The Crystal Structure of the *Methanocaldococcus jannaschii* Multifunctional L7Ae RNA-Binding Protein Reveals an Induced-Fit Interaction with the Box C/D RNAs[†,‡]

Jimmy Suryadi,[§,||] Elizabeth J. Tran,[⊥,#] E. Stuart Maxwell,[⊥] and Bernard A. Brown, II*[,§,||]

Department of Chemistry, Wake Forest University, Winston-Salem, North Carolina 27109, Center for Structural Biology, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, and Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, North Carolina 26795

Received March 29, 2005; Revised Manuscript Received May 14, 2005

ABSTRACT: Archaeal ribosomal protein L7Ae is a multifunctional RNA-binding protein that recognizes the K-turn motif in ribosomal, box H/ACA, and box C/D sRNAs. The crystal structure of *Methanocaldococcus jannaschii* L7Ae has been determined to 1.45 Å, and L7Ae's amino acid composition, evolutionary conservation, functional characteristics, and structural details have been analyzed. Comparison of the L7Ae structure to those of a number of related proteins with diverse functions has revealed significant structural homology which suggests that this protein fold is an ancient RNA-binding motif. Notably, the free *M. jannaschii* L7Ae structure is essentially identical to that with RNA bound, suggesting that RNA binding occurs through an induced-fit interaction. Circular dichroism experiments show that box C/D and C′/D′ RNA motifs undergo conformational changes when magnesium or the L7Ae protein is added, corroborating the induced-fit model for L7Ae−box C/D RNA interactions.

The most prevalent nucleotide modifications of ribosomal RNA are the methylation of ribose 2′-hydroxyl moieties and the isomerization of uridines to pseudouridines. Although their biological significance is unclear, most of these modifications are relegated to distinct functional regions and thus suggest an important role in ribosome function (*1*−*9*). Failure to properly modify rRNA can impair the efficiency of protein synthesis (*6*, *7*). In eukaryotic cells, the small nucleolar RNAs (snoRNAs)[1] guide the 2′-O-methylation of approximately 100 nucleotides and the pseudouridylation of a similar number of nucleotides (*10*). The mechanisms specifying the target sites for RNA modification are remarkably similar between Archaea and Eucarya (*11*−*13*). Comparable snoRNA-like RNAs (sRNA) also guide the 2′-O-methylation and pseudouridylation of RNA in Archaea. It appears that in both kingdoms strong evolutionary pressure has maintained the function of these guide RNAs, which likely have common and thus ancient origins prior to the divergence of eukaryotes and archaea, some 2 billion years ago (*1*, *3*, *9*, *12*, *14*).

Modification of ribosomal and other RNAs is directed by small nucleolar ribonucleoprotein complexes (RNPs) that consist of a guide RNA bound by a set of RNP core proteins. The box C/D family of guide RNAs specifically targets 2′-O-methylation of both eukaryotic and archaeal RNAs (*8*, *12*, *15*−*18*). Box C/D RNAs possess C and D *box* sequences located near their 5′- and 3′-termini, respectively, and often contain sequence-related C′ and D′ boxes located internally (Figure 1) (*16*, *19*−*23*). The C/D and C′/D′ boxes are short, conserved nucleotide sequences that fold to form specific core protein-recognition sites. Regions that guide methylation are positioned upstream of boxes D and D′ and are complementary to the target RNA(s) where modification occurs. The target nucleotide of the snoRNA−target duplex positioned five nucleotides upstream of box D or D′ is the substrate nucleotide for 2′-O-methylation (*24*, *25*).

The box C/D core motif is positioned at the termini of most box C/D RNAs and situates the conserved C and D motifs in proximity for base pairing and core protein binding. Recent studies have shown that box C/D motifs adopt a folded RNA structure, termed the "kink-turn" or "K-turn" (Figure 1) (*23*, *26*). In the resulting helix−loop−helix motif, the internal loop is always asymmetric and typically has three unpaired nucleotides on one strand and none on the other. The resulting helix−loop−helix motif exhibits a sharp turn that bends the RNA helices at an ∼65° angle. Critical for K-turn formation are two tandemly sheared G·A base pairs that hydrogen bond across the asymmetric loop (*26*). For most known K-turns, the first helical stem ends with two Watson−Crick base pairs (typically C·G pairs), and the second helical stem starts with the two sheared G·A base pairs stacked on stem II. The sheared G·A base pairs are

[†] This research was supported by grants from the Wake Forest University Science Research Fund and the American Cancer Society (Grant IRG 093-035-09) to B.A.B. and the National Science Foundation (Grant MCB 0215545) to E.S.M.

[‡] Coordinates and structure factors for *M. jannaschii* L7Ae have been deposited in the Protein Data Bank as entries 1RA4 and 1XBI.

* To whom correspondence should be addressed. E-mail: brownba@wfu.edu. Phone: (336) 758-5514. Fax: (336) 758-4656.

[§] Wake Forest University.

[||] Wake Forest University School of Medicine.

[⊥] North Carolina State University.

[#] Current address: Department of Cell and Developmental Biology, Vanderbilt University Medical Center, Nashville, TN 37232.

[1] Abbreviations: *Af*, *Archaeoglobus fulgidus*; *Ap*, *Aeropyrum pernix*; CD, circular dichroism; Gdn·HCl, guanidine hydrochloride; *Hm*, *Haloarcula marismortui*; *Mj*, *Methanocaldococcus jannaschii*; *Pa*, *Pyrococcus abyssii*; rmsd, root-mean-square deviation; *Sc*, *Saccharomyces cerevisiae*; snoRNAs, small nucleolar RNAs; sRNA, small RNA; sRNP, small ribonucleoprotein complex; *Tc*, *Thermococcus celer*.

10.1021/bi050568q CCC: $30.25
Published on Web 06/24/2005

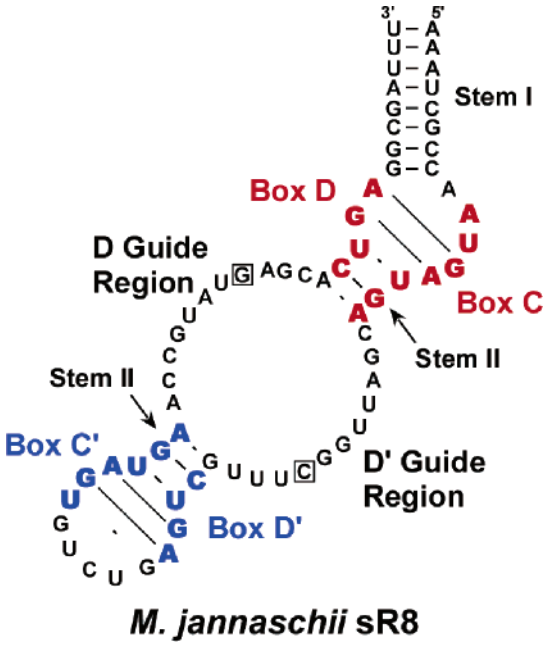

FIGURE 1: *M. jannaschii* sR8. The C/D (blue) and C′/D′ (red) box motifs are shown in bold colors, and the guide regions complementary to the target RNAs are indicated. The nucleotide complementary to the methylation target is boxed. Note that the box C/D and C′/D′ motifs have similar sequences but form different structures. Box C′/D′ has no stem I but rather a five- or six-nucleotide loop.

characteristic of this motif and are essential for tertiary folding. Examples of this conserved RNA motif have been identified in structures of the *Haloarcula marismortui* and *Thermus thermophilus* ribosomal subunits, archaeal and eukaryotic box C/D RNAs, archaeal box H/ACA RNAs, eukaryotic snRNAs, bacterial antitermination sequences, and recently a group I intron (*26*−*37*). Divalent cations, particularly Mg^{2+}, have been shown to stabilize K-turn motifs in solution (*38*, *39*). Cations are presumably also involved in protein−K-turn RNA interactions, although definitive identification has proven to be elusive. Four calcium ions were observed in an L7Ae−box H/ACA−duplex complex structure, yet they interacted with a bulged region of the duplex, and not directly with the K-turn motif (*35*). The exact structure of the internal box C′/D′ motif has yet to be determined. It has box C- and D-related sequences, yet typically includes RNA loops in place of stem I. We propose that the box C′/D′ motif folds into a structure similar to, but distinct from, box C/D. This hypothesis is supported by the demonstration that the core proteins differentially recognize the box C/D and C′/D′ motifs.

In eukaryotes, 15.5kD and Nop58−fibrillarin core proteins bind the box C/D RNA region, while the Nop56−fibrillarin dimer binds the C′/D′ motif. In archaea, L7Ae, Nop56/58, and fibrillarin collectively bind to both box C/D and C′/D′ motifs (*40*−*42*). The sRNA motifs are thought to undergo a conformational change upon protein binding (*30*, *42*, *43*). Because they are differentially recognized by core proteins, the box C/D and C′/D′ motifs must have different RNA secondary structures (*40*−*42*). The L7Ae protein (15.5kD/Snu13p homologue) is particularly interesting, functioning as a ribosomal protein and as both a box C/D and box H/ACA core protein (*34*). This finding indicates that L7Ae plays at least three roles in archaeal organisms and may enable cells to coordinate rRNA processing with ribosome assembly (*9*, *14*, *23*, *42*). Remarkably, archaeal L7Ae interacts with both box C/D and C′/D′, whereas the 15.5kD/Snu13p protein apparently interacts with only box C/D.

Since the L7Ae RNA-binding motif was originally identified (*44*), this protein fold has surfaced in a number of diverse RNA-related functions. L7Ae has notable sequence and structural similarity with eukaryotic 15.5kD/Snu13p, Nhp2p, L7a, L30, S12, and possibly other proteins, such as the selenocysteine insertion sequence binding protein, SBP2 (*33*, *45*). Several groups have studied the 15.5kD, archaeal L7Ae proteins (*26*, *30*, *35*, *36*, *46*, *47*), and the structurally similar ribosomal L30 proteins (*27*, *48*−*52*), but primarily as RNA-bound protein complexes. We were interested in examining the structure of *Methanocaldococcus jannaschii* (*Mj*) L7Ae in the absence of RNA and comparing and contrasting it to the RNA-bound form (*35*) and other homologues. Here, we present a detailed structural analysis of *Mj* L7Ae and biophysical evidence that it stabilizes box C/D K-turn motifs through an induced-fit mechanism that only slightly alters the protein's conformation.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. The *Mj* L7Ae gene was cloned from *Mj* genomic DNA (ATCC) via PCR and ligated into the pET28a plasmid (Novagen) as described by Tran and colleagues (*42*). This vector system produces an N-terminal His_6-tagged protein that can be cleaved by thrombin, leaving only three vector-encoded amino acids before the wild-type methionine. The L7Ae protein was overexpressed in the Rosetta(DE3) pLysS cell line in LB medium by IPTG induction (Novagen). Cells were harvested by centrifugation and lysed using a French press, and the lysate was clarified by centrifugation at 25000*g*. Afterward, the recombinant L7Ae was purified from the cleared lysate by nickel affinity chromatography as suggested by the manufacturer (Qiagen). Peak fractions were pooled and treated with thrombin to cleave the His_6 tag, and the mixture was dialyzed extensively against 20 mM Tris·HCl (pH 8.0), 20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, and 1 mM PMSF (Q-anion exchange buffer). The resulting solution was further purified by FPLC anion exchange chromatography with a RESOURCE-Q column (Amersham Biosciences). Following extensive washing, a linear NaCl concentration gradient from 20 mM to 1 M in Q-anion exchange buffer was used to elute the protein. This step primarily serves to remove contaminating nucleic acids. The purified L7Ae protein was then dialyzed against 10 mM HEPES·KOH (pH 7.5) and 20 mM NaCl and concentrated with centrifugal filters. Purified stock solutions of L7Ae were filtered, dispensed as 1 mL aliquots, flash-frozen in liquid N_2, and stored at −20 °C (the L7Ae protein is quite susceptible to crystallization when stored at 4 °C under low-salt conditions). Yields were typically 35−40 mg of purified L7Ae protein per liter of *Escherichia coli* culture.

RNA Purification. RNA oligonucleotides were purchased from Dharmacon Inc. (Lafayette, CO) and were deprotected according to the manufacturer's instructions (*53*). RNAs were dialyzed against Q-anion exchange buffer and purified using FPLC anion exchange chromatography with a RESOURCE-Q column (Amersham Biosciences). A linear NaCl concentra-

tion gradient from 20 mM to 1 M in 20 mM Tris·HCl (pH 8.0) and 0.5 mM EDTA was used to elute the RNAs. Peak fractions were pooled and dialyzed against buffer containing 10 mM HEPES·KOH (pH 7.5), 10 mM NaCl, 10 mM KCl, and 5 mM $MgCl_2$. The RNAs were then denatured at 65 °C, slowly cooled to 4 °C under dilute conditions (<1 μM) to prevent hairpin dimerization, concentrated, filtered, and stored at −20 °C.

Crystallization. Samples of L7Ae protein [1 mM ≈ 13 mg/mL in 10 mM HEPES·KOH (pH 7.5) and 20 mM NaCl] were screened for crystallization conditions by the hanging-drop vapor-diffusion technique using several commercial sparse-matrix screens (Hampton Research). Typically, 3 μL of an L7Ae protein solution was mixed with 1 or 2 μL of crystallization solution on a siliconized coverslip, inverted, and placed over 1 mL of crystallization solution in VDX crystallization plates (Hampton Research) sealed with Corning Release Compound. A number of long needles appeared in several conditions after approximately 3 days. While conditions were being optimized, a large single crystal (~0.1 mm × 0.1 mm × 0.7 mm) was observed 12 days after the initial screening in condition 41 of Hampton Research's Crystal Screen I (100 mM HEPES·NaOH, pH 7.5 buffer containing 10% 2-propanol and 20% polyethylene glycol 4000). Due to its striking size and quality, this crystal was selected for X-ray analysis without further optimization. Data collection, processing, structure determination, and refinement of this initial low-resolution structure (1.86 Å) are described in the Supporting Information; the details of the subsequent high-resolution structure (1.45 Å) are reported below.

Other lead conditions were iteratively optimized to produce well-diffracting crystals with cryoprotectant qualities. Final crystallization conditions for high-resolution data collection were obtained by mixing 2 μL of 1 mM L7Ae with 1 μL of 20% (w/v) polyethylene glycol 3350, 50 mM HEPES·KOH (pH 7.5), 10% 2-propanol, and 5% ethylene glycol. Crystals were apparent after 1 day as long, intermeshed needles and reached their maximum size after 3−5 days.

Data Collection and Processing. A single crystal was harvested from the mother liquor with a nylon loop, mounted on a goniometer, and flash-frozen in a liquid N_2 vapor stream at 100 K. X-ray diffraction data were collected at beamline X26c (National Synchrotron Light Source, Brookhaven National Laboratory, Upton, NY) at a wavelength of 0.900 Å using an ADSC Quantum 4 detector with a crystal-to-detector distance of 120 mm. Images were collected as 20 s exposures in 1° oscillations; 180 images with 146 021 reflections were colleted in the resolution range of 27.63−1.45 Å. Data were indexed and scaled using HKL 2000 (*54*), resulting in 21 176 unique reflections (19 072 above 1σ). Data reduction statistics and unit cell parameters are summarized in Table 1.

Structure Determination and Refinement. The structure of L7Ae was determined by the molecular replacement method with EPMR (*55*) using the lower-resolution *Mj* L7Ae structure as the search model. The solution of this structure is described in the Supporting Information. One solution was straightforwardly identified with a correlation coefficient of 0.713. Refinement was initially carried out using CNS with ~5% of the data excluded for the R_{free} calculation (*56*). Rigid body refinement of the molecular replacement solution produced a model with an R_{cryst} of 32.03% and an R_{free} of 33.38% in the 27−1.5 Å resolution range. One round of simulated annealing significantly improved the model (R_{cryst} = 28.15%; R_{free} = 29.68%), and manual fitting was initiated using XtalView (*57*). Cross-validated, sigma-A-weighted composite omit maps or sigma-A-weighted $2F_o - F_c$, $5F_o - 3F_c$, and $F_o - F_c$ maps calculated in CNS were used for fitting and model building. Density modification, implemented in CNS, was also used to reduce model bias and improve maps. Approximately 20 iterations of manual fitting and refinement in the resolution range of 27−1.5 Å were performed, yielding a model with an R_{cryst} of 18.61% and an R_{free} of 21.32% (Table 1 in the Supporting Information).

Table 1: Crystallographic Data Statistics for *Mj* L7Ae (1XBI)[a]

space group	$P2_12_12_1$
unit cell dimensions	
a (Å)	43.67
b (Å)	48.87
c (Å)	51.91
$\alpha = \beta = \gamma$ (deg)	90
Mathews coefficient (Å^3/Da)	2.13
solvent content (%)	41.7
resolution range (Å)	27.63−1.45
no. of reflections	
total	146021
unique	21176 (1249)[b]
completeness (%)	100.0 (70.4)
redundancy	7.1 (7.0)
R_{merge} (%)[c]	7.9 (83.9)
$I/\sigma(I)$	31.5 (3.2)
Wilson *B*-factor (Å^2)	19.8

[a] Data were collected on beamline X26c at the NSLS. [b] Highest-resolution bin. [c] $R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i|I_i(hkl)|$, where $I(hkl)$ is the observed intensity of the *i*th measurement of reflection *hkl* and $\langle I(hkl)\rangle$ is the mean intensity of reflection *hkl*.

High-Resolution Refinement. High-resolution refinement was carried out with REFMAC5 (Table 1 in the Supporting Information) (*58*, *59*). Ten cycles of translation, libration, and screw-rotation (TLS) refinement, combined with four cycles of individual atomic refinement, were performed on the L7Ae protein and a 4-(2-hydroxyethyl)-1-piperazineethanesulfate (HEPES) buffer molecule located in the solvent region near H60. Subsequently, eight cycles of restrained refinement with isotropic *B*-factors were carried out. Riding hydrogens were added to the protein model, and 10 additional cycles of restrained refinement were performed. Finally, individual anisotropic *B*-factors were refined, including contributions from riding hydrogens (six cycles), which produced a model with an R_{cryst} of 13.53% and an R_{free} of 19.30% (Table 2). Refinement progress is summarized in Table 1 in the Supporting Information. The final model was analyzed using SFCHECK, PROCHECK, and MolProbity (*58*, *60*−*62*), and statistics are given in Table 2.

Structural and Sequence Analyses. The *Mj* L7Ae secondary structure was assigned using DSSP (*63*), and intermolecular interactions, solvent accessibility, and molecular volume were analyzed using the CCP4 programs *CONTACT*, *NCONT*, *ACT*, *SURFACE*, and *VOLUME* (*58*). Hydrogen bonding was analyzed with *HBPLUS* (*64*). Isotropic temperature factors were examined using *MOLEMAN* (*65*), and anisotropic displacement factors were analyzed with Parvati (*66*). The surface electrostatic potential was calculated with the Delphi module of Accelrys DS Modeling 1.1 using a

Table 2: Refinement Statistics for *Mj* L7Ae (1XBI)

resolution (Å)	27.63−1.45
R_{cryst}[a] (%)	13.5 (13.1)[b]
R_{free}[c] (%)	19.3 (24.7)
estimated coordinate error (Å)	0.074
Luzzati coordinate error (Å)	0.15
Luzzati coordinate error R_{free} (Å)	0.18
Luzzati Sigma A (Å)	0.09
Luzzati Sigma A R_{free} (Å)	0.14
$F_o - F_c$ correlation coefficient	0.973
$F_o - F_{cfree}$ correlation coefficient	0.957
rmsd from ideal geometry	
bond lengths (Å)	0.063
bond angles (deg)	4.590
Ramachandran plot analysis	
most favored regions	93.4% (99 of 106)
additional allowed regions	6.6% (7 of 106)
glycine or proline residues	8.5% (10 of 118)
terminal residues	1.7% (2 of 118)
B-factors (Å^2)	
average	18.52
protein backbone	12.79
protein side chains	18.81
solvent molecules	33.46
total no. of atoms[d]	2393
no. of protein atoms	1859
no. of waters	173

[a] $R_{cryst} = \sum_{hkl}||F_o| - |F_c||/\sum_{hkl}|F_o|$. [b] Highest-resolution bin. [c] R_{free} was monitored with 962 reflections out of 18 451 (~5%) excluded from refinement. [d] Assuming single occupancy for all protein side chain atoms.

finite difference solution of the Poisson−Boltzmann equation (*67*). Protein sequence data used in generating Figure 3 were based on NCBI gene bank entries 15669389 (*M. jannaschii*), 11498370 (*Archaeglobus fulgidus*), 14601646 (*Aeropyrum pernix*), 14520882 (*Pyrococcus abyssii*), 15825948 (*Haloarcula marismortui*), 4826860 (*Homo sapiens*), 132877 (*Thermococcus celer*), 6321408 (*Saccharomyces cerevisiae*), and 6980844. Sequences were aligned using ClustalW at the EMBL-EBI server (*68*), and aligned sequences were compared using the Sequence Manipulation Suite (*69*). ConSurf analyses were performed with the ConSurf server (*70*). Structural alignments, superpositions, and comparisons were carried out using the VAST and Dali servers and with DS Modeling 1.1 (*71*, *72*). Structural figures were created using XtalView (*57*), Molscript (*73*), Raster3D (*74*), Accelrys DS ViewerPro 6.0, and DS Modeling 1.1, and then annotated and rendered in Canvas X (ACD Systems).

Coordinates. Coordinates and structure factors have been deposited in the Protein Data Bank as entries 1RA4 (1.86 Å) and 1XB1 (1.45 Å).

Circular Dichroism Experiments. Wavelength experiments were carried out by monitoring the ellipticity in 1 nm steps from 200 to 350 nm using an Aviv 215 spectrometer. Typically, samples were prepared in 20 mM Na_2HPO_4 or sodium cacodylate (pH 7.0) and 20 mM NaCl (1 mL volumes) in 0.2 or 0.5 cm quartz cuvettes at 25 °C. L7Ae and RNA concentrations were 5 μM, unless otherwise specified.

Circular dichroism thermal denaturation experiments were carried out by first collecting a wavelength spectrum at 25 °C. The sample was then cooled to 2 °C and slowly heated from 2 to 110 °C in 2 °C steps, while the ellipticity was monitored at 222 nm. After the temperature was held constant at 110 °C for 2 min, the sample was recooled to 2 °C in 2 °C increments while the ellipticity was monitored at 222 nm. Finally, the sample was returned to 25 °C, and a wavelength spectrum was collected. These last two steps tested the reversibility of the thermal denaturation. The heating and cooling rates were 20 °C/min, with 0.5 min equilibrations at each temperature. After temperature equilibration, ellipticity data were collected for a 10 s period. Melting experiments were performed in triplicate, and the ellipticity values at each temperature were averaged. The melting temperature (T_m) and the van't Hoff enthalpy of denaturation (ΔH_{vH}) were determined from a plot of the averaged ellipticity at 222 nm versus temperature (in kelvin). The melting curve was fit to a six-variable equation (assuming a two-state transition) implemented in ORIGIN 7.0 (Microcal) (*75*, *76*).

Urea and guanidine hydrochloride (Gdn·HCl) denaturation experiments were carried out by preparing individual samples containing 5 μM L7Ae in 10 mM sodium cacodylate (pH 7.0) buffer containing 20 mM NaCl, 0.1 mM EDTA, and the desired concentration of denaturant. Denaturants were made fresh on the day of experimentation, and concentrations ranged from 0 to 10 M for urea and from 0 to 6 M for Gdn·HCl. Samples were equilibrated at 25 °C for at least 30 min before spectra were acquired in the 200−350 nm range. Experiments were performed in duplicate, and the ellipticity values were averaged at each denaturant concentration. A plot of the averaged ellipticity at 222 nm versus denaturant concentration was constructed, and the free energy of denaturation (ΔG_D) was obtained by fitting the curve with an equilibrium two-state denaturation model using ORIGIN 7.0 (*77*).

RNA Titrations. Circular dichroism titration experiments were used to monitor the effects of magnesium and L7Ae

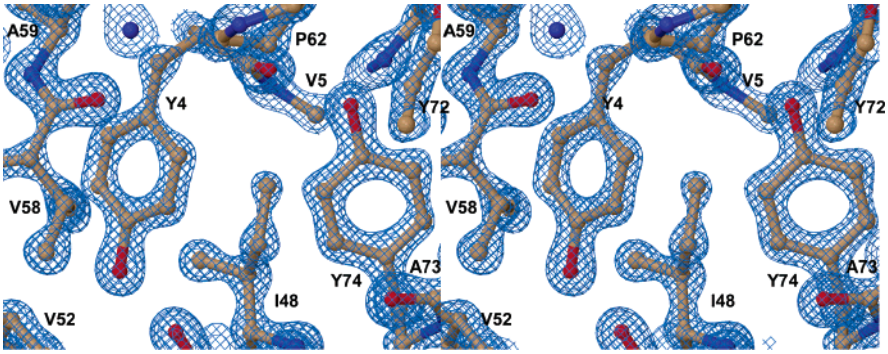

FIGURE 2: Fine structure of *M. jannaschii* L7Ae. Stereoview of a sigma-A-weighted $2mF_o - DF_c$ electron density map contoured at 3σ surrounding the aromatic trefoil in *Mj* L7Ae. This region consists of residues from loop 1 and β1 and forms a hydrophobic barrier on this side of the protein.

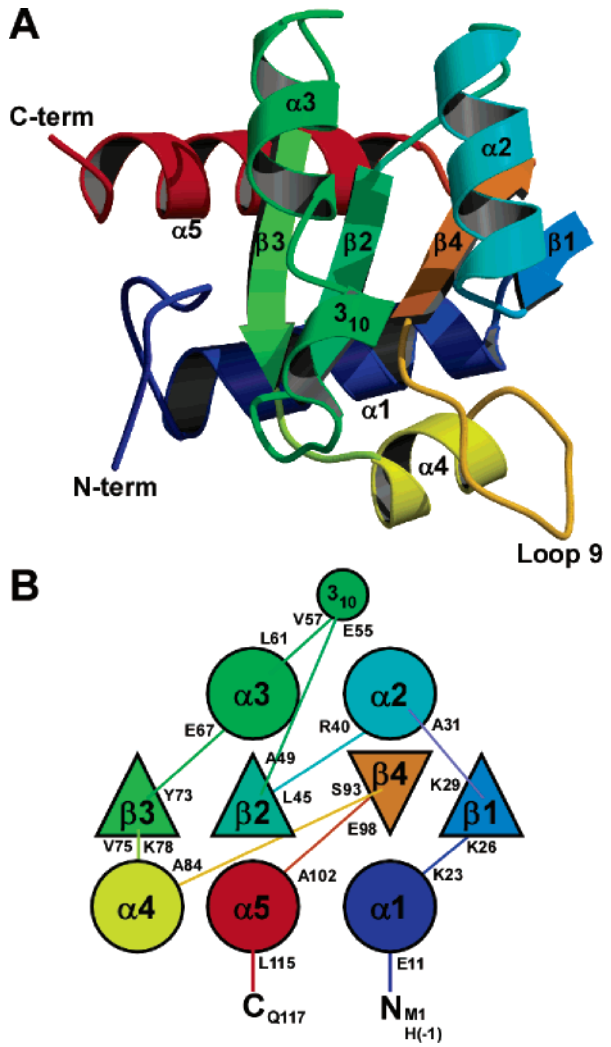

FIGURE 3: Schematic presentation of *M. jannaschii* L7Ae tertiary structure. (A) Ribbon diagram of the L7Ae structure. On one side of the structure, helices α1, α4, and α5 are roughly parallel to each other but are approximately perpendicular to the α2, 3_{10}, and α3 helices on the opposite side. Loop 9 forms a prominent scaffold for interacting with RNA. (B) Topological map of the L7Ae structure clearly showing the α−*β*−α sandwich motif. Colors in the topology diagram correspond to secondary structural elements in the ribbon structure.

on the structure of the box C/D and C′/D′ RNA motifs. RNA samples (5 μM in ~1 mL) were equilibrated in 10 mM sodium cacodylate (pH 7.0) and 20 or 150 mM NaCl. Aliquots of $MgCl_2$ or L7Ae (~1 μL additions of varied concentrations) were added stepwise to the RNA sample; the solution was mixed by inversion, and a CD spectrum in the 350−200 nm region was acquired after a 10 min equilibration at 25 °C. This procedure was repeated until the desired titrant concentration was achieved, usually in fewer than eight additions. The volume and concentrations of the initial RNA samples were prepared so that the final sample volume at the conclusion of the titration was 1.0 mL and the protein and RNA concentrations were both 5 μM (e.g., 995 μL of RNA solution and 5 μL of L7Ae). Dilution of the initial RNA concentration by the titrant was kept below 1% in all experiments (i.e., <10 μL of titrant was added to the initial RNA sample).

RESULTS AND DISCUSSION

L7Ae Structure Determination. To evaluate the differences in the structure of free L7Ae versus RNA-bound protein and to gain insight into the structure basis of L7Ae's broader RNA specificity versus the eukaryotic 15.5kD protein homologue, we crystallized and determined the structure of *Mj* L7Ae. Data were collected at beamline X26c at the National Synchrotron Light Source using a single crystal (Table 1). The structure was determined using the molecular replacement method based on a lower-resolution structure (see the Supporting Information for a complete description). Refinement was initially performed with CNS (*56*), and then REFMAC5 was used to refine anisotropic *B*-factors and riding hydrogens (*59*, *78*). The final model was refined to a resolution of 1.45 Å, yielding exceptional crystallographic statistics and stereochemical parameters (Table 2).

Quality of the Structural Model. The recombinant *Mj* L7Ae protein contains 120 amino acids, including three N-terminal plasmid-encoded amino acids (G−S−H...) that remain after thrombin cleavage of the His_6 tag. The final crystallographic model contains 118 residues, beginning with His −1 and continuing to the wild-type C-terminal residue, Gln118. Electron density was observed for virtually all of the wild-type protein residues, but was not apparent for the vector-encoded Gly −2 and Ser −3, indicating that they were disordered in the crystallographic lattice. A fully traceable density envelope was apparent at the beginning of refinement but during the latter stages of refinement became striking, fully encompassing virtually every amino acid and water, even at 3−4σ contouring (Figure 2). Alternative conformations were modeled for 13 residues (K6, Q25, N32, E39, R40, I47, K53, E55, E56, V57, I70, E103, and E104), and a HEPES buffer molecule, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonate, apparent in the solvent region near H60 and Y63, was included in refinement. The final model contains 1859 non-hydrogen protein atoms and 173 water molecules. Analysis of the Ramachandran plot showed that 99 of 106 amino acids, which were not glycine, proline, or terminal residues, were in the most favored regions of the plot (93.4%) and seven residues (6.6%) were in additionally allowed regions (Table 2). Average isotropic displacement (*B*) factors were 18.5 Å^2, with 12.8 Å^2 for the peptide backbone, 18.8 Å^2 for side chains, and 33.6 Å^2 for solvent molecules (Figure 1 in the Supporting Information). Analysis of anisotropic displacement factors (U^{ij}) was accomplished using PARVATI (*66*). Protein atoms had mean anisotropic factors of 0.618 ± 0.137 and 0.483 ± 0.137 for solvent atoms.

This protein structure was determined and refined at the highest resolution among L7Ae homologues or structurally related proteins, to date, permitting detailed analyses of intermolecular contacts and interactions with solvent molecules or ions. We have been able to thoroughly analyze the effects of RNA binding on the protein structure and compare it with the RNA-bound form (*35*), other homologues (*26*, *30*, *36*, *46*, *47*), and structurally related proteins (*27*, *50*, *52*).

Structure Description. The L7Ae tertiary structure consists of an α−*β*−α sandwich fold with alternating α-helices and *β*-sheets separated by coiled regions (Figure 3A). Helices α1, α4, and α5 are parallel to one another and orthogonal to the roughly parallel α2, α3, and 3_{10} helices on the opposite

side of the globular domain (Figure 3A,B). The central core of the α−β−α sandwich is composed of three parallel β-strands (β1−β3) and one antiparallel β-strand (β4) topologically spaced between the β1 and β2 strands; together, these four β-strands form a β-sheet, with β1 approximately 120° out of plane compared to β3 (Figure 5A). In the *Mj* L7Ae structure, there are two type I turns (A2−V5 and A49−V52), two type IV turns (A43−V46 and V52−E55), and four type VIII turns (V3−I6, A24−I27, I96−E99, and E99−A102). The N-terminus contains a 10-residue coiled region containing two turns (loop 1). It lies perpendicular to the central β-sheet axis and forms a hydrophobic barrier on this side of the protein. A second prominent six-residue coiled region (loop 9) extends from the distal end of α4 to the proximal end of β4. It forms a distinct cradle for interacting with K-turn RNA motifs and is buttressed from below by hydrophobic contacts from α4. Residues in loop 9 are thought to be responsible for establishing the RNA binding specificity of the L7Ae family of proteins (*35*).

All of the L7Ae α-helices are amphipathic, with hydrophobic residues extending inward toward the core with a considerable number of acidic and basic residues exposed on the solvent faces. Three β-strands (β2−β4) comprise the hydrophobic core and, not surprisingly, contain almost entirely hydrophobic or nonpolar residues. However, the β1 strand lies on the periphery of the hydrophobic core and roughly alternates between hydrophobic and hydrophilic residues. All of the hydrophilic β1 residues are lysine and extend toward the RNA-binding interface. On the opposite side of the protein, loop 1 makes a hydrogen-bonded turn and a loose bend, buttressing Y4 against the core side of the 3_{10} helix. It then forms a perpendicular brace against β3 before ending with P10 at the proximal end of α1. A striking feature of loop 1 is that F7 forms a three-residue aromatic trefoil with Y72 and Y74 of β3, effectively sealing off the hydrophobic core on this side of the domain (Figure 2). It is interesting that four of the five aromatic residues in L7Ae are involved in interactions with or by loop 1. The number of aromatic residues in the L7Ae family is highly conserved, and most homologues have four or five aromatic residues, typically in the same tertiary region as *Mj* L7Ae. In the archaeal L7Ae homologues, tyrosine is the predominant aromatic residue, but phenylalanine is most common in eukaryotes.

Mj L7Ae contains approximately 151 hydrogen bond donor moieties and 176 hydrogen bond acceptor groups. The secondary structural elements of L7Ae are involved in approximately 150 intraprotein hydrogen bonding interactions; the majority of these (i.e., 77) are $O_i \rightarrow H-N_j$ interactions common in β-sheets, and the second greatest class (40) is made up of $O_i \rightarrow H-N_{i+4}$ interactions, typical of α-helices. The remainder of the hydrogen bonds are in parallel bridges (5), antiparallel bridges (11), and $O_i \rightarrow H-N_{i+3}$ (11) or $O_i \rightarrow H-N_{i+5}$ (3) helical interactions. Approximately 140 hydrogen bonds are made with water molecules surrounding the protein, and in most of these cases (101), the protein is the hydrogen bond acceptor. The solvent shell surrounding the L7Ae protein (173 waters) is involved in at least 186 H_2O-H_2O hydrogen bonding interactions.

Mj L7Ae contains nine ion pairs between lysine and glutamate or aspartate residues within 3.6 Å of each other. Seven of these salt bridges are tertiary interactions. They interlink the secondary structural elements to each other and stabilize the tertiary structure (K6B−E66, K23−E104, K44−E99, K53−E55, K68−E39, K78−D51, and K111−E16). The other two are intrahelical and involve residues one turn around an α-helix from one another. One such ion pair exists in α1 between K15 and D19, and the other is found in α2 between E33 and K36. When the range of ionic interactions is extended out to 6 Å, the number of potential ion pairs doubles (total of 18), and at an 8 Å radius, there are at least eight additional interactions, for a total of 26 ion pairs (*79*). Of 36 charged residues in *Mj* L7Ae (19 acidic and 17 basic), only six are not involved in some type of ion pair interaction (K26, K29, E56, E67, E88, and E110). Of these six, three are highly conserved (K29, E56, and E67) and four are involved in intermolecular interactions. K26, K29, and E88 contact RNA, and E56 is possibly involved in interactions with ribosomal protein L15, based on similar interactions seen with *Hm* L7Ae in the 50S ribosome subunit (Figure 4) (*46*). Ostensibly, ionic interactions are important for maintaining L7Ae's stability at *Mj*'s high growth temperature and interaction with L7Ae's binding partners. Overall, 3474 intramolecular contacts are within 3.6 Å of each other in the L7Ae structure, emphasizing the compact and extremely stable globular domain.

The *Mj* L7Ae protein is composed primarily of nonpolar amino acids (68), including five aromatic residues. However, the protein has 49 polar amino acids with 19 acidic and 17 basic residues (Figure 4). The calculated isoelectric point is 5.59, and the protein is estimated to have a charge of −2 at pH 7.0, based on the naïve assumption that all residues have pK_a values equivalent to the individual amino acids. *Mj* L7Ae has 55 hydrophobic amino acids that are offset by 44 hydrophilic and 18 neutral residues. The structural hydrophobic core is comprised of 51 residues, 23 of which are completely buried, and the rest are partially exposed (Figure 4). Of the hydrophobic core residues, 11 are neutral amino acids and two are hydrophilic. L7Ae's solvent-exposed surface consists of 66 residues, seven of which are neutral amino acids and 17 hydrophobic. The surprising number of hydrophobic surface residues (∼25%) may indicate regions that are involved in hydrophobic interactions with other proteins. Proteins from thermophilic organisms typically minimize exposure of nonpolar residues and increase the number of charged surface residues as a means of augmenting thermal stability (*79*). The solvent-exposed surface area of *Mj* L7Ae is 6450.7 $Å^2$, and the molecular volume is 17 437.1 $Å^3$, providing a surface area:volume ratio of 0.37 $Å^{-1}$. The molecular density is 1.55 g/cm^3. One small internal cavity of approximately 2.6 $Å^3$ exists between the distal end of α1 and the proximal end of α5 and is lined by A20, K23, E99, G100, D101, E104, and L105. Collectively, these results indicate that *Mj* L7Ae is tightly packed, and this may facilitate a number of favorable van der Waals interactions and contribute to its thermal stability.

Mj L7Ae's surface is punctuated with charged residues, particularly lysines and glutamates (Figure 5B). The electrostatic potential calculated by Delphi is 11 217.8 kcal/mol (46 966.77 kJ/mol or 18 942.93 kT) with a solvation energy of −1275.46 kcal/mol (−5340.09 kJ/mol or −2153.80 kT) (Figure 5C). At least 20 surface amino acids have been identified on *Mj* L7Ae or homologous proteins that contact RNA. They contact RNA either directly or through hydrogen

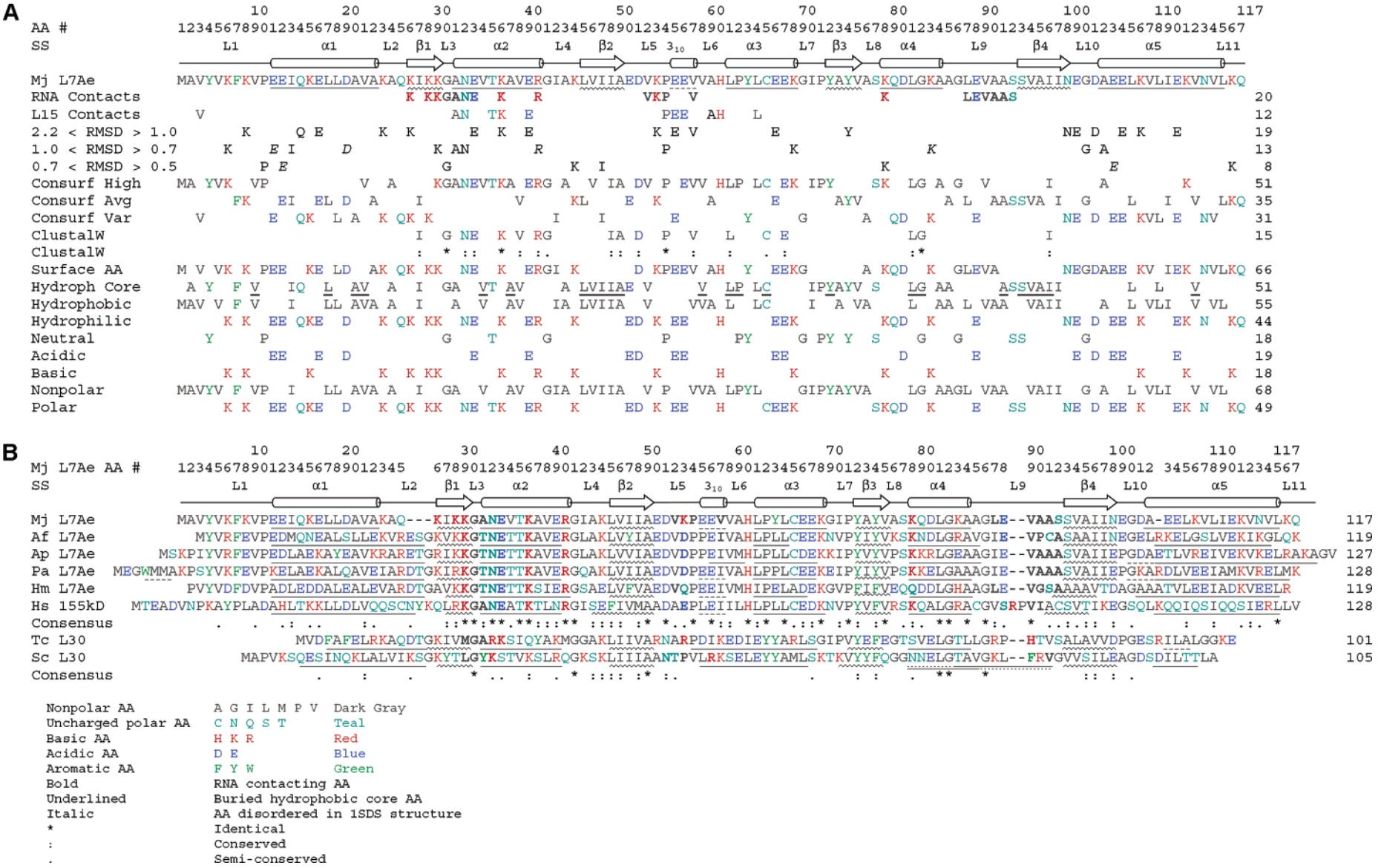

FIGURE 4: Analysis of the secondary structure and functionality of *M. jannaschii* L7Ae amino acid residues. (A) Primary and secondary structure of *Mj* L7Ae and functional attributes of amino acid residues. Amino acids are colored according to their characteristics as described in the figure. RNA-contacting residues are in bold type; underlined amino acids are hydrophobic core residues that are completely inaccessible to solvent, and italic residues are amino acids that were disordered in the *Mj* L7Ae−box H/ACA duplex structure (*35*). Residues that differ between the free and RNA-bound *Mj* L7Ae structures are classified by their rmsds. Conservation of amino acids by ConSurf and ClustalW alignments is indicated (*68*, *70*). (B) Sequence and secondary structural alignment of L7Ae homologues. The sequences of the determined structures of homologous L7Ae and L30 proteins have been aligned and the secondary structures indicated. Residues in α-helices are represented by straight lines, those in 3_{10} helices by dashed lines, and those in *β*-sheets by saw-tooth lines. Loops, turns, and coils are indicated by blank spaces. Residues that contact RNA are in bold type. ClustalW sequence alignment conservation is indicated as follows: asterisk, identical residues; colon, strong conservation; and period, semiconservation.

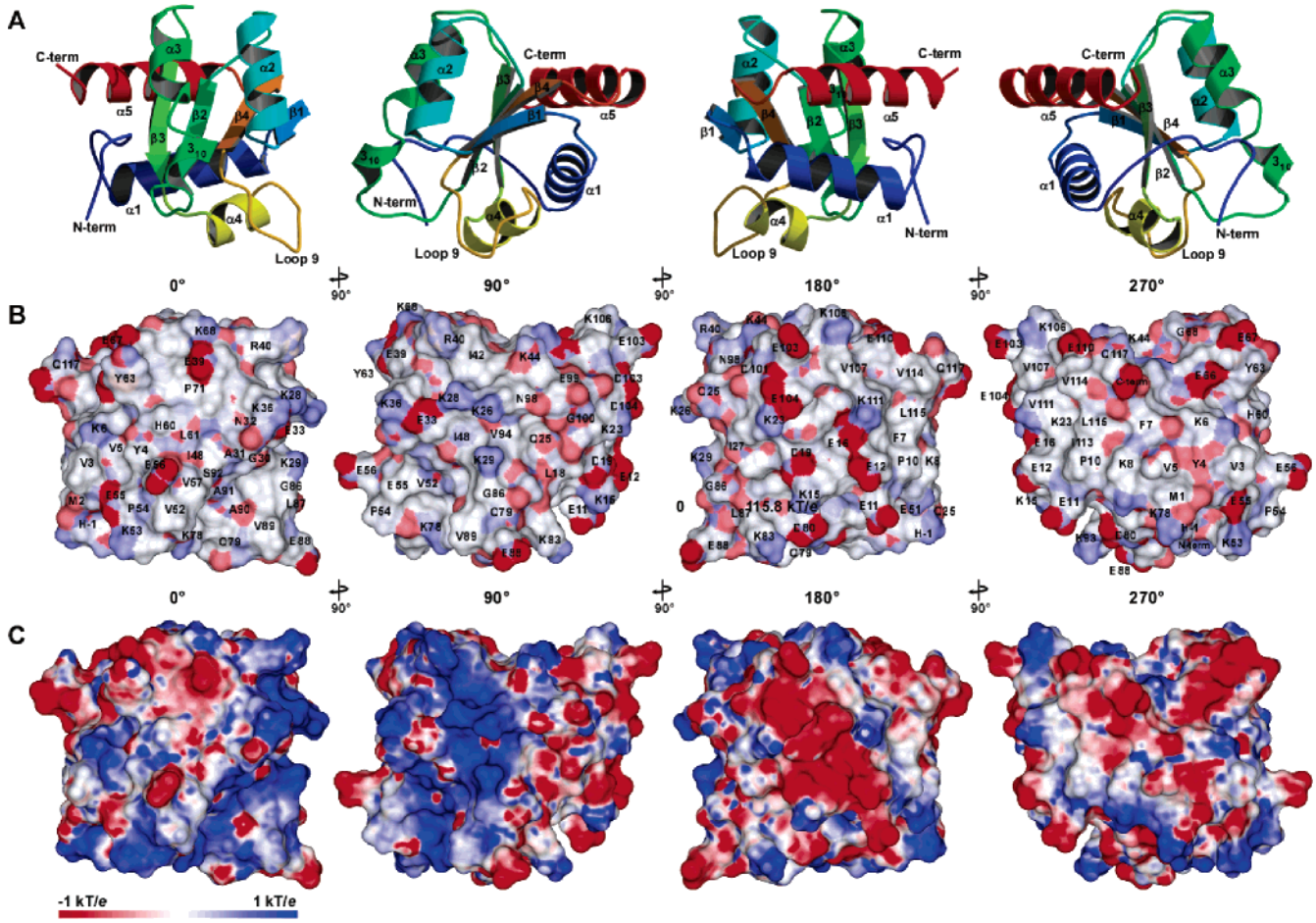

FIGURE 5: Secondary structure, solvent-exposed residues, and charged surface of *M. jannaschii* L7Ae. (A) The complete rotation of the ribbon structure of L7Ae in 90° increments shows the architectural arrangement of secondary structural elements. Rotations directly correspond to the diagrams shown in panels B and C. (B) Solvent-exposed surface maps of *Mj* L7Ae indicating specific charges. Electronegative residues are colored red and positively charged residues blue. The positions of amino acid residues are indicated. (C) Electrostatic potential calculated using Delphi mapped on the solvent-exposed surface of *Mj* L7Ae (*67*). Colors correspond to −1 (red), 0 (white), and 1 kT/e (blue).

bonding and water-mediated interactions (Figure 4) (*30*, *35*, *36*, *47*). RNA-contacting residues are relegated to three distinct regions: the β1−loop 3−α2 region (K26, K28, K29, G30, A31, N32, E33, K36, and R40), the loop 5−3_{10} helix region (V52, K53, P54, and V57), and loop 9 (L87, E88, V89, A90, A91, and S92). Residues that contact RNA or are involved in secondary structural features of the binding region form a roughly 20 Å × 25 Å area on one side of the protein. K78, which makes specific contacts with the extruded uridine of typical K-turn RNAs (*35*), is located at the proximal end of α4 and protrudes toward the binding site. A significant band of positive electrostatic potential exists around the RNA-binding interface and is likely responsible for attracting RNA targets toward the binding site (Figure 5C). At least eight basic residues are involved in this focus of positive potential, including K26, K28, K29, K36, K44, K68, R40, and H60. On the opposite side of *Mj* L7Ae, there is a diagonal swatch of negative surface potential containing E11, E12, E16, D19, D80, D101, E103, and E104. This region is likely involved in electrostatic interactions with its protein binding partners.

Analysis of the *Mj* L7Ae sequence classifies it as a member of Pfam 01246 (ribosomal L7Ae, L30e, S12, Gadd45 family), COG 1358 (ribosomal protein HS6 type [S12, L30, L7a]), and COG 1911 (ribosomal protein L30). ClustalW sequence alignments performed with similar proteins obtained from BLAST searches indicated that there are 19 residues that are phylogenically conserved (Figure 4; Figure 3 in the Supporting Information) (*68*). Of these 19 amino acids, four are universally conserved (G30, K36, P54, and G82), 13 are strongly conserved (I27, N32, E33, V38, R40, I48, A49, D51, V57, L61, E67, L81, and I96), and two are semiconserved (G41 and C65). Note that two of the highly conserved resides are located in critical loop regions and are essential components of the protein fold (G30, loop 3; P54, loop 5). Lysine 36 is part of α2, forms a salt bridge with E33, is involved in RNA interactions, and thus has likely been conserved due to its RNA binding capacity. Glycine 82 exists in α4 directly under loop 9 and is perhaps conserved for steric reasons.

ConSurf analysis of the *Mj* L7Ae structure was also performed. This method maps conserved residues onto the three-dimensional structure through an integrated PSI-BLAST search, ClustalW alignment, and position-based scoring using Bayesian or maximum likelihood algorithms. ConSurf identified 51 highly conserved residues (grades 9-7; 9:11; 8:21; 7:19), 35 with average conservation, and 31 variable residues (Figure 4A). Most of the highly conserved residues identified in the ClustalW alignment were also identified by ConSurf (18 of 19) with the exception of I27,

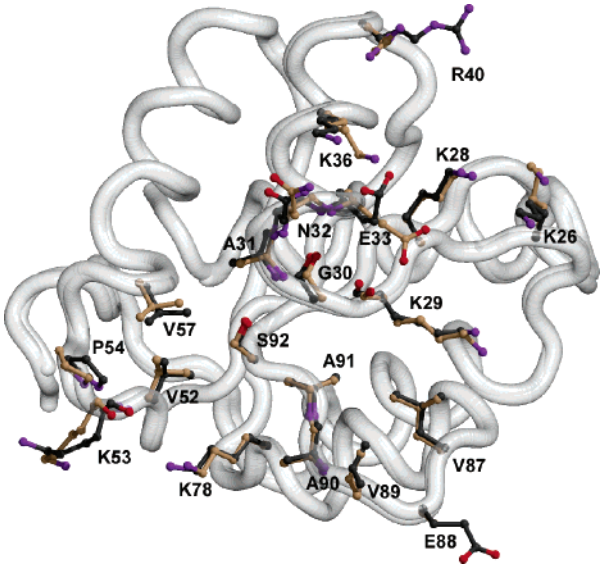

FIGURE 6: Superposition of the free and RNA-bound structures of *M. jannaschii* L7Ae. The rmsd between 115 Cα atoms is 0.355 Å. For all 456 backbone atoms, the rmsd is 0.346 Å, and for all of the 825 atoms in the superposition, the rmsd is 0.835 Å. Specific amino acids that contact RNA are designated; backbone atoms are also shown if they are involved in RNA interactions. The free L7Ae carbon residues are colored dark gray, and the RNA-bound carbons are colored tan. Individual root-mean-square deviations between all RNA-contacting residues are summarized in Table 3. Note the slight shifts of the backbone in regions that contact RNA. The most notable differences between the two structures are seen in residues E33, R40, P54, and V57.

which was classified as having an average level of conservation. ConSurf classified as highly conserved most of the residues which have been shown to be involved in RNA interactions at the distal end of β1 and loop 3 (K29, G30, A31, N32, and E33), as well as R40, K78, and residues in the loop 5−3_{10} helix region (V52, P54, and V57). However, the proximal end of β1, K53 in loop 5, and most of the RNA-contacting residues in loop 9 were classified as having average (K53, L87, A90, A91, and S92) or variable levels of conservation (K26, K28, and E88); the exception was V89, which was classified as highly conserved. This finding is interesting because a phylogenic trend has been observed in the loop 9 region of L7Ae homologues, where archaeal L7Aes have the consensus sequence A-G-L/I-X-V-X′-A-A, and many eukaryotic L7Ae orthologs have the C-G-V-S-R-P-V-I sequence. Upon examination of the sequences, it appears that more variation exists in this region with the protist homologues (i.e., *Trypanosoma*, *Plasmodium*, *Euglena*, *Giardia*, etc.) (Figure 4B and Figure 3 in the Supporting Information). This region of loop 9 contacts the K-turn RNA motif and is likely the basis of RNA target specificity (*35*).

Comparison of the Free and RNA-Bound Mj L7Ae Structures. The *Mj* L7Ae structure experiences minimal conformational changes when binding RNA. Comparison of the free structure of *M. jannaschii* L7Ae to that bound to a K-turn motif from an archaeal box H/ACA-like RNA (*35*) revealed that the overall structure does not significantly change when it interacts with RNA (Figure 6). Several groups had alluded to this possibility in studies of homologous L7Ae proteins, but this study is the first structural comparison of the bound and free forms of the same L7Ae protein (*35*, *36*, *47*). The root-mean-square deviation (rmsd) between 115 Cα atoms in residues 2−116 is 0.355 Å (Table 2 in the Supporting Information); the rmsd is 0.346 Å for all backbone atoms (456 atoms), and for all the atoms in 115 residues (825 atoms), the rmsd is 0.835 Å. Clearly, these structures are extremely similar.

Table 3: *Mj* L7Ae Residues that Interact with RNA and Their rmsds in the Free Protein

residue	rmsd (Å)[a]	rmsd (Å)[b]	contacting atoms
K26	1.76	0.58	H-bond
K28	0.35	0.14	H-bond, H_2O
K29	0.78	0.66	O
G30	0.58	0.06	Cα, C
A31	0.71	0.05	H-bond
N32	0.81	0.33	H-bond
E33	2.16	1.00	N, Cγ, Oϵ2
K36	1.28	0.62	H-bond, H_2O
R40	0.82	0.59	H-bond
V52	0.43	0.05	Cα
K53	1.37	0.84	N, C, O
P54	0.83	0.27	Cδ
V57	1.45	1.43	Cγ1
K78	0.65	0.31	Nζ
L87	0.22	0.07	Cδ2
E88	0.41	0.11	H-bond, H_2O
V89	0.49	0.08	Cβ, Cγ1, Cγ2
A90	0.22	0.06	O
A91	0.31	0.04	O, Cα, Cβ
S92	0.31	0.08	N, Cβ, Oγ

[a] Global rmsd per residue. [b] Per-residue rmsd.

However, some differences appear in specific amino acid residues, particularly those involved in RNA interactions: 19 have rms deviations greater than 1 Å, and 21 have rmsd values between 0.5 and 0.7 Å (Table 3, Figure 4, and Figure 2 in the Supporting Information). Most are found in four distinct regions, two of which are key RNA binding sites: α1, the β1−loop 3−α2 region, the loop 5−3_{10} region, and the loop 10−α5 region. We were most interested in the conformational changes occurring upon RNA binding, so we focused on these residues. Of the 20 residues that contact RNA or are involved in hydrogen bonding or water-mediated interactions with RNA, at least 12 have differences with rms deviation values greater than 0.8 Å between the two *Mj* L7Ae structures (K26, K29, G30, A31, N32, E33, K36, R40, K53, P54, V57, and K78; Table 3, Figures 4 and 6, and Figure 2 in the Supporting Information). The most notable conformational differences are observed for E33, R40, P54, and V57. Smaller structural changes were noted with K26, N32, and K36, where the side chain direction is consistent in both structures, but the side chains have alternative conformations or rotamers.

In the free *Mj* L7Ae structure, E33 is parallel to the α2 helix axis and participates in a salt bridge with K36. However, in the L7Ae−box H/ACA RNA complex, E33 reorients toward loop 9 and establishes hydrogen bonding interactions with the RNA (Figure 6). This ∼120° conformational change was also noted by Charon and colleagues in the free *Pa* L7Ae structure compared to the 15.5kD−U4 snRNA and *Hm* L7Ae−23S rRNA structures (*26*, *30*, *47*). We see the same conformation of E38 in the structure of the free *A. pernix* L7Ae (J. Suryadi and B. A. Brown, II, manuscript in preparation), indicating that reorientation of this glutamate residue occurs upon RNA binding. Initially, in the free protein, E33 forms a salt bridge with K36, but upon RNA binding, E33 apparently reorients and both E33 and K36 establish hydrogen bonding interactions with the

RNA. Note that E33 and K36 are highly conserved amino acids and are probably both essential for productive RNA interactions.

The R40 A conformer also differed between the bound and free *Mj* L7Ae structures (Figure 6 and Table 3). In the free *Mj* L7Ae structure, R40 juts out of the distal end of the α2 helix and points away from the RNA-binding interface, parallel with the α1 and α5 helices. The Cγ, Cδ, and Cζ atoms form a 120° angle, bending the arginine ~65° upward from the position typically observed in other L7Ae homologues. Arginine 40 was truncated at Cγ in the L7Ae−box H/ACA complex, presumably due to disorder (*35*). We were able to model the general direction of this side chain based on the Cγ trajectory, and this positioning is consistent with all of the other homologous structures, including the unbound L7Ae structures from *Pa* and *Ap* (*47*). In the *Mj* L7Ae−box H/ACA RNA complex, R40 faces outward toward the RNA and participates in electrostatic interactions with the RNA phosphodiester backbone (Figure 6), as in the *Af* L7Ae−box C/D RNA complex (*35*, *36*). The R40 B conformer in the free *Mj* structure is oriented in the same direction typical of the homologous structures, suggesting that R40 is mobile until it becomes involved in RNA interactions. Analysis of the R40 *B*-factor in the free L7Ae protein corroborates this idea (Figure 1 in the Supporting Information).

Proline 54 lies on *Mj* L7Ae's surface and forms a bend in loop 5 adjacent to the 3_{10} helix (Figure 6). In the *Mj* L7Ae−box H/ACA duplex structure, Cδ of P54 can make van der Waals contact with the extruded uridine of the K-turn motif. The position of P54 varies by 0.83 Å (rmsd) between the free *Mj* L7Ae structure and the L7Ae−RNA complex. When the rmsd values for P54 of each structure are compared (rather than the global rmsd differences), these residues differ only by 0.27 Å (Table 3). In the free complex, the P54 C−Cα−N angle is greater by ~1.4°, causing Cγ of P54 to shift by 1.2 Å toward the RNA-binding interface. Valine 57 is also shifted toward loop 9 by approximately 0.5 Å in the free structure, resulting in an rmsd of 1.45 Å between the residues in the two structures (Figure 6). In fact, most of the residues in loop 5 through the 3_{10} helix (from Cα of D51 to C of V58) are shifted out of plane by several degrees (0.79 Å rmsd for these residues). Lysine 53 also appears to reorient, turning toward the RNA to make hydrogen bonding interactions. The loop 5−3_{10} helix region also shifts slightly away from the binding interface upon RNA binding. Similar shifts are seen in the amino acids in loop 3 and the proximal end of α2 (residues G30−A37; 0.86 Å rmsd) and, much less frequently, in loop 9 (residues L87−S91; 0.14 Å rmsd; Figure 6). These results collectively suggest that small shifts, on the order of 0.8 Å, do occur in discrete secondary structural regions of the protein backbone upon RNA binding, although the net effects on the global structure are minimal.

Structural Comparison to Homologous Proteins. Comparison to Archaeal and Eukaryotic L7Ae Homologues. Among other L7Ae homologues, the sequence (Figure 4B) and structure of *Mj* L7Ae are highly conserved (Figure 7 and Table 2 in the Supporting Information). Superposition of all determined archaeal L7Ae homologues from *H. marismortui* (*26*, *46*), *A. fulgidus* (*36*), *P. abyssii* (*47*), *A. pernix* (J. Suryadi and A. B. Brown, II, manuscript in preparation) and the human 15.5kD protein (*30*) demonstrates that the global L7Ae structure is practically invariant, even

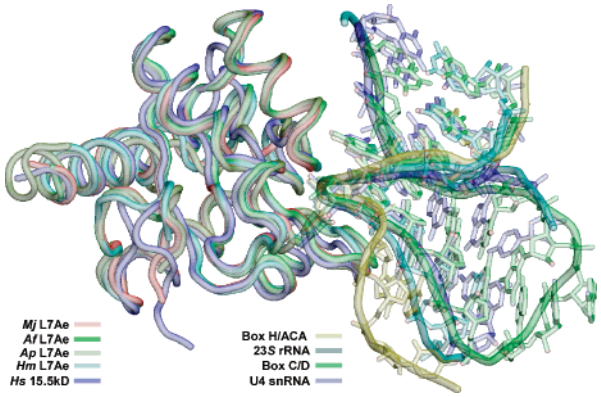

FIGURE 7: Superposition of *M. jannaschii* L7Ae with other L7Ae homologues. Structures are shown for *M. jannaschii* L7Ae (free) modeled with the box H/ACA duplex, *Ar. fulgidus* L7Ae bound to a box C/D RNA motif (0.656 Å Cα rmsd), *A. pernix* L7Ae (free protein; 0.923 Å Cα rmsd), *H. marismortui* bound to KT15 in the 23S rRNA (0.728 Å Cα rmsd), and the human 15.5kD protein bound to U4 snRNA (1.104 Å Cα rmsd). Note the similarity in the protein structures which indicates that the L7Ae structure is evolutionarily conserved and does not undergo significant change upon RNA binding.

when bound to RNA (Figure 7). The rmsds for the homologous proteins range from 0.723 Å for *Pa* L7Ae to 1.104 Å for human 15.5kD (Table 2 in the Supporting Information). The lengths of the secondary structural elements, particularly strands β2 and β3, are highly conserved among the L7Ae homologues, although the lengths of the N- and C-terminal loops and the penultimate α-helices vary to some extent.

Sequence alignment of *Mj* L7Ae with other homologues indicates that *Mj* L7Ae (and its archaeal relative *Methanococcus maripaludis*) is apparently missing three residues in loop 2 between Q25 and K26 and one amino acid in α5 between L105 and K106 (Figure 4B and Figure 3 in the Supporting Information). Structural comparisons of the α1 region show that *Mj* L7Ae makes an abrupt bend from helix 1 into β1, where typical homologues have approximately one additional helical turn in α1. The missing residue in α5, based on the sequence alignment, appears to shorten loop 10, where the *Mj* L7Ae protein makes a sharp three-residue bend into α5. The *Hm* and *Af* L7Ae proteins have a four-residue loop 10 that makes a more gradual transitions into α5. In contrast, the *Pa* and *Ap* L7Ae homologues have a proline residue coupled with a three-residue 3_{10} helix to make this turn. The human 15.5kD protein has a seven-residue loop 10 (I106−L112) that extends outward, making a type II turn. Sequence comparison and modeling studies with other 15.5kD homologues suggest that this extended loop 10 is a common feature of the eukaryotic 15.5kD/Snu13p proteins.

As mentioned above, all of the archaeal L7Ae proteins differ from eukaryotic homologues in loop 9. The consensus sequence for eukaryotic 15.5kD/Snu13p is C-G-V-S-R-P-V-I, whereas archaeal organisms typically have loop 9 sequences of A-G-L/I-X-V-X′-A-A, where X represents E, N, Q, G, or A and X′ represents A, S, G, or, to a lesser extent, P. These differences in sequence were noted by Hamma and Ferré-D'Amaré, and proposed as the basis of RNA target specificity, primarily due to steric hindrance from the highly conserved arginine at the fifth position in loop 9 (*35*). Eukaryotic homologues containing this arginine are incapable of interacting with the box C′/D′ sRNA motif (*42*,

47). We propose that proline at position 6 in loop 9 may also contribute to RNA target specificity.

Comparison to Archaeal and Eukaryotic L30 Homologues. The L7Ae proteins also are similar in structure and sequence to the eukaryotic L30 family of ribosomal proteins. The structures of both protein families are α−*β*−α sandwiches, and they interact with K-turn RNA targets. *Mj* L7ae has 25 identical residues (32.2% similar) and a 1.37 Å rmsd with *S. cerevisiae* L30 and a 1.91 Å rmsd with 21 residues identical (33.9% similar) to *T. celer* L30 (Figure 4B and Table 2 in the Supporting Information). The L30 ribosomal proteins differ from the L7Ae proteins in that they do not have a 3_{10} helix−loop 6 region between *β*2 and α-helix 3. Instead, they have a long continuous helix (α3) between loops 5 and 6 (Figure 4B). The L30 proteins also have shorter N- and C-termini compared to L7Ae homologues. Typically, L7Ae homologues have a long ∼10 residue loop 1 before beginning α-helix 1. At the N-terminus, *S. cerevisiae* L30 has a six-amino acid coiled region before beginning a short three-turn α-helix (∼12 residues), and *T. celer* L30 has only a three-residue loop 1, but a similar short α-helix which begins around the second turn in *Mj* L7Ae's α-helix 1 (residue E16). The C-termini are also different. L7Ae homologues have long four-turn α-helices (∼15 residues) with short one- or two-residue loop 10 regions at their C-termini. The L30 structures have shorter loop 10 regions which transition into a rough helix, formed by type II and IV hydrogen-bonded turns, terminating near the second turn from the end in the *Mj* L7Ae structure. These differences may permit structural differentiation between the similar L30 and L7Ae proteins by their respective binding partners.

Another interesting difference is the structural flexibility of the loop 7−α4 region in the eukaryotic L30 protein. *S. cerevisiae* L30 interacts with K-turn RNAs like L7Ae does, but through a mutually induced-fit binding event, where both the protein and the RNA adopt different structures upon binding (*48*, *49*, *80*). In the free L30 protein, residues N74−G78 form loop 7 and residues T79−G88 comprise α4 (Figure 4B). When RNA binds, a structural transition occurs, and residues N74−V81 become loop 7 and G82−G88 α4. Accompanying these structural transitions is a 13 Å (∼76°) shift of the loop 7 vertex from under the protein's core to the canonical position of loop 9 among the L7Ae/L30 homologues. Crystal and NMR structures of *T. celer* L30 have demonstrated that this archaeal protein does not undergo conformational transitions and is apparently preformed for interactions with RNA, like the archaeal L7Ae homologues (*50*, *51*). On the basis of sequence comparisons, it appears that glycine residues at positions 72, 73, and 88 in *Sc* L30 are partly responsible for the structural transition upon RNA binding. Both the archaeal L30 and L7Ae homologues do not conserve glycine residues at these positions and are not prone to conformational changes; the amino acid residues at other positions within this region are fairly conserved in terms of function and size.

Proteins with Structural Homology to L7Ae. Dali and VAST database searches (*71*, *72*) indicated that the L7Ae family is structurally homologous to a number of diverse proteins, including domain 3 of *H. sapiens* eukaryotic release factor 1 [1DT9 (*81*)], *T. thermophilus* RNA 2′-*O*-ribose methyltransferase [1IPA (*82*)], *E. coli* 23S RNA G2251 2′-*O*-ribose methyltransferase Rlmb [1GZO (*83*)], and *Bacillus subtilis* cytoplasmic ribose-binding protein Rbsd [1OGD (*84*)]. Structurally, these proteins have rmsds between 1.90 and 2.40 Å with the L7Ae proteins (Table 2 in the Supporting Information). All have subdomains with alternating α-helices and *β*-strands and at least three- or four-strand *β*-sheet cores. The structural similarity is primarily due to the *β*-sheet cores surrounded by two to three α-helices, although the topologies differ considerably. Many of these proteins are involved in RNA modification or translation, and it is particularly notable that the two 2′-*O*-ribose methyltransferases have structural features and functions similar to those of L7Ae, suggesting that this protein fold has been involved in RNA recognition since early in evolutionary time.

Mj L7Ae Is an Extremely Stable Protein. We have performed thermal and chemical denaturation studies to gain insight into the stability of the L7Ae protein. The *Mj* L7Ae protein has an apparent melting temperature of 107.3 ± 5.0 °C (380.5 K), which corresponds to a van't Hoff denaturation enthalpy (ΔH_{vH}) of 70.4 ± 3.2 kcal/mol (294.8 ± 13.3 kJ/mol) in 20 mM Na_2HPO_4 (pH 7.0), 20 mM NaCl, and 0.1 mM EDTA (Figure 8A). Denaturation is reversible with minimal hysteresis, and the protein can interact with RNA upon cooling. In fact, *Mj* L7Ae incubated at 100 °C for at least 30 min is still capable of binding box C/D RNA motifs when renatured!

Mj L7Ae's thermal stability should not be surprising, considering that *Mj*'s optimal growth temperature is around 86 °C. The melting temperature is approximately 20 °C greater than the growth temperature, which makes its denaturation temperature more reasonable in comparison to mesophilic proteins (i.e., $T_m \sim 50$ °C for 25−30 °C growth optimums). In comparison, the melting temperature of the *A. pernix* L7Ae homologue is >110 °C and cannot be accurately measured using our standard conditions; *A. pernix*'s optimal growth temperature is 90−95 °C. *Mj* L7Ae has seven salt bridges that interconnect most of the secondary structural elements and probably contributes to the enhanced thermal stability. In addition, the large number of intraprotein hydrogen bonds and the tightly packed hydrophobic core may also contribute to stability. Similar characteristics were noted in the L7Ae homologue from *P. abyssii* and in the *T. celer* L30 structure (*47*, *51*).

We also performed isothermal urea and guanidine hydrochloride (Gdn·HCl) denaturation experiments with *Mj* L7Ae. As with thermal denaturation, *Mj* L7Ae is quite resistant to chemical denaturation (Figure 8B). The midpoint for the Gdn·HCl denaturation titration at 25 °C was at ∼3.25 M Gdn·HCl, which corresponds to a Gibbs free energy of denaturation ($\Delta G_{D,H_2O}$) of 9.7 ± 1.4 kcal/mol (40.8 ± 5.8 kJ/mol) and an *m* value of 3.0 ± 0.42 kcal mol^{-1} M^{-1} (12.7 ± 1.8 kJ mol^{-1} M^{-1}). For the urea-induced denaturation, the midpoint was at ∼6.9 M urea, yielding a ΔG_D of 8.7 ± 2.5 kcal/mol (36.6 ± 10.5 kJ/mol) and an *m* value of 1.3 ± 0.4 kcal mol^{-1} M^{-1} (5.3 ± 1.6 kJ mol^{-1} M^{-1}). Both chemical denaturations were reversible, as the wavelength spectra resembled the native-state spectra after dilution or dialysis to remove the denaturant. The average $\Delta G_{D,H_2O}$ between the urea and Gdn·HCl denaturation experiments is 9.2 ± 2.0 kcal/mol (38.7 ± 8.2 kJ/mol).

Overall, these studies support the finding that *Mj* L7Ae is an extremely stable protein. Our results are similar to those of studies performed with the *T. celer* L30 protein (*51*). *T.*

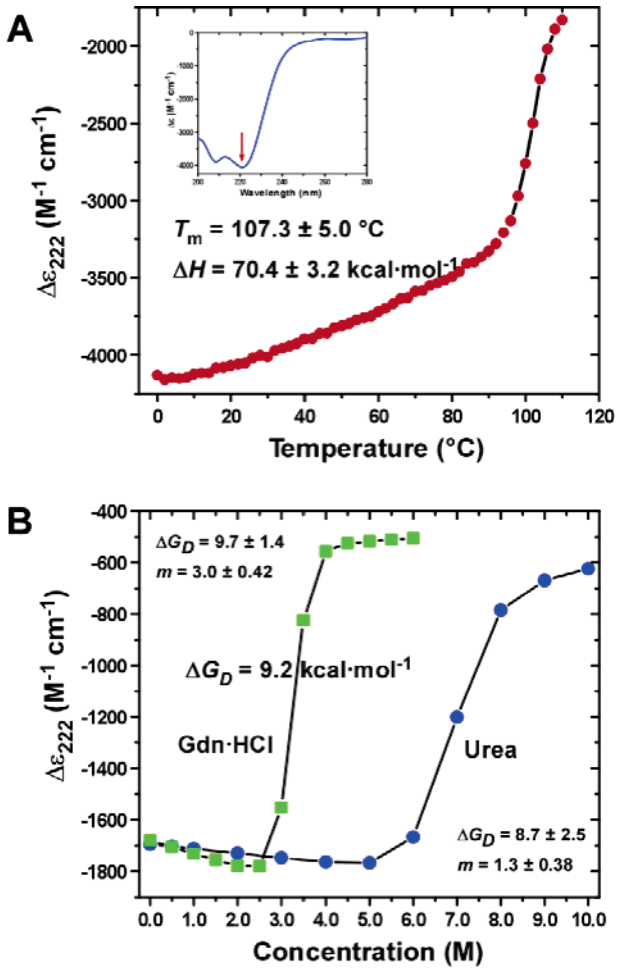

FIGURE 8: Thermal and chemical denaturation of *M. jannaschii* L7Ae. (A) Thermal denaturation of *Mj* L7Ae monitored by CD at 222 nm. L7Ae has an apparent melting temperature of 107.3 ± 5.0 °C which corresponds to a van't Hoff enthalpy of denaturation (ΔH_{vH}) of 70.4 ± 5.0 kcal/mol (294.8 ± 13.3 kJ/mol). The inset shows the *Mj* L7Ae CD wavelength spectrum from 200 to 280 nm, and the arrow denotes 222 nm, the signature wavelength for α-helical secondary structure. (B) Chemical denaturation of L7Ae. Guanidine hydrochloride (green squares) or urea (blue circles) induced denaturation of L7Ae. The midpoints for the titrations were 3.25 M Gdn·HCl and 6.9 M urea. Analyses of the individual denaturation transitions yielded a $\Delta G_{D,H_2O}$ of 9.7 ± 1.4 kcal/mol and an *m* value of 3.0 ± 0.42 kcal mol^{-1} M^{-1} for Gdn·HCl and a $\Delta G_{D,H_2O}$ of 8.7 ± 2.5 kcal/mol and an *m* value of 1.3 ± 0.4 kcal mol^{-1} M^{-1} for urea. The average $\Delta G_{D,H_2O}$ between the urea and Gdn·HCl denaturation experiments is 9.2 ± 2.0 kcal/mol (38.7 ± 8.2 kJ/mol).

celer is also a thermophilic archaeon with an optimal growth temperature near that of *M. jannaschii* (~85 °C). Interestingly, *Tc* L30 has a lower melting temperature than *Mj* L7Ae (94 vs ~107 °C) but an increased resistance to chemical denaturation. *Tc* L7Ae denatures at 4.45 M Gdn·HCl, whereas *Mj* L7Ae denatures at 3.25 M Gdn·HCl. This may be due to the large number of ionic interactions that stabilize *Mj* L7Ae, masked by the high concentrations of Gdn·HCl, which is an ionic denaturant. *Tc* L30 has four short-range salt bridges, ~12 long-range (6−8 Å) ionic interactions, and ~75−80 intraprotein hydrogen bonds. In contrast, *Mj* L7Ae has nine short-range salt bridges (≤4 Å), ~17 long-range ionic interactions (between 4 and 8 Å), and ~150 intraprotein hydrogen bonds. Thus, *Mj* L7Ae would be expected to be more destabilized in the high ionic strength of Gdn·HCl

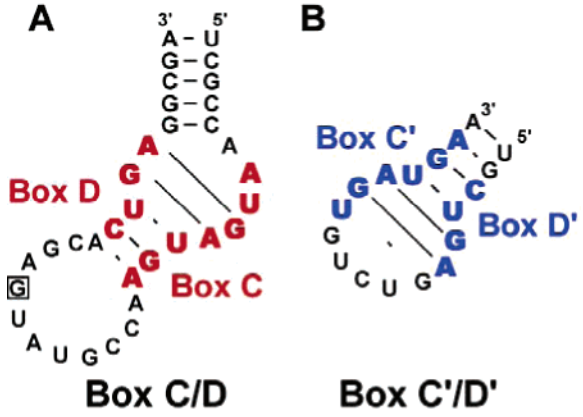

FIGURE 9: *M. jannaschii* sR8 halfmer constructs. (A) Box C/D and (B) C′/D′ halfmer constructs used in circular dichroism experiments (Figure 10). These RNAs recapitulate the box C/D and C′/D′ motifs in isolation (compare to Figure 1). Both motifs have previously been shown to bind the *Mj* box C/D sRNP core proteins. The box C/D halfmer sRNP complex is active in methylation assays, but the box C′/D′ sRNP is not (*42*).

relative to *Tc* L7Ae, and this inference is consistent with our results.

L7Ae Induces RNA Conformational Changes upon Binding. Structural analyses have revealed that *Mj* L7Ae undergoes minimal conformational changes when binding to RNA. We were curious to see whether RNA targets showed conformational changes when L7Ae interacted with them. In these experiments, we used independent RNA constructs representing the box C/D or C′/D′ motifs of the *Mj* sRNA sR8 (Figure 9). The box C/D halfmer construct can assemble sRNP complexes and catalyze 2′-O-methylation in vitro (*42*). The box C′/D′ motif can also bind all three core proteins, but is not enzymatically active in methylation assays (*42*). Our box C′/D′ construct replicates the sR8 box C′/D′ motif and contains an extra A·U base pair at the beginning of stem II for greater stability. Gel mobility shift analyses have demonstrated that L7Ae interacts with both the box C/D and C′/D′ constructs at concentrations used in the CD studies described below (results not shown).

Circular dichroism spectroscopy was used to monitor conformational transitions in the RNA halfmer constructs, while aliquots of magnesium chloride were added to the samples as an initial analysis of structure. In fluorescence energy resonance transfer (FRET) experiments, a KT-7 RNA construct of a K-turn motif found in helix 7 of 23S rRNA of the *H. marismortui* 50S ribosomal subunit undergoes conformational changes in response to magnesium (*26*, *39*). These experiments validated our circular dichroism studies as a way of observing conformational changes induced by L7Ae. When samples of either the box C/D (Figure 10A) or C′/D′ (Figure 10B) halfmers (5 μM RNA) were titrated with $MgCl_2$ from 20 μM to 2 mM, changes in the molar ellipticity were observed. They were greatest at the peak molar ellipticity around 270 nm (268 nm for the box C/D halfmer and 272 nm for the box C′/D′ halfmer), and smaller changes were observed in the 220 nm region. RNA CD spectra arise from $\pi \rightarrow \pi^*$ electronic transitions and are sensitive to nucleotide base stacking interactions and ribose pseudorotation (sugar pucker). When the molar ellipticity values at around 270 nm are plotted against the concentration of magnesium titrated, curves resembling binding isotherms are

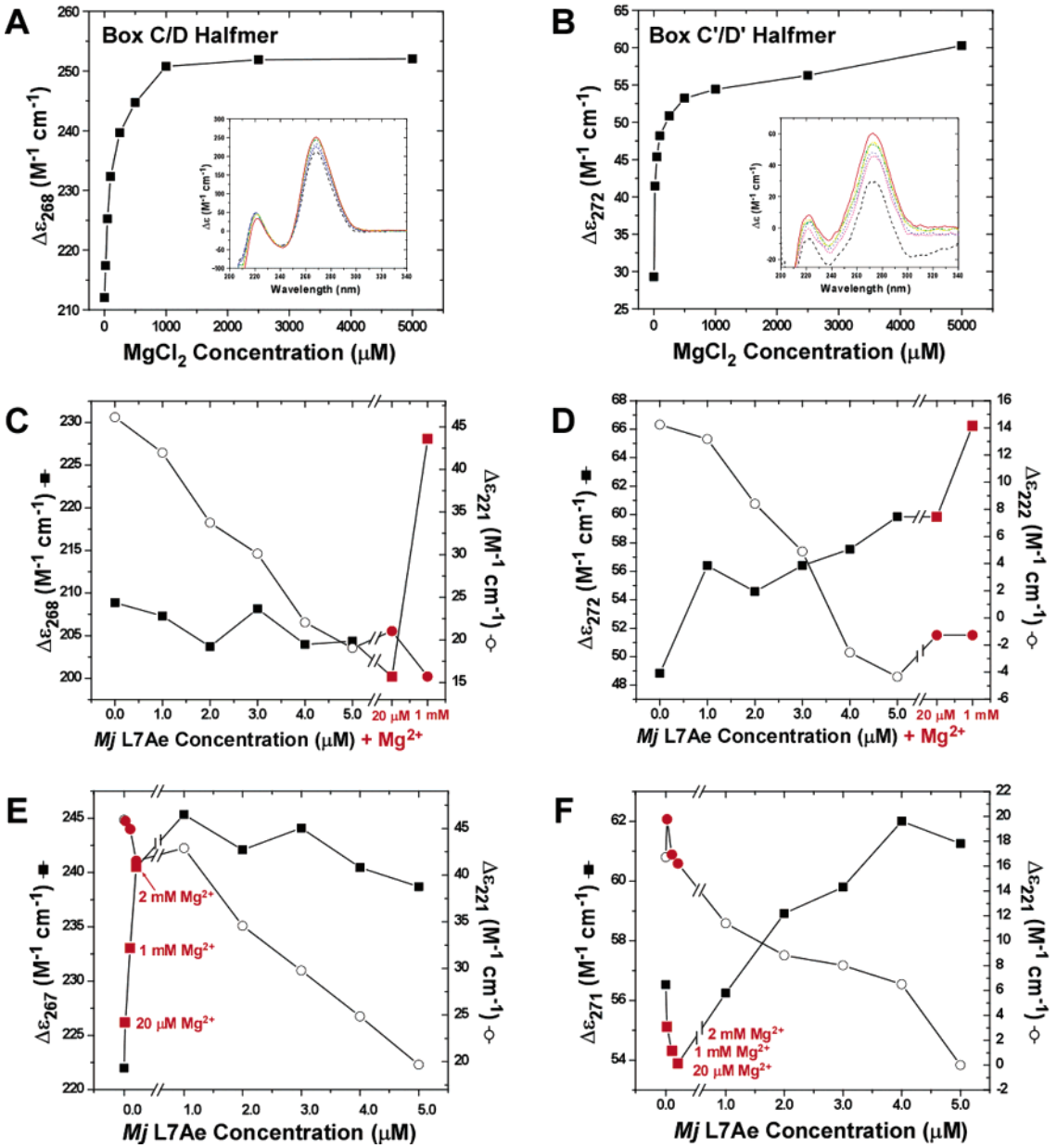

FIGURE 10: Conformational changes in box C/D and C′/D′ RNA structures induced by magnesium and L7Ae binding monitored by circular dichroism difference spectroscopy. Panels A, C, and E depict data for the box C/D halfmer, and panels B, D, and F depict data for the box C′/D′ halfmer. Difference ellipticity values at peak wavelengths, 268 nm (A) and 272 nm (B), are plotted against magnesium concentration. Panels C and D show the difference ellipticity as *Mj* L7Ae is titrated with either the box C/D (C) or box C′/D′ (D) halfmer at an L7Ae concentration from 1 to 5 μM, at a constant 5 μM RNA concentration. Additions of magnesium to 20 μM and 1 mM after L7Ae has been titrated to the equivalent stoichiometry are also indicated in plots C and D as filled red squares and circles, respectively. Panels E and F show results when magnesium is added in varying concentrations (filled red squares or circles, 20 μM, 1 mM, and 2 mM) to the RNA in 150 mM NaCl, and then L7Ae is titrated with stoichiometric amounts (5 μM protein and 5 μM RNA).

obtained (Figure 10A,B). The magnesium-induced structural transitions were saturated at concentrations greater than 1 mM $MgCl_2$ for both RNA halfmers, although the C′/D′ motif appears to saturate at lower concentrations. Overall, this result is in good agreement with findings obtained with KT-7, where the noncooperative magnesium-induced structural transition midpoint occurred at 1 mM Mg^{2+} (*39*). Full wavelength spectra of all panels in Figure 10 are included in the Supporting Information (Figure 4).

Changes in molar ellipticity are different when either the box C/D or box C′/D′ constructs are sequentially titrated with L7Ae (from 1 to 5 μM) in low-ionic strength buffer solutions without Mg^{2+} present [20 mM NaH_2PO_4 (pH 7.0) and 20 mM NaCl]. In these experiments, the CD contributions from the L7Ae protein have been subtracted, revealing only spectral changes resulting from the RNA motifs. The box C/D halfmer has a relatively constant molar ellipticity at 267 nm, but it steadily decreases at 221 nm (Figure 10C). In contrast, the box C′/D′ halfmer exhibits a steady increase in molar ellipticity at 272 nm with a decrease in ellipticity at 222 nm (Figure 10D). When magnesium is supplemented in the samples at a concentration of 1 mM, significant changes occur [last two (red) points in panels C and D of Figure 10]. Both box C/D halfmers exhibit positive ellipticity increases in the 270 nm regions of their spectra, but have modest changes around 220 nm. These results indicate that L7Ae can induce structural changes in both box C/D halfmers, but they require Mg^{2+} (or other divalent cations)

for favorable transition of the RNA. The effects of magnesium are more pronounced for the C/D halfmer, suggesting that this motif may require divalents more than its box C′/D′ counterpart.

When similar titration experiments are performed in physiological-like buffers containing 150 mM NaCl and 2 mM $MgCl_2$, the results differ drastically. In these experiments, the RNA was equilibrated in the buffer, and then magnesium was added stepwise to a final concentration of 2 mM [the first three (red) points in panels E and F of Figure 10], followed by titration of L7Ae from 1 to 5 mM. The box C/D halfmer had a significant increase in molar ellipticity at 267 nm when magnesium was added but exhibited only modest increases in ellipticity when titrated with L7Ae (Figure 10E). The results agree with the previous results (Figure 10C) where L7Ae produced only modest effects in the absence of magnesium, but more pronounced effects in its presence. The results with the box C′/D′ halfmer were similar: when magnesium was added to the RNA sample, ellipticity decreased modestly at 271 nm and produced virtually no change at 221 nm (Figure 10F). However, when L7Ae was titrated into the box C′/D′ sample, both wavelengths showed ellipticity changes corresponding to the additions; ellipticity at 271 nm increased, and that at 221 nm decreased. These results virtually mirrored those in Figure 10D, where magnesium was added to the sample after addition of L7Ae.

Collectively, these experiments demonstrate that L7Ae and magnesium can both induce structural changes in the box C/D and C′/D′ halfmer motifs. The dependence on magnesium is more pronounced with the box C/D halfmer than with the box C′/D′ halfmer. This disparity may reflect the structural and topological differences between the two constructs. The box C′/D′ motif contains a five- or six-nucleotide loop instead of stem I, which may be more flexible in transitioning into a K-turn motif and not require magnesium to the extent that the box C/D does. Studies with the wild-type sR8 molecule may provide more insight into the cation requirements and L7Ae-induced conformation changes occurring upon binding.

CONCLUSIONS

The high-resolution structure of *Mj* L7Ae has permitted detailed structural analysis of the free and RNA-bound protein. The structure of the free protein is remarkably similar to that bound to RNA, although several amino acids change conformation and the backbone slightly shifts when the RNA−protein complex is formed. Our detailed analyses have also shown that the L7Ae fold is highly conserved, although subtle amino acid changes and secondary structural regions distinguish the different homologues in this diverse protein family. Most importantly, we have shown that L7Ae induces conformational changes in box C/D sRNA motifs and "locks" the RNA in the K-turn conformation. K-Turn motifs are often dimorphic, existing in both a tightly kinked turn and a loosely bent form, depending on the concentration of divalent metal ions (*39*). We propose that L7Ae binds the loosely bent form and stabilizes the tightly kinked form in the presence of magnesium or other divalent ions. Conformational changes occurring upon binding are often observed in protein−RNA interactions (*80*). However, it is interesting that archaeal L7Ae homologues apparently utilize protein-induced RNA folding, or RNA structure stabilization, while yeast L30 interacts through a mutually induced-fit mode of RNA binding, where both the protein and RNA change conformations. Further studies with additional eukaryotic L7Ae homologues are necessary to determine if this is a general trend.

L7Ae is a multifunctional protein having at least three roles in archaeal organisms; these functions include being a core protein for both the box C/D and box H/ACA RNP complexes as well as a protein of the large 50S ribosome subunit. Archaeal L7Ae and its homologues have a strikingly wide range of RNA ligands with their binding specificity primarily being determined by the ability of the bound RNA to fold into the K-turn motif. A large number of diverse RNA ligands which contain K-turn motifs have been identified either by immunoaffinity chromatography of L7Ae−RNA complexes (*85*) or by direct cloning of cDNAs (*86*) from *Sulfolobus solfataricus*. Many of these noncoding sRNAs were sense-strand, antisense, or from intergenic regions and thus may have regulatory roles. Consequently, L7Ae or its homologues may be involved in such activities. In contrast, eukaryotic L7Ae homologues appear to have become more specialized in both functionality and RNA target recognition. For example, the 15.5kD/Snu13p protein appears to exclusively bind box C/D sn(o)RNAs and does not interact with the C′/D′ motif; Nhp2p only interacts with box H/ACA snoRNAs, and a variety of homologous ribosomal proteins exist with different rRNA K-turn binding sites (e.g., L7a, L30, S12, and HS6). Collectively, these findings suggest that the L7Ae protein fold and its K-turn RNA ligands are ancient structural motifs which have been maintained and have evolved into specialized roles in a variety of RNA-related biological processes in both Archaea and Eucarya.

NOTE ADDED IN PROOF

A fluorescence resonance energy transfer (FRET) study by the Lilley group has independently corroborated that L7Ae induces conformational changes when binding to kink-turn RNAs (*87*).

ACKNOWLEDGMENT

We thank Matthew G. Fulton, Zhaoli Zhou, Eric C. Jewett, David V. Lieberman (Wake Forest University), Keith T. Gagnon, and Xinxin Zhang (North Carolina State University) for experimental assistance and productive discussions. High-resolution crystallographic data were collected on beamline X26c at the National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the U.S. Department of Energy, Division of Materials Sciences and Division of Chemical Sciences, under Contract DE-AC02-98CH10886. Additional thanks to Ulrich Bierbach (Wake Forest University), Traci Hall, Jeffery Vargason (National Institute of Environmental and Health Sciences, National Institutes of Health, Research Triangle Park, NC), and Anthony Russell (Dalhousie University, Halifax, NS) for comments on the manuscript.

SUPPORTING INFORMATION AVAILABLE

A description of protein crystallization, data collection and processing, determination and refinement, and requisite

statistics of the original *Mj* L7Ae structure at 1.86 Å used at the molecular replacement model for the high-resolution structure. Two tables and four supplemental figures referenced in the text are also included. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Maxwell, E. S., and Fournier, M. J. (1995) The small nucleolar RNAs, *Annu. Rev. Biochem. 64*, 897−934.
2. Kiss, T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs, *EMBO J. 20*, 3617−3622.
3. Kiss, T. (2002) Small nucleolar RNAs: An abundant group of noncoding RNAs with diverse cellular functions, *Cell 109*, 145−148.
4. Bachellerie, J. P., Cavaille, J., and Huttenhofer, A. (2002) The expanding snoRNA world, *Biochimie 84*, 775−790.
5. Terns, M. P., and Terns, R. M. (2002) Small nucleolar RNAs: Versatile trans-acting molecules of ancient evolutionary origin, *Gene Expression 10*, 17−39.
6. Decatur, W. A., and Fournier, M. J. (2002) rRNA modifications and ribosome function, *Trends Biochem. Sci. 27*, 344−351.
7. King, T. H., Liu, B., McCully, R. R., and Fournier, M. J. (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center, *Mol. Cell 11*, 425−435.
8. Henras, A. K., Dez, C., and Henry, Y. (2004) RNA structure and function in C/D and H/ACA s(no)RNPs, *Curr. Opin. Struct. Biol. 14*, 335−343.
9. Tran, E., Brown, J., and Maxwell, E. S. (2004) Evolutionary origins of the RNA-guided nucleotide-modification complexes: From the primitive translation apparatus? *Trends Biochem. Sci. 29*, 343−350.
10. Maden, B. E. (1990) The numerous modified nucleotides in eukaryotic ribosomal RNA, *Prog. Nucleic Acids Res. 39*, 241−303.
11. Gaspin, C., Cavaille, J., Erauso, G., and Bachellerie, J. P. (2000) Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: Lessons from the *Pyrococcus* genomes, *J. Mol. Biol. 297*, 895−906.
12. Omer, A. D., Lowe, T. M., Russell, A. G., Ebhardt, H., Eddy, S. R., and Dennis, P. P. (2000) Homologs of small nucleolar RNAs in Archaea, *Science 288*, 517−522.
13. Dennis, P. P., Omer, A., and Lowe, T. (2001) A guided tour: Small RNA function in Archaea, *Mol. Microbiol. 40*, 509−519.
14. Kuhn, J. F., Tran, E. J., and Maxwell, E. S. (2002) Archaeal ribosomal protein L7 is a functional homolog of the eukaryotic 15.5kD/Snu13p snoRNP core protein, *Nucleic Acids Res. 30*, 931−941.
15. Tollervey, D. (1996) Trans-acting factors in ribosome synthesis, *Exp. Cell Res. 229*, 226−232.
16. Kiss-Laszlo, Z., Henry, Y., and Kiss, T. (1998) Sequence and structural elements of methylation guide snoRNAs essential for site-specific ribose methylation of pre-rRNA, *EMBO J. 17*, 797−807.
17. Tycowski, K. T., You, Z. H., Graham, P. J., and Steitz, J. A. (1998) Modification of U6 spliceosomal RNA is guided by other small RNAs, *Mol. Cell 2*, 629−638.
18. Jady, B., and Kiss, T. (2001) A small nucleolar guide RNA functions both in 2′-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA, *EMBO J. 20*, 541−551.
19. Tyc, K., and Steitz, J. A. (1989) U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus, *EMBO J. 8*, 3113−3119.
20. Caffarelli, E., Fatica, A., Prislei, S., De Gregorio, E., Fragapane, P., and Bozzoni, I. (1996) Processing of the intron-encoded U16 and U18 snoRNAs: The conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA, *EMBO J. 15*, 1121−1131.
21. Cavaille, J., Nicoloso, M., and Bachellerie, J. P. (1996) Targeted ribose methylation of RNA in vivo directed by tailored antisense RNA guides, *Nature 383*, 732−735.
22. Watkins, N. J., Leverette, R. D., Xia, L., Andrews, M. T., and Maxwell, E. S. (1996) Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D, *RNA 2*, 118−133.
23. Watkins, N. J., Segault, V., Charpentier, B., Nottrott, S., Fabrizio, P., Bachi, A., Wilm, M., Rosbash, M., Branlant, C., and Lührmann, R. (2000) A common core RNP structure shared between the small nucleolar box C/D RNPs and the spliceosomal U4 snRNP, *Cell 103*, 457−466.
24. Kiss-Laszlo, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996) Site-specific ribose methylation of preribosomal RNA: A novel function for small nucleolar RNAs, *Cell 85*, 1077−1088.
25. Bachellerie, J. P., and Cavaille, J. (1997) Guiding ribose methylation of rRNA, *Trends Biochem. Sci. 22*, 257−261.
26. Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2001) The kink-turn: A new RNA secondary structure motif, *EMBO J. 20*, 4214−4221.
27. Mao, H., White, S. A., and Williamson, J. R. (1999) A novel loop-loop recognition motif in the yeast ribosomal protein L30 autoregulatory RNA complex, *Nat. Struct. Biol. 6*, 1139−1147.
28. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution, *Science 289*, 905−920.
29. Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath, A. (2000) Structure of functionally activated small ribosomal subunit at 3.3 Å resolution, *Cell 102*, 615−623.
30. Vidovic, I., Nottrott, S., Hartmuth, K., Lührmann, R., and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment, *Mol. Cell 6*, 1331−1342.
31. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit, *Nature 407*, 327−339.
32. Winkler, W. C., Grundy, F. J., Murphy, B. A., and Henkin, T. M. (2001) The G•A motif: An RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs, *RNA 7*, 1165−1172.
33. Allmang, C., Carbon, P., and Krol, A. (2002) The SBP2 and 15.5 kD/Snu13p proteins share the same RNA binding domain: Identification of SBP2 amino acids important to SECIS RNA binding, *RNA 8*, 1308−1318.
34. Rozhdestvensky, T. S., Tang, T. H., Tchirkova, I. V., Brosius, J., Bachellerie, J. P., and Huttenhofer, A. (2003) Binding of L7Ae protein to the K-turn of archaeal snoRNAs: A shared RNA binding motif for C/D and H/ACA box snoRNAs in Archaea, *Nucleic Acids Res. 31*, 869−877.
35. Hamma, T., and Ferré-D'Amaré, A. R. (2004) Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 Å resolution, *Structure 12*, 893−903.
36. Moore, T., Zhang, Y., Fenley, M. O., and Li, H. (2004) Molecular basis of box C/D RNA-protein interactions: Cocrystal structure of archaeal L7Ae and a box C/D RNA, *Structure 12*, 807−818.
37. Strobel, S. A., Adams, P. L., Stahley, M. R., and Wang, J. (2004) RNA kink turns to the left and to the right, *RNA 10*, 1852−1854.
38. Matsumura, S., Ikawa, Y., and Inoue, T. (2003) Biochemical characterization of the kink-turn RNA motif, *Nucleic Acids Res. 31*, 5544−5551.
39. Goody, T. A., Melcher, S. E., Norman, D. G., and Lilley, D. M. (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependent, *RNA 10*, 254−264.
40. Cahill, N. M., Friend, K., Speckmann, W., Li, Z. H., Terns, R. M., Terns, M. P., and Steitz, J. A. (2002) Site-specific cross-linking analyses reveal an asymmetric protein distribution for a box C/D snoRNP, *EMBO J. 21*, 3816−3828.
41. Szewczak, L. B., DeGregorio, S. J., Strobel, S. A., and Steitz, J. A. (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP, *Chem. Biol. 9*, 1095−1107.
42. Tran, E. J., Zhang, X., and Maxwell, E. S. (2003) Efficient RNA 2′-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C′/D′ RNPs, *EMBO J. 22*, 3930−3940.
43. Omer, A. D., Ziesche, S., Ebhardt, H., and Dennis, P. P. (2002) In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex, *Proc. Natl. Acad. Sci. U.S.A. 99*, 5289−5294.
44. Koonin, E. V., Bork, P., and Sander, C. (1994) A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res. 22*, 2166−2167.

45. Chavatte, L., Brown, B. A., and Driscoll, D. M. (2005) Ribosomal protein L30 is a component of the UGA-selenocysteine recoding machinery in eukaryotes, *Nat. Struct. Mol. Biol. 12*, 408−416.
46. Klein, D. J., Moore, P. B., and Steitz, T. A. (2004) The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit, *J. Mol. Biol. 340*, 141−177.
47. Charron, C., Manival, X., Clery, A., Senty-Segault, V., Charpentier, B., Marmier-Gourrier, N., Branlant, C., and Aubry, A. (2004) The archaeal sRNA binding protein L7Ae has a 3D structure very similar to that of its eukaryal counterpart while having a broader RNA-binding specificity, *J. Mol. Biol. 342*, 757−773.
48. Mao, H., and Williamson, J. R. (1999) Local folding coupled to RNA binding in the yeast ribosomal protein L30, *J. Mol. Biol. 292*, 345−359.
49. Chao, J. A., Prasad, G. S., White, S. A., Stout, C. D., and Williamson, J. R. (2003) Inherent protein structural flexibility at the RNA-binding interface of L30e, *J. Mol. Biol. 326*, 999−1004.
50. Chen, Y. W., Bycroft, M., and Wong, K. B. (2003) Crystal structure of ribosomal protein L30e from the extreme thermophile *Thermococcus celer*: Thermal stability and RNA binding, *Biochemistry 42*, 2857−2865.
51. Wong, K. B., Lee, C. F., Chan, S. H., Leung, T. Y., Chen, Y. W., and Bycroft, M. (2003) Solution structure and thermal stability of ribosomal protein L30e from hyperthermophilic archaeon *Thermococcus celer*, *Protein Sci. 12*, 1483−1495.
52. Chao, J. A., and Williamson, J. R. (2004) Joint X-ray and NMR refinement of the yeast L30e-mRNA complex, *Structure 12*, 1165−1176.
53. Scaringe, S. A. (2000) Advanced 5′-silyl-2′-ortho ester approach to RNA oligonucleotide synthesis, *Methods Enzymol. 317*, 3−18.
54. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray data in collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
55. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) Rapid automated molecular replacement by evolutionary search, *Acta Crystallogr. D55*, 484−491.
56. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
57. McRee, D. E. (1999) XtalView/Xfit: A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
58. Collaborative Computational Project, No. 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
59. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240−255.
60. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) SFCHECK: A unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model, *Acta Crystallogr. D55*, 191−205.
61. Laskowsky, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check stereochemical quality of protein structure coordinates, *J. Appl. Crystallogr. 26*, 283−291.
62. Lovell, S. C., Davis, I. W., Arendall, W. B., III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) Structure validation by Cα geometry: ϕ, ψ, and Cβ deviation, *Proteins 50*, 437−450.
63. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577−2637.
64. McDonald, I. K., and Thornton, J. M. (1994) Satisfying hydrogen bonding potential in proteins, *J. Mol. Biol. 238*, 777−793.
65. Kleywegt, G. J. (1997) Validation of protein models from Cα coordinates alone, *J. Mol. Biol. 273*, 371−376.
66. Merritt, E. A. (1999) Expanding the model: Anisotropic displacement parameters in protein structure refinement, *Acta Crystallogr. D55*, 1109−1117.
67. Nicholls, A., and Honig, B. (1991) A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson−Boltzmann equation, *J. Comput. Chem. 12*, 435−445.
68. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
69. Stothard, P. (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences, *BioTechniques 28*, 1102−1104.
70. Glaser, F., Pupko, T., Paz, I., Bell, R. E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003) ConSurf: Identification of functional regions in proteins by surface-mapping of phylogenetic information, *Bioinformatics 19*, 163−164.
71. Gibrat, J. F., Madej, T., and Bryant, S. H. (1996) Surprising similarities in structure comparison, *Curr. Opin. Struct. Biol. 6*, 377−385.
72. Holm, L., and Sander, C. (1996) Mapping the protein universe, *Science 273*, 595−603.
73. Kraulis, P. J. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
74. Merritt, E. A., and Bacon, D. J. (1997) Raster3D: Photorealistic Molecular Graphics, *Methods Enzymol. 277*, 505−524.
75. Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves, *Biopolymers 26*, 1601−1620.
76. Puglisi, J. D., and Tinoco, I., Jr. (1989) Absorbance melting curves of RNA, *Methods Enzymol. 180*, 304−325.
77. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol. 131*, 266−280.
78. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT, *Acta Crystallogr. D55*, 247−255.
79. Szilagyi, A., and Zavodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: Results of a comprehensive survey, *Structure 8*, 493−504.
80. Williamson, J. R. (2000) Induced fit in RNA-protein recognition, *Nat. Struct. Biol. 7*, 834−837.
81. Song, H., Mugnier, P., Das, A. K., Webb, H. M., Evans, D. R., Tuite, M. F., Hemmings, B. A., and Barford, D. (2000) The crystal structure of human eukaryotic release factor eRF1: Mechanism of stop codon recognition and peptidyl-tRNA hydrolysis, *Cell 100*, 311−321.
82. Nureki, O., Shirouzu, M., Hashimoto, K., Ishitani, R., Terada, T., Tamakoshi, M., Oshima, T., Chijimatsu, M., Takio, K., Vassylyev, D. G., Shibata, T., Inoue, Y., Kuramitsu, S., and Yokoyama, S. (2002) An enzyme with a deep trefoil knot for the active-site architecture, *Acta Crystallogr. D58*, 1129−1137.
83. Michel, G., Sauve, V., Larocque, R., Li, Y., Matte, A., and Cygler, M. (2002) The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot, *Structure 10*, 1303−1315.
84. Kim, M.-S., and Oh, B.-H. (2003) RCSB Protein Data Bank 1OGD.
85. Zago, M. A., Dennis, P. P., and Omer, A. D. (2005) The expanding world of small RNAs in the hyperthermophilic archaeon *Sulfolobus solfataricus*, *Mol. Microbiol. 55*, 1812−1828.
86. Tang, T. H., Polacek, N., Zywicki, M., Huber, H., Brugger, K., Garrett, R., Bachellerie, J. P., and Huttenhofer, A. (2005) Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus*, *Mol. Microbiol. 55*, 469−481.
87. Turner, B., Melcher, S. E., Wilson, T. J., Norman, D. G., and Lilley, D. M. J. (2005) Induced fit of RNA on binding the L7Ae protein to the kink-turn motif. *RNA 11* (in press).

BI050568Q