A Role for Aromatic Amino Acids in the Binding of *Xenopus* Ribosomal Protein L5 to 5S rRNA†

Jonathan P. DiNitto and Paul W. Huber*

Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556

Received July 10, 2001; Revised Manuscript Received August 22, 2001

ABSTRACT: The formation of the *Xenopus* L5−5S rRNA complex depends on nonelectrostatic interactions. Fluorescence assays with 1-anilino-8-naphthalenesulfonate demonstrate that a hydrophobic region on L5 becomes exposed upon removal of bound 5S rRNA by treatment with ribonucleases. Several conserved aromatic amino acids, mostly tyrosines, were identified by comparative sequence analysis and changed individually to alanine. Substitution with alanine at any of three positions, Y86, Y99, or Y226, essentially abolishes RNA-binding activity, whereas those made at Y95 and Y207 have more modest effects. Replacement with phenylalanine at Y86 and Y226 does not change binding affinity, indicating that the aromatic ring of the side chain, not the hydroxyl group, is the critical functionality for binding. Alternatively, the phenolic hydroxyls at Y99 and Y207 do contribute to binding. The structural integrity of the mutant proteins was assessed using thermal denaturation and limited digestion with proteases. The T_m of Y99A is 10 °C lower than that of the wild-type protein, and there are some differences in the protease digestion patterns that together indicate the structure of this mutant has been significantly perturbed. The structures of the other variants are not detectably different from the wild-type protein. These results provide evidence that intermolecular stacking interactions involving at least two tyrosine residues, Y86 and Y226, are necessary for formation of the L5−5S rRNA complex and can account, at least in part, for the contribution nonelectrostatic interactions make to the free energy of binding.

L5 is the principal ribosomal protein associated with 5S rRNA in eukaryotic ribosomes. This complex appears to be added to the ribosome intact and provides a controlled means by which 5S rRNA is delivered to the nucleolus (*1*). During the early stages of oogenesis in *Xenopus*, 5S rRNA is stored in the cytoplasm bound to either transcription factor IIIA (TFIIIA) or the related zinc finger protein, p43 (*2*−*5*). During vitellogenesis, much of this pool of 5S rRNA becomes associated with L5, which is necessary for transport of the RNA into the nucleus for ribosome assembly (*6*−*8*). Variants of 5S rRNA that cannot bind to L5 remain localized in the cytoplasm (*9*).

There is no identifiable RNA-binding motif in L5. Assays with deletion mutants have revealed that nearly the entire length of L5 is required for high-affinity binding to 5S rRNA (*10*), which is reflected in the unusually large footprint of the protein on the RNA (*11*). Thus, the interaction of L5 with 5S rRNA may be complex and involve a particularly large surface area on each component. The amino acid composition of L5 is typical of ribosomal proteins with a preponderance (20%) of basic amino acids that presumably contribute to nucleic acid binding through ionic and hydrogen-bonding interactions. However, the association of *Xenopus* L5 with 5S rRNA is remarkably insensitive to ionic strength, indicating that nonelectrostatic interactions make a significant contribution to the free energy of binding (*12*). It is notable, then, that L5 also contains a large number of aromatic amino acids, including 19 tyrosines.

The major identity elements for L5 are located in the hairpin comprised of helix III and loop C (*12*). Helix III contains a 2-nucleotide bulge, and the 12-nucleotide loop C has several non-Watson−Crick base pairs. Both the bulge and the loop are targeted by the metal complex $Rh(phen)_2(phi)^{3+}$,[1] which binds to sites on RNA where the major groove is opened and accessible to stacking interactions with the phenanthrenequinone diimine ligand (*13*). We observed a correlation between intercalative binding of $Rh(phen)_2(phi)^{3+}$ and binding of L5 to mutants of 5S rRNA and proposed that stacking interactions play a critical role in the formation of the L5−5S rRNA complex (*12*). We have now tested the role of several conserved aromatic amino acids in the binding of L5 to 5S rRNA and have identified three essential residues: Y86, Y99, and Y226. Whereas an alanine substitution at Y99 appears to affect the stability of the protein, mutations at Y86 and Y226 do not, suggesting that the latter fail to bind to 5S rRNA because important contacts to the RNA have been eliminated.

EXPERIMENTAL PROCEDURES

Plasmids and Nucleic Acids. A cDNA clone (L5b) of the gene encoding *Xenopus laevis* ribosomal protein L5 (*14*) was

† Supported by a grant (GM38200) from the National Institutes of Health.

* Address correspondence to this author at the Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556. Phone: 219-631-6042. FAX: 219-631-6652. E-mail: huber.1@nd.edu.

[1] Abbreviations: $Rh(phen)_2(phi)^{3+}$, bis(phenanthroline)(phenanthrenequinone diimine)rhodium(III); ANS, 1-anilino-8-naphthalenesulfonate; MBP, maltose-binding protein; fL5, the chimeric protein composed of MBP fused to L5; CD, circular dichroism; T_m, melting temperature.

10.1021/bi011439m CCC: $20.00
Published on Web 09/21/2001

amplified by the polymerase chain reaction (PCR) and inserted into the vector plasmid, pET23b(+) (Novagen), at the *Bam*HI and *Xho*I restriction sites. The resulting plasmid, pET23-L5b, encodes the L5 protein preceded by a 14-amino acid T7 epitope tag and followed by the sequence $LE(H)_6$ at the carboxyl terminus. The QuikChange kit (Stratagene) was used for site-directed mutagenesis. Native *Xenopus* oocyte 5S rRNA was prepared from 7S RNP particles isolated from immature oocytes (*15*). The expression and purification of the maltose-binding protein−L5 fusion (fL5) and the synthesis of internally radiolabeled 5S rRNA by runoff transcription using T7 RNA polymerase and [α-^{32}P]CTP followed procedures described previously (*12*).

RNA-Binding Assays. Binding of L5 to 5S rRNA was measured using mobility shift assays (*12*). To compare the binding affinity of L5 mutants, it was necessary to use a method that is independent of the amount of active protein in any given preparation. Thus, the data from the binding assays were processed as Woolf plots with lines fit by least-squares linear regression analysis (*16*). Binding reactions (10 μL) contained 1 nM internally radiolabeled 5S rRNA, 40 nM fL5, and an increasing amount of unlabeled 5S rRNA ranging from 0 to 60 nM. Binding reactions were kept at room temperature for 4 h to ensure equilibrium was reached. Samples were then loaded on 8% polyacrylamide gels; electrophoresis was run at 120 V. Laser densitometry was used to quantitate free and bound RNA on the resulting autoradiographs. Protein concentrations were determined by the method of Bradford using bovine serum albumin as standard (*17*). The concentration of RNA samples was determined spectrophotometrically at 260 nm using an extinction coefficient of 22.2 $(mg/mL)^{-1}$ cm^{-1}.

Fluorescence Spectroscopy. The L5−5S rRNA complex was formed using 4 μM L5 and 5 μM 5S rRNA in a buffer containing 20 mM HEPES, pH 8.0, 30 mM NH_4Cl, 200 mM KCl, 0.5 mM $MgCl_2$, 4% glycerol, 2 mM DTT. This buffer differs from the standard binding buffer in that BSA and NP-40 were omitted and the KCl concentration was increased from 100 mM. All samples were analyzed on mobility shift gels (stained with silver) to ensure that the RNP complex had formed and that the fluorescent probe ANS does not cause any dissociation. After formation of the RNP complex, ANS was added (50 μM), and spectra were acquired using an excitation wavelength of 355 nm. To degrade the 5S rRNA in the complex, RNase A (1:52 wt:wt ratio RNase A:5S rRNA) and RNase T1 (0.076 unit/μL) were added, and emission spectra were taken at increasing time intervals. Separate assays established that these concentrations of ribonuclease do not contribute to the fluorescence emission. Spectra were recorded at 21 °C using an SLM Aminco model 8100 spectrofluorometer. Data were collected every 1 nm with a scan rate of 60 nm/min. Fluorescence emission is plotted as relative fluorescence (F_S/F_R) where F_S and F_R are the intensities of the sample and the internal reference, respectively. The emission spectra were not smoothed.

Circular Dichroism Spectroscopy. Samples were dialyzed against buffer containing 10 mM sodium phosphate (pH 7.5) and 350 mM NaCl. Ellipticity was measured on an AVIV model 62DS/202 series spectropolarimeter. The thermal denaturation spectra were recorded at 222 nm from 20 to 90 °C with an equilibration time of 1 min/°C. Data were acquired using a 1 nm bandwidth and a 1 min signal averaging time. The T_m for each protein sample was taken as the maximum of the first derivative of the melting curve using the program Table Curve 2D (Jandel Scientific).

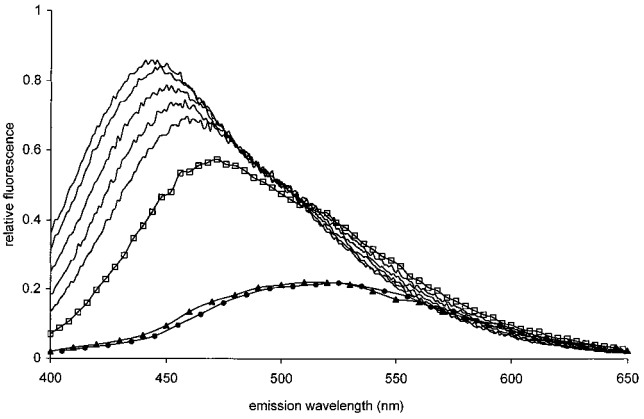

FIGURE 1: Fluorescence emission spectra of ANS in the presence of the L5−5S rRNA complex. The complex was formed by incubating 4 μM L5 and 5 μM 5S rRNA. ANS was added to a final concentration of 50 μM, and the emission spectrum of the intact complex (□) was measured using an excitation wavelength of 355 nm. Following the addition of ribonucleases A and T1, spectra were measured at 20, 30, 40, 50, and 60 min (curves with increasing emission, respectively). The spectra of ANS alone (●) or in the presence of 5 μM 5S rRNA (▲) are also shown.

Proteolytic Digests. Samples of wild-type and variant fL5 were digested with protease at the following mass ratios: V8, 1:24; trypsin and chymotrypsin, 1:100. Each sample of fL5 was diluted with 10 mM Tris-HCl (pH 7.5) to a final concentration of 0.57 μg/μL, protease was added, and the digestions were run at room temperature for the following times: trypsin, 45 min; V8 and chymotrypsin, 1 h. Reactions were quenched by adding SDS loading solution and incubating the samples at 100 °C for 10 min prior to loading on SDS−polyacrylamide gels.

RESULTS

Exposure of Hydrophobic Residues on the Surface of L5. The binding of L5 to 5S rRNA is notably insensitive to ionic strength (*12*). An increase in monovalent cation concentration from 0.1 to 0.6 M KCl results in only a 2-fold decrease in binding affinity. Thus, it appears that nonelectrostatic interactions make an important contribution to the binding free energy. Despite the large number of basic amino acids in L5, there is also a considerable number of nonpolar and aromatic amino acids. The fluorescence emission of 1-anilino-8-naphthalenesulfonate (ANS) is markedly increased in hydrophobic environments with a concomitant shift of the emission maximum to shorter wavelengths. We used this external fluorescent probe to detect the exposure of hydrophobic sites on L5 upon removal of bound 5S rRNA by treatment with ribonucleases (Figure 1).

The L5−5S rRNA complex itself increases the fluorescence emission intensity of ANS by approximately 3-fold, which indicates that hydrophobic sites on the protein are accessible to the probe even when 5S rRNA is bound. Mobility shift assays of the samples used in these experiments confirmed that most of the protein was bound to 5S rRNA (data not shown). Thus, the observed increase in ANS fluorescence cannot be attributed to free L5. Upon addition of ribonucleases A and T1, there is a progressive increase

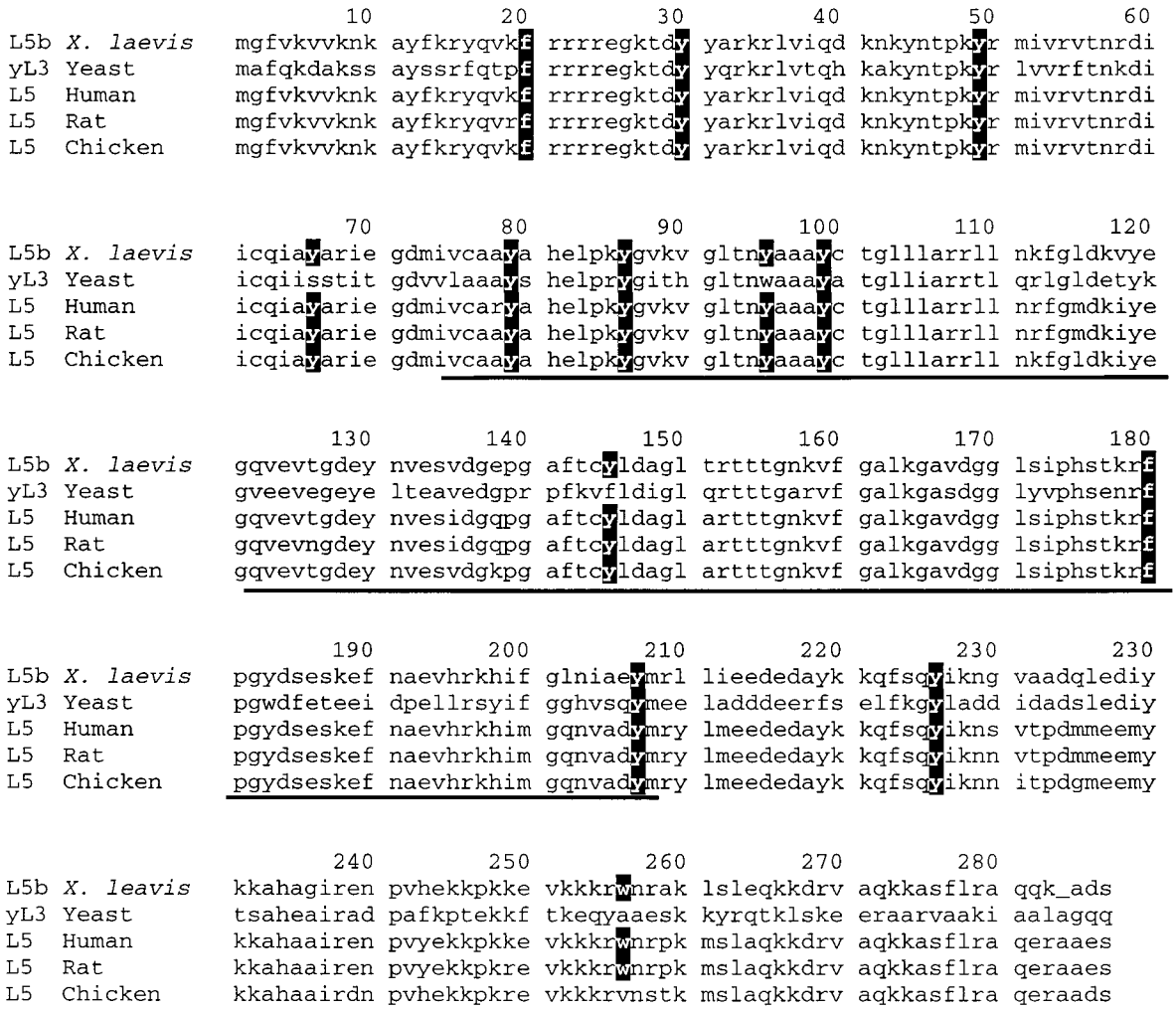

FIGURE 2: Alignment of selected eukaryotic L5 sequences. The positions designated with background are conserved aromatic residues that were selected for alanine mutagenesis. The line indicates the cyanogen bromide fragment of L5 that contains at least two tyrosines that can be iodinated in the native protein.

in fluorescence emission and the expected shift of the emission maximum to shorter wavelengths, which indicate that ANS is binding to hydrophobic sites that become exposed as the 5S rRNA is eliminated by digestion. These results indicate that hydrophobic regions on L5 are proximate to 5S rRNA and could be involved in direct contacts with the nucleic acid.

Role of Conserved Aromatic Amino Acids in Binding 5S rRNA. The major identity elements for L5 are located in the helix III−loop C hairpin, although the protein interacts with additional parts of the RNA (*11*, *12*). There are at least two sites within the hairpin that are targeted by the intercalative probe, $Rh(phen)_2(phi)^{3+}$, establishing that this region of 5S rRNA can participate in stacking interactions with external ligands (*13*). The solution structure of the two-nucleotide bulge in helix III has been solved by NMR spectroscopy (*18*). The two unpaired adenosines are stacked upon each other in the minor groove, which results in a widened major groove exposing the base of G48. Intermolecular stacking interactions, therefore, may represent the principal nonelectrostatic contacts that contribute to the free energy of binding in the complex. Additionally, the pH profile of the apparent association constant of the complex exhibits an ionization with a pK_a of 9.1, which could represent the ionization of a tyrosine residue (*12*). Iodination of L5 indicates that at least two tyrosine residues, located in the cyanogen bromide fragment that extends from amino acids 74 to 208, are accessible to solvent (S. S. Ganser and P. W. Huber, unpublished results). We have now directly tested the importance of aromatic residues in the formation of the L5−5S rRNA complex using site-directed mutagenesis.

Alignment of the amino acid sequences of L5 from several eukaryotes reveals the positions of several conserved aromatic residues (Figure 2). In this initial round of mutagenesis, we chose to focus on these positions and tyrosine residues in particular. The selected residues were mutated to alanine, and binding of the mutant protein to 5S rRNA was measured using mobility shift assays (Figure 3). For these assays, the variants of L5 were expressed as a fusion protein with maltose-binding protein (MBP). We have shown that the MBP domain does not influence the RNA-binding activity of L5, but that this domain does suppress aggregation of the protein (*12*).

To compare the L5 mutants directly with each other, it was necessary to measure binding affinity using a method that is independent of the percent of active protein in each sample. Binding assays contained a constant amount of L5 and radiolabeled 5S rRNA and increasing amounts of unlabeled 5S rRNA. Autoradiographs of the mobility shift gels were scanned with a laser densitometer to quantitate

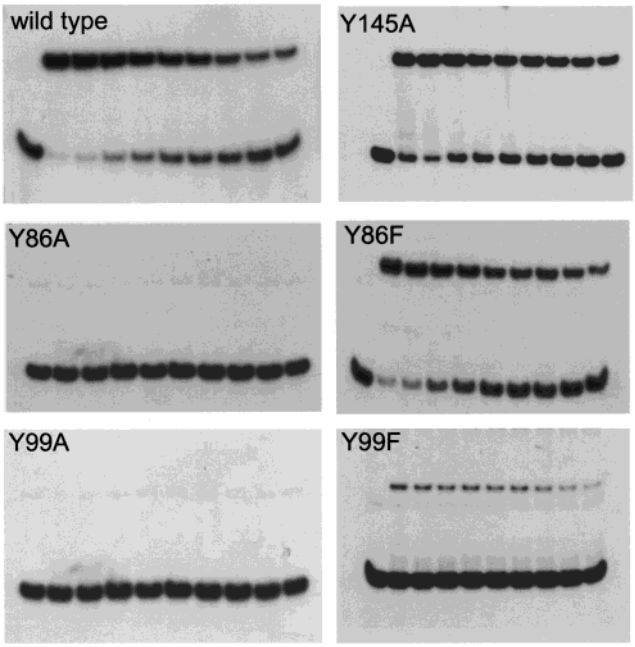

FIGURE 3: Mobility shift assays for binding of fL5 mutants to 5S rRNA. Autoradiographs for a selection of mutants are presented. In each assay, 1 nM internally labeled 5S rRNA was incubated with 40 nM fL5. The first lane in each case is 5S rRNA alone. The subsequent lanes contain 0, 4.4, 9.8, 13, 18, 25, 33, 45, and 60 nM unlabeled 5S rRNA.

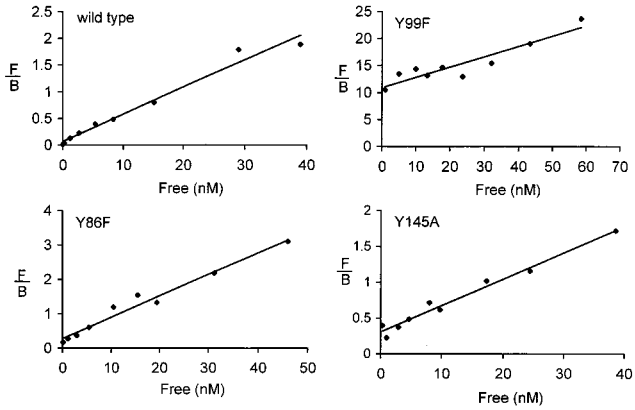

FIGURE 4: Woolf plots derived from RNA mobility shift assays. The autoradiographs of the mobility shift gels were scanned with a laser densitometer to quantitate the amounts of free and bound RNA. Exposures were in the linear response range of the film. The ratio of free to bound RNA is plotted as a function of the concentration of free RNA. Curves were fit to the data by least-squares regression analysis. The value of K_d is taken from the *x*-intercept.

the amount of free and bound 5S rRNA and the data used to construct Woolf plots (Figure 4), which have certain statistical advantages over Scatchard plots (*16, 19*). The K_d of the complex is obtained from the intercept on the abscissa of these plots. The K_d of the wild-type protein, determined from the Woolf plot (2 nM), is identical to that measured previously by direct titration of the RNA with increasing concentrations of L5.

Substitution with alanine at Y86, Y99, or Y226 essentially eliminates binding of L5 to 5S rRNA within the range of this assay (Table 1). Substitutions at Y95 and Y207 also have some effect, exhibiting 6.6- and 10-fold decreases in affinity, respectively, relative to wild-type protein. Notably, substitution of the sole tryptophan by alanine has an insignificant impact on binding affinity.

Table 1: Dissociation Constants for the Binding of fL5 Mutants to 5S rRNA

mutant	$K_d \pm$ SDM (nM)	determinations	$\Delta\Delta G^a$ (kcal/mol)
wild type	2.5 ± 0.87	5	
F20A	5.6 ± 3.1	2	
Y30A	5.5 ± 2.0	2	
Y49A	8.0 ± 5.6	2	
Y66A	3.0 ± 1.1	2	
Y79A	2.0 ± 1.2	4	
Y86A	nd[b]	3	
Y86F	1.9 ± 1.7	4	
Y95A	16.4 ± 4.1	2	−1.0
Y99A	nd	2	
Y99F	55.4 ± 2.1	2	−1.8
Y145A	4.6 ± 3.5	3	
F180A	0.74 ± 0.16	2	
Y207A	25.8 ± 3.6	2	−1.3
Y207C	21.6 ± 6.3	4	−1.2
Y207F	11.7 ± 0.1	2	−0.9
Y226A	nd	2	
Y226F	1.6 ± 1.1	4	
W266A	10.7 ± 3.6	3	

[a] $\Delta\Delta G = \Delta G_{\text{wild type}} - \Delta G_{\text{mutant}}$ where $\Delta G = -RT \ln(1/K_d)$. [b] nd: no detectable binding.

The importance of the phenolic hydroxyl groups on Y86, Y99, and Y226 was tested by making mutants with phenylalanine at these positions. In the case of positions 86 and 226, the phenylalanine variants exhibit wild-type binding affinity, indicating that only the aromatic ring and not the hydroxyl group is critical for function. The phenylalanine substitution at position 99 results in a 22-fold increase in the K_d of the complex. This translates to a 1.8 kcal/mol change in the free energy of binding, which is within the range of a hydrogen bond. Thus, both the aromatic ring and the hydroxyl moiety of Y99 contribute to the ability of L5 to bind to 5S rRNA. The $\Delta\Delta G$ of binding for Y207F indicates some type of hydrogen bonding involving the side chain hydroxyl group of this residue.

Thermal Stability of Mutant Proteins. In the absence of a high-resolution structure, it is impossible to know the consequences of a particular amino acid substitution. Alanine mutants at three positions in L5 have considerable effects on binding to 5S rRNA. Within the limits of the mobility shift assay, binding affinity was reduced by at least 100-fold. These individual substitutions either eliminate an important contact necessary for formation of the L5−5S rRNA complex, markedly alter the structure of L5, or both. Since aromatic residues can be especially important for packing interactions in the interior of proteins, which contribute significantly to stability, it is difficult to determine the origin of the defective RNA-binding activity. Samples of L5 at concentrations required for NMR spectroscopy were unstable and formed insoluble aggregates, which precluded analysis of mutants by this method. Consequently, we turned to thermal denaturation to assess the structural integrity of the alanine mutants that abolish the RNA-binding activity of L5. Protein unfolding was measured by circular dichroism spectroscopy at 222 nm.

MBP−L5 fusion proteins were used in these experiments, since this form of L5 is much less susceptible to aggregation in buffer lacking detergent. NP-40, which is a normal

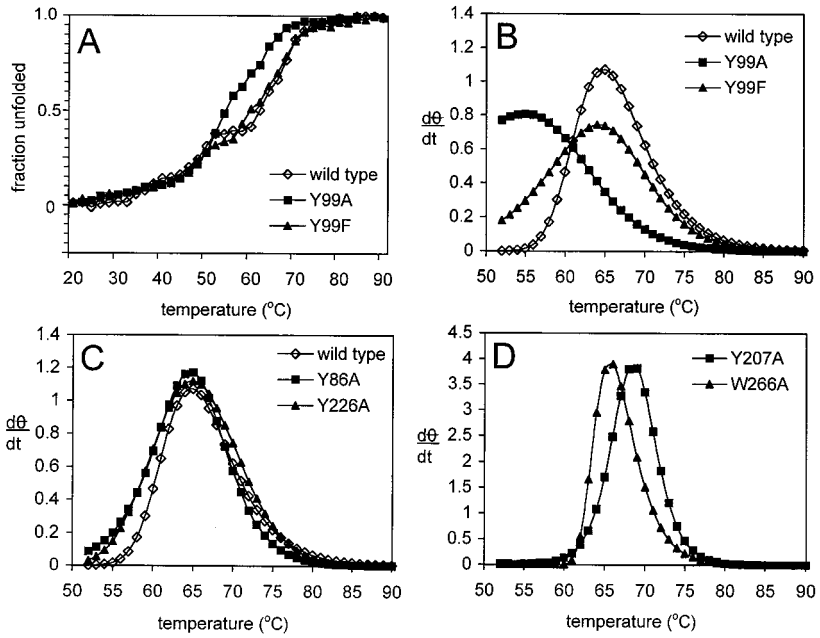

FIGURE 5: Thermal denaturation of fL5 mutants. (A) The ellipticity of each protein sample (4.5 μM) was measured at 222 nm as a function of temperature and is plotted as the fraction unfolded. The first transition corresponds to the maltose-binding domain (T_m = 50 °C) of the fusion protein and the second to L5. (B−D) Data for the second transition in the melting curves are plotted as the first derivative in order to determine the T_m for each variant of L5.

Table 2: Melting Temperatures of fL5 Mutants

fL5	T_m (°C)	determinations
wild-type	64.4 ± 0.8	2
Y86A	65.0 ± 0.6	2
Y99A	53.9 ± 1.3	4
Y99F	64.9 ± 1.0	2
Y207A	68.6 ± 0.1	2
Y226A	65.9 ± 1.2	2
W266A	65.9 ± 0.5	2

component of binding buffer, cannot be used in CD experiments due to light scattering in the UV region. The legitimacy of studying the folding of a protein within the context of a fusion containing MBP has been established in experiments using differential scanning calorimetry (*20*). The thermodynamic behavior of the two protein domains was uncoupled in all of the fusion proteins examined. Indeed, the unfolding profile of the MBP−L5 fusion protein exhibits two distinct transitions (Figure 5A). First-derivative plots of the melting curves were used to determine the T_m of wild-type and mutant variants of L5 (Figure 5B−D). The two inflection points for the wild-type protein are at 49.7 and 64.4 °C. The T_m for maltose-binding protein in these conditions is 53.8 °C, establishing that the first transition in the melting curve of the fusion protein corresponds to the MBP domain. A decrease in the T_m of MBP in the context of fusion proteins has been noted before (*20*).

Of the three mutants that lost 5S rRNA binding activity, the T_m values of Y86A and Y226A are essentially the same as wild-type L5 (65.0 and 65.9 °C, respectively), whereas Y99A is shifted by more than 10 °C to 53.9 °C (Table 2). Thus, the failure of Y99A to bind 5S rRNA likely arises from a perturbation in the structure of the protein. On the other hand, the substitutions at Y86 and Y226 appear to be structurally benign, admitting the possibility that these mutants fail to bind because the substitution with alanine eliminates an essential contact between the protein and 5S rRNA. The fact that phenylalanine can replace tyrosine at these positions without any change in binding affinity indicates that these are most likely stacking interactions involving the aromatic ring of the side chain.

The phenylalanine substitution at Y99 reduces (~20-fold), but does not eliminate, RNA binding. The T_m of Y99F (64.9 °C) is the same as wild-type protein. Since there is no difference in stability to account for the $\Delta\Delta G$ in binding free energy for this mutant, the reduced affinity of Y99F for 5S rRNA likely arises from the loss of a hydrogen bond to the RNA. Thus, the tyrosine at position 99 provides an aromatic ring that is essential for the structure of the protein and also a hydroxyl group that forms an important hydrogen bond, possibly to 5S rRNA. Substitutions at Y207 to alanine, cysteine, or phenylalanine have moderate reductions in binding affinity ranging from 5- to 10-fold, indicating that the hydroxyl group of this residue, likewise, forms a hydrogen bond to the RNA. The T_m of Y207A is actually increased about 4 °C relative to wild-type L5 (Table 2). Thus, it is also possible, in the case of this substitution, that a conformation is stabilized in the protein that has a reduced affinity for 5S rRNA.

Alanine substitution for the sole tryptophan in L5 has little effect on RNA-binding activity. In accord with this result, the T_m of this mutant protein is nearly the same as that of the wild-type protein. Fluorescence quenching experiments indicate that the tryptophan is on the surface of L5 and binding of 5S rRNA does not change its environment (J. P. DiNitto and P. W. Huber, in preparation).

Limited Protease Digestions of Mutant Proteins. Proteolytic digestions with V8, trypsin, or chymotrypsin provided additional information on the structural integrity of the L5 mutants (Figure 6). None of the digestion patterns is substantially different from the wild-type protein. Some differences can be seen in the digests of Y99A (Figure 6A, lane 11, and 6B, lane 5) that are less apparent or absent in

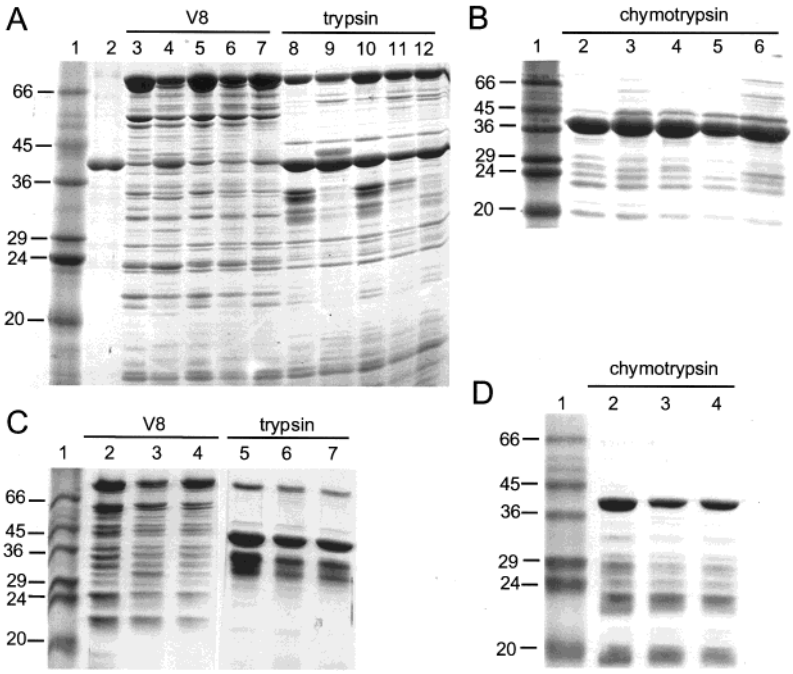

FIGURE 6: Limited protease digestions of fL5 mutants. Digestions were run as described under Experimental Procedures and analyzed by SDS−polyacrylamide gel electrophoresis. The protease is indicated above the lanes of the gels. (A) The lanes contained the following samples: lane 1, molecular weight standards of the indicated mass; lane 2, maltose-binding protein; lanes 3−7, V8 digests of wild-type, Y86A, Y86F, Y99A, Y99F, respectively; lanes 8−12, trypsin digests of wild-type, Y86A, Y86F, Y99A, Y99F, respectively. (B) Lane 1, molecular weight standards; lanes 2−6, chymotrypsin digests of wild-type, Y86A, Y86F, Y99A, Y99F, respectively. (C) Lane 1, molecular weight standards; lanes 2−4, V8 digests of wild-type, Y226A, Y226F, respectively; lanes 5−7, trypsin digests of wild-type, Y226A, Y226F, respectively. (D) Lane 1, molecular weight standards; lanes 2−4, chymotrypsin digests of wild type, Y226A, Y226F, respectively.

the digests of Y99F (Figure 6A, lane 12, and 6B, lane 6). These results are in accord with the lower stability of the Y99A mutant detected in the thermal denaturation assays. Some differences are also apparent in the trypsin digest of Y86A (Figure 6A, lane 9), but not in the digests of this mutant with either V8 or chymotrypsin. It must be noted that digestions with trypsin were significantly more variable than those with V8 and chymotrypsin for all proteins tested. The digests of both Y226 mutants are the same as wild-type L5.

Notwithstanding some differences in the trypsin digest of Y86A, these assays support the results from the thermal denaturation experiments. Apparent structural changes in the Y99A mutant make it impossible at this time to decide whether this amino acid makes an essential contact to 5S rRNA through the aromatic ring of the side chain. However, the reduced affinity of the Y99F mutant, for which there is no evidence for structural perturbations, suggests that the hydroxyl group at this position may form a hydrogen bond to the RNA. On the other hand, alanine replacements at tyrosines 86 and 226 markedly diminish RNA-binding activity without substantially changing the structure of L5, making these residues good candidates for putative stacking interactions with 5S rRNA.

DISCUSSION

Stacking interactions involving aromatic amino acids appear to be a common mode of contact with RNA. RNA, as opposed to DNA, is particularly well suited for recognition in this way, because the complex folds of the molecule, resulting from secondary and tertiary structure, can present the bases of nucleotides that are otherwise inaccessible in a canonical DNA helix. The importance of intermolecular stacking interactions is reflected in the frequency at which aromatic amino acids occur within consensus RNA-binding motifs. The best, and most studied, example is the RNP domain, which contains multiple, highly conserved, aromatic residues that are involved in stacking interactions in the structure of all RNP domain complexes determined to date (*21*−*24*). However, critical stacking interactions are seen in the structures of several other RNA−protein complexes, including anticodon interactions with tRNA synthetases (*25*, *26*), MS2 viral coat protein bound to operator RNA (*27*, *28*), λ N protein bound to *boxB* RNA (*29*), and ribosomal protein L30 bound to an autoregulatory RNA element (*30*).

The metal complex $Rh(phen)_2(phi)^{3+}$ binds to RNA on the basis of shape selection at sites in RNA where the major groove is opened. It is sterically excluded from double-stranded regions of RNA, because the major groove of an A-type helix is too narrow, nor does the probe bind to unstructured single-stranded sites that do not permit stacking interactions (*31*). $Rh(phen)_2(phi)^{3+}$, which cleaves nucleic acids upon photoactivation, targets loop E and the helix III−loop C hairpin in *Xenopus* 5S rRNA (*13*). NMR spectroscopy has been used to determine the conformation of loop E, which, indeed, has a major groove opened by a base triple structure involving a reverse Hoogsteen A·U pair (*32*). The intercalative probe also cleaves opposite the two-nucleotide bulge in helix III and at the helix III−loop C junction. These are exactly the positions, identified in mutagenesis experiments, that encompass the major identity elements for L5 (*12*). Mutations in 5S rRNA that affect binding of L5 also change the cleavage profile of $Rh(phen)_2(phi)^{3+}$ in the hairpin.

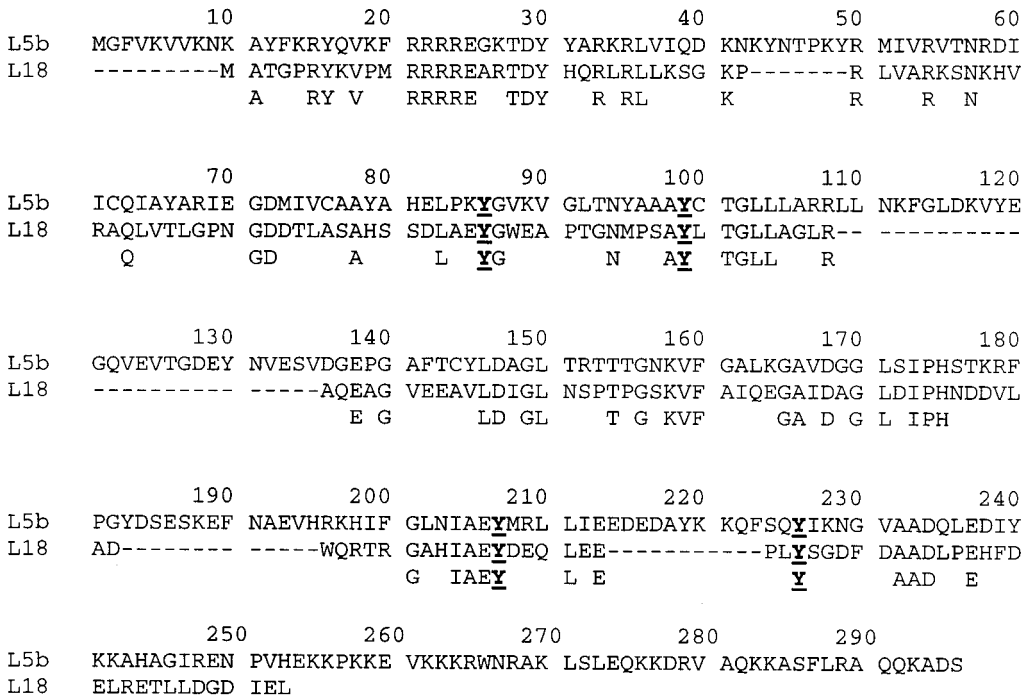

FIGURE 7: Sequence alignment of *Xenopus* L5 and *H. marismortui* L18. Residues conserved in both organisms and identified by alanine mutagenesis as essential for binding to 5S rRNA are in boldface and underlined.

The solution structure of helix III in *Xenopus* oocyte 5S rRNA has been determined by NMR spectroscopy (*18*). The two unpaired adenosines are stacked upon the 3′ guanosine in the minor groove. This configuration opens the major groove and exposes the base of the 5′ guanosine that flanks the bulge (G48), creating a site that can potentially accept an external ligand in a stacking interaction. Imino resonances in preliminary NMR experiments indicate that loop C contains non-Watson−Crick base pairs, which supports an earlier suggestion that the loop may be folded to form a GUCU tetraloop (*12*). Interestingly, a selection experiment for RNA aptamers that can bind to phenylalanine repeatedly identified a core sequence motif, CUCGUGU, that matches a sequence (positions 34−40) in loop C at 6 of 7 nucleotides (*33*). Secondary structure prediction indicates that this core sequence is part of a loop structure, which was verified in cleavage assays with lead. In addition, the sequence AUU commonly occurred at the 3′ end of the aptamer loops, which compares well with the AUC sequence at the 3′ end of loop C. Thus, loop C appears to have structural elements found in RNAs selected for binding an aromatic amino acid.

The evidence that the identity elements for L5 are sites capable of binding aromatic ligands taken together with the identification of at least two positions in L5 that require an aromatic side chain for binding suggests that intermolecular stacking interactions provide a key recognition event in the formation of the L5−5S rRNA complex. A tyrosine in the MS2 coat protein stacks on the base of C-5 in the operator RNA (*27*). Replacement of this aromatic residue with alanine results in a 3 kcal/mol change in binding free energy (*34*), which is comparable to the Y86A and Y226A mutants in L5 (minimum $\Delta\Delta G$ of ~2.5 kcal/mol). However, it has been suggested that in addition to a substantial thermodynamic contribution, stacking interactions are also important for conferring specificity, in some cases by stabilizing protein-induced conformational changes in the RNA that permit additional intermolecular contacts (*35*, *36*). In fact, *Xenopus* L5 induces a conformational change in 5S rRNA detected by CD spectroscopy (J. P. DiNitto and P. W. Huber, in preparation).

Deletion mutagenesis has been used in attempts to identify the regions in metazoan L5 required for binding to 5S rRNA (*10*, *37*−*39*). Although there are some discrepancies in these data, the consensus indicates that regions near the amino and carboxyl ends of L5 are necessary for binding to 5S rRNA. Polypeptides encompassing internal regions of L5 do not form complexes with the RNA. These results, in general, parallel those for the yeast homologue of L5, which has been more thoroughly characterized (*40*, *41*). In addition to sequences at both termini of the yeast protein, internal sequences are also required for binding to 5S rRNA. These results indicate that the organization of the RNA-binding domain of L5 is complex. The N- and C-termini of L5 contain a notable preponderance of basic amino acids, whereas the tyrosines identified in this study are located near the center of the primary sequence. The restricted location of identity elements within the much larger region of 5S rRNA contacted by L5 may reflect the organization of the protein. The nonelectrostatic contacts mediated by stacking interactions in the helix III−loop C hairpin may be supplemented by structure-specific interactions to additional sites on the RNA mediated by the basic residues at either end of L5.

The crystal structure of the large ribosomal subunit from *H. marismortui* has been determined at 2.4 Å resolution (*42*). However, inferences from the structure of the archaeal ribosome are difficult to apply to the eukaryotic L5−5S rRNA complex. Whereas there appears to be one primary protein associated with 5S rRNA in eukaryotes, there are three in *H. marismortui* (L5, L18, and L30). *H. marismortui* L18 has only 35% identity to *Xenopus* L5, and the former (186 amino acids) is considerably smaller than the latter (296

amino acids). *H. marismortui* L18, however, is situated along helix III, indicating that it engages the region of 5S rRNA that contains the principal identity elements for the eukaryotic protein.

Despite having modest sequence identity, an alignment of *Xenopus* L5 and *H. marismortui* L18 reveals a striking conservation of the three tyrosine residues (Y86, Y99, and Y226) that are necessary for binding of the former protein to 5S rRNA (Figure 7). In addition, Y207, which also has some role in binding 5S rRNA, is also conserved in *H. marismortui* L18. What is the location of these amino acids relative to 5S rRNA in the crystal structure? Since the positions of side chains are not yet available, only a general description can be given, but it is valuable nonetheless. The α-carbon of Y226 (Y160 in *H. marismortui*) is ~8 Å from the C8 position of the first adenosine (A51) comprising the two-nucleotide bulge in helix III, and the α-carbon of Y207 (Y151 in *H. marismortui*), which appears to be involved in a hydrogen bond interaction, is ~7 Å from the 2′ oxygen of G49. Thus, the α-carbons of these amino acids are near 5S rRNA in the *H. marismortui* ribosome, but it is not clear that the side chains are in contact with the nucleic acid.

The alanine substitution at Y99 shifted the T_m of the mutant protein by 10 °C, indicating a substantial destabilization of the protein that would account for the loss of RNA-binding activity. This conclusion is supported by the crystal structure, which shows that the equivalent residue (Y83) in *H. marismortui* L18 is part of a hydrophobic core.

The Y86A mutation substantially reduced binding to 5S rRNA, but had essentially no effect on the T_m of the mutant protein and just minor differences in the trypsin digestion profile. The equivalent tyrosine (Y70) is in a loop on the surface of *H. marismortui* L18; however, it is quite far away from 5S rRNA. Thus, it is not apparent in this case why this substitution has such a marked effect on RNA binding. Interestingly, a comparable situation has been reported for *B. stearothermophilus* ribosomal protein L5, which is one of two proteins that bind to 5S rRNA in this organism (*43*). The crystal structure of *B. stearothermophilus* L5 revealed that its folding topology is related to the RNP domain, which was confirmed by results from site-directed mutagenesis. The folds of *B. stearothermophilus* L5 and its homologue, *H. marismortui* L5, are similar; yet, a phenylalanine located in an exposed loop, that is important for binding of the former protein to 5S rRNA, is far removed from the RNA in the *H. marismortui* structure. It was suggested that the phenylalanine residue in *B. stearothermophilus* L5 may form an important, but transient, contact during the formation of the complex. The same may be true for Y86 in *Xenopus* L5. Alternatively and, perhaps, more likely, the disparity could simply reflect a bone fide difference between the structures of eukaryotic and archaeal ribosomes. The important lesson here is that it may be difficult to extrapolate from the archaeal subunit structure to eukaryotic ribosomes. This may be particularly true in the case of 5S rRNA where the role of three proteins appears to have been taken over by one.

ACKNOWLEDGMENT

We are grateful to Dr. Thomas Nowak for use of his spectrofluorometer, and to Dr. Michael Mossing and Dr. Mary Prorok for helpful suggestions and discussions.

REFERENCES

1. Steitz, J. A., Berg, C., Hendrick, J. P., La Branche-Chabot, H., Metspalu, A., Rinke, J., and Yario, T. (1988) *J. Cell Biol. 106*, 545−556.
2. Picard, B., and Wegnez, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 241−245.
3. Honda, B. M., and Roeder, R. G. (1980) *Cell 22*, 119−126.
4. Pelham, H. R. B., and Brown, D. D. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4170−4174.
5. Joho, K. E., Darby, M. K., Crowford, E. T., and Brown, D. D. (1990) *Cell 61*, 293−300.
6. Guddat, U., Bakken, A. H., and Pieler, T. (1990) *Cell 60*, 619−628.
7. Allison, L. A., Romaniuk, P. J., and Bakken, A. H. (1991) *Dev. Biol. 144*, 129−144.
8. Murdoch, K. J., and Allison, L. A. (1996) *Exp. Cell Res. 227*, 332−343.
9. Rudt, F., and Pieler, T. (1996) *EMBO J. 15*, 1383−1391.
10. Clausen, M., Rudt, F., and Pieler, T. (1999) *J. Biol. Chem. 274*, 33951−33958.
11. Huber, P. W., and Wool, I. G. (1986) *J. Biol. Chem. 261*, 3002−3005.
12. Scripture, J. B., and Huber, P. W. (1995) *J. Biol. Chem. 270*, 27358−27365.
13. Chow, C. S., Hartmann, K. M., Rawlings, S. L., Huber, P. W., and Barton, J. K. (1992) *Biochemistry 31*, 3534−3542.
14. Wormington, W. M. (1989) *Mol. Cell. Biol. 9*, 5281−5288.
15. Huber, P. W., Blobe, G. C., and Hartmann, K. M. (1991) *J. Biol. Chem. 266*, 3278−3286.
16. Keightley, D. D., and Cressie, N. A. (1980) *J. Steroid Biochem. 13*, 1317−1323.
17. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
18. Huber, P. W., Rife, J. P., and Moore, P. B. (2001) *J. Mol. Biol. 312*, 823−832.
19. Keightley, D. D., Fischer, R. J., and Cressie, N. A. (1983) *J. Steriod Biochem. 19*, 1407−1412.
20. Novokhatny, V., and Ingham, K. (1997) *Protein Sci. 6*, 141−146.
21. Oubridge, C., Ito, N., Evans, P. R., Teo, C.-H., and Nagai, K. (1994) *Nature 372*, 432−438.
22. Price, S. R., Evans, P. R., and Nagai, K. (1998) *Nature 394*, 645−650.
23. Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) *Nature 398*, 579−585.
24. Deo, R. C., Bonanno, J. B., Sonenberg, N., and Burley, S. K. (1999) *Cell 98*, 835−845.
25. Cavarelli, J., Rees, B., Ruff, M., Thierry, J. C., and Moras, D. (1993) *Nature 362*, 181−184.
26. Rould, M. A., Perona, J. J., and Steitz, T. A. (1991) *Nature 352*, 213−218.
27. Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Liljas, L. (1994) *Nature 371*, 623−626.
28. Valegård, K., Murray, J. B., Stonehouse, N. J., van den Worm, S., Stockley, P. G., and Liljas, L. (1997) *J. Mol. Biol. 270*, 724−738.
29. Legault, P., Li, J., Mogridge, J., Kay, L. E., and Greenblatt, J. (1998) *Cell 93*, 289−299.
30. Mao, H., White, S. A., and Williamson, J. R. (1999) *Nat. Struct. Biol. 6*, 1139−1147.
31. Chow, C. S., Behlen, L. S., Uhlenbeck, O. C., and Barton, J. K. (1992) *Biochemistry 31*, 972−982.
32. Wimberly, B., Varani, G., and Tinoco, I. (1993) *Biochemistry 32*, 1078−1087.
33. Zinnen, S., and Yarus, M. (1995) *Nucleic Acids Symp. Ser. 33*, 148−151.
34. LeCuyer, K. A., Behlen, L. S., and Uhlenbeck, O. C. (1996) *EMBO J. 15*, 6847−6853.
35. Draper, D. E. (1999) *J. Mol. Biol. 293*, 255−270.
36. Moras, D., and Poterszman, A. (1995) *Curr. Biol. 5*, 249−251.

37. Michael, W. M., and Dreyfuss, G. (1996) *J. Biol. Chem. 271*, 11571−11574.
38. Rosorius, O., Fries, B., Stauber, R. H., Hirschmann, N., Bevec, D., and Hauber, J. (2000) *J. Biol. Chem. 275*, 12061−12068.
39. Lin, E., Lin, S. W., and Lin, A. (2001) *Nucleic Acids Res. 29*, 2510−2516.
40. Deshmukh, M., Stark, J., Yeh, L. C., Lee, J. C., and Woolford, J. L. J. (1995) *J. Biol. Chem. 270*, 30148−30156.
41. Yeh, L. C., Deshmukh, M., Woolford, J. L., and Lee, J. C. (1996) *Biochim. Biophys. Acta 1308*, 133−141.
42. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science 289*, 905−920.
43. Nakashima, T., Yao, M., Kawamura, S., Iwasaki, K., Kimura, M., and Tanaka, I. (2001) *RNA 7*, 692−701.

BI011439M