Crystal structure of an RNA helix recognized by a zinc-finger protein: An 18-bp duplex at 1.6 Å resolution

SUSANA LIMA,^{1,*} JAYNE HILDENBRAND,^{1,*} ANDREI KOROSTELEV,¹
STANLEY HATTMAN,² and HONG LI¹

¹Department of Chemistry and Biochemistry, Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, USA

²Department of Biology, University of Rochester, Rochester, New York 14627-0211, USA

ABSTRACT

The crystal structure of the 19-mer RNA, 5'-GAAUGCCUGCGAGCAUCCC-3' has been determined from X-ray diffraction data to 1.6 Å resolution by the multiwavelength anomalous diffraction method from crystals containing a brominated uridine. In the crystal, this RNA forms an 18-mer self-complementary double helix with the 19th nucleotide flipped out of the helix. This helix contains most of the target stem recognized by the bacteriophage Mu Com protein (control of *mom*), which activates translation of an unusual DNA modification enzyme, Mom. The 19-mer duplex, which contains one A•C mismatch and one A•C/G•U tandem wobble pair, was shown to bind to the Com protein by native gel electrophoresis shift assay. Comparison of the geometries and base stacking properties between Watson–Crick base pairs and the mismatches in the crystal structure suggest that both hydrogen bonding and base stacking are important for stabilizing these mismatched base pairs, and that the unusual geometry adopted by the A•C mismatch may reveal a unique structural motif required for the function of Com.

Keywords: RNA mismatch; RNA–protein interactions; translational activation; zinc-finger protein

INTRODUCTION

RNA molecules are multifunctional biological molecules with diverse three-dimensional structures. RNA functions as messengers (mRNA), adaptors (tRNA), and in large ribonucleoprotein enzymes such as the ribosome (rRNA) and the spliceosome (snRNA). It is now known that the RNA component of the ribosome catalyzes peptide-bond formation during protein synthesis (Ban et al., 2000). A similar catalytic role may likely apply to spliceosomal RNA during splicing reactions (Staley & Guthrie, 1998; Valadkhan & Manley, 2001). In addition, noncoding RNA sequences have been discovered with a diverse range of functions that include RNA modification (snoRNA; Weinstein & Steitz, 1999), RNA degradation (exosome; Allmang et al., 2000), and protein translocation (Lutcke, 1995).

RNA structures are also important in gene regulation, as exemplified by the translational activation mechanism of the bacteriophage Mu *mom* gene (reviewed in Hattman, 1999). The bacteriophage protein, Com (con-

trol of *mom*), regulates the expression of the *mom* gene by a mechanism that requires a unique structural rearrangement of RNA. The *mom* gene codes for an unusual DNA modification enzyme, which renders the viral DNA resistant to a variety of bacterial DNA restriction endonucleases. The Com protein contains only 62 amino acids, and it was shown to be a zinc-binding protein (Witkowski et al., 1995). Thus it is the smallest in size among all known zinc-finger proteins. The structural zinc is coordinated by four cysteine residues in the pattern of Cys⁶-(x)₂-Cys⁹-(x)₁₆-Cys²⁶-(x)₂-Cys²⁹ (Cys₄), reminiscent of erythroid transcription factor GATA-1 zinc-finger protein (Omichinski et al., 1993). It has been suggested that the Com protein destabilizes an mRNA stem (translation initiation structure, or TIS) that sequesters both the Shine–Dalgarno (SD) sequence and the *mom* start codon by binding to an adjacent hairpin (Com-binding structure, or CBS) in an energy-independent fashion. The unique chemical–physical nature of how the Com protein interacts with its target RNA is important for understanding the *mom* translational regulation mechanism.

In an effort to cocrystallize the Com protein with its hairpin RNA-binding target, we obtained crystals of the alternative duplex structure of a 19-mer RNA oligo (in

Reprint requests to: Hong Li, Department of Chemistry and Biochemistry, Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, USA; e-mail: hongli@sb.fsu.edu.

*Both authors contributed equally to the work.

the absence of the Com protein). The 19-mer oligo forms the Com-binding hairpin under dilute concentrations, but forms a self-complementary duplex at the concentration (10 mg/mL) used for crystallization. The 19-mer RNA crystals diffracted to 1.6 Å resolution and the structure includes an A•C mismatch and a tandem G•U/A•C mismatch. This structure allows us to examine base pairing, base stacking, and hydration properties of these mismatches and to compare them with those of other RNA base pairs.

Because the duplex form of the 19-mer oligo shares the same stem region as that of its hairpin form (Fig. 1A,B), perhaps not surprisingly, we could consistently show that Com protein binds specifically to the duplex form of the 19-mer oligo by gel electrophoresis (Fig. 1C). This result underscores the importance of the observed 19-mer duplex structure to our understanding of Com protein interaction with its RNA target, although a more thorough understanding of this interaction awaits a cocrystal structure of the Com protein with its RNA target.

In general, RNA molecules fold in three dimensions into specific shapes through the assembly of secondary structures. The major form of RNA secondary structure is the double-stranded helix. Although base-pairing and base-stacking interactions dominate this form of secondary structure, non-Watson–Crick base pairs are frequently found in RNA duplexes. Thermodynamic stabilities of mismatches depend on the sequence of the pair and on the adjacent Watson–Crick base pairs (Wu et al., 1995). This is likely to be due to subtle differences in hydrogen bonding geometry and base-stacking

interactions. Structural analysis of these base pair geometries and their base-stacking environment at high resolution is of great importance in understanding RNA structures, as well as RNA interaction with proteins.

RESULTS AND DISCUSSION

The 19-mer RNA interacts with the Com protein

Figure 1 shows the secondary structures of the 19-mer oligo. The 19-mer RNA forms a hairpin under dilute condition that represents the minimal CBS as shown in Figure 1A. At the concentration required for crystallization, the 19-mer RNA forms a self-complementary duplex in which the 19th nucleotide is not base paired (Fig. 1B). The gel shift results of the 19-mer by the Com protein are shown in Figure 1C. Under the conditions used for protein–RNA binding and electrophoresis (see Materials and Methods), we have observed a repeatable shift of the duplex 19-mer by the Com protein. We also show that Com protein binds specifically to the 19-mer duplex by the following evidence. Com protein does not bind to a nonrelated 17-mer duplex (Fig. 1C, lanes 1 and 2); yeast tRNA (2 μM) was unable to compete for Com binding to the 19-mer (Fig. 1C, lane 5); an increasing amount of unlabeled 19-mer competes for Com protein binding to radioactively labeled 19-mer (Fig. 1C, lanes 6–10). These results suggest an important role for the stem region in the CBS for Com–RNA recognition.

FIGURE 1. Oligonucleotide designs used in crystallization studies. Bold nucleotides are those protected from chemical or nuclease cleavage in the presence of the Com protein (from Hattman, 1999). **A:** The 19-mer forms a hairpin at low concentrations. **B:** The 19-mer RNA forms a self-complementary duplex at high concentrations. This is the form observed in crystal. **C:** Native gel electrophoresis shift experiments show that the Com protein binds specifically to the duplex 19-mer RNA. Com did not generate a shift in mobility of the hairpin form of the oligo. This could be due either to a lower stability of the hairpin itself or to the binary com-hairpin comp in electrophoresis.

Overall Structure

Figure 2A shows an overall structure of the 19-mer RNA where ribbons follow the phosphate backbones. The 19-mer RNA crystallizes in the space group $P6_322$ with a single 19-mer strand in each asymmetric unit and is arranged in a duplex of type-A form parallel to the *c*-axis with its crystallographic symmetry mate (*x*, *x*-*y*, -*z*). The crystallographic twofold axis along the *a*-axis coincides with the molecular twofold axis that relates the two self-complementary strands. The overhanging base C19 is flipped out of the helix and is not visible in electron density maps (Fig. 2B). As a result, the duplex is an 18-mer A-form helix that stacks head to tail, with the duplex related by the crystallographic twofold screw symmetry axis parallel to the *c*-axis. This pseudo-continuous helix of 36 bp is the repeating unit in the crystal along the *c*-axis with an average rise per base pair of $95.12 \text{ \AA}/36 = 2.64 \text{ \AA}$, a value smaller than that of a standard A-form helix ($\sim 2.7 \text{ \AA}$). The major reduction in rise occurs at the base step, C17/C18 ($1.72 \text{ \AA}$), as calculated by the helix geometry analysis program CURVES (Lavery & Sklenar, 1989; Ravishanker et al., 1989). C17 forms a C•A base pair with A2 of the other strand and also has a large tilt angle from C18 (11.86°). The unusual geometry of the C17•A2 base pair is dissimilar to that for the other A•C base pair, A12•C7, in the same duplex, which exhibits normal rise and tilt values. Thus the geometry of C17•A2

may be the re-sult of end-to-end stacking interactions mediated by the next base pair, C18G1.

The geometry at the junction of the two helices also deviates from those in base stacking of a standard A-form helix. The G1 nucleotide of the stacking strand has a large negative slide ($\Delta y = -6.6 \text{ \AA}$), shifting away from the C18 nucleotide but towards the G1 nucleotide of the bottom duplex, such that it can best stack with the bottom G1 nucleotide (Fig. 3). As a result, the C18 nucleotide interacts with the ribose ring of G1 (of the stacking strand), which is perpendicular to the cytosine of C18, with the O4' atom of G1 ribose closest to the cytosine ring of C18. The pseudo-GpC step between the two stacked duplexes has a rise of $2.27 \text{ \AA}$ and a roll value of 16.6° . These parameters are different from those of a type-A duplex, indicating a discontinuity in helical form.

Crystallographic packing in the *a*-*b* plane primarily involves backbone-to-backbone interactions through the O2' of the ribose moiety. In two cases, the side-by-side packing also involves ribose from one duplex and the shallow grooves of the neighboring duplex. Two such interactions are illustrated in Figure 3. The 2'-OH group of the nucleotide from one duplex, one nearby water molecule, O2 of the pyrimidine base, and N4 of the purine base form a network of hydrogen bonds that hold the parallel duplexes together.

Hydration

A total of 38 water molecules were identified for each strand of the 19-mer crystal. The majority of these were found to be located in the major groove side except for the G•U wobble pair. In the high-resolution crystal structure of the COPY18 RNA duplex, an even pattern of average hydration on both sides of GC and AU base pairs was identified (Klosterman et al., 1999). Although no water molecules are associated with any positions in bases G1 through U4, bases G5 through C18 have different patterns of hydration. In every base from G5 to G9, and in C17, a single water molecule was found to be hydrogen bonded to some base atom (N7 in purine, O4 or N4 in pyrimidines), which is then connected to another water molecule that is hydrogen bonded to the O1P atom of the phosphodiester. The water molecules connected to the phosphodiester oxygen also hydrogen bond with the O1P of the preceding phosphate group. In the remaining bases, water molecules associated with the N7 or O/N4 position do not make a connection to the phosphate oxygen atoms. Water molecules associated only with backbone atoms were observed for G5, C6, and A12. In all these cases, the water molecules participate in crystallographic packing interactions that involve the ribose moiety of the three bases. Only four bases, G5, G11, U8, and A15, have minor groove water hydrogen bonded to their N3 or O2 position. The G•U wobble pair, G8•U11, shares a

FIGURE 2. The final model in ribbon diagram (A) and the experimental map after solvent modification at 1σ superimposed with the final model of the 19-mer RNA (B). Water molecules are depicted as red balls.

FIGURE 3. Crystallographic packing of the 19-mer RNA. Sixfold axes are indicated. Only the twofold axis along the *a*-axis is indicated. The atomic details at two major sites of crystallographic interactions are shown.

single water molecule in their minor groove site that mediates crystallographic packing interactions with C7 (Fig. 3).

Geometries of non-Watson–Crick base pairs

Although the Watson–Crick pairs between complementary bases are necessary to form the helical framework of a complex RNA, the non-Watson–Crick pairs are pivotal in RNA–RNA and RNA–protein recognition. We describe below the structural features observed in the one G•U and two A•C wobble pairs.

There are two A•C mismatches, A2•C17 and C7•A12. Two hydrogen bonds are observed for each A•C pair (Fig. 4). It can be inferred from the geometries of both pairs that the N1 of adenosine is protonated in order to donate its proton to the acceptor O2 of cytosine. The *pK_a* value for N1 is approximately 4 in free adenosine

(Saenger, 1984), but it can reach 5.5 in the context of oligonucleotides (Brown et al., 1990). Thus the protonation status of N1 is consistent with the fact that the 19-mer crystals were only obtained at pH 5.6. This form of A•C mismatch, A(+)•C pair, is isosteric with the G•U wobble pair, as indicated by the similar C1'–C1' and P5'–O3' distances between U8•G11 and C7•A12 base pairs (Fig. 4). Figure 4 also lists hydrogen bonding geometry values for the C•A and U•G pairs.

There is one G•U wobble pair in the 19-mer RNA crystal structure, U8•G11. G•U is the most common non-Watson–Crick base pair, accounting for about 50% of non-Watson–Crick base pairs and mismatches (Gautheret et al., 1995). As shown in Figure 4, two hydrogen bonds are observed for this wobble pair, as is expected for a G•U wobble pair (Leontis & Westhof, 1998). When G•U wobble occurs in the context of 5'-GU-3'/3'-UG-5', only one hydrogen bond is observed

FIGURE 4. Geometries of non-Watson-Crick base pairs.

(Chen et al., 2000), which may help to explain the lowest thermostability of 5'-GU-3'/3'-UG-5' among the three possible G•U tandem wobble pairs.

Base-stacking properties of non-Watson-Crick base pairs

To assess the degree of base stacking in each base step of the 19-mer RNA, we have computed buried interstrand base or intrastrand base areas of the 19-mer duplex. Buried surface area for a particular base is defined as the net decrease in solvent accessible area in the presence of another base when using a solvent molecule with a 1.4-Å probe radius. In Figure 5, the two numbers printed under each base denote the buried intrastrand base (closer to the edge) and the buried interstrand base (closer to the center) area respectively.

In RNA, a 5'-pyrimidine-purine-3' step normally involves interstrand purine stacking. There are two examples of such base stacking in the 19-mer duplex

structure, 5'-U4G5-3', and 5'-U8G9-3'. In both cases, the buried surface area by the interstrand purines is significantly larger than that of other buried interstrand bases (126.6 Å² and 136.5 Å²; Fig. 5). Cross-strand purine stacks frequently form ion-binding sites (Cate et al., 1996; Correll et al., 1997), although no ions were found at these sites in the 19-mer RNA structure. This is most likely due to the high concentration of LiSO₄ used in crystallizing the 19-mer RNA. The second example, the 5'-U8G9-3' step, which contains the G•U wobble pair, also maintains this pattern of cross-strand purine stacking. In fact, the cross-strand purines in both steps dominate the base-stacking interactions, because the buried surface areas between the cross-strand pyrimidines are negligible (9.9 Å² and -2.4 Å²).

The wobble pairs, G•U and the N1 protonated A(+•)C, are isosteric with each other. The 19-mer RNA duplex contains a G•U wobble followed by a A(+•)C wobble, forming a 5'-UC-3'/3'-GA-5' motif. A high-resolution structure of this arrangement has not been

FIGURE 5. Buried surface areas between interstrand (e.g., between A2 and C18; values printed towards the center of the helix) and intrastrand (e.g., between A2 and A3; values printed along the edge of the helix) bases and base pairs (e.g., between A2-C17 and G1-C18; values printed outside the helix).

reported previously. The tandem 5'-UG-3'/3'-GU-5' is a frequent motif found in natural RNA molecules (Gautheret et al., 1995), and is thermodynamically more stable than the reverse orientation 5'-GU-3'/3'-UG-5' (He et al., 1991; Wu et al., 1995). The stability is largely due to favorable cross-strand purine stacking (Deng & Sundaralingam, 2000) and a favorable number of hydrogen bonding interactions (Chen et al., 2000). The tandem 5'-CA-3'/3'-AC-5' motif has been crystallized

(Jang et al., 1998) and found to be isosteric with the symmetric tandem UG motif. In contrast, the 5'-UU-3'/3'-GG-5' motif does not contain cross-strand purine stacking (Deng & Sundaralingam, 2000). Despite the absence of favorable cross-strand purine stacking interactions, the 5'-UU-3'/3'-GG-5' tandem wobble pair occurs more frequently in natural RNAs than the symmetric tandem wobble 5'-GU-3'/3'-UG-5' (Gautheret et al., 1995). This order of statistical occurrence is in agreement with measured thermodynamic stabilities (Wu et al., 1995). As with the 5'-UU-3'/3'-GG-5', the 5'-UC-3'/3'-GA-5' in our 19-mer crystal structure does not include cross-strand base stacking. But, the overall base stacking interaction is similar to that of tandem AU base pairs, as judged by the buried surface areas (Fig. 5).

A G•U or G•T base pair tends to stack to the base 3' to G. This is also observed in the 5'-UC-3'/3'-GA-5' tandem wobble repeat in the 19-mer RNA duplex. G11 stacks to the A12 within the same strand rather than the 5' nucleotide, C10. As a result, the buried surface area between base pairs U8•G11 and G8•C10 (219.2 Å²) is smaller than that between U8•G11 and C7•A12 (237.2 Å²). Overall, the average buried surface area of the three steps surrounding the tandem 5'-UC-3'/3'-GA-5' wobble pairs (230.0 Å²) is slightly less than that for the entire duplex (231.3 Å²), suggesting an overall stabilized structure.

Implications for Com binding

The fact that Com binds to the synthetic 19-mer duplex as present in the crystal structure suggests an important role for the CBS stem in Com-RNA interactions. We show in this report that, despite the fact that wobble pair geometries can deviate from those of the standard A-form helix, the overall helix stability is maintained through hydrogen bonding and base stacking interactions between wobble pairs. The presence of an A•C mismatch affects the structure of the duplex in several ways that might be exploited in specific RNA recognition. The recognition of the 19-mer hairpin stem by Com protein could involve direct read-out of the bases including the A•C mismatch, or a change in the conformation of the backbone, or the recognition of solvent that is ordered in a sequence specific way.

The ability of an A•C mismatch to introduce structural flexibility and diversity of A-RNA helices has been reported for the structures of the signal recognition particle RNA helix 6 (Wild et al., 1999) and a self-complementing octamer (Jang et al., 1998). Both structures were obtained at pH 5.5. In these cases, the imino proton of adenosine can act as proton donor in forming hydrogen bonds, thus stabilizing the A•C mismatch. It was proposed that an A•C wobble pair presents an opportunity for a conformational switch based on environmental factors modifying the adenosine protonation (Jang et al., 1998).

An A•C wobble pair has been found in several biologically important RNA motifs, including a ribozyme cleavage site and ribosomal RNA motifs. A•C is also isosteric to the wobble G•U pair and can functionally substitute for the G•U pair in a group I intron (Doudna et al., 1989) and tRNA^{ala} (Gabriel et al., 1996). Adenosine protonation may act as a switch as the pH changes with the protonation status of N1. Although the formation of the A•C base pair depends on local pH, the cellular function and its implication for Com recognition remain to be explored.

MATERIALS AND METHODS

RNA synthesis, crystallization, and X-ray data collection

The 19-mer RNA oligonucleotides, 5'-rGAAUGCCUGCGAG CAUCCC-3' (hereafter referred to according to positions 1 through 19), were synthesized by Dharmacon, Inc. The oligos were deprotected, desalted using a C18 column (Waters), and used in crystallization directly. Dried RNA oligos were resuspended to 40 mg/mL in an annealing buffer containing 50 mM cacodylic acid, pH 6.0, 2.5 mM MgCl₂. Annealing was carried out by heating the RNA solution to 70 °C for 15 min in a water bath and then cooling to room temperature over a period of 3 h. The RNA solution was then diluted to 10 mg/mL in 5–10 mM CrSO₄ or 5–10 mM CrCl₂ before subjecting it to crystallization. Crystals of the 19-mer of size 0.3 × 0.4 × 0.6 mm were obtained at 22 °C by the vapor diffusion method using 4 μL drops. Each drop contained 5 mM RNA, 50 mM cacodylic acid, pH 6.5, 5 mM MgCl₂, 2.5–5 mM CrCl₂ (or CrSO₄), 0.8–1.1 M LiSO₄. The well solution contained 100 mM cacodylic acid, pH 6.5, 5 mM MgCl₂, 1.9–2.2 mM LiSO₄. Although addition of CrCl₂ to the RNA solution was critical to obtain diffractable crystals of the unmodified 19-mer, CrCl₂ precipitated the 19-mer RNAs that contained Br5U modification at positions 16 and 8. Thus Br-RNA crystals were grown in the absence of CrCl₂.

Crystals of the 19-mer RNA are in the space group P6₃22, as indicated by the clear absence of $l = 2n + 1$ reflections. Cell dimensions are $a = b = 45.34$ Å, and $c = 95.13$ Å. Multiple wavelength dispersion (MAD) data from the K edge of the Br atom were collected at the beamline X4A of the National Synchrotron Light Source, Brookhaven National Laboratory. Crystals could be cryoprotected in a cryosolution containing 100 mM cacodylic acid, pH 6.5, 10 mM MgCl₂, 2.2 M LiSO₄, and 5% Xylitol. Preliminary diffraction data indicated that the sixfold symmetry axis coincided with the longest edge of the crystals; thus, the highest symmetry axis could be placed perpendicular to the incident X-ray beam. Before data collection after each beam injection, a fluorescence scan centered around the Br K edge was taken to accurately tune the X-ray wavelength to three wavelengths; those at the inflection point (edge), the absorption maximum (peak), and ~200 eV above the peak wavelength (remote). Three nearly complete diffraction data sets were taken on the crystal at edge wavelength (λ_1), the peak wavelength (λ_2) of the Br atom, and the remote wavelength (λ_3). The three wavelengths were alternated after each sweep of 25 degrees in

phi with a 1° oscillation. A total of 100° were collected for each wavelength. Diffraction data were indexed and integrated with the HKL package (Otwinowski & Minor, 1997). Table 1 summarizes the statistics of the X-ray diffraction data.

Structure determination

The single Br atom was initially located by the program SOLVE (www.solve.lanl.gov); however, the electron density map calculated with the phases from SOLVE (Terwilliger & Berendzen, 1999) did not yield an interpretable electron density map. A previously collected data set (using the in-house X-ray source at 1.54 Å) on a separate crystal of a 19-mer RNA brominated at position 16 was combined with the three-wavelength Br-MAD data set into a four-wavelength data set to maximize the observed dispersive differences. The four-wavelength MAD data set was treated as a special case of MIRAS with a pseudo heavy atom that has zero f_o , but anomalous structural factors (f'' , f') = ($\Delta f''$, $\Delta f'$) where Δ denotes the difference between the “derivative wavelength” and the reference wavelength (Ramakrishnan & Biou, 1997). Phases were calculated from the program MLPHARE (Otwinowski, 1993) in the CCP4 program suite (Collaborative Computational Project, 1994) using the Br position determined from SOLVE. Initial electron densities using the MLPHARE phases were clearly interpretable and the density modification procedure as implemented in the program SOLOMON (Abrahams & Leslie, 1996) further improved their quality (Fig. 3B). Phasing and density modification statistics are shown in Table 1.

Model building and refinement

An initial model of the 19-mer RNA was built by fitting four ideal nucleotides into the experimental densities. There was no observable density for the cytosine ring of the C19 nucleotide; thus, only the phosphate group is included for this residue in the final model. U16 was included as a 5-Br uridine residue. The initial model was refined using the real-space refinement program RSREF without manual intervention (Chen et al., 1999). R factors dropped from 0.45 to 0.3. Subsequent refinement was carried out against the observed structure factors including anomalous difference of the remote wavelength at 0.9998 Å. This model was refined further by the simulated annealing method using torsion angle dynamics followed by several cycles of individual positional and B -factor refinements. Thirty-eight water oxygen atoms were identified using the combination of 2σ peaks in $F_o - F_c$ difference map and the hydrogen bonding environment information. The crystallographic residual values, R/R_{free} , were reduced to 25.6/28.8. At this point, a two-conformer model was introduced, which further reduced R/R_{free} to 23.0/24.7.

Electrophoresis mobility assay

Com protein was purified as described by Hattman et al. (1991). The 19-mer was labeled by the T4 polynucleotide kinase (NEB) in the presence of γ -³²P-ATP. ³²P-labeled RNA was phenol/CHCl₃ extracted, precipitated, and desalted on a C18 column (Waters). RNA samples were annealed in 30 mM cacodylic acid, pH 6.0, 2.5 mM MgCl₂ by heating at 80 °C for

TABLE 1. Data collection, phasing, and refinement statistics.

Data collection statistics												
Data sets	Wavelength (Å)	Resolution (Å)	Measured reflections	Unique reflections	R_{sym} ^a (%)	$\langle I \rangle / \langle \sigma(I) \rangle$	Completeness	Diffraction ratios ^b				
								λ_1	λ_2	λ_3	λ_4	
λ_1 (edge)	0.9202	50.0–1.9 (1.93–1.90)	107,677	8681	0.056 (0.358)	33.6 (6.1)	99.5 (100)	0.045	0.019	0.037	0.10	
λ_2 (peak)	0.9196	50.0–1.9 (1.93–1.90)	107,653	8680	0.057 (0.400)	33.1 (5.9)	99.6 (100)		0.049	0.030	0.10	
λ_3 (remote1)	0.9061	50.0–1.9 (1.93–1.90)	108,395	8703	0.053 (0.498)	33.1 (3.8)	99.5 (92.5)			0.046	0.093	
λ_4 (remote2)	1.5418	50.0–2.8 (1.93–1.90)	16,671	1684	0.091 (0.338)	26.4 (4.2)	99.2 (98.7)				0.042	
Refinement statistics								Phasing statistics				
Resolution (Å)	R_{sym} ^a (%)	R_{cryst} ^c (%)	R_{free} ^c (%)	# of nonhydrogen atoms	# of water molecules	RMSD bond (Å)	RMSD angle (°)	Wavelength				
								λ_1	λ_2	λ_3	λ_4	
50–1.6	5.6	26.0	28.8	41	38	0.008	1.211	Phasing power ^d	—	1.34	1.03	0.80
								R_{Cullis}^{iso} ^e	—	0.74	0.87	0.87
								R_{Cullis}^{ano} ^e	—	0.39	0.37	0.80
								Figure of merits		0.804		

^a $R_{sym} = \sum_{hkl} \langle |I_{hkl}| - \langle I_{hkl} \rangle \rangle / \sum_{hkl} \langle I_{hkl} \rangle$, where I_{hkl} is the i th measured diffraction intensity and $\langle I_{hkl} \rangle$ is the mean of the intensity for the miller index (hkl).

^bDiffraction ratios are fined as $\langle \Delta|F|^2 \rangle^{1/2} / \langle |F|^2 \rangle^{1/2}$, where $\Delta|F|^2$ are taken between the current and the reference wavelength (λ_1) for dispersive ratios and between matched Bijvoet pairs for anomalous ratios.

^c $R_{cryst} = \sum_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \sum_{hkl} |F_{obs}(hkl)|$; $R_{free} = R_{cryst}$ for a test set of reflections (10%) not included in refinement.

^dPhasing power for the i th wavelength = $\sqrt{\sum_{hkl} F_H^2 / \sum_{hkl} (F_{\lambda_1}^{obs} - F_{\lambda_i}^{obs})^2}$.

^e $R_{Cullis}^{iso} = \sum_{hkl} ||F_{\lambda_i}| \pm |F_{\lambda_{ref}}| - |F_H| / \sum_{hkl} ||F_{\lambda_i}| - |F_{\lambda_{ref}}||$; R_{Cullis}^{ano} is calculated similarly with anomalous differences.

15 min and cooling to room temperature over the period of 3 h. A nonrelated 17-mer self-complementary duplex RNA was prepared identically and used as a negative control (Fig. 1C, lanes 1 and 2). Each binding reaction consisted of 1 nM RNA and 4 μ M Com protein and increasing amount of unlabeled 19-mer oligo (Fig. 1C, lanes 6–10) or 2 μ M yeast tRNA (Fig. 1C, lane 5); incubation was carried out for 1.5 h on ice before loading on a 15% native acrylamide gel in 0.5 $\times$ Tris-borate-EDTA (TBE) buffer. The separated products were visualized using a phosphorimager.

Protein Data Bank accession number

The coordinates and structure factors have been deposited in the Nucleic Acid Database (deposit number 1KFO).

ACKNOWLEDGMENTS

We thank Craig Ogata for access to beamline X4A at Brookhaven National Laboratory and Michael Chapman for helpful discussions in refinement. Support by an American Chemical Society Petroleum Research Fund #37150-G4 (to H.L.) and a U.S. Public Health Service Grant GM29227 (to S.H.) are gratefully acknowledged.

Received November 27, 2001; returned for revision January 4, 2002; revised manuscript received April 29, 2002

REFERENCES

- Abrahams JP, Leslie AGW. 1996. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D* 52:30–42.
- Allmang C, Mitchell P, Petfalski E, Tollervey D. 2000. Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Res* 28:1684–1691.
- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution [see comments]. *Science* 289:905–920.
- Brown T, Leonard GA, Booth ED, Kneale G. 1990. Influence of pH on the conformation and stability of mismatch base-pairs in DNA. *J Mol Biol* 212:437–440.
- Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science* 273:1678–1685.
- Chen X, McDowell JA, Kierzek R, Krugh TR, Turner DH. 2000. Nuclear magnetic resonance spectroscopy and molecular modeling reveal that different hydrogen bonding patterns are possible for G.U pairs: One hydrogen bond for each G.U pair in r(GGC GUGCC)(2) and two for each G.U pair in r(GAGUGCUC)(2). *Biochemistry* 39:8970–8982.
- Chen Z, Blanc E, Chapman MS. 1999. Real-space molecular-dynamics structure refinement. *Acta Cryst D Biol Crystallogr* 55: 464–468.
- Collaborative Computational Project. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Cryst D* 50:760–763.
- Correll CC, Freeborn B, Moore PB, Steitz TA. 1997. Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell* 91:705–712.
- Deng JP, Sundaralingam M. 2000. Synthesis and crystal structure of an octamer RNA r(guguuuac)/r(guaggcac) with G center dot G/U center dot U tandem wobble base pairs: Comparison with other tandem G center dot U pairs. *Nucleic Acids Res* 28:4376–4381.

- Doudna JA, Cormack BP, Szostak JW. 1989. RNA structure, not sequence, determines the 5' splice-site specificity of a group I intron. *Proc Natl Acad Sci USA* 86:7402–7406.
- Gabriel K, Schneider J, McClain WH. 1996. Functional evidence for indirect recognition of G.U in tRNA(Ala) by alanyl-tRNA synthetase. *Science* 271:195–197.
- Gautheret D, Konings D, Gutell RR. 1995. G.U base pairing motifs in ribosomal RNA. *RNA* 1:807–814.
- Hattman S. 1999. Unusual transcriptional and translational regulation of the bacteriophage Mu mom operon. *Pharmacol Ther* 84:367–388.
- Hattman S, Newman L, Murthy HM, Nagaraja V. 1991. Com, the phage Mu mom translational activator, is a zinc-binding protein that binds specifically to its cognate mRNA. *Proc Natl Acad Sci USA* 88:10027–10031.
- He L, Kierzek R, SantaLucia J Jr, Walter AE, Turner DH. 1991. Nearest-neighbor parameters for G.U mismatches: 5'GU3'/3'UG5' is destabilizing in the contexts CGUG/GUGC, UGUA/AUGU, and AGUU/UUGA but stabilizing in GGUC/CUGG. *Biochemistry* 30:11124–11132.
- Jang SB, Hung LW, Chi YI, Holbrook EL, Carter RJ, Holbrook SR. 1998. Structure of an RNA internal loop consisting of tandem C-A+ base pairs. *Biochemistry* 37:11726–11731.
- Klosterman PS, Shah SA, Steitz TA. 1999. Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution. *Biochemistry* 38:14784–14792.
- Lavery R, Sklenar H. 1989. Defining the structure of irregular nucleic acids: Conventions and principles. *J Biomol Struct Dyn* 6:655–667.
- Leontis NB, Westhof E. 1998. Conserved geometrical base-pairing patterns in RNA. *Q Rev Biophys* 31:399–455.
- Lutcke H. 1995. Signal recognition particle (SRP), a ubiquitous initiator of protein translocation. *Eur J Biochem* 228:531–550.
- Omichinski JG, Clore GM, Schaad O, Felsenfeld G, Trainor C, Appella E, Stahl SJ, Gronenborn AM. 1993. NMR structure of a specific DNA complex of Zn-containing DNA binding domain of GATA-1. *Science* 261:438–446.
- Otwiniński Z. 1993. Oscillation and data reduction program. In: Sawyer L, Isaacs N, Bailey S, eds., *Proc. CCP4 Study Weekend: Data Collection and Processing*. Warrington, UK: Daresbury Laboratory. pp 56–62.
- Otwiniński Z, Minor W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Ramakrishnan V, Biou V. 1997. Treatment of multiwavelength anomalous diffraction data as a special case of multiple isomorphous replacement. *Methods Enzymol* 276:538–557.
- Ravishanker G, Swaminathan S, Beveridge DL, Lavery R, Sklenar H. 1989. Conformational and helicoidal analysis of 30 PS of molecular dynamics on the d(CGCGAATTCGCG) double helix: "Curves", dials and windows. *J Biomol Struct Dyn* 6:669–699.
- Saenger W. 1984. *Principles of nucleic acid structure*. New York: Springer-Verlag.
- Staley JP, Guthrie C. 1998. Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* 92:315–326.
- Terwilliger TC, Berendzen J. 1999. Automated structure solution for MIR and MAD. *Acta Crystallogr D* 55:849–861.
- Valadkhan S, Manley JL. 2001. Splicing-related catalysis by protein-free snRNAs. *Nature* 413:701–707.
- Weinstein LB, Steitz JA. 1999. Guided tours: From precursor snoRNA to functional snoRNP. *Curr Opin Cell Biol* 11:378–384.
- Wild K, Weichenrieder O, Leonard GA, Cusack S. 1999. The 2 Å structure of helix 6 of the human signal recognition particle RNA. *Structure Fold Des* 7:1345–1352.
- Witkowski RT, Hattman S, Newman L, Clark K, Tierney DL, Penner-Hahn J, McLendon G. 1995. The zinc coordination site of the bacteriophage Mu translational activator protein, Com. *J Mol Biol* 247:753–764.
- Wu M, McDowell JA, Turner DH. 1995. A periodic table of symmetric tandem mismatches in RNA. *Biochemistry* 34:3204–3211.