in the early stages of structure determination, when the random starting structures bear little resemblance to RNA.

Isotopic labeling was critical to obtaining this large number of constraints. In addition to its value in assigning the RNA described in the previous section, 409 of the 820 constraints derived from NOESY spectra were obtained from the ^{13}C-edited spectrum (Figure 4). Many dihedral angles required estimates of carbon−phosphorus couplings and proton−phosphorus couplings, which would have been impossible to obtain without carbon editing.

Structure Determination. The experimental constraints obtained from the NMR data were used in restrained molecular dynamics simulations carried out with the program X-PLOR 3.1 as described in Experimental Procedures and schematically represented in Figure 5. Of the 50 starting structures with randomized backbone torsion angles, 18 had potential energies so high that the simulated annealing protocol was unable to find convergence. These structures literally blew apart in the first few steps of simulated annealing when covalent bonds were broken. The remaining 32 structures were put through the full calculation protocol. At the end of the minimization, 21 were selected on the basis of total energy of the final structures. Among the 21 low energy structures, the average number of NOE violations greater than 0.1 Å was 2.7. Among the high energy structures, the average number of NOE violations greater than 0.1 Å was 19. No NOE violations were greater than 0.3 Å for any of the converged structures. One can select structures on the basis of two other criteria: the NOE energy of the final structures and the number of violations in the final structures. If one chooses 21 structures on the basis of either criterion, the set of final structures is almost identical, with the substitution of only one structure.

In summary, 40% of the random structures converge. This number is smaller than that for protein molecules of comparable molecular weight but seems to be characteristic of RNA structures analyzed in this laboratory (Aboul-ela et al., 1995; Allain & Varani, 1995). A convergence rate of 67% among the structures which finish the first stage of the dynamics is comparable between proteins and nucleic acids using X-PLOR-based molecular dynamics protocols and randomized starting structures. What seems different is the large number of RNA structures which do not complete the first stage of the molecular dynamics. It is not clear why the structures that crash do so—structures which crash with one set of constraints may very well finish the calculation and even converge given a different constraint set.

Agreement with the NMR Data. After completion of a first set of calculations, a list of all proton pairs close to each other in the RNA structures (i.e., <5 Å apart) but without a corresponding entry in the constraint list was generated. The NOESY spectra were then reanalyzed to

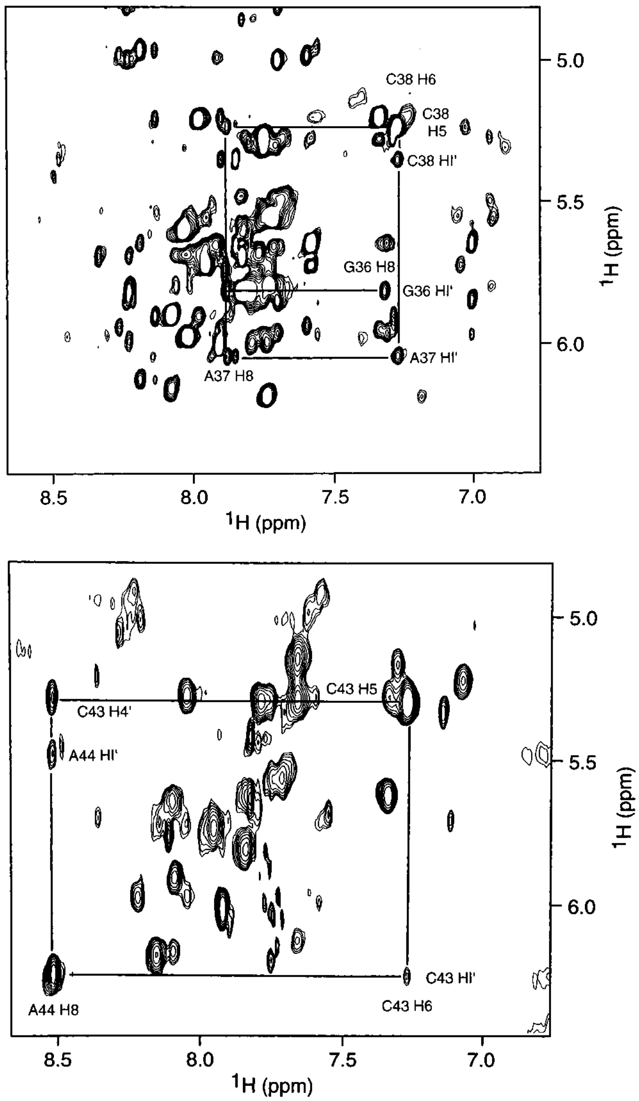

FIGURE 3: (A, top) Portion of a NOESY spectrum of the free 3′-UTR RNA at 200 ms mixing time recorded in D_2O, including NOEs between H8, H6, and H2 protons (horizontal axis) and H5 and H1′ protons. (B, bottom) The same region of the NOESY spectrum at 100 ms mixing time for the RNA−protein complex.

Table 3: Statistics of Experimental Constraints used in the Structure Determination

hydrogen bonds	
two for A-U pairs and three for G-C pairs	25
dihedral angle restraints	
$\alpha, \beta, \gamma, \delta, \epsilon$, and ζ	126
NOE distance constraints	
from data in D_2O	708
from data in H_2O	112
total number of constraints	971
constraints per residue	32
NOE constraints by range in sequence	
intranucleotide	477
sequential	260
long/medium range	83
NOE constraints by intensity	
strong (0−3 Å)	50
medium (0−4 Å)	182
weak (0−5 Å)	315
very weak (0−6 Å)	234
extremely weak (0−7 Å)	87

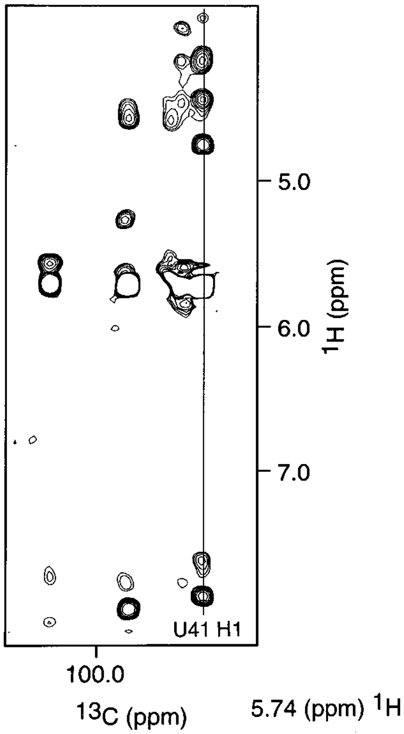

FIGURE 4: Plane from a 3D ^{13}C-edited NOESY spectrum at 200 ms mixing time of the free 3′-UTR RNA shows the excellent spectral editing obtainable with ^{13}C labeling.

examine whether the absence of an observed constraint was due to spectral overlap or an inconsistent constraint list. Reexamination of the NOESY spectra led to the observation of 70 additional distance constraints that could be identified unambiguously at this stage of the structure determination protocol. When the final structures were calculated, the list of close protons was again generated. In the tetraloop and the stems, the great majority of proton pairs closer than 5 Å but without a corresponding constraint in the restraint list refer to resonances overlapped even in the 3D spectra or to broadened resonances (for example, amino resonances from G and A residues). In the stems and the tetraloop, the proton pairs which are close but not constrained are largely the same between different structures. In contrast, the proton pairs which were close yet unconstrained in the internal loop were highly variable between structures. The differences in the structures of the internal loop are sufficiently large to generate many different sets of predicted NOE cross peaks. We have not attempted to use non-NOE constraints to improve the precision of the internal loop structure. The presence of conformational flexibility within this region (see below) would make this approach highly misleading by providing a more precise but less accurate picture of the structure.

Overall, the average number of pairs of protons which are within 5 Å of one another, and for which we have no NOE derived constraint, is 526. Since the total number of NOE constraints is 820, the data set is 61% complete (the ratio of assigned NOE cross peaks to total number of close proton pairs, observed and unobserved). Breaking this down by different portions of the structure, the data set is 58% complete for the stems, 83% complete for the tetraloop, and 52% complete for the internal loop. The total of 126 dihedral angle constraints, for a total of 176 possible backbone dihedral angles in the molecule, represents a 72% complete set of dihedral angle constraints.

Precision of the Structure. This extended RNA molecule does not have a well defined global structure, but the local

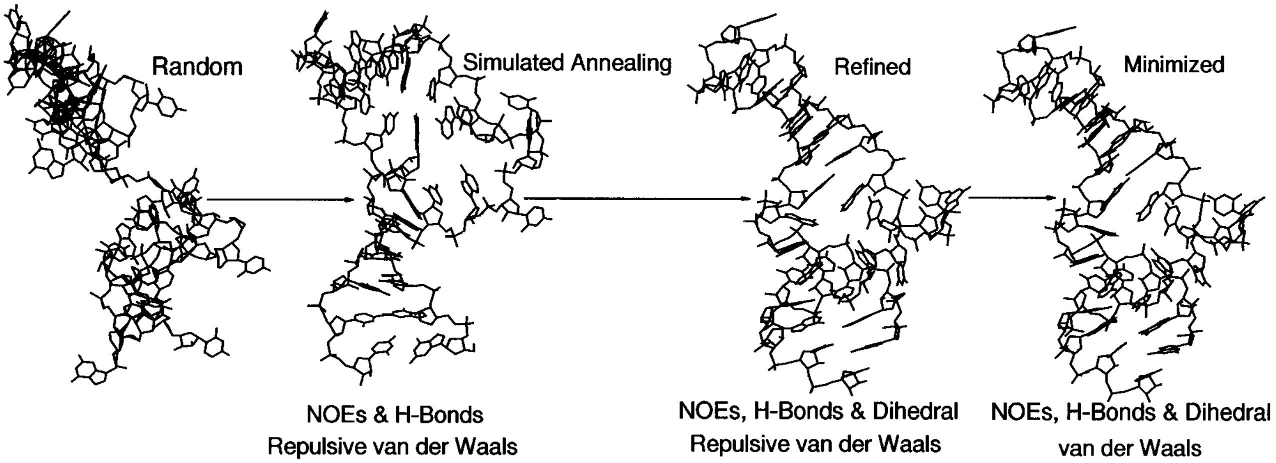

FIGURE 5: Illustration of the calculation protocol used in determining the structure of the RNA molecule. Initial randomized structures are subjected to (1) simulated annealing, (2) refinement, and (3) minimization.

Table 4: Local RMSD's from the Average Structure for Heavy Atoms only for the Final Energy Minimized Structures Calculated As Described in Experimental Procedures

stem 2	(G19−G23,C46-C50)	**0.78** Å
stem-loop 3	(G25−C38)	**0.75** Å
stem 3	(G25−C28,G35−C38)	**0.67** Å
tetraloop	(U29−G34)	**0.52** Å
internal loop	(A39−C45)	**3.44** Å
AUU	(A39−U41)	**2.38** Å

structure is very well defined, as detailed in the superposition statistics presented in Table 4. The overall rmsd for both stems, 3.5 Å, gives an idea of the global structural definition: the orientation of the helices is not random, yet not well defined in the absence of any long range constraint that would precisely define the relative position of the two helical regions. Most NMR observables are within a single nucleotide, or from one nucleotide to the next, and there are no NMR observables which directly link the ends of each helix. The high precision of local elements of structure reflects the wealth of data available from spectra of isotopically labeled RNA. As was expected, the tetraloop structure is particularly well defined due to its unusual compactness and to the favorable spectral dispersion. The two stems are similarly well defined: stem 3 in conjunction with the UUCG tetraloop resembles the P1 helix structure previously determined in this laboratory (Allain & Varani, 1995). The stems and the tetraloop display small rmsd's from the local average structure, as is expected for such compact elements of structure. The all atom rmsd's for these regions are comparable to those seen in NMR-derived protein structures. An overall view of the structure and superpositions of individual elements of the RNA structure are shown in Figures 6 and 7.

The rmsd from the average for the internal loop (3.4 Å, Table 4) reflects the wide range of structures which are consistent with the data. The rmsd is relatively large even for the first three nucleotides, A39−U41, which display as many NOEs as seen within each strand of double-helical RNA. The high density of NOE constraints makes it unlikely that the poor definition of the loop structure is the consequence of lack of experimental information.

In general, it is difficult to distinguish between the lack of experimental information and genuine conformational flexibility. However, the analysis of the scalar coupling patterns clearly indicate that significant conformational averaging occurs in the internal loop region. The conformation of the sugar ring of A39 is predominantly C3′-*endo*, but, for U40, U41, and G42, cross peaks in homonuclear and heteronuclear spectra indicate couplings of 3−6 Hz between both H1′−H2′ and H3′−H4′ resonances as well as between H2′−H3′ protons. The H2′−H3′ coupling is 4−6 Hz regardless of sugar conformation and therefore informationally redundant, but the 4−6 Hz coupling for both H1′−H2′ and H3′−H4′ indicates either an unusual (O4′-*endo*) conformation or the presence of an approximately equimolar population of C2′-*endo* and C3′-*endo* conformers (Altona, 1982). The analysis of the pattern of scalar couplings in the backbone strongly indicates that this second interpretation is more likely to be correct. $^{1}H-^{31}P$ and $^{13}C-^{31}P$ couplings can be used to analyze the conformation of the backbone torsion angles β and ϵ. Three couplings, P−H5′, P−H5″, and P−C4′ (all three are intranucleotide connections), give direct information regarding the β torsion angle (Altona, 1982). Examination of the Karplus equations (Mooren et al., 1994) reveals that no single region of conformational space is consistent with a nonvanishing value for all three angles. Thus, observation of resolved couplings (4−6 Hz for a molecule of this size) in $^{1}H-^{31}P$ correlated experiments indicates that conformational averaging is occurring. This situation is present for both U41 and G42, but not for U40, indicating the preservation of a relatively well defined conformation 5′ to U40. The ϵ dihedral angle can be treated similarly. In this case, the three connections which are directly related to the torsion angle are C4′−P, C3′−P and H3′−P (these are internucleotide connectivities) (Altona, 1982). Both ϵ for U41 and G42 are averaged on a fast (sub microsecond) time scale.

Precision of the Dihedral Angles and Helical Parameters. The precision of the dihedral angles and helical parameters reflects the precision of each structural building block of this RNA ((Figure 8). In the stems and the tetraloop, the uncertainty of the measured angles and helical parameters is small, with the exception of the terminal bases in the stems, and χ for U30. This angle is poorly defined by only a few NOEs and confirms the limited role of this base in the packing of the tetraloop (Cheong et al., 1990; Woese et al., 1990). The precision of the measured parameters in the internal loop is poor, reflecting the flexibility of this region of the structure.

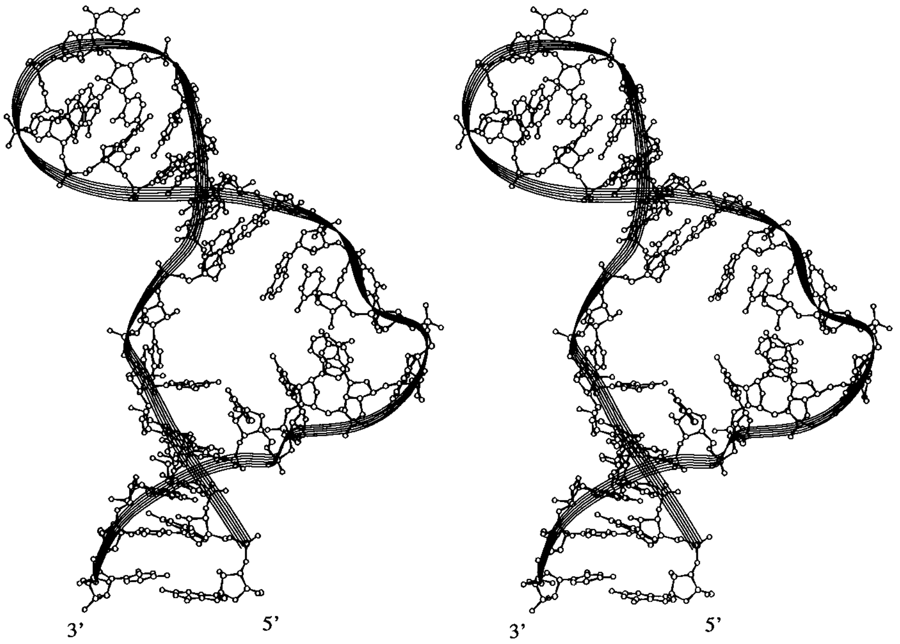

FIGURE 6: Stereoview of a low energy 3′-UTR RNA structure.

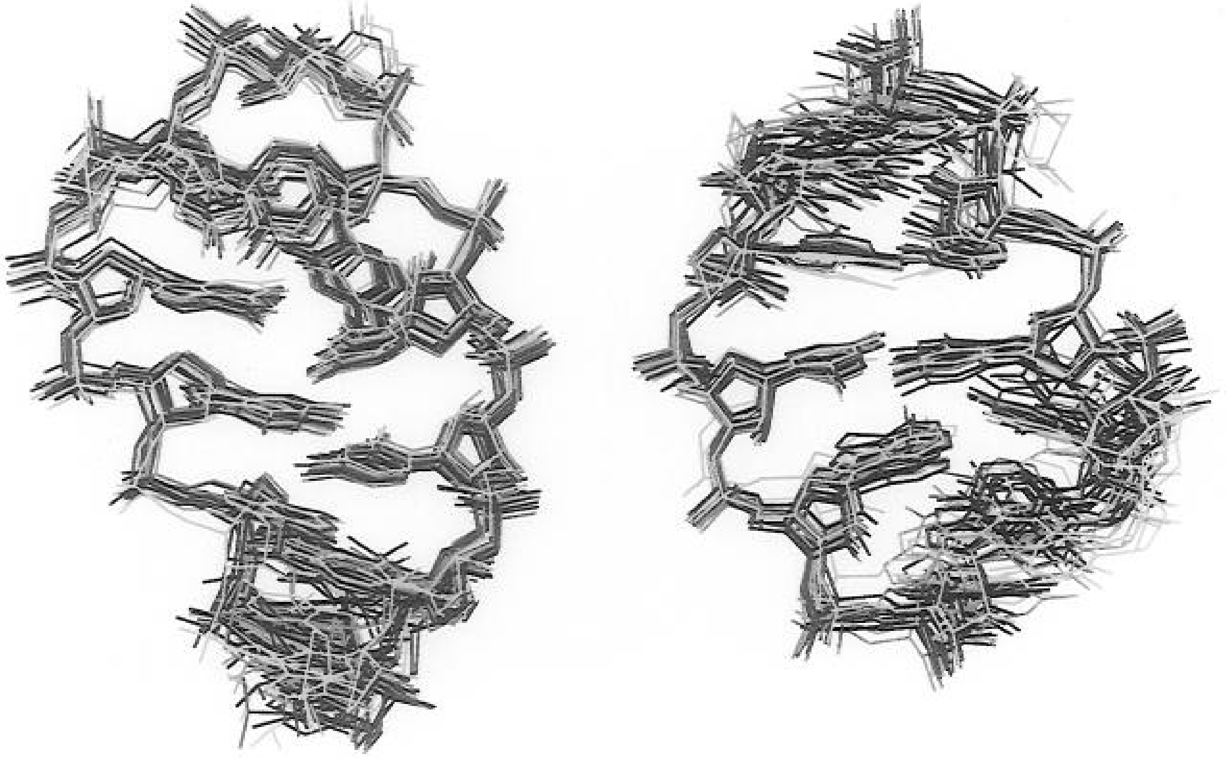

FIGURE 7: Superposition of nucleotides from stem 2 (a, left) and from stem 3 (b, right) show that the helical elements of structure are well defined. For each of the superpositions, all 21 converged structures are shown. Protons have been omitted for clarity.

Analysis of the Complex between the U1A N-Terminal RNP Domain and the 3′-UTR RNA. Only the RNA component of the complex will be described here. The first step in the assignment of the 3′-UTR RNA in complex with the U1A protein N-terminal RNP domain was based on the comparison of NOESY spectra for the free and bound RNA (Figure 2 panel A vs B Figure 3 panel A vs B). Spectra were acquired at different temperatures (7−27 °C) to facilitate the identification of overlapped resonances. The observation of the characteristically shifted G and U imino resonances in the tetraloop indicated that this part of the structure is unperturbed upon protein binding. Using available assignments for the free RNA, it was possible to link together all imino resonances from both stems, indicating the preservation of the secondary structure of the RNA upon protein binding. As with the free RNA, amino resonances were then assigned from the correlation to the corresponding imino resonances. Two-dimensional ^{1}H−^{15}N correlation spectra allowed identification of the ^{15}N chemical shift for the imino and some amino resonances. Identification of the AH2 and most H1′−H5 resonances from the spectra in water provided crucial starting points for the assignments of the

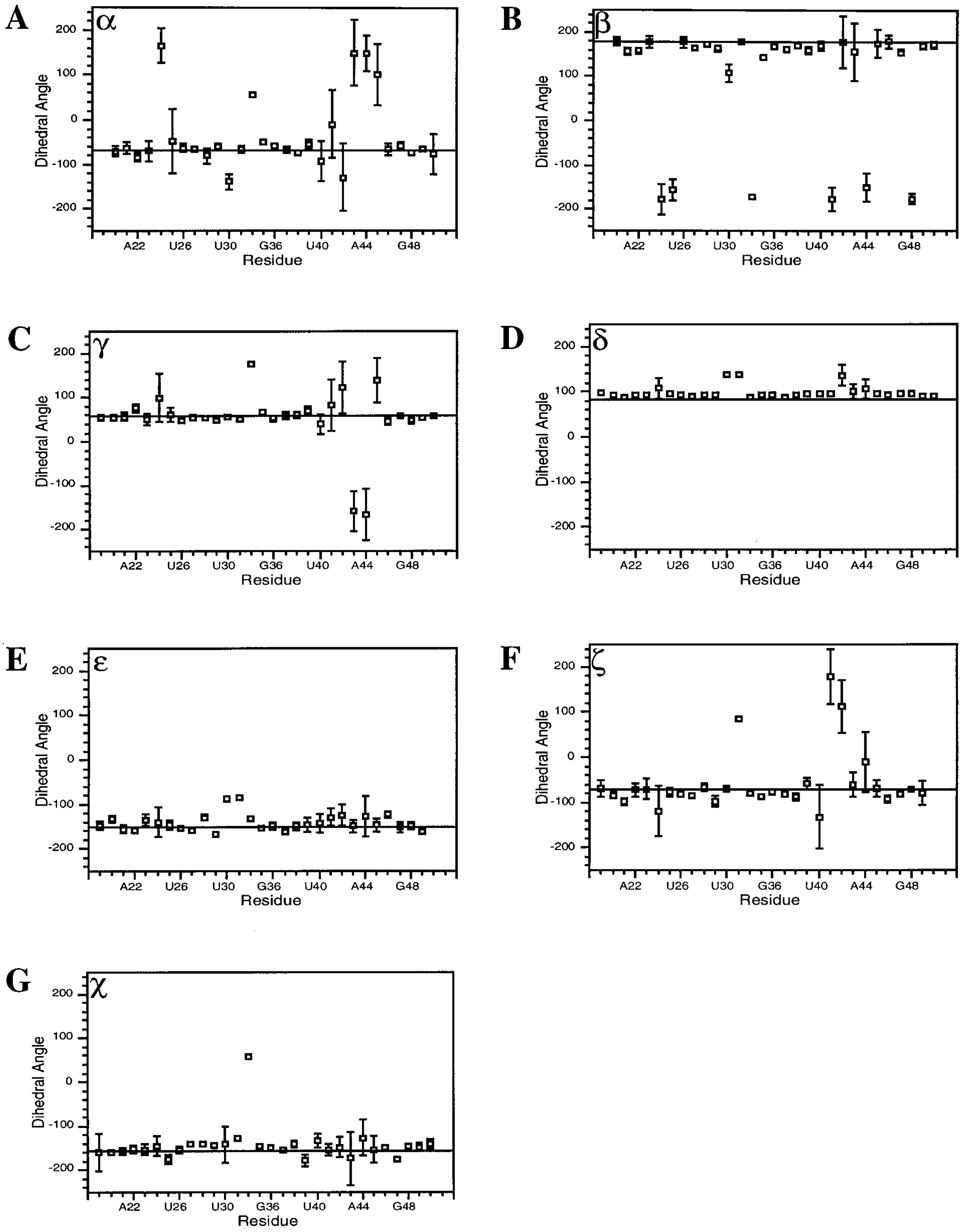

FIGURE 8: (A−G) Plots of the dihedral angles in the structures of the free RNA: the angles α, *β*, *γ*, *δ*, *ϵ*, *ζ*, and *χ* (in degrees) were calculated from the final converged structures and plotted with the standard deviation from the mean represented by the error bars. The canonical A-form values are represented by the unbroken horizontal lines.

helical stems in the complex by means of two-dimensional NOESY spectra acquired at different mixing times (40−120 ms) with unlabeled RNA and either ^{15}N- or ^{13}C−^{15}N−labeled protein. The comparison of spectra obtained for the ^{13}C-labeled and unlabeled protein and obtained without ^{13}C decoupling allowed the attribution of some ambiguous NOEs as intermolecular rather than intramolecular contacts. Obviously, resonances from ^{13}C-attached resonances are split by the characteristic one-bond coupling and significantly broadened due to increased relaxation. Since many of the most informative RNA resonances lie in regions where there are few protein resonances, it was possible to assign part of the RNA spectrum from 2D experiments alone using the available assignments for the free RNA as a guide. Resonance assignments were then completed using a set of 3D ^{13}C-edited experiments, including primarily ^{13}C-edited NOESY and HCCH-COSY spectra.

When the U1A protein is bound, spectral assignments and the RNA structure within the internal loop region and at the two flanking base pairs are substantially altered (Tables 1 and 2). On one side of the internal loop (G23−G25), spectral assignments could nevertheless be obtained straightforwardly due to the conservation of helical stacking between A22, G23, and A24 and between G25 and U26. The A24 H2 resonance was assigned from its NOE contact to G25 H1′, analogous to what observed in the free 3′-UTR RNA. In the seven-nucleotide stretch representing the main U1A protein binding site, assignments were very difficult to obtain due to the increased line width of most RNA resonances in ^{13}C-edited experiments. Furthermore, minor peaks from unbound RNA or small populations of misfolded RNA molecules appear with relatively high intensity given the much sharper line width of these molecular species. Nevertheless, A39 and U40 could be straightforwardly assigned due to the conservation of regular helical stacking as observed in the free RNA molecule. A44 and C45 could also be identified as a purine residue stacked on a pyrimidine residue: the unambiguous identification of the A44 H2 allowed us to distinguish the A44−C45 step from the G42−C43 step. C46 could also be assigned from its regular A-form connectivities to its base paired pattern G23 and to the following nucleotide, U47. Identification of the U41 H5−H6 resonances was possible, since these represented the only remaining unassigned uracil resonances, and U C5 resonates at characteristic chemical shifts. A connection between U41 H1′ and G42 H8 allowed the assignment of G42, despite the very unusual shifts of its H2′ and H3′ resonances (5.26 and 5.14 ppm, respectively). The C43 H1′ resonance was identified as the last remaining unassigned H1′ resonance from its very strong NOE connectivities to protein side chain resonances, leading to the identification of the C43 H5 and H6 resonances and most sugar resonances.

We were suspicious of the possible misidentification of resonances from free RNA as bound RNA resonances, particularly in some of the correlated experiments where the much smaller line width of the free RNA presents an obvious advantage. However, this is unlikely for three reasons: (1) the residual concentration of free RNA in the samples of the complex is less than 5−10% of the total RNA concentration, as the samples were prepared with only a slight RNA excess to enhance solubility, (2) the line width of the unperturbed resonances is clearly greater than the line width of resonances in spectra of the free RNA, and (3) the resonances which are unchanged in the complex are linked by NOE to resonances which are displaced from the free RNA chemical shifts, and which show intermolecular NOEs to the protein. Observation of NOEs to the A102 protein resonances was a critical defining criterion to distinguish major conformers from the RNA−protein complex from minor species from free RNA. For example, U40, U41, and G42 NH resonate at very unusual chemical shifts, but their identification was straightforward from the intermolecular NOE contacts to assigned protein resonances.

The assignments for the RNA resonances substantially shifted in the presence of the protein are shown in Tables 1 and 2. Assignments are nearly complete with the exception of most C5′-H5′-H5″ and exocyclic NH_2 resonances within the protein binding site. Failure of triple-resonance correlation experiments at the large molecular weight of this complex and the general weakness of NOE interactions involving these resonances prevented their reliable assignment at this stage.

DISCUSSION

The spectral assignment of the RNA component in a complex of 22 kDa demonstrates that RNA−protein complexes as well as RNAs of such substantial size can be studied by NMR. Clearly, using labeled RNA and unlabeled protein is a great advantage at such high molecular weight. The preparation of RNA samples that are selectively labeled in certain parts of the molecule will produce spectra of equivalent spectral dispersion and line width to those presented in this work. Spectral assignments and high resolution structure determination should therefore be possible for RNAs of up to at least 25 kDa, or approximately 80 nucleotides.

Protein-Induced RNA Folding. In this report, we have described the solution structure of a conserved element from the 3′-UTR of the human U1A pre-mRNA (van Gelder et al., 1993). The structure contains two helical stems in nearly ideal A-form geometry capped by a UUCG tetraloop and interrupted by an internal loop. The internal loop, recognized with high affinity and specificity by the U1A protein itself (van Gelder et al., 1993), contains well defined local elements of structure, but it is generally flexible in the absence of the protein ligand. Spectral assignments have been obtained both for the free and bound form of the RNA, and the characterization of the conformation and dynamic behavior of this RNA element in both its free and bound forms allows a detailed description of the conformational changes and folding transitions induced by U1A binding.

The overall structure of the seven nucleotides of the internal loop is flexible in the absence of the U1A protein, and this is reflected in a large local rms deviation between structures (≈3.4 Å, Table 4). However, well defined local structural features are clearly inferred from the NMR data and can be observed in the structures themselves. The stacking pattern for the seven nucleotides in the single-stranded region and the bases which flank it is the following: C38:A39:U40:U41, G42, C43, A44:C45:C46 where a colon denotes a stacking interaction. Thus, A39, U40, and U41 form a continuation of the upper stem, and A44 and C45 form a continuation of the lower stem. This can be seen in the overall view of the structure presented in Figure 6. The greater part of the conformational distortion of the loop is in the flexible "hinge" nucleotides G42 and C43. However, the analysis of the scalar coupling patterns indicates that the conformational deformation is not a static distortion localized at G42 and C43 but rather a continuous dynamic deformation throughout most of the single-stranded region. On the other side of the loop, A24 stacks on G23, but a clear break in the stacking interactions occurs between A24 and G25.

The stacking pattern in the bound 3′-UTR RNA is different from that observed in the free RNA. In the complex, the stacking pattern in the internal loop region is C38:A39:U40, U41:G42, Tyr13:C43,Phe56:A44:C45,C46 where a colon denotes a stacking interaction and intermolecular stacking with three key protein side chains is also included. Thus, in both the free and bound RNA the 5′-end of the internal loop forms a continuation of the helical stem, but the interaction with the protein completely distorts the structure within the remainder of the single-stranded region. On the other side of the loop, the stacking pattern remains unchanged despite the presence of many contacts with protein side chains. These differences indicate that the RNA undergoes a significant structural rearrangement upon protein binding to a conformation that is similar (at least at this qualitative level of analysis) to that seen for the equivalent nucleotides in the crystal structure of the U1A stem−loop II complex (Oubridge et al., 1994). Although the unbound RNA is clearly flexible in solution, calculations of the structure of the protein−RNA complex indicate that the rms deviation within the same seven nucleotides is reduced from 3.4 Å in the free RNA to 1 Å in the bound RNA (Allain et al., manuscript in preparation). Thus, protein binding is accompanied not only by a local conformational rearrangement but also by a significant reduction in the flexibility of the RNA structure.

The well defined surface of the four-stranded β-sheet of the RNP domain provides a structural scaffold against which the RNA element folds into a more static conformation. This situation can be compared to what has been observed with polypeptides from the HIV-1 transcriptional enhancer Tat (Aboul-ela et al., 1995; Puglisi et al., 1993). In the Tat-TAR interaction, protein binding induces a conformational change and a stabilization of the RNA structure as observed for the U1A−3′-UTR interaction. At the same time, however, the polypeptide also folds into a more ordered conformation upon binding its RNA target (Aboul-ela et al., 1995). It will be of great interest to investigate in detail the dynamic properties of these folding rearrangements and study by thermodynamic methods how these processes contribute to the specificity and affinity of the RNA−protein interaction (Spolar & Record, 1994).

The 3′-UTR RNA Structure. As evident from the general appearance of the helices, the values of the backbone dihedral angles, and the helical parameters, both stems are essentially ideal A-form helices. The average and standard errors for the backbone dihedral angles are depicted in Figure 8, along with ideal A-form parameters. The standard A-form values for the dihedral angles are generally within the error of the measured dihedral angles for nucleotides in the stems. The dihedral angles for other parts of the structure deviate substantially from A-form values, and the precision also varies with the elements of structure. In the stems and the tetraloop, the standard deviation of the mean value is small, but in the internal loop the error is substantially larger. Among helical parameters, the displacement, which is a clear discriminant of A- and B-form helices, lies between −0.8 and −4.7 Å. The average value for A-form RNA is −3.8 Å, whereas the average displacement for B-form helices is 0.3 Å. This may indicate that the helices are intermediate between A-form and B-form. However, base pairs near the termini of the helices have displacements nearer to the B-form values, but base pairs near the center of the helices have displacements nearer to the A-form values. Since the structures are more poorly defined near the termini of the helices, it may be appropriate to disregard the values derived from these base pairs.

The structure of the UUCG tetraloop is virtually indistinguishable from the structure recently redetermined (Allain & Varani, 1995). Its most notable feature is the unusual U·G pair, in which the O6 of G34 makes a hydrogen bond to the 2′-OH proton of U29. The similarity of the tetraloop structure between this work and the previous study and the similar precision indicate that RNA structures can be determined to equally high levels of precision up to at least 10 kDa, and we are confident that this limit can be significantly increased.

Several NMR observables support for the existence of conformational flexibility within the internal loop. The relative paucity of NOEs and breadth of some ^{1}H and ^{13}C resonances in some portions of the loop, particularly G42 and C43, are indicative of conformational exchange. The C43 resonances are significantly broadened indicating the occurrence of conformational exchange (perhaps stacking/destacking motions) on the millisecond time scale. Since the analysis of scalar coupling interactions indicated the presence of conformational averaging, NOE cross peaks in the internal loop were interpreted more conservatively than in the rest of the structure (in practice, strong constraints were never used). We believe that this procedure reflects the presence of conformational flexibility by resulting in a poorly defined structure. Had tighter NOE constraints been used, mutually inconsistent pairs of NOEs would have probably been found, and the resulting structures would have been more precise (due to the minimization of NOE violations) but less accurate. We emphasize that only the analysis of the scalar coupling interactions allowed us to conclude that the relatively low precision of the structure is not the result of the lack of experimental information, but rather of genuine conformational flexibility.

The steric requirements of such an asymmetric loop probably require bending of the RNA structure. The overall angle between the two helices in the structure cannot be defined precisely by NMR, as a consequence of the short range character of NMR data which provide no constraints to tie the ends of the helices together. Nevertheless, the outcome of the calculations suggest that the structure has a considerable degree of bending, a result that is supported by band shift analysis (A. Murchie and D. Lilley, personal communication). The RNA contains a considerable degree of bending in its bound form as well (F. Allain et al., manuscript in preparation). Protein-induced RNA bending may have a role in bringing into close proximity regulatory elements that are distant in primary sequence.

Identification of Intermolecular Contacts. Analysis of the pattern of chemical shift changes upon protein binding provides a fine footprint of the protein on the RNA structure. Most resonances from the double-helical regions and the tetraloop are unperturbed upon protein binding, with the exception of the U47•A22 base pair and the two base pairs closing the internal loop, G23•C46 and G25•C38. In addition, the G23•C46 base pair is clearly more stable in the presence of the protein. Thus, the A102 protein interacts with the internal loop region, the two loop closing base pairs and also with the two base pairs below the loop in Figure 1b. The interaction with the two base pairs below the loop comprise both direct contacts between protein side chains and the G23•C46 base pair and with A24 as well as (presumably) electrostatic contacts between side chains of basic residues and the phosphodiester backbone. Similarly, the interaction with the G25•C38 base pair comprises both direct base−protein side chain contacts and electrostatic

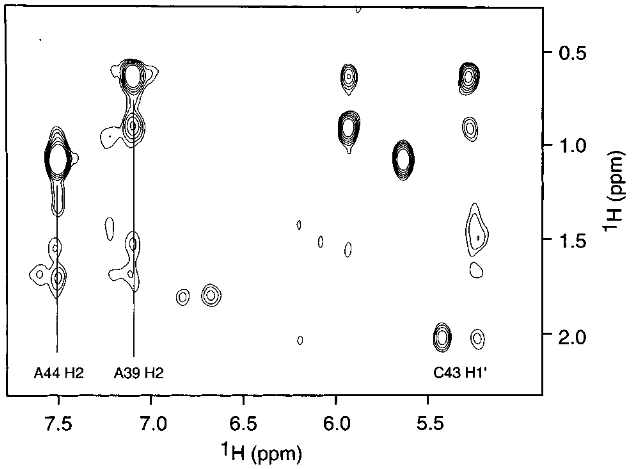

FIGURE 9: Portion of the ω_1-half-filtered NOESY (Otting & Wüthrich, 1990) spectrum at 100 ms mixing time of the RNA–protein complex. In this experiment, only intermolecular NOEs are observed since only the protein component was ^{13}C labeled.

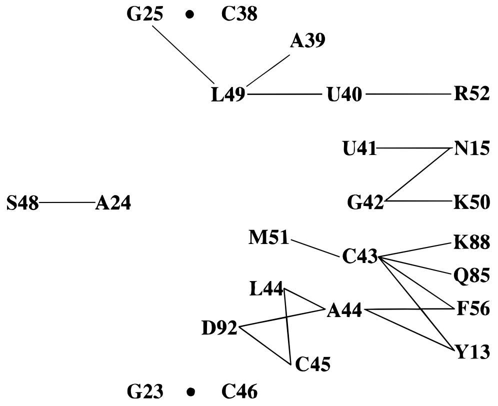

FIGURE 10: Schematic representation of the intermolecular contacts observed in the RNA–protein complex. Each connection represents one or more intermolecular NOE contact: a total of ≈100 such NOEs have been conclusively assigned.

interactions with the RNA backbone. Disruption of the base-paired stems results in severely decreased protein binding to almost nonspecific levels and loss of function in inhibition of polyadenylation (van Gelder et al., 1993). This set of interactions highlights the importance of the RNA structural context in protein binding.

In addition to large and widespread chemical shift changes, many protons in the internal loop show intermolecular NOE contacts to protein resonances (Figures 9 and 10). For example, G25 and U40 both show extensive NOEs to leucine 49, and U40 is also connected to arginine 52. A44 and C45 show NOEs to leucine 44, lysine 88, and methionine 51. C43 shows NOEs to methionine 51, alanine 87, and phenylalanine 56. A24 interacts with serine 48 and arginine 47. Altogether, we have identified and assigned ≈100 intermolecular NOEs in the complex. The pattern of intermolecular NOEs and chemical shift changes is consistent with the genetic and biochemical evidence (van Gelder et al., 1993) and demonstrates that the RNA is recognized by amino acids located on the surface of the β-sheet and in three variable loops connecting protein secondary structural elements (β1–helix A, β2–β3 and β4–helix C loops). The structure of the complex and the comparison with the existing crystal structure of the related hairpin complex will reveal the structural basis for recognition of such diverse RNA substrates by the U1A protein. The work presented in this report forms the basis for a formal determination of the structure of the full RNA–protein complex.

ACKNOWLEDGMENT

It is a pleasure to thank Dr. Iain Mattaj (EMBL, Heidelberg), Dr. Brian Wimberly (Scripps Research Institute), Dr. Fareed Aboul-ela, and Mr. Frederic Allain for many helpful suggestions. Mr. Peter Howe prepared the U1A protein N-terminal RNP domain that was used in the study of the complex. Mr. Chris Oubridge suggested the use of the double mutant protein. We are also grateful to Dr. Ian Mattaj, Dr. Alistair Murchie, and Prof. David Lilley (University of Dundee) for communicating unpublished results from their laboratories. Dr. Marla Babcock (Rutgers University) kindly provided a copy of the program RNA. C.C.G. was supported by predoctoral fellowships from the Churchill Foundation of the United States and the Howard Hughes Medical Institute.

REFERENCES

Aboul-ela, F., Karn, J. and Varani, G. (1995) *J. Mol. Biol. 253*, 313–332.

Allain, F. H.-T., & Varani, G. (1995a) *Nucleic Acids Res. 23*, 341–350.

Allain, F. H.-T., & Varani, G. (1995b) *J. Mol. Biol. 250*, 333–353.

Altona, C. (1982) *Rec. Trav. Chem. Pays-Bas 101*, 413–433.

Archer, S. J., Baldisseri, D. M., & Torchia, D. A. (1992) *J. Magn. Reson. 97*, 602–606.

Avis, J., Allain, F. H. T. A., Howe, P. W. A. H., Varani, G., Neuhaus, D., & Nagai, K. (1996) *J. Mol. Biol.* (in press).

Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. D., & Williamson, J. R. (1992) *Nucleic Acids Res. 20*, 4515–4523.

Boelens, W. C., Jansen, E. J. R., van Venrooij, W. J., Stripecke, R., Mattaj, I. W., & Gunderson, S. I. (1993) *Cell 72*, 881–892.

Brünger, A. T. (1990) X-PLOR, Yale University, New Haven, CT.

Burd, C. G., & Dreyfuss, G. (1994) *Science 265*, 615–621.

Cheong, C., Varani, G., & Tinoco, I., Jr. (1990) *Nature 346*, 680–682.

Diamond, R. (1992) *Protein Sci. 1*, 1279–1287.

Flickinger, T. W., & Salz, H. K. (1994) *Genes Dev. 8*, 914–925.

Gemmecker, G., Olejniczak, E. T., & Fesik, S. W. (1992) *J. Magn. Reson. 96*, 199–204.

Gerchman, S. E., Graziano, V., & Ramakrishnan, V. (1994) *Protein Express. Purif. 5*, 242–251.

Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications*, Academic Press, New York.

Griesinger, C., & Eggenberger, U. (1992) *J. Magn. Reson. 97*, 426–434.

Gronenborn, A. M., Bax, A., Wingfield, P. T., & Clore, G. M. (1989) *FEBS Lett. 243*, 93–98.

Gunderson, S. I., Beyer, K., Martin, G., Keller, W., Boelens, W. C., & Mattaj, I. W. (1994) *Cell 76*, 531–541.

Hall, K. B. (1994) *Biochemistry 33*, 10076–10088.

Hall, K. B., & Stump, W. T. (1992) *Nucleic Acids Res. 20*, 4283–4290.

Hines, J. V., Landry, S. M., Varani, G., & Tinoco, I., Jr. (1994) *J. Am. Chem. Soc. 116*, 5823–5831.

Hoffman, D. W., Query, C. C., Golden, B. L., White, S. W., & Keene, J. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2495–2499.

Howe, P. W. A., Nagai, K., Neuhaus, D., & Varani, G. (1994) *EMBO J. 13*, 3873–3881.

Jahnke, W., Baur, M., Gemmecker, G., & Kessler, H. (1995) *J. Magn. Reson. B 106*, 86–88.

Jessen, T. H., Oubridge, C., Teo, C. H., Pritchard, C., & Nagai, K. (1991) *EMBO J. 10*, 3447–3456.

Kambach, C., & Mattaj, I. W. (1992) *J. Cell Biol. 118*, 11–21.

Kay, L. E., Ikura, M., & Bax, A. (1990) *J. Am. Chem. Soc. 112*, 888–889.

Kellogg, G. W. (1992) *J. Magn. Reson. 98*, 176–182.

Lutz, C. S., & Alwine, J. C. (1994) *Genes Dev. 8*, 576−586.

Majumdar, A., & Zuiderweg, E. R. P. (1993) *J. Magn. Reson. B 102*, 242−244.

Marino, J. P., Schwalbe, H., Anklin, C., Bermel, W., Crothers, D. M., & Griesinger, C. (1994) *J. Am. Chem. Soc. 116*, 6472−6473.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.

Mattaj, I. W. (1993) *Cell 73*, 837−840.

Messerle, B. A., Wider, G., Otting, G., Weber, C., & Wüthrich, K. (1989) *J. Magn. Reson. 85*, 608−613.

Milligan, J. F., Groebe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783−8789.

Mooren, M. M. W., Wijmenga, S. S., van der Marel, G. A., van Boom, J. H., & Hilbers, C. W. (1994) *Nucleic Acids Res. 22*, 2658−2666.

Nagai, K., Oubridge, C., Jessen, T. H., Li, J., & Evans, P. R. (1990) *Nature 348*, 515−520.

Nagai, K., Oubridge, C., Ito, N., Avis, J., & Evans, P. (1995) *Trends Biochem. Sci. 20*, 235−240.

Nickonowicz, E. P., Sirr, A., Legault, P., Jucker, F. M., Baer, L. M., & Pardi, A. (1992) *Nucleic Acids Res. 20*, 4507−4513.

Otting, G., & Wüthrich, K. (1990) *Q. Rev. Biophys. 23*, 39−96.

Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., & Nagai, K. (1994) *Nature 372*, 432−438.

Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.

Puglisi, J. D., Chen, L., Frankel, A. D., & Williamson, J. R. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 3680−3684.

Scherly, D., Boelens, W., Dathan, N. A., van Venrooij, W. J., & Mattaj, I. (1990) *Nature 345*, 502−506.

Scherly, D., Kambach, C., Boelens, W., van Venrooij, W. J., & Mattaj, I. W. (1991) *J. Mol. Biol. 219*, 577−584.

Schwalbe, H., Marino, J. P., King, G. C., Wechselberger, R., Bermel, W., & Griesinger, C. (1994) *J. Biomol. NMR 4*, 631−644.

Shaka, A. J., Barker, P., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547−552.

Sklénar, V., Miyashiro, H., Zon, G., & Bax, A. (1986) *FEBS Lett. 208*, 94−98.

Sklénar, V., Peterson, R. D., Rejante, M. R., & Feigon, J. (1994) *J. Biomol. NMR 4*, 117−122.

Spolar, R. S., & Record, M. T., Jr. (1994) *Science 263*, 777−784.

Stonehouse, J., Shaw, G. L., Keeler, J., & Laue, E. D. (1994) *J. Magn. Reson. A 107*, 178−184.

Stump, W. T., & Hall, K. B. (1995) *RNA 1*, 55−63.

Szewczak, A. A., Kellogg, G. W., & Moore, P. B. (1993) *FEBS Lett. 327*, 261−264.

Tsai, D. E., Harper, D. S., & Keene, J. D. (1991) *Nucleic Acids Res. 19*, 4931−4936.

van Gelder, C. W. G., Gunderson, S. I., Jansen, E. J. R., Boelens, W. C., Polycarpou-Schwartz, M., Mattaj, I. W., & van Venrooij, W. J. (1993) *EMBO J. 12*, 5191−5200.

Varani, G., Aboul-ela, F., Allain, F. H.-T., & Gubser, C. C. (1995) *J. Biomol. NMR 5*, 315−320.

Woese, C. R., Winker, S., & Guttell, R. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8467−8471.

Zuiderweg, E. R. P., McIntosh, L. P., Dahlquist, F. W., & Fesik, S. W. (1990) *J. Magn. Reson. 86*, 210−216.

BI952319F