Cell, Vol. 91, 705–712, November 28, 1997, Copyright ©1997 by Cell Press

Metals, Motifs, and Recognition in the Crystal Structure of a 5S rRNA Domain

Carl C. Correll,*‡ Betty Freeborn,† Peter B. Moore,*† and Thomas A. Steitz*†‡§

*Department of Molecular Biophysics and Biochemistry

†Department of Chemistry

Yale University

‡Howard Hughes Medical Institute

New Haven, Connecticut 06520-8114

Summary

Two new RNA structures portray how non-Watson-Crick base pairs and metal ions can produce a unique RNA shape suitable for recognition by proteins. The crystal structures of a 62 nt domain of E. coli 5S ribosomal RNA and a duplex dodecamer encompassing an internal loop E have been determined at 3.0 and 1.5 Å, respectively. This loop E region is distorted by three "cross-strand purine stacks" and three novel, water-mediated noncanonical base pairs and stabilized by a four metal ion zipper. These features give its minor groove a unique hydrogen-bonding surface and make the adjacent major groove wide enough to permit recognition by the ribosomal protein L25, which is expected to bind to this surface.

Introduction

The principles that govern the folding of RNAs into complex globular structures are still emerging due to the paucity of known RNA structures. Large RNA molecules such as those in the ribosome, the spliceosome, and other ribonuclear protein complexes show secondary structures that have only short stretches of Watson-Crick duplex connected frequently by "loops" and bulges whose secondary structures cannot now be accurately predicted. The repeated occurrence of certain motifs—such as the GNRA tetraloop—suggest that it may be possible to predict the non-Watson-Crick secondary structure if the "language" of the RNA motifs is understood. It has also been long known that Mg^{2+} ions are important for the activity, stability, and presumably structures of RNAs and their complexes. A glimpse of the important role that magnesium plays in the formation and stabilization of RNA tertiary structure has come from seeing a metal ion core at the heart of the p4-6 fragment of tetrahymena group I intron (Cate et al., 1997). However, what role, if any, Mg^{2+} plays in the formation and stabilization of secondary structure, particularly in RNA loops, is less well established.

E. coli ribosomal 5S RNA (5S rRNA), which contains 120 nucleotides (nt), forms part of the 50S ribosomal subunit and binds three proteins—L25, L18, and L5 (Moore, 1996), has long been investigated by those interested in RNA conformation. Like all structured RNAs, it has internal non-Watson-Crick base-paired regions of "loops," and one of them, loop E, adopts structure only in the presence of millimolar Mg^{2+} (Leontis et al., 1986). Mild nuclease digestion of 5S rRNA yields a 62 nt fragment I, which includes helices I, IV, and loop E (Douthwaite et al., 1979). The ribosomal protein L25 binds to both 5S rRNA and fragment I and protects helix IV and loop E from chemical modification (Douthwaite et al., 1982; Huber and Wool, 1984; Toukifimpa et al., 1989).

Results and Discussion

Structure Determination and Overview

Although crystals of fragment I were obtained in 1983 (Abdel-Meguid et al., 1983), we were unsuccessful in solving its structure by multiple isomorphous replacement, in part because heavy atom binding produced nonisomorphous crystals (Kim, 1992). A combination of two multiwavelength anomalous dispersion (MAD) experiments using chemically modified RNA molecules (Correll et al., 1997) has now yielded interpretable electron density maps. MAD phases and phases from single isomorphous replacement that incorporated anomalous scattering were obtained using a phosphorothioate variant of fragment I derivatized with ethylmercury phosphate combined with MAD phases obtained using a bromine derivatized variant (Table 1A). The RNA phosphodiester backbone was positioned in a 5 Å resolution map obtained after cycles of solvent flattening. Iterative cycles of phase-restrained, torsion angle dynamics refinement (Rice et al., submitted), followed by structure factor averaging (Brünger et al., 1997; Shamoo et al., 1997) at 3.5 Å resolution, permitted the base sequence to be added to the structure (Figure 1C). Data to 3 Å resolution obtained from one of the variant RNA molecules were used for the final cycles of maximum likelihood coordinate refinement (Pannu and Read, 1996; Adams et al., 1997) (Table 1B).

Crystals of an RNA duplex containing the minimal 11 bp required for formation of the loop E duplex (Figure 1A) and binding of L25 yielded data to 1.5 Å resolution. The loop E duplex contained a 3′ single-stranded G overhang on one strand and a complementary 3′ single-stranded C on the other. The molecule crystallized with end-to-end stacking of helices that was dictated by complementary interactions between the overhanging nucleotides of successive helices. The structure of the resulting dodecamer was solved by molecular replacement using the corresponding loop E structure from fragment I as the search model and refined at 1.5 Å resolution using maximum likelihood coordinate refinement (Pannu and Read, 1996; Adams et al., 1997) and SHELX-97 (Sheldrick and Schneider, 1997) (Table 1B).

Fragment I forms a linear structure about 94 Å long, which is roughly the radius of the E. coli ribosome. Helix I is not coaxially stacked onto the loop E duplex, to

§To whom correspondence should be addressed.

Table 1. Phasing and Refinement

A. Phasing Data

Data Set	Resolution (Å)	$P6_122$ a (Å)	$P6_122$ c (Å)	Reflection Observations (Unique)	Completeness Overall (%)	R_{sym} (%)[e] All	R_{sym} (%)[e] Bijvoet	Phasing Power (Acentric)[a] 20–5Å	Phasing Power (Acentric)[a] 5–3.5Å	<m>[b] 20–5Å	<m>[b] 5–3.5Å
I. MAD-Hg										0.47	0.16
λ_1 = 1.0092 Å	20–3.5	58.7	248.9	29667 (3602)							
λ_2 = 1.0059 Å				29033 (3615)				0.48	0.17		
λ_3 = 0.9933 Å				29344 (3611)	99.5	5.7[d]	5.7[f] (2.3[g])	0.95	0.24		
II. MAD-Br										0.44	0.11
λ_1 = 0.9203 Å	20–3.5	58.8	252.4	26626 (3582)							
λ_2 = 0.9198 Å				30778 (3592)				0.66	0.19		
λ_3 = 0.9067 Å				30256 (3582)	98.1	7.7[d]	7.3[f] (4.0[g])	1.05	0.36		
III. Hg-Der	20–3.8	58.7	250.8	13172 (4582)[c]	95.7	7.3		1.85	0.69[h]	0.51	0.18[h]
IV. Parent	20–4.2	58.7	252.1	8186 (2235)	98.9	8.4					
Combined										0.72	0.25

B. Refinement Statistics

Data Set	Resolution	N Reflections (% Complete)	R_{sym} (%)	Number of Atoms (nt)	Waters, Mg^{2+}	R (R_{free})	Bond rms Deviations Length (Å)	Bond rms Deviations Angle (°)
I. MAD-Hg	10–3.5 Å	See above	See above	1310 (60)	0, 8[o]	0.260 (0.312[k]) 0.255 (0.305)[l]	0.008	1.4
V. Frag-I[i]	20–3.0 Å	4950 (82.5)	4.7	1310 (60)	4, 12	0.272 (0.324[k]) 0.264 (0.318)[l]	0.007	1.2
VI. Dodecamer[j]	20–1.5 Å	12295 (90.7)	5.5	516 (24)	124, 5	0.202 (0.228[k]) 0.167 (0.224[m])[n]	0.010 0.012[n]	1.2 1.7[n]

(A) For each of the MAD data sets, intensity data were collected at three energies: λ_1, minimum f′, λ_2 maximum f′′, λ_3 remote point. MLPHARE refinement statistics from the two MAD data sets are shown.
[a] Phasing power is defined as rms $<F_H>$ / rms closure error reported for acentric data.
[b] <m> is the figure of merit, defined as $\cos<\sigma(\Delta\varphi)>$.
[c] Number reflects both Bijvoet pairs.
[d] After local scaling with NEWLSC.
[e] $R_{sym} = \Sigma|I - <I>| / \Sigma I$
[f] Acentric data.
[g] Centric data.
[h] 5–4 Å.
(B) Refinement statistics are shown.
[i] Space group is $P6_122$: a = b = 59.0 Å, c = 254.1 Å.
[j] Space group is C2: a = 71.2 Å, b = 37.7 Å, c = 32.9 Å, and β = 105.9°.
[k] ~10% of data.
[l] Data with $F > 2\sigma$.
[m] ~5% of data.
[n] SHELX-97.
[o] Also includes one Hg^{2+}.

which it is linked by two single-stranded bases. This linker allows the helix I axis to translate and rotate relative to loop E so that the terminal nucleotides produced by enzymatic cleavage of 5S rRNA 11 and 70 can dock with the missing stem that is formed by nucleotides 12 through 69 and is missing from fragment I (Figure 1B). The overall structure of this fragment in solution should be similar to that seen here, since the radius of gyration calculated from the crystal structure is only 2% larger than that measured in solution (Leontis and Moore, 1984). Furthermore, gel electrophoresis and transient electric birefringence measurements (Shen and Hagerman, 1994) show that helix I, loop E, and helix IV are colinear in 5S rRNA. Previous solution NMR experiments showed that all loop E bases are paired, although the precise nature of that pairing was not then fully determined (Dallas et al., 1995).

Cross-Strand Purine Stacks

In the crystal structure, loop E forms a double helix that is severely distorted by its 7 non-Watson-Crick bp (Figures 1 and 2). It contains a pair of identical, two-fold-related, 3 bp motifs that are separated by 3 noncanonical bp. Both motifs consist of a Watson-Crick G•C bp followed by a sheared A•G bp and a reversed-Hoogsteen A•U bp (Figure 2A). We call this motif a "cross-strand A stack" because the A of the sheared G•A pair stacks on the A from the reversed Hoogsteen pair, which comes from the opposite strand. This motif is stabilized in part by intra- and interstrand hydrogen bonds involving a 2′ OH and a water molecule. A severe kink in the backbone at the A of the reversed-Hoogsteen bp positions the 2′ OH of the G in the A•G bp to form secondary structure contacts. Further, a water molecule forms a bridge between the two strands by H-bonding as an obligate

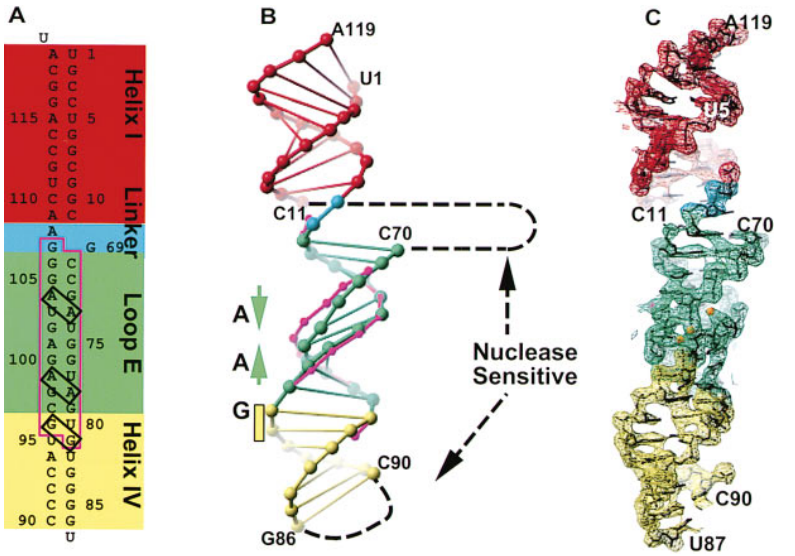

Figure 1. Fragment I Structure Overview

Helix I is drawn in red, linker in cyan, loop E in green, and helix IV in yellow.

(A) The sequence of fragment I in which the loop E dodecamer sequence is outlined in magenta and the cross-strand purine stacks are boxed.

(B) A phosphate backbone representation (with paired nucleotides represented by a line between their phosphates) that shows the two regions cleaved by nuclease as dashed lines. The superimposed backbone of the loop E dodecamer in magenta has a wider major groove between the two cross-strand A stacks whose position and directionality are represented by As and arrows.

(C) A 3.5 Å resolution $2F_o - F_c$ map calculated using phases from simultaneous composite omit and mix target experimental phase restraint refinement (Shamoo et al., 1997) is superimposed on the refined model. The mercury ion is magenta and the magnesium ions are yellow.

acceptor to an N2 proton from G72 and a donor to the two pro-R_p oxygen atoms on the other strand. These interactions may compensate for the loss of stabilization energy that results from the failure of the G in the sheared A•G pair to stack with a neighboring base on the 3′ side. Unstacking of this G creates a pocket in the major groove that here accommodates a metal ion and could accommodate a base in other contexts.

The three cross-strand purine stacks in loop E and helix IV significantly alter the shapes of both the major and minor grooves (Figures 1B, 3B, and 3C). The kinks in the cross-strand A stacks mentioned above, which involve the β and γ backbone angles between the G (72 or 98) and the A (73 or 99), widen the major groove in the direction 5′ to the G and narrow it in the direction 3′ to the G. The effect is magnified by the close proximity of the two kinks. The major groove of loop E is 2.1 Å narrower than A-form RNA in the dodecamer and 6 Å narrower in fragment I (Figures 1B and 3C). The minor groove is correspondingly expanded by 2.2 Å in the dodecamer. Helix IV contains a cross-strand G stack formed by two adjacent G•U wobble bp. The major groove of helix IV is up to 7 Å wider than in A form (as wide as in B form) due to the additive effect of the kink at A99 (in loop E) and the cross-strand G stack. Thus, combinations of cross-strand purine stacks can either increase or decrease the widths of the major and minor grooves of RNA.

The three noncanonical base pairs that lie between the two cross-strand A stacks of loop E (Figure 3A) have not previously been observed and differ from the textbook (Saenger, 1984) non-Watson-Crick base pairs, which all contain two direct interbase hydrogen bonds. These loop E pairs contain either a single, direct interbase hydrogen bond or a bifurcated hydrogen bond between a carboxyl oxygen of one base and the imino and an exocyclic amino group of its partner. Further, they are all stabilized by bridging water molecules.

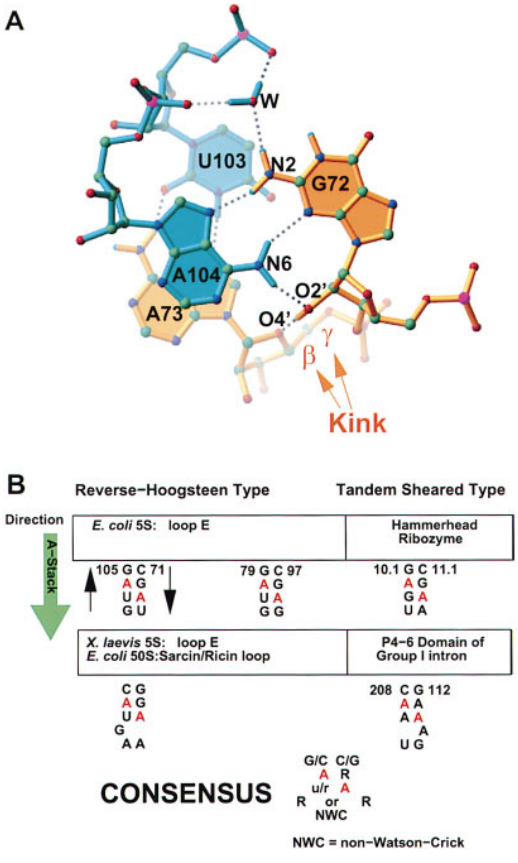

Figure 2. The Cross-Strand Purine Stack

(A) A cross-strand A stack showing A104 from one strand in blue stacked on A73 of the other strand in yellow. The position of the kink at the β and γ angles of the yellow backbone is indicated.

(B) Nucleotide sequences of the six observed cross-strand A stacks (Wimberly et al., 1993; Pley et al., 1994; Szewczak and Moore, 1995; Cate et al., 1996) and the consensus for bulged and nonbulged A stacks.

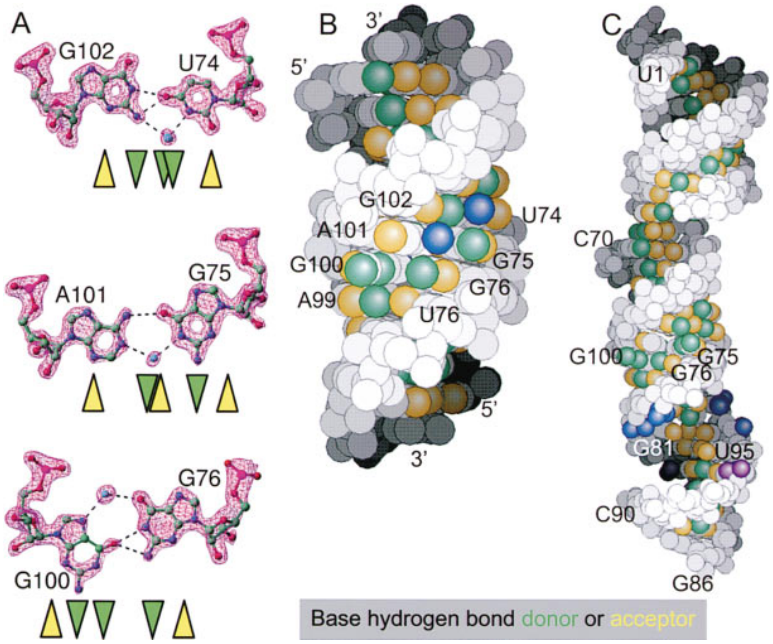

Figure 3. Possible L25 Recognition Surface

Hydrogen bond donors are colored green while acceptors are yellow.

(A) The middle three noncanonical base pairs of loop E superimposed on a 1.5 Å resolution $F_o - F_c$ electron density map obtained using phases calculated with the atoms shown omitted and the other atoms refined after simulated annealing. The hydrogen bond donors and acceptors exposed in the minor groove are indicated by green and yellow arrowheads, respectively.

(B) A van der Waals surface representation of the dodecamer showing a complex H-bonding array in the minor groove, including the bridging structural waters colored blue.

(C) A van der Waals surface representation of fragment I showing the nearly A-form helix I on top, loop E in the middle, and helix IV with its widened major groove on the bottom. The regions of fragment I that the binding of L25 protects from ethylation (Toukifimpa et al., 1989) are blue and from RNase IV hydrolysis (Douthwaite et al., 1982) are purple.

Metal Ion Stabilization

Metal ions bound in the major groove stabilize these three central, noncanonical bp in loop E and contribute to the narrowing of the major groove, particularly in fragment I. While the positions of some metal ions can be plausibly inferred in the 3 Å and 3.5 Å resolution fragment I structures (Figure 4C), the details of their interactions and hydration state are unambiguous in the 1.5 Å resolution map of the dodecamer (Figures 4A and 4B). Magnesium ions and specifically bound water molecules have different coordination geometry (octahedral versus tetrahedral) and different bond lengths (2.1 Å versus 2.7 Å). The bound Mg^{2+} ions are either hexahydrated (A and D) or pentahydrated (B, C, and E), if they are making one direct interaction with the RNA. The major groove of loop E binds five metal ions, all of which are coordinated by base groups and nonbridging phosphate oxygens. As observed in other RNA crystal structures (Pan et al., 1993; Pley et al., 1994; Scott et al., 1995; Cate and Doudna, 1996), the metal ions contact primarily purines, especially the N7 and O6 of guanine (Figure 5). Two magnesium ions (B and C) are involved in a novel binuclear cluster in which three water molecules bridge the two metal ions. The fully hydrated metal ion D stabilizes the duplex by interacting with the backbones of both strands, bases on one strand, and the binuclear cluster (B and C) associated primarily with the other strand. Four of these metal ions lying in the center of loop E (B–E) form a "metal zipper" that may facilitate the narrowing of the major groove.

The even narrower major groove of loop E in fragment I is stabilized by an additional metal ion, D′. Two metal ions, D and D′, bridge between the phosphoryl groups of G75 and U74 of one strand and that of A99 of the other through direct inner sphere coordination (Figure 4C). The difference between the major groove widths of loop E in the dodecamer and fragment I as well as this extra metal ion may be a consequence of the 1.5 M Mg^{2+} concentration in fragment I crystals as compared with 10 mM in the dodecamer crystals. Alternatively, the differing (though very similar) requirements of crystal packing may allow a narrower major groove in one case than in the other.

Implications for rRNA

Based on structural and phylogenetic comparisons (R. R. Gutell, personal communication), it is clear that structural features of loop E in 5S rRNA are not conserved across kingdoms. All bacterial loop Es are likely to contain 3 non-Watson-Crick bp sandwiched between two cross-strand A stacks, as is the case in the E. coli 5S rRNA. In contrast, the loop E regions from eukaryotic 5S rRNAs have a single cross-strand A stack with a bulged G located in the major groove, similar to the structure of the X. laevis loop E (Wimberly et al., 1993). These differences could account for the failure of eukaryotic 5S rRNA to substitute for the bacterial 5S rRNA in large subunit reconstitutions (Hartmann et al., 1988).

Although fragment I is the first RNA in which multiple cross-strand purine stacks have been found, single cross-strand stacks have been observed before. So far, two types of A stacks have been seen. Both are closed at one end by a Watson-Crick G•C pair, which is followed by a sheared purine•purine pair (Figure 2B). Since the third base pair can be different, it distinguishes the two types: either the reversed-Hoogsteen A stack, reported here, or a tandem sheared-pair A stack, as seen in the group I intron (Cate et al., 1996) and the hammerhead ribozyme (Pley et al., 1994; Scott et al., 1995). The third base pair in an A stack can be followed by a bulged base that reaches across the major groove. In the cases of the X. laevis loop E (Wimberly et al., 1993) and the rat Sarcin/Ricin loop, a bulged G lies in the major groove (Szewczak and Moore, 1995). Bulged bases provide alternative recognition opportunities, and in the Sarcin/Ricin loop this bulged G is a recognition element for

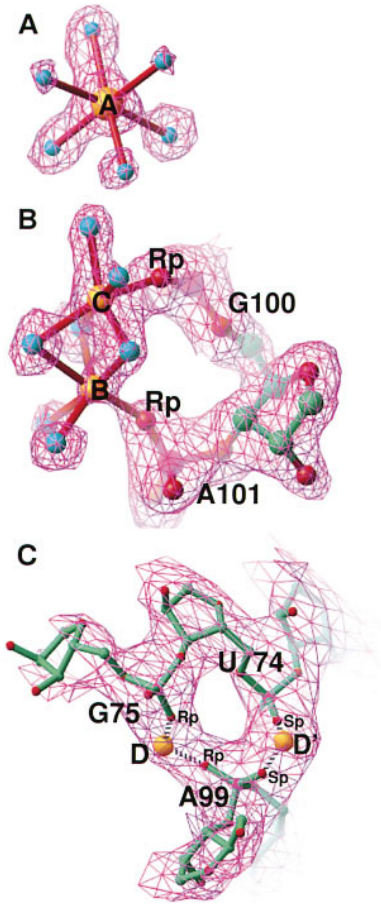

Figure 4. Electron Density of Metal Ions

The metal ions are colored gold, waters are cyan, and inner-sphere bonds between RNA and metal ions are red.

(A) An $F_o - F_c$ map at 1.5 Å resolution calculated as in Figure 3A showing hydration of the A metal ion site.

(B) The B and C metal ion sites. The atoms drawn were omitted from the simulated annealing refinement.

(C) A 3.5 Å resolution map of fragment I calculated as in Figure 1C. The very narrowed major groove allows inner-sphere coordination between metal ion D′ and the pro-S_p oxygens of U74 and A99, as well as between metal ion D and the pro-R_p oxygens of G75 and A99.

α-Sarcin (Gluck and Wool, 1996) and possibly for elongation factors G and Tu (Moazed et al., 1988). The cross-strand G stack seen here is also expected to be a common motif. The sequence, which is a tandem, symmetric, repeated GU, is the most stable tandem, symmetric mismatch (Wu et al., 1995) and also the most common in rRNA. This RNA motif is found in fragment I and a recent octamer crystal structure (Biswas et al., 1997) and was modeled (Gautheret et al., 1995) based on the structure of a DNA G•T tandem pair (Rabinovich et al., 1988).

A model for 5S rRNA structure was proposed previously based on extensive chemical and enzymatic probing of E. coli 5S rRNA; however, it differs significantly from the crystal structure of fragment I. Given the prevalence of water-mediated interactions in nonstandard base pairings and the important roles of metal ions in the structure of loop E, it is perhaps not surprising that none of the seven non-Watson-Crick base pairs was correctly diagnosed by these methods (Brunel et al., 1991).

Implications for RNA Recognition

Perhaps the major functionally important consequence of the non-Watson-Crick base pairs and the cross-strand purine stacks in loop E–helix IV is the creation of a distinctive surface that can be recognized by protein L25. While the accessible minor groove of A-form RNA varies only by the presence or absence of the exocyclic N2 of guanine, the minor groove of loop E is much wider, and its central three noncanonical base pairs present an unusual array of hydrogen bond donors and acceptors (Figure 3). The out of base plane directionality of the lone pairs of the obligate acceptor water molecule mediating U74 and G102 (Figure 3A) adds a geometric component to the varied array of donors and acceptors presented to the minor groove edge of these three base pairs.

The major groove of helix IV, which is adjacent to the loop E minor groove, is wide enough to permit access to an interacting protein. Because of its juxtaposed A and G stacks, it is wide enough to accommodate an α helix, a phenomenon noted previously with RNA structures only in connection with bulged nucleotides (Battiste et al., 1996) and at the ends of A-form duplex (Weeks and Crothers, 1993). Thus, these motifs provide a third way of widening the narrow A-form major groove. In contrast to the normal Watson-Crick base pairs, the wobble base pairs G96•U80 and U95•G81 both present three hydrogen bond acceptors in this widened major groove, thus providing further recognition diversity.

The varied minor groove side of loop E and the adjacent, enlarged major groove are likely to be the surfaces of 5S rRNA that are recognized by the specific binding of ribosomal protein L25. L25 protection of 5S rRNA from ethylation and nuclease digestion positions the L25 binding site predominantly on the two-part surface consisting of the loop E minor groove and the adjacent major groove of helix IV (Figures 3C). Thus, the non-Watson-Crick base pairs, which include three cross-strand purine stacks and three novel base pairs, can augment the unique features to be recognized by proteins in both the major and minor grooves.

The side of the RNA helix orthogonal to these adjacent minor and major groove protein recognition surfaces may be important in the tertiary interactions between RNA helices (Figures 3C and 6). A lattice contact is observed in both the dodecamer and 5S rRNA fragment I crystals at the juxtaposition of the A and G cross-strand stacks. This intermolecular interaction involves the minor groove of loop E and a rather flat backbone (Figure 6). It is stabilized by 14 base–ribose and 8 ribose–ribose hydrogen bonds and buries 560 $Å^2$ of molecular surface per molecule. Thus, in the ribosome, one side of loop E–helix IV may interact with L25 and nearby, perhaps, with 23S rRNA.

Experimental Procedures

Data Collection

Data sets I and II (Table 1A) were collected using phosphorimaging plates (beamline X4A) at Brookhaven National Laboratories; data sets III and IV were collected on an R-axis II at Yale University; data

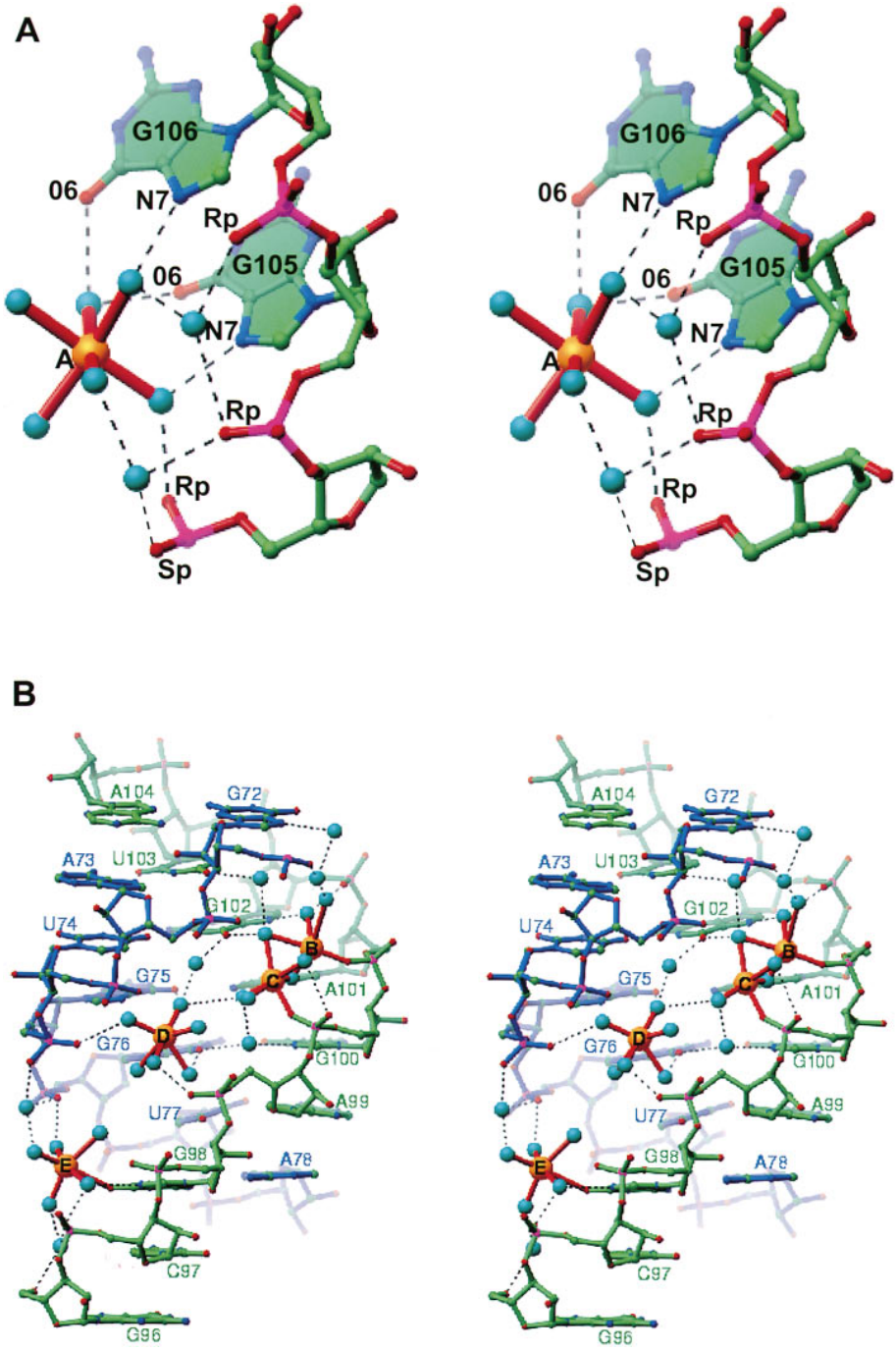

Figure 5. Stereo Diagrams of Metal Sites A through E

(A) A frequently observed G•G metal ion binding site showing the involvement of water in outer-sphere coordination.
(B) The "metal ion zipper" in the narrow major groove of loop E. The two strands are in green and blue.

sets V and VI used for refinement reported in Table 1B were collected using a 2K charge-coupled device (CCD) on beamline A1 at Cornell High Energy Synchrotron Source (CHESS). Bijvoet pairs were measured by the inverse beam method.

Fragment I

Fragment I of 5S rRNA was prepared and crystallized from solutions containing 1.35 M $MgSO_4$ as described previously (Abdel-Meguid et al., 1983; Kim, 1992; Correll et al., 1997). The chemical modifications and preparation of RNAs contained in the fragment I data sets from Table 1 have been previously described (Correll et al., 1997). MAD-Hg, Hg-Der, and Parent crystals contain RNA with a Sp phosphorothioate linkage between bases 4 and 5 and a deoxyribose at base 4. The MAD-Br variant is similar with 5-Bromine at base 4. Frag-I crystals contain RNA with a phosphorothioate linkage between bases 3 and 4 and a deoxyribose at base 3. The Parent RNA was soaked overnight in 1 mM $CdCl_2$; MAD-Hg RNA was soaked overnight in 0.2 mM ethylmercury phosphate, and Frag-I RNA was cocrystallized with a 3-fold molar excess of methylmercury (II) chloride.

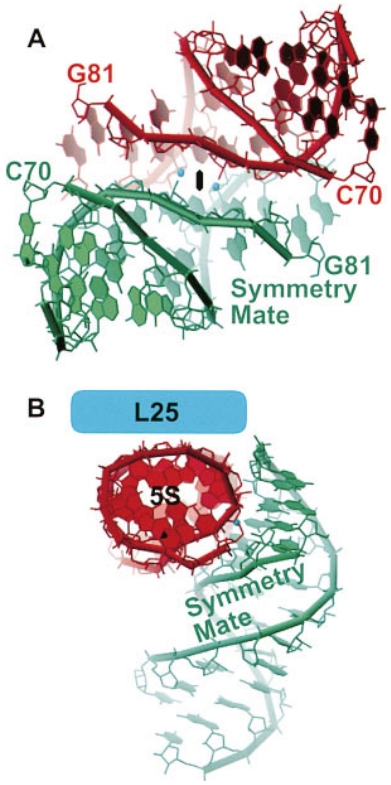

Figure 6. Putative RNA Tertiary Interface

(A) An extensive contact surface between two (red and green) diad axis–related dodecamer molecules in the crystal is at the position of the kink that straightens the backbone.

(B) A view down one loop E helix axis shows the relationship between the putative L25 binding surface and the lattice RNA contact, which do not overlap.

Two factors that were critical for refinement of the fragment I structure at 3.5 Å resolution were first, the use of mixed target refinement (Rice et al., submitted), which incorporates experimental phases during refinement, and second, the use of the R_{free} (Brünger, 1992a) to monitor the progress of the refinement. The register of the nucleotide chain in the map was confirmed by the bromine position at C4. Density peaks were identified as metal ions if their position was less than 2.4 Å from an RNA oxygen atom or if they corresponded to a magnesium ion site seen in the dodecamer structure. Before each cycle of refinement, bulk solvent and anisotropic B corrections were applied to the applitudes. Large anisotropy and thermal diffuse scatter in the fragment I data sets possibly contribute to the high R factor.

Loop E Dodecamer

The dodecamer was synthesized chemically by the Yale Keck Microchemical facility, then deprotected and purified as described previously (Correll et al., 1997). Prior to crystallization, the concentration of the duplex was adjusted to 0.2 mM, and it was annealed (Correll et al., 1997) in 5 mM $MgCl_2$ and 50 mM Na•cacodylate (Cac•6) (pH 6.0). The annealed dodecamer was then concentrated to ~1 mM using a centricon-5 from Amicon and mixed with an equal volume of a reservoir solution containing 5%–25% 2–4 methyl pentane diol (MPD), 5–25 mM $MgCl_2$, and Cac•6. Crystals were grown by vapor diffusion by mixing 2λ of this RNA containing solution with 2λ of well solution and was equilibrated against the reservoir at 20°C. Crystals were transferred to a cryostabilizing solution containing 25% MPD, 10 mM $MgCl_2$, and Cac•6 for 5 min prior to flash-freezing in propane cooled to its freezing point by liquid nitrogen.

The dodecamer structure was solved by molecular replacement using the corresponding portion of the fragment I structure as the search model, which included C71–U80, G96–G105, and four metal ions whose occupancy was set to 0.4. A direct rotation search after Patterson correlation refinement produced a 7σ peak. A translation search produced a 3.4σ peak. Although these peaks were only modestly higher than the next highest, the solution led directly to the refined structure. After rigid body refinement of the search model, the four nucleotides not initially included were visible in difference Fourier maps, and one round of simulated annealing refinement reduced the R_{free} to 0.315. Further, the deoxyribose at 105 and the phosphorothioate between 105 and 106 were detected in difference maps. Difference density peaks were identified as magnesium ions if they were octahedrally coordinated with at least four of the ligand distances being less than 2.4 Å. Before each cycle of refinement, bulk solvent and anisotropic B corrections were applied to the applitudes. The final refinement statistics are shown in Table 1B.

Software

All integration and merging of intensities were calculated using DENZO and SCALEPACK (Otwinowski and Minor, 1997). For each MAD data set, local scaling with NEWLSC written by A. M. Friedman (Purdue University) provided a slight improvement in the data quality. MLPHARE (CCP4, 1994) was used to determine the MAD and SIRAS phases, while SIGMAA (CCP4, 1994) was used to combine the phase sets. DM (CCP4, 1994) was used for solvent flattening. The models were built with the program O (Jones et al., 1991), and X-PLOR (Brünger, 1992b) was utilized to refine the models. Final stages of refinement for fragment I structures (Table IB) employed the developmental program Crystallography and NMR System (A. T. Brünger, personal communication), and SHELX-97 was also employed for the dodecamer structure. All figures were generated with RIBBONS (Carson, 1991), except Figure 3B and 3C, which employed GRASP (Nicholls et al., 1993).

Acknowledgments

This work was supported by American Cancer Society postdoctoral fellowship-3840 to C. C. C., National Institutes of Health (NIH) grant GM-54216 to P. B. M., and NIH grant GM-22778 to T. A. S.

Received September 22, 1997; revised October 20, 1997.

References

Abdel-Meguid, S.S., Moore, P.B., and Steitz, T.A. (1983). Crystallization of a ribonuclease-resistant fragment of *Escherichia coli* 5S ribosomal RNA and its complex with protein L25. J. Mol. Biol. *171*, 207–215.

Adams, P.D., Pannu, N.S., Read, R.J., and Brünger, A.T. (1997). Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement. Proc. Natl. Acad. Sci. USA *94*, 5018–5023.

Battiste, J.L., Mao, H., Rao, N.S., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D., and Williamson, J.R. (1996). α helix-RNA major groove recognition in an HIV-1 Rev peptide-RRE RNA complex. Science *273*, 1547–1551.

Biswas, R., Wahl, M.C., Ban, C., and Sundaralingam, M. (1997). Crystal structure of an alternating octamer r(GUAUGUA)dC with adjacent G•U wobble pairs. J. Mol. Biol. *267*, 1149–1156.

Brunel, C., Romby, P., Westhof, E., Ehresmann, C., and Ehresmann, B. (1991). Three-dimensional model of *Escherichia coli* ribosomal 5S RNA as deduced from structure probing in solution and computer modeling. J. Mol. Biol. *221*, 293–308.

Brünger, A.T., Adams, P.D., and Rice, L.M. (1997). New applications of simulated annealing in X-ray crystallography and solution NMR. Structure *5*, 325–336.

Brünger, A.T. (1992a). The free R value: a novel statistical quality for assessing the accuracy of crystal structures. Nature *355*, 472–474.

Brünger, A.T. (1992b). X-PLOR (Version 3.1): A System for X-Ray Crystallography and NMR (New Haven, CT: Yale University).

Carson, M. (1991). Ribbons 2.0. J. Appl. Crystallogr. *24*, 958–961.

Cate, J.H., and Doudna, J.A. (1996). Metal-binding sites in the major groove of a large ribozyme domain. Structure *4*, 1221–1229.

Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R., and Doudna, J.A. (1996). Crystal structure of a group I intron ribozyme domain: principles of RNA packing. Science *273*, 1678–1685.

Cate, J.H., Hanna, R.L., and Doudna, J.A. (1997). A magnesium ion core at the heart of a ribozyme domain. Nat. Struct. Biol. *4*, 553–558.

CCP4. (1994). Collaborative computing project number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. *D50*, 760–763.

Correll, C.C., Freeborn, B., Moore, P.B., and Steitz, T.A. (1997). Use of chemically modified nucleotides to determine a 62-nucleotide RNA crystal structure: a survey of phosphorothioates, Br, Pt, and Hg. J. Biomol. Struct. Dyn. *15*, 165–172.

Dallas, A., Rycyna, R., and Moore, P.B. (1995). A proposal for the conformation of loop E in *Escherichia coli* 5S rRNA. Biochem. Cell Biol. *73*, 887–897.

Douthwaite, S., Garrett, R.A., Wagner, R., and Feunteun, J. (1979). A nuclease-resistant region of 5S RNA and its relation to the RNA binding sites of proteins L18 and L25. Nucleic Acids Res. *6*, 2453–2470.

Douthwaite, S., Christensen, A., and Garrett, R.A. (1982). Binding sites of ribosomal proteins on prokaryotic 5S rRNAs: a study with ribonucleases. Biochemistry *21*, 2313–2320.

Gautheret, D., Konings, D., and Gutell, R.R. (1995). G•U base pairing motifs in ribosomal RNA. RNA *1*, 807–814.

Gluck, A., and Wool I.G. (1996). Determination of the 28 S ribosomal RNA identity element (G4319) for α-sarcin and the relationship of recognition to the selection of the catalytic site. J. Mol. Biol. *256*, 838–848.

Hartmann, R.K., Vogel, D.W., Walker, R.T., and Erdmann, V.A. (1988). *In vitro* incorporation of eubacterial, archaebacterial and eukaryotic 5S rRNAs into large ribosomal subunits of *Bacillus stearothermophilus*. Nucleic Acids Res. *16*, 3511–3524.

Huber, P.W., and Wool, I.G. (1984). Nuclease protection analysis of ribonucleoprotein complexes: use of the cytotoxic α-sarcin to determine the binding sites for *Esherichia coli* ribosomal proteins L5, L18, and L25 on 5S rRNA. Proc. Natl. Acad. Sci. USA *81*, 322–326.

Jones, T.A., Zou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building models in electron density maps and the location of errors in these models. Acta Crystallogr. *A47*, 110–119.

Kim, J.L. (1992). X-ray crystallographic studies of a ribonuclease-resistant fragment of E. coli 5S RNA. PhD dissertation, Yale University, New Haven, CT.

Leontis, N.B., and Moore, P.B. (1984). A small angle X-ray study of a fragment derived from *E. coli* 5S RNA. Nucleic Acids Res. *12*, 2193–2203.

Leontis, N.B., Ghosh, P., and Moore, P.B. (1986). Effect of magnesium ion on the structure of the 5S RNA from *Escherichia coli*. An imino proton magnetic resonance study of the helix I, IV, and V regions of the molecule. Biochemistry *25*, 7386–7392.

Moazed, D., Robertson, J.M., and Noller, H.F. (1988). Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S RNA. Nature *334*, 362–364.

Moore, P.B. (1996). The structure and function of 5S ribosomal RNA. In Ribosomal RNA: Structure, Evolution, Processing, and Function in Protein Biosynthesis, R.A. Zimmermann and A. E. Dahlberg, eds. (Boca Raton, Florida: CRC Press), pp. 199–236.

Nicholls, A., Bharadwaj, R., and Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. Biophys. J. *64* (Part 2), A166.

Otwinowski, Z., and Minor W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Meth. Enzymol. *276*, 307–326.

Pan, T., Long, D.M., and Uhlenbeck, O.C. (1993). Divalent metal ions in RNA folding and catalysis. In The RNA World, R.F. Gesteland and J.F. Atkins, eds. (Plainview, New York: Cold Spring Harbor Laboratory Press), pp. 271–302.

Pannu, N.S., and Read, R.J. (1996). Improved structure refinement through maximum likelihood. Acta Crystallogr. *A52*, 659–668.

Pley, H.W., Flaherty, K.M., and McKay, D.B. (1994). Three-dimensional structure of a hammerhead ribozyme. Nature *372*, 68–74.

Rabinovich, D., Haran, T., Eisenstein, M., and Shakked, Z. (1988). Structures of the mismatched duplex d(GGGTGCCC) and one of its Watson-Crick analogues d(GGGCGCCC). J. Mol. Biol. *200*, 151–161.

Saenger, W. (1984). Forces stabilizing associations between bases: hydrogen bonding and base stacking. In Principles of Nucleic Acid Structure (New York, NY: Springer-Verlag New York Inc.), pp. 120–121.

Scott, W.G., Finch J. T., and Klug A. (1995). Crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA catalytic cleavage. Cell *81*, 991–1002.

Shamoo, Y., Krueger, U., Rice, L.M., Williams, K.R., and Steitz, T.A. (1997). Crystal structure of the two RNA binding domains of Human hnRNP A1 at 1.75 Å resolution. Nat. Struct. Biol. *4*, 215–222.

Sheldrick, G.M., and Schneider, T.R. (1997). SHELXL: high-resolution refinement. Meth. Enzymol. *277*, 319–343.

Shen, Z., and Hagerman, P.J. (1994). Conformation of the central, three-helix junction of the 5S ribosomal RNA of *Sulfolobus acidocaldarius*. J. Mol. Biol. *241*, 415–430.

Szewczak, A.A., and Moore, P.B. (1995). The sarcin/ricin loop, a modular RNA. J. Mol. Biol. *247*, 81–98.

Toukifimpa, R., Romby, P., Rozier, C., Ehresmann, C., Ehresmann, B., and Mache, R. (1989). Characterization and footprint analysis of two 5S rRNA binding proteins from spinach chloroplast ribosomes. Biochemistry *28*, 5840–5846.

Weeks, K.M., and Crothers, D.M. (1993). Major groove accessibility of RNA. Science *261*, 1574–1577.

Wimberly B., Varani, G., and Tinoco, I., Jr. (1993). The conformation of loop E of eukaryotic 5S ribosomal RNA. Biochemistry *32*, 1078–1087.

Wu, M., McDowell, J.A., and Turner, D.H. (1995). A periodic table of symmetric tandem mismatches in RNA. Biochemistry *34*, 3204–3211.

Nucleic Acid Data Base Accession Numbers

The coordinates for the 3.0 and 3.5 Å resolution fragment I structures as well as the dodecamer structure have been deposited with the Nucleic Acid Data Base as entry numbers URL065, URL066, and URL064, respectively.