doi:10.1006/jmbi.2001.4454 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2001) 307, 229–246

JMB

Structural and Energetic Analysis of RNA Recognition by a Universally Conserved Protein from the Signal Recognition Particle

Robert T. Batey, M. Bidya Sagar and Jennifer A. Doudna*

Department of Molecular Biophysics and Biochemistry and Howard Hughes Medical Institute, Yale University P.O. Box 208114, New Haven CT 06520-8114, USA

*Corresponding author

The signal recognition particle (SRP) is a ribonucleoprotein complex responsible for targeting proteins to the endoplasmic reticulum in eukarya or to the inner membrane in prokarya. The crystal structure of the universally conserved RNA-protein core of the *Escherichia coli* SRP, refined here to 1.5 Å resolution, revealed minor groove recognition of the 4.5 S RNA component by the M domain of the Ffh protein. Within the RNA, nucleotides comprising two phylogenetically conserved internal loops create a unique surface for protein recognition. To determine the energetic importance of conserved nucleotides for SRP assembly, we measured the affinity of the M domain for a series of RNA mutants. This analysis reveals how conserved nucleotides within the two internal loop motifs establish the architecture of the macromolecular interface and position essential functional groups for direct recognition by the protein.

Keywords: RNA structure; RNA-protein interactions; X-ray crystallography; signal recognition particle; protein translocation

Introduction

Targeting of proteins for localization within cellular compartments or membranes, or secretion to the extracellular environment, involves several evolutionarily conserved pathways in which a cytosolic factor recognizes specific proteins for delivery to the membrane-bound translocational machinery (reviewed by Schatz & Dobberstein, 1996). In eukaryotic cells, the signal recognition particle (SRP) binds to an amino-terminal signal sequence of nascent polypeptides destined for the endoplasmic reticulum, arresting elongation and mediating the docking of the associated ribosome with receptor proteins on the membrane in a GTP-dependent process (reviewed by Walter, 1995). The eukaryotic SRP is a ribonucleoprotein (RNP) complex comprising a 220-nucleotide RNA and six proteins. At its core lies SRP54 (Figure 1), a protein composed of the NG and M domains. The NG domain is a ras-type GTPase, and along with the SRP receptor, carries out GTP-dependent regulation of translocation (Miller *et al*., 1993). RNA recognition and signal sequence binding activities of SRP54 are mediated by the M domain (methionine-rich domain) (Zopf *et al*., 1990). In *Escherichia coli*, the SRP is composed of the Ffh protein, a homolog of SRP54 (Poritz *et al*., 1990; Ribes *et al*., 1990), and the 4.5 S RNA containing a region homologous to domain IV of the 7 S RNA (Poritz *et al*., 1988; Struck *et al*., 1988) (Figure 1). The evolutionary conservation of these components is underscored by the ability of human SRP54 to bind with high affinity to *E. coli* 4.5 S RNA (Poritz *et al*., 1990), the human 7 S RNA to rescue 4.5 S RNA deficiency in *E. coli* (Ribes *et al*., 1990), and for Ffh to replace SRP54 in a chimeric mammalian SRP that is capable of translational arrest and signal sequence recognition (Bernstein *et al*., 1993). Thus, it appears that the bacterial SRP is a minimized structural and functional homolog of the eukaryotic SRP, making it attractive for structural studies of SRP function.

As part of an effort to understand the function of the universally conserved RNP core of the SRP, we determined its crystal structure (Batey *et al*., 2000). In the structure, two alpha helices within the M-domain interact with the minor groove of the symmetric internal loop as well as a conserved adenosine of the asymmetric internal loop of the RNA (Figure 2). The RNA-protein interface is entirely defined by amino acid residues and nucleotides that are invariant across the three king-

Abbreviations used: SRP, signal recognition particle; RNP, ribonucleoprotein.

E-mail address of the corresponding author: doudna@csb.yale.edu

0022-2836/01/010229–18 $35.00/0

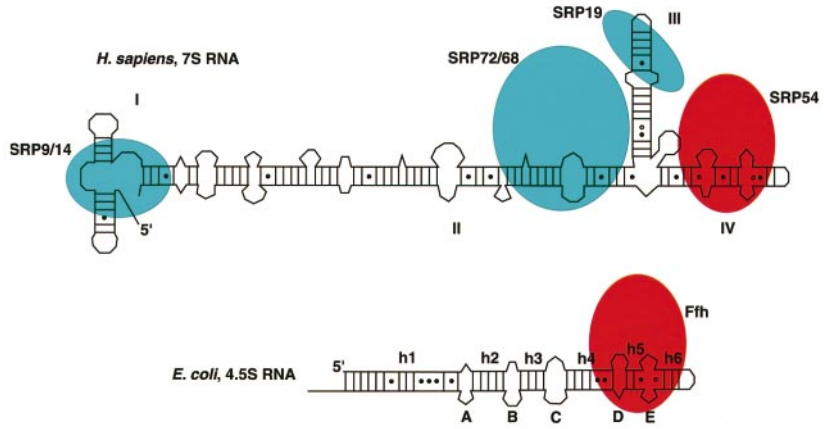

Figure 1. Representation of the human and *E. coli* signal recognition particles. The universally conserved SRP54/Ffh protein is shown in red and the other proteins are shown in light blue. The proteins are placed on their binding sites on the SRP RNA. The RNA domains are labeled with Roman numerals for the human SRP RNA, and as helices h1-h5 and internal loops A-E for the *E. coli* 4.5 S RNA.

doms of life, indicating that this mode of RNA recognition is utilized by all SRP54/Ffh variants. Strikingly, this fragment of SRP RNA is capable of rescuing a lethal knockout of the 4.5 S RNA in *E. coli*, indicating that it contains all of the elements necessary and sufficient for *in vivo* function (Batey *et al.*, 2000). The role of the RNA in SRP function may involve recognition and binding of signal sequences (Batey *et al.*, 2000) as well as recruitment of the receptor protein FtsY to the SRP (Peluso *et al.*, 2000).

Here, we describe the design and analysis of an RNA-protein complex for crystallization, resulting in crystals that diffracted to high resolution. The structure of the SRP ribonucleoprotein core has now been fully refined using data extending to 1.5 Å resolution, revealing a novel mode of RNA recognition by a protein. Guided by this structure, a series of RNAs containing single functional group substitutions that delete specific hydrogen bonds at the macromolecular interface was tested for binding to the M domain. This study reveals that four nucleotides within the two RNA internal loops create a series of thermodynamically coupled hydrogen bonds with the protein to form an exceptionally high affinity protein-RNA interaction. Flanking sequences play a critical role in establishing the architecture of the RNA at the macromolecular interface, while only making minor energetic contributions to direct protein-RNA contacts.

Results and Discussion

Chemical footprinting of the 4.5 S RNA-Ffh complex revealed that the Ffh binding site lies in the region of helices 4-6 of the 4.5 S RNA (Lentzen *et al.*, 1996). Furthermore, a small RNA containing nucleotides (nt) 32-74 of the 4.5 S RNA interacts with Ffh and Ffh-M domain with equal affinity (Schmitz *et al.*, 1996). These studies indicated that protein-RNA recognition in the *E. coli* SRP was localized to the M domain and domain IV of the SRP RNA. In addition, expression of a 49 nucleotide domain IV construct *in vivo* supports growth of a strain of *E. coli* conditionally deficient in the essential 4.5 S gene (Batey *et al.*, 2000). These results guided the design of a functional core complex for X-ray crystallography.

Determination of a minimal 4.5 S RNA binding domain of Ffh-M domain

Our initial M-domain construct included amino acid residues 296-453 of Ffh (FfhM-1), considered to comprise the entire RNA binding domain (Althoff *et al.*, 1994). To assess whether this fragment could be further truncated, we used a combination of limited proteolysis in the presence of RNA followed by MALDI-TOF mass spectrometry. Mild treatment of FfhM-1 with trypsin in the absence of RNA resulted in rapid proteolysis of the entire protein within five minutes (Figure 3(a)). In the presence of RNA, however, large fragments

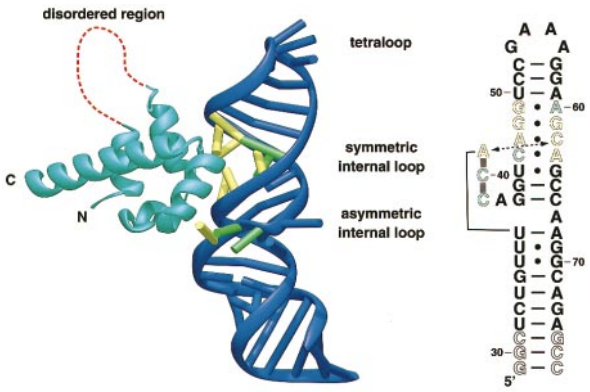

Figure 2. Crystal structure of the universally conserved ribonucleoprotein core of the *E. coli* signal recognition particle. The M domain is represented as a light blue ribbon with the disordered loop shown as a broken line. Nucleotide bases within the RNA (blue) are shown as rods, with universally conserved and highly conserved nucleotides depicted as yellow and green, respectively. To the right is the secondary structure of the 4.5 S RNA, as observed in the crystal structure. A tertiary contact between the asymmetric and symmetric loop is shown as a broken arrow.

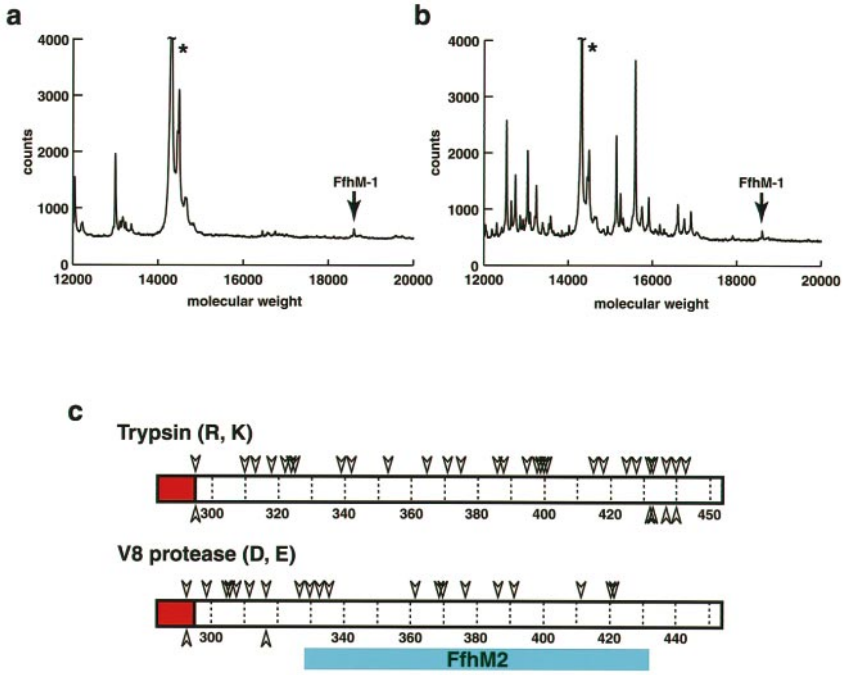

Figure 3. MALDI-MS analysis of the 4.5 S RNA-M domain interaction. (a), (b) Mass spectrum of the M domain after a five minute exposure to trypsin in the absence and presence of a stoichiometric quantity of 4.5 S RNA, respectively. (c) Summary of cleavages induced by trypsin and V8 proteases observed in FfhM-1 in the presence of RNA. Arrowheads on the top represent potential cleavage sites within the protein, while those on the bottom represent observed cleavages. The red bar indicates the N-terminal His_6 tag and Factor Xa cleavage site that precedes the native M domain sequence. The light blue bar delineates the part of the M domain that was present in FfhM-2 that was used for crystallography.

of the M-domain remain protease-resistant (Figure 3(b)), most of which are stable to proteolysis for over an hour under these conditions. The smallest of the stable fragments corresponds to amino acid residues 296-432. Similar results were obtained using V8 protease, which yielded a stable fragment comprising amino acid residues 328-453. This suggested a core RNA binding domain that consists of residues 328-432. An M domain construct corresponding to this region (FfhM-2) retains the ability to bind to the 4.5 S RNA with the same affinity as wild-type Ffh (data not shown).

The FfhM-2 construct tended to dimerize under oxidizing conditions due to the presence of a single cysteine residue. This amino acid residue, C406, resides in the middle of the conserved RNA binding domain and is a serine in all other SRP54/Ffh homologs. Oxidation of this residue leads to the complete loss of RNA binding activity, indicating its importance in forming a stable protein-RNA complex (Keenan *et al.*, 1998). Therefore, C406 was mutated to serine in wild-type Ffh and FfhM-2, and both resulting variants retain full 4.5 S RNA binding activity (data not shown). FfhM-2(C406S) was utilized throughout subsequent crystallization trials of the SRP protein-RNA complexes.

Determination of a minimal M-domain binding RNA fragment

To map the boundaries of a minimal RNA binding site in the 4.5 S RNA, we employed alkaline hydrolysis ladder selection (Query *et al.*, 1989). A pool of RNA length variants, generated by hydroxide ion mediated strand scission, was assayed for high-affinity binding to FfhM-1 and analyzed on polyacrylamide sequencing gels. Analysis of full length 4.5 S RNA revealed a sharp break in the sequencing ladder in 5′ and 3′ end-labeled RNA located within helix 4 (Figure 4(a)). This demonstrated that only nucleotides 33 to 72 are required for high affinity binding of the M domain, in agreement with the minimal site utilized in NMR studies of domain IV of the *E. coli* SRP RNA (Schmitz *et al.*, 1999) and with the minimal functional SRP RNA required *in vivo* (Batey *et al.*, 2000).

Nucleotides within the 4.5 S RNA necessary for M domain binding were probed using chemical (Conway & Wickens, 1989) and nucleotide analog interference analysis (Ryder *et al.*, 2000). By using a variety of nucleotide analogs that target adenosine (Figure 4(b)) and guanosine bases, and hydrazine, which creates abasic sites at cytosines and uracil bases, the entire sequence of the 4.5 S RNA was analyzed. Since only nucleotides within two internal loops of the 4.5 S RNA were recognized by the protein (Figure 4(c)), sequences outside these loops are either required to present these loops in the correct structural context or play other roles in SRP function, as may be the case for the conserved GGAA tetraloop. Consistent with these observations, both wild-type Ffh and the FfhM-2(C406S) bind to the 4.5 S RNA and a minimal RNA element comprising nucleotides 34 to 72 with the same apparent equilibrium dissociation constant (K_d) of 30 pM.

The ability of the M-domain and full-length Ffh to bind to the 4.5 S RNA with the same affinity does not preclude the presence of contacts made between the NG-domain of Ffh and the 4.5 S RNA. In the Fe^{2+}-EDTA footprinting analysis of the Ffh-4.5 S RNA interaction, additional strong protections were seen at nucleotides that do not interact with the M domain (Lentzen *et al.*, 1996). This

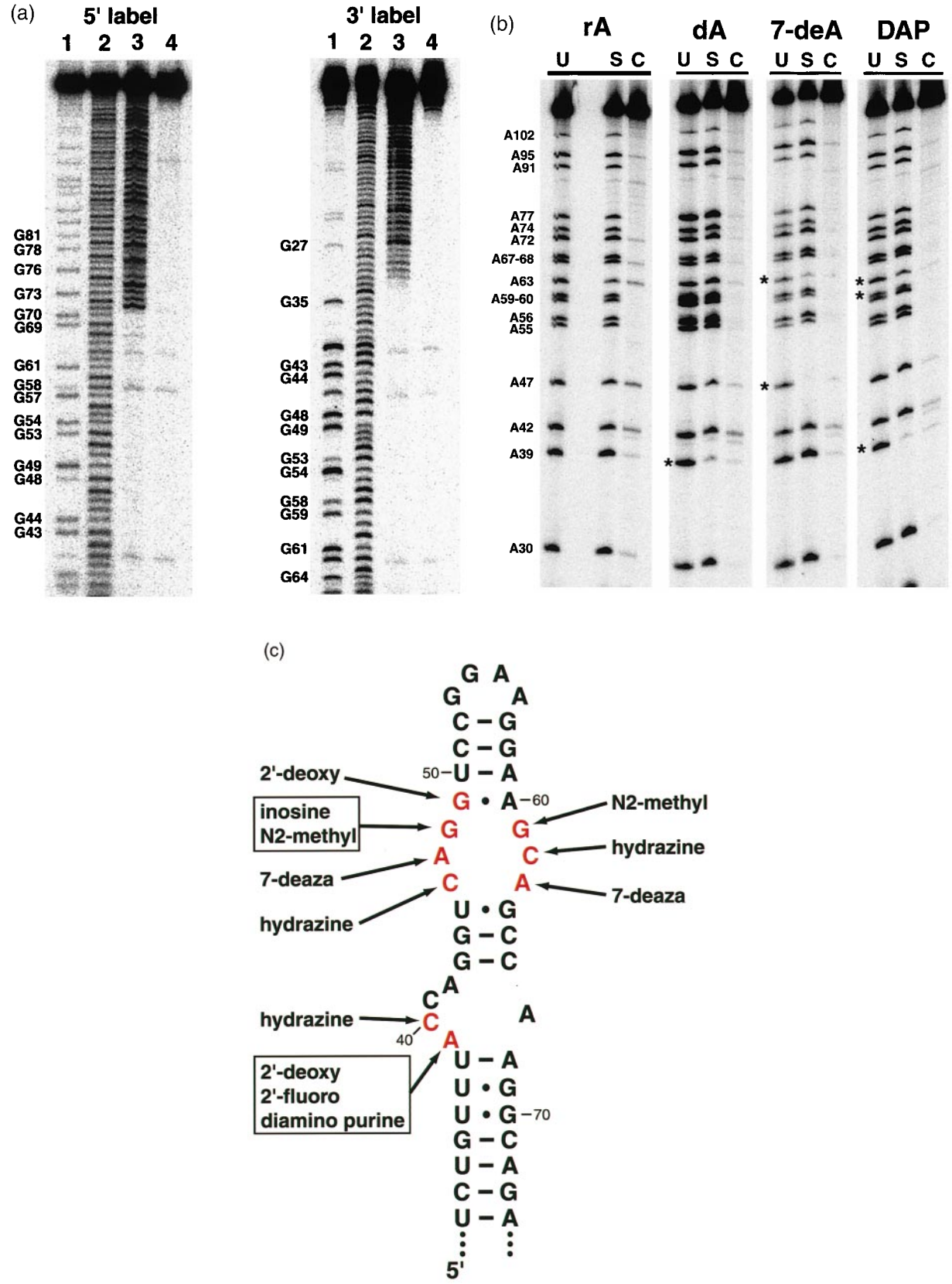

Figure 4. Selection analysis of the 4.5 S RNA-M domain interaction. (a) 5′ (left) and 3′ (right) end-labeled alkaline hydrolysis ladder selection. Lanes are labeled as: 1, RNase T_1 sequencing ladder; 2, alkaline hydrolysis ladder; 3, RNA from the alkaline hydrolysis ladder capable of binding RNA; 4, unhydrolyzed control RNA brought through the selection assay. (b) Nucleotide analog analysis using α-S adenosine, adenosine 2′-deoxyribonucleotide, 7-deaza adenosine and diaminopurine. Lanes are labeled as: 1, unselected RNA; 2, selected RNA; 3, selected RNA not subjected to iodine-mediated strand scission. An asterisk indicates the sites of interference. (c) Summary of functional groups and bases whose modification interferes with the ability of the M domain to bind the 4.5 S RNA.

caveat of the minimization approach to structural studies of protein-nucleic acid complexes is illustrated in biochemical and structural studies of the ribosomal protein S15. The minimal 16 S ribosomal RNA binding site of S15 in *Bacillus stearothermophilus* (Batey & Williamson, 1996) and *Thermus thermophilus* (Serganov *et al.*, 1996) is an RNA containing a three-way junction element and an internal loop. However, recent crystal structures have revealed that S15 makes a number of contacts to the 16 S rRNA outside its biochemically defined binding site (Agalarov *et al.*, 2000; Wimberly *et al.*, 2000), consistent with RNase and Fe^{2+}-EDTA probing of the S15-16 S rRNA complex (Powers & Noller, 1995).

Using the results of the biochemical assays, a series of minimal RNAs were constructed, employing two strategies to improve the potential crystallizability of the protein-RNA complex. The first was to change the wild-type GGAA tetraloop to a GAAA tetraloop. In crystals of the hammerhead ribozyme, the GAAA tetraloop contributed to intermolecular contacts that established the lattice (Pley *et al.*, 1994). Also, this tetraloop mediates intramolecular interactions within group I and group II introns (Michel & Westhof, 1990; Costa & Michel, 1997), and RNase P RNA (Brown *et al.*, 1996). Based on these observations, it has been proposed that engineering GAAA tetraloops and their receptors into RNA constructs is useful for promoting crystallization (Ferré-D'Amaré *et al.*, 1998b). In the SRP constructs, tandem G-C pairs were introduced at the blunt end of the RNA to promote an interaction with the GAAA tetraloop (Figure 5; Pley *et al.*, 1994; Costa & Michel, 1995, 1997). The second strategy was to systematically vary the length of the first helix (Figure 5), a technique used in the crystallization of protein-DNA complexes (Jordan *et al.*, 1985; Joachimiak & Sigler, 1991) as well as several protein-RNA complexes (Price *et al.*, 1995; Ferré-D'Amaré & Doudna, 2000).

Crystallization of the protein-RNA complex

The protein-RNA complex aggregated reversibly in monovalent cation concentrations below 300 mM, indicating a salt-dependent solubility behavior that could be exploited to drive crystallization. To form a soluble complex under low salt conditions prior to crystallization, the protein-RNA complex was denatured in 8 M urea and slowly reconstituted by dialysis against 10 mM K-Hepes (pH 8.0). This reconstitution procedure was critical for obtaining diffraction quality crystals of the complex. Urea-mediated refolding was used to overcome excessive RNA dimerization in the crystallization of the U2A′/U2B″/U2snRNA complex (Price *et al.*, 1998). The *Tetrahymena thermophila* group I intron has been shown to more efficiently evade kinetic traps during the refolding process in the presence of low urea concentrations (1-2 M) (Pan & Woodson, 1997; Rook *et al.*, 1998). The ability of urea to allow this large RNA to access alternative folding pathways may result in more homogeneous populations of RNA, potentially improving its crystallizability. Thus, the use of urea as an additive for refolding both small and large RNAs may be an important avenue to explore during the search for suitable crystals.

Each FfhM-2(C406S)-RNA complex was subjected to two commonly used sparse matrices of crystallization conditions (Jancarik & Kim, 1991; Scott *et al.*, 1995). Within a series of RNA constructs, protein complexes with LM11 or LM12 (Figure 5) readily yielded crystals under a broad range of conditions, ranging from birefringent precipitate to single, diffraction quality crystals. Crystals of the LM12 complex diffracted X-rays to

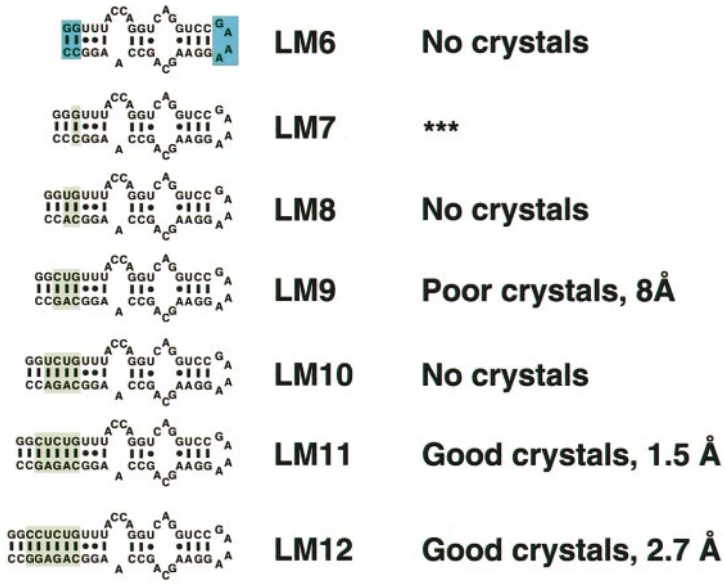

Figure 5. The series of RNA constructs that were tested with FfhM-2(C406S) for crystallization using the ''magic-50'' sparse matrix (Jancarik & Kim, 1991) and an RNA-directed sparse matrix (Scott *et al.*, 1995). The construct LM7 in the series was not tested because the 3′ hepatitis delta virus ribozyme did not cleave to yield the correct product. The light blue boxes on LM6 correspond to the tandem G-C pairs added to the blunt end of the RNA and the GAAA tetraloop, features that were retained in all of the tested constructs. Sequences highlighted in green were segments that were added to the minimal LM6 RNA to generate the length variant series.

2.7 Å resolution at a synchrotron source and had a high mosaic spread (1.2-1.5°), while crystals of the LM11 complex diffracted beyond 1.5 Å and had significantly lower mosaicity (0.3-0.5°). Therefore, all further crystallographic analysis of the complex was performed using the LM11 construct.

To solve the phase problem, crystals containing selenomethionine derivatized M domain were used for multiwavelenth anomalous dispersion (MAD) phasing (Batey *et al.*, 2000). However, the presence of selenomethionine changed the crystallization properties of the complex. While the native complex readily yielded single, diffraction quality crystals, selenomethionine labeled FfhM-2(C406C)/LM11 consistently generated highly polymorphic crystals that were unsuitable for diffraction studies. Upon rescreening, it was found that single crystals of the derivatized complex could be readily obtained under almost identical conditions except for the addition of detergent. While usable crystals were obtained with a number of non-ionic detergents, it was empirically found that C-HEGA 10 yielded the highest quality crystals.

Refinement and overview of the structure

Using data that extended to 1.5 Å resolution from a native FfhM-2(C406)/LM11 complex, the structure of the ribonucleoprotein core of the *E. coli* SRP was determined by molecular replacement (MR) with the previous structure, using AMoRe (Navaza, 1994) and initially refined with CNS (Brünger *et al.*, 1998). After the R_{free} had dropped to 0.231, the model was refined in SHELX-97 (Sheldrick & Schneider, 1997) using anisotropic temperature factor refinement to yield a final conventional *R*-factor of 0.154 and an R_{free} of 0.198 (Table 1). The resulting model has several minor differences from that previously described (Batey *et al.*, 2000). Seven additional amino acids were built according to electron density in $2F_o - F_c$ maps. Furthermore, a potassium ion that was described at the protein-RNA interface was changed to a water molecule. This change was confirmed using a thallium soak (Basu *et al.*, 1998) and refining the individual thermal *B*-factor using a potassium, sodium, or water at this position. Other metal ions that were previously described have been confirmed through a combination of monovalent and divalent cation soaks.

The 4.5 S RNA fragment folds into a hairpin structure containing two internal loop motifs with a GAAA tetraloop at the end. The symmetric internal loop is defined by five consecutive non-canonical base-pairs (Figure 2) in which all nucleotides participate in pairing interactions. In contrast, none of the nucleotides within the asymmetric loop are paired. Instead, the single adenosine, A67, remains stacked into the helix, while on the other side of this loop the four nucleotides are extruded into the solvent, such that the ribose-phosphate backbone faces the interior of the helix. Three stacked nucleotides, A39-C41, present A39 to the protein for extensive recognition. Two alpha helices of the M domain interact with the RNA along the minor groove face of the symmetric internal loop centered about the reverse-Hoogsteen A47-C62 pair.

Recognition of the minor groove is accomplished primarily using carbonyl functional groups in the protein backbone rather than side-chains. The extensive use of the protein backbone was observed in the complex of ribosomal protein L11 with its rRNA target (Conn *et al.*, 1999; Wimberly *et al.*, 1999). In both cases a critical glycine residue allows for the close approach of the protein and RNA. However, unlike most of the structurally characterized protein-RNA interactions, electrostatic contacts play virtually no role in recognition by the M domain; the only such interaction observed in the structure is between the non-conserved Lys386 and non-bridging phosphate

Table 1. Crystallographic data and refinement statistics

A. *Data statistics*	
Space group	*C*2
Cell dimensions	$a = 136.5$ Å, $b = 78.3$ Å, $c = 32.9$ Å $\alpha = \gamma = 90°$, $\beta = 96.2°$
Synchrotron beamline	ALS, 5.02
Number of crystals used	3
Resolution (Å)	40-1.52
Completeness	98.2 % (89.3 %)[a]
No. of observations	1,178,959
No. of unique reflections	51,967
Redundancy	10.8
R_{merge} (%)[b]	7.0 (11.5)[a]
I/σ	35.3 (1.6)[a]
B. *Molecular replacement*	
Rf AMoRe (%)	31.5
Correlation coefficient	69.6
C. *Refinement statistics*	
Resolution limits (Å)	10-1.52
Number of reflections in working set	38,752
Number of reflections in free set	2031
R-value (%)[c]	15.1
R_{free} (%)	19.9
Number of atoms in model	
Protein	606
RNA	1054
Metal ions	7
Water	287
RMSD from ideal values	
Bond length (Å)	0.01
Angle distance (Å)	0.03
Ramachandran plot	
Most favored region (%)	89.7
Additional allowed region (%)	7.4
Generously allowed regions (%)	2.9
Disallowed regions	0.0

[a] In parentheses for the highest resolution shell (1.57-1.52 Å).
[b] $R_{merge} = \Sigma|\langle I\rangle - Ih|/\Sigma\langle I\rangle \times 100$; Ih is the intensity of a measurement and $\langle I\rangle$ is the average of the measurements for a reflection h.
[c] $R = \Sigma|F_o - kF_c|/\Sigma F_o \times 100$, where F_o and F_c are observed and calculated structure amplitudes, respectively.
[d] R_{free} is R for a 5 % subset of all the reflections, which was not used in the crystallographic refinement.

oxygen atoms of U38 and A39. The relative unimportance of this interaction is underscored by the insensitivity of the protein-RNA complex to extremely high salt concentrations (data not shown).

RNA conformation within the 4.5 S RNA

Following torsion angle refinement, Cartesian space refinement was used to produce the final model. Thus, rather than restraining the sugar pucker in the RNA to C2′- or C3′-*endo* conformation, we were able to accurately refine the torsion angles within the ribose sugars. In the refined model, all of the ribose sugars adopt the common C3′-*endo* conformation or the related $^{3}_{2}$T conformation except for A42, which adopts the C2′-*endo* related $^{2}_{1}$T conformation (Saenger, 1984). Analysis of the individual torsion angles reveals that they all lie within favorable regions of the conformational map of the ribose-phosphate backbone for RNA (Murthy *et al.*, 1999). Thus, the architecture of the M domain binding site is achieved using standard backbone torsion angles despite the formation of an unusual helical structure involving the participation of the backbone in several non-canonical base-pairs and the extrusion of nucleotides from the asymmetric internal loop (Figure 10(a), (c)). This is consistent with observations of other high resolution RNA structures (Murthy *et al.*, 1999; Westhof & Fritsch, 2000).

RNA-RNA and RNA-protein interactions as crystal contacts

The arrangement of the complex in the crystal lattice is dominated by RNA-RNA interactions that establish molecular packing within all three dimensions. The first interaction involves the coaxial stacking of helices to form tail-to-tail dimers (Figure 6). This arrangement underscores the importance of using ribozymes to eliminate 5′- and 3′-end heterogeneity, which might have had a significant deleterious effect on formation of this contact. The presence of a 2′-3′ cyclic phosphate, a byproduct of the cleavage of the hepatitis delta virus ribozyme at the 3′-end did not disrupt the ability of two RNA helices to coaxially stack.

A principal strategy in engineering the RNA was the use of a GAAA tetraloop in conjunction with tandem G-C pairs at the opposite end. In the structure, however, the tetraloop bound to the minor groove face of a 5′G/CA3′ dinucleotide step (Figure 7(a)), instead of interacting with the tandem G-C pairs as observed in crystal of the hammerhead ribozyme (Pley *et al.*, 1994). Two adenosine bases from the tetraloop form minor groove triples using their Watson-Crick face (Figure 7(b)), unlike most known examples of this interaction in which adenosine bases use their minor groove face (Pley *et al.*, 1994; Cate *et al.*, 1996; Ferré-D'Amaré *et al.*, 1998a; Conn *et al.*, 1999; Wimberly *et al.*, 1999; Ban *et al.*, 2000). At the site

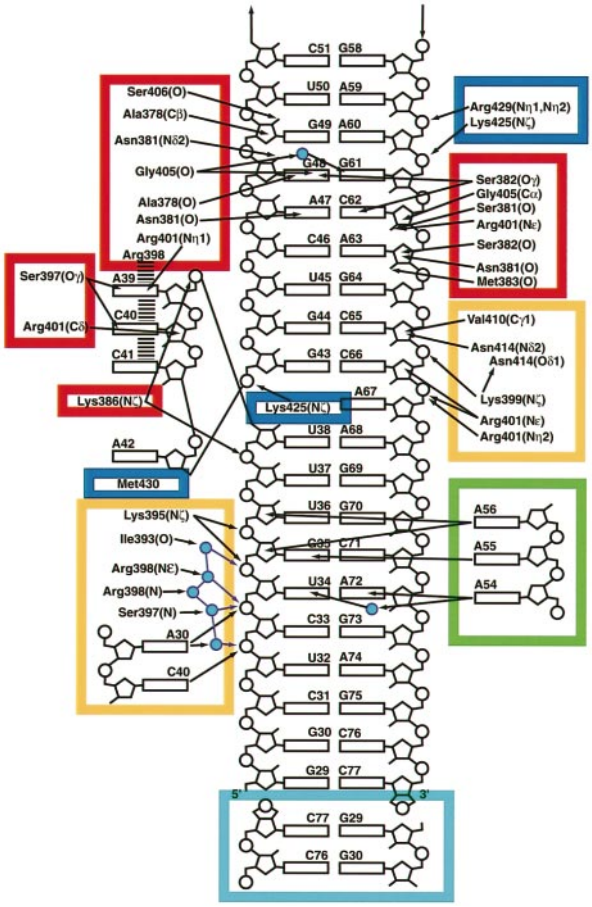

Figure 6. Summary of the M domain-RNA interactions and crystal contacts observed. The intermolecular contacts made by protein and RNA that are boxed in red. Crystal contacts made by neighboring complexes in the crystal are boxed in light blue, purple, orange, and green, with each color representing a physically distinct molecule. The thick broken lines between Arg398 and A39, C40 and C41 represent stacking interactions and blue spheres denote water-mediated interactions.

of interaction, the helix is slightly overwound such that the two purines form a partial cross-strand stack that facilitates direct recognition by the tetraloop. This mode of adenosine-minor groove interaction is observed within the 16 S rRNA in a slightly different context (Wimberly *et al.*, 2000). Here, three phylogenetically conserved adenosine bases within an internal loop (A607-609 using the *E. coli* 16 S rRNA numbering scheme) form a virtually identical set of contacts with the G309-C291 (A55-like, Figure 7(b)) and C308-G292 (A54-like) base-pairs. Since A54·U34-A72 does not involve the C2 of A72 (Figure 7(b)), a U-A or C-G pair support this interaction, but not a G-C or A-U pair. Thus, this mode of adenosine-mediated recognition of the minor groove of a double-stranded RNA helix is distinct from the more common adenosine-minor groove triple motif (Doherty *et al.*, 2001).

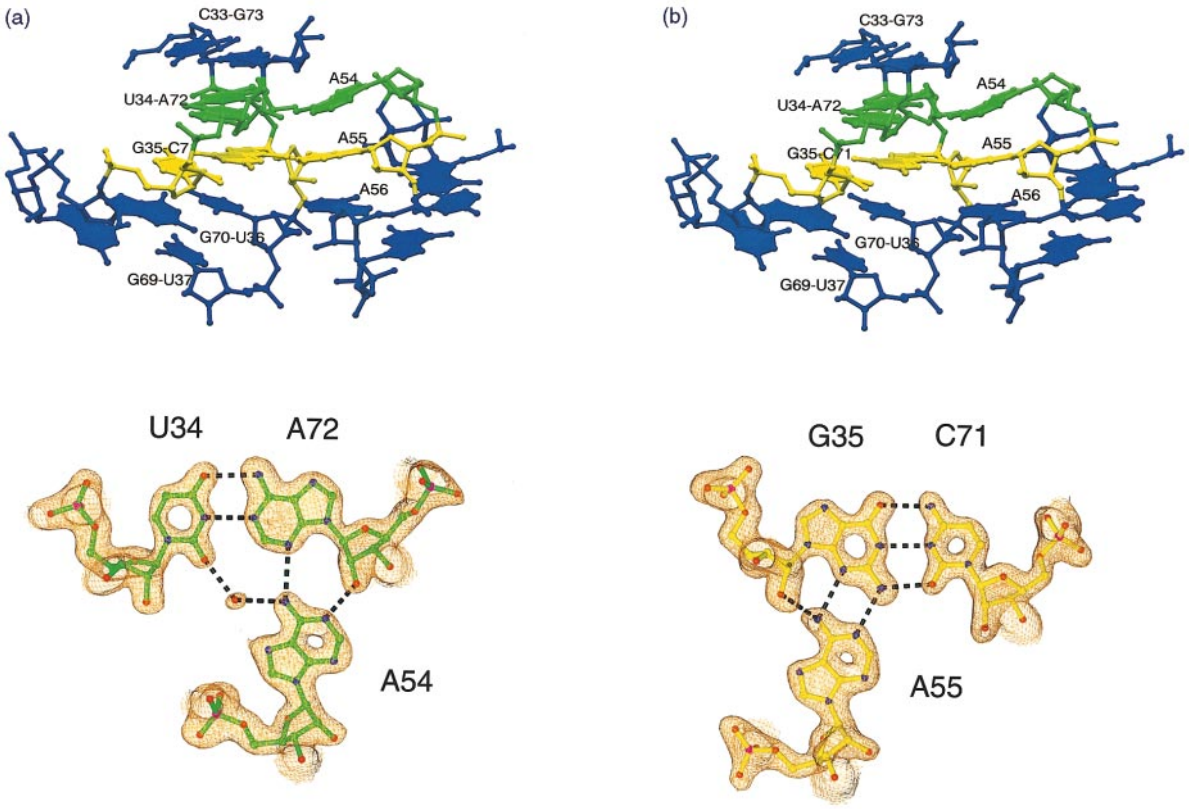

Figure 7. Tetraloop-minor groove interaction. (a) Stereo representation of the GAAA tetraloop of one molecule interacting with the minor groove of an adjacent RNA. (b) Base triples formed by this contact, with the final $2F_o - F_c$ electron density map contoured at 1.6σ superimposed. The hydrogen bond between A54 and U34 is mediated by a solvent molecule.

The second series of intermolecular contacts involves interactions between the Watson-Crick face of bases in the asymmetric loop and basic residues in the M-domain with the ribose-phosphate backbone of an adjacent RNA molecule. On one strand, this involves a series of RNA-, protein-, and water-mediated interactions to the phosphate backbone. The other side of the major groove provides three more phosphates that contact the protein. Further contacts to the RNA are also made by the extreme C terminus of the M domain. Met380 stacks upon A42, while Lys425 and Arg429 interact with non-bridging phosphate oxygen atoms of A60 and G61. Thus, most of the crystal contacts made between the protein and the RNA are electrostatic in nature, which is typical of non-specific protein-nucleic acid interactions. This may also explain the observation that the complex does not crystallize under high salt conditions (>300 mM monovalent ion), since increasing ionic strength would destabilize interactions needed to form crystal contacts.

Structure of the Ffh-M domain

The Ffh M domain is a bundle of five alpha helices arranged around a small hydrophobic core, creating the binding site for both the 4.5 S RNA and signal sequences. Superimposing the *E. coli*, *T. aquaticus* (Keenan *et al.*, 1998) and human Ffh/SRP54-M domain crystal structures (Clemons *et al.*, 1999) (Figure 8(a)) reveals that the signal sequence and RNA binding sites behave in fundamentally different ways. The RNA binding face of the protein, centered about helices two and three, appears to be rigid; even amino acid side-chains essential for RNA recognition are in virtually identical conformations (Figure 8(b)). Along with the observation that the coordinates of the free and bound symmetric internal loop of the *E. coli* 4.5 S RNA superimpose well (Jovine *et al.*, 2000), this suggests these molecules bind mostly through a rigid body association, in contrast to many protein-RNA interactions that involve an induced fit mechanism of binding (Draper, 1999; Frankel, 2000). This may explain why, despite the relatively small binding interface, this protein-RNA interaction is particularly tight as it may pay a minimal entropic penalty for conformational ordering upon binding.

Despite the presence of RNA, however, a 33-amino acid residue segment (338-370) within the M-domain was disordered within the original co-crystal structure (Batey *et al.*, 2000). During the refinement procedure, we attempted to account for this missing region in the electron density map,

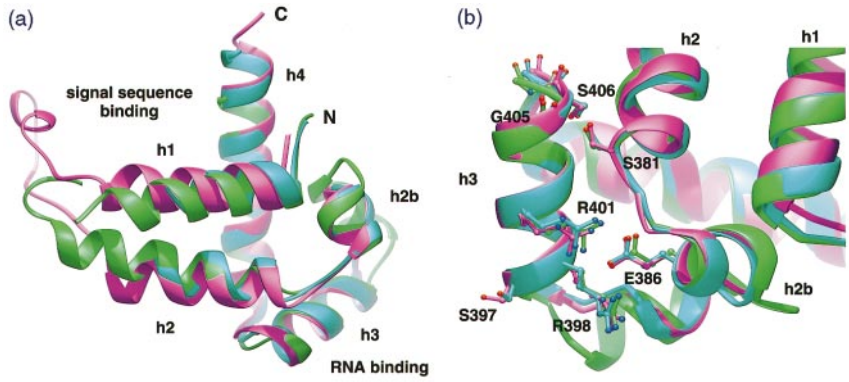

Figure 8. (a) Superposition of the *E. coli* (light blue), *T. aquaticus* (magenta), and human (green) M domains. In the human SRP54-M domain structure, the protein crystallized as a dimer in which helix 1 was swapped between the two molecules (Clemons *et al.*, 1999). Despite the SRP54-M domain helix 1 belonging to physically distinct molecule, it superimposes quite well with helix 1 bacterial variants. The superposition was performed using LSQMAN (Kleywegt & Jones, 1994). (b) Close-up view of the RNA binding face of the M domain with critical side-chains for RNA binding shown.

but were only able to build an additional three amino acid residues in helix 1 and two amino acid residues in helix 2. Additionally, we visually inspected maps at different resolutions to see whether there was electron density observed at lower resolution that becomes lost when higher resolution data is included. In the *Haloarcula marismortui* 50 S ribosomal subunit, for example, the L1 protein is visible in low resolution maps (9 Å), but the density is lost at higher resolution (5.5 and 2.4 Å) (Ban *et al.*, 2000). While the peptide binding region is ordered in the *T. aquaticus* Ffh-M domain, likely due to the intermolecular crystal contacts involving part of this segment, it adopted different conformations in different crystal forms, implying a degree of flexibility in this part of the protein (Keenan *et al.*, 1998). The homologous region within the human SRP54 M domain is similarly involved in intermolecular interactions corresponding to a swapping of helix 1 between adjacent molecules (Clemons *et al.*, 1999). This, along with biochemical evidence (Zheng & Gierasch, 1997), underscores the conformational flexibility of the signal sequence binding site, which may be critical for its ability to recognize a variety of signal sequences without requiring a specific amino acid sequence.

To further address signal sequence recognition, we attempted to co-crystallize or soak crystals with a peptide corresponding to the LamB signal sequence, which is recognized by the *E. coli* SRP (Miller *et al.*, 1994). While diffraction-quality crystals were obtained in each case, these experiments did not reveal any significant additional electron density in either the signal sequence binding groove or conformational ordering of the disordered region of the M-domain. Additionally, we fused to the amino terminus of the M-domain construct several different signal sequences (LamB and leader peptidase (de Gier *et al.*, 1996)) followed by a flexible linker. While all of these constructs yielded crystals that diffracted to better than 2.5 Å resolution, the resulting electron density maps again did not reveal any additional ordering of the groove or binding of the signal peptide. This may suggest that the NG domain is required in part for efficient signal sequence binding by the M domain (Newitt & Bernstein, 1997). Another possibility is that the signal sequence bearing peptide is binding to the groove, but that binding is weak and involves significant rotational and translational disorder, making it difficult to observe crystallographically. In a recent NMR structure of the Tom20 mitochondrial import receptor bound to a peptide containing a mitochondrial signal sequence, the bound peptide exhibited significant conformational flexibility, which was interpreted as important for its ability to recognize a variety of signal sequences (Abe *et al.*, 2000). Structural plasticity of sites of protein-protein interactions has been observed to be a critical feature for developing binding sites capable of recognizing a variety of ligands (Sundberg & Mariuzza, 2000).

Specific RNA-protein recognition

The crystal structure revealed in detail the interaction between RNA and protein functional groups at the molecular interface of the SRP complex. To determine their energetic contributions to complex stability, a series of single functional group deletions and base mutations in the RNA were created and tested for their ability to interact with FfhM-2(C406S) (Table 2). This was accomplished in part by dividing LM10 at the tetraloop, creating a two-piece construct that binds to the Ffh-M domain protein with wild-type affinity (Figure 9(a), (b)).

Many of the base substitutions that we describe will certainly have significant affects upon the RNA structure, resulting in local rearrangements in base-pairs, which has been observed in the P4-P6

Table 2. Binding of mutations

Mutation[a]	K_d (pM)[b]	$\Delta\Delta G$ (kcal/mol)
LM10, wild-type	33.3 ± 8.9	n.a.
LM10A + LM10B	34.4 ± 9.5	0.0
Asymmetric internal loop		
dA39	710 ± 180	−1.8
C40U	80 ± 16	−0.5
U45-G64 pair		
U45C	304 ± 70	−1.3
G64A	86.8 ± 37	−0.55
U45C/G64A	42.2 ± 7.0	−0.12
C46-A63 pair		
dC46 (*a*)[c]	24.6 ± 8.1	+0.20
C46U (*b*)	237 ± 60	−1.1
C46G	346 ± 44	−1.4
dA63 (*c*)	47.3 ± 17	−0.19
A63P (*d*)	62.5 ± 26	−0.35
A63G	1370 ± 520	−2.2
A47-C62 pair		
A47P (*e*)	1720 ± 430	−2.3
dC62 (*f*)	269,000 ± 190,000	−5.3
C62U (*g*)	85,200 ± 6700	−4.6
G48-G61 pair		
dG48 (*h*)	34.9 ± 11	−0.01
G48I (*i*)	96.2 ± 26	−0.61
G48(2AP) (*j*)	200,000 ± 88,000	−5.1
G48P	220,000 ± 120,000	−5.2
dG61 (*k*)	229 ± 69	−1.1
G61I (*l*)	91.6 ± 30	−0.58
G61(2AP) (*m*)	153 ± 47	−0.88
G61P	448 ± 150	−1.52
G49-A60 pair		
dG49 (*n*)	160 ± 41	−0.91
G49P	38.1 ± 16	−0.06
G49A	85.6 ± 8.0	−0.54
G49U	3090 ± 650	−2.7
A60G	553 ± 43	−1.6
A60P	118 ± 35	−0.73
A60C	69,900 ± 20,000	−4.5
G49A/A60G	102 ± 68	−0.64

[a] Nucleotide analogs are abbreviated as follows: dN, 2′-deoxyribonucleotide; P, purine base; I, inosine; 2AP, 2-amino purine.

[b] Each measurement is the average of at least three independent experiments with errors given as the 90% confidence limits.

[c] Italicized letters correspond to hydrogen bonding interactions shown in Figure 10.

domain of the *T. thermophila* group I intron (Szewczak & Cech, 1997). As a result, the presumed removal of specific hydrogen bonds within the RNA or at the protein-RNA interface as strictly interpreted from the structure may be compensated by the formation of alternate interactions. However, many of the mutants that we have used to interpret the energetics of the M domain-4.5 S interaction involve single functional group deletions in the RNA. These substitutions have been presumed to only introduce minor perturbations in RNA structure, although minor structural rearrangements cannot be completely eliminated (Silverman & Cech, 1999). For interpretation of the energetics of these functional group deletions, we have assumed that the RNA structure is the same as that observed in the crystal structure. Nonetheless, as described below, these mutations provide insights into which elements of RNA structure play pivotal roles in M domain recognition.

The central feature of the 4.5 S RNA-M domain interface is a network of hydrogen bonds between the A47-C62 pair, A39 of the asymmetric loop and a universally conserved salt bridge between Glu387 and Arg401 and (Figure 10(a)). This is supported by the observation that the C62G mutation within the 4.5 S RNA is lethal in *E. coli* (Wood *et al.*, 1992). Removal of any functional group in the A47-C62 pair results in significant loss of M domain binding affinity. This is especially evident in the 2′-deoxy C62 mutant, which results in a $\Delta\Delta G$ of −5.3 kcal/mol; removal of this single functional group results in the most significant loss of affinity of any mutation tested. Functional group substitutions in A39(C2) show similarly strong negative affects on M domain recognition (Doherty *et al.*, 2001). The symmetric and asymmetric internal loops contact each other through a single tertiary interaction between the 2′ hydroxyl of A39 and a non-bridging oxygen of A63, which serves to create a cleft in the RNA in which the E386-R401 salt bridge rests. Deletion of this hydrogen bond leads to a moderate reduction in binding affinity, but is less severe than removal of hydrogen bonds between the protein and the RNA.

The principal recognition elements of the M domain within the symmetric internal loop lie in the G48-G61 and A47-C62 pairs. Single functional group deletions in G61 that establish the proper hydrogen bonding contacts with G48 and the backbone have moderate effects upon protein binding (Figure 10(b)). Significantly, removal of the single hydrogen bond between the bases only weakly affects protein recognition. However, deletion of 2′-hydroxyl of G61, which contacts the protein through three extremely well ordered water molecules (Figure 10(b)) has a $\Delta\Delta G = -1.1$ kcal/mol, indicating that water-mediated contacts along the protein-RNA interface likely stabilize the interaction. Removal of the G48(N2) and N1 imino proton seriously disrupts protein recognition with a $\Delta\Delta G$ of −5.1 kcal/mol. This energetic penalty is far greater than that expected for removal of two hydrogen bonds to the protein (approximately 1.2-2.0 kcal/mol (Fersht *et al.*, 1985; Turner *et al.*, 1987)). Thus, at least part of the hydrogen bonding network between the protein and the RNA is thermodynamically coupled, with removal of the G48(N1)-Gly405(O) hydrogen bond leading to the destabilization of other interactions along the interface. This is similar to observations made for some protein-protein interactions, where a subset of intermolecular interactions at the center of the interface contribute most significantly to binding affinity (Conte *et al.*, 1999).

Non-canonical base-pairs flanking the critical nucleotides within the symmetric internal loop contribute indirectly to complex formation in the SRP. The U45-G64 wobble base-pair is not directly recognized by the M domain, but forms part of a binding site for a potassium ion (Batey *et al.*, 2000).

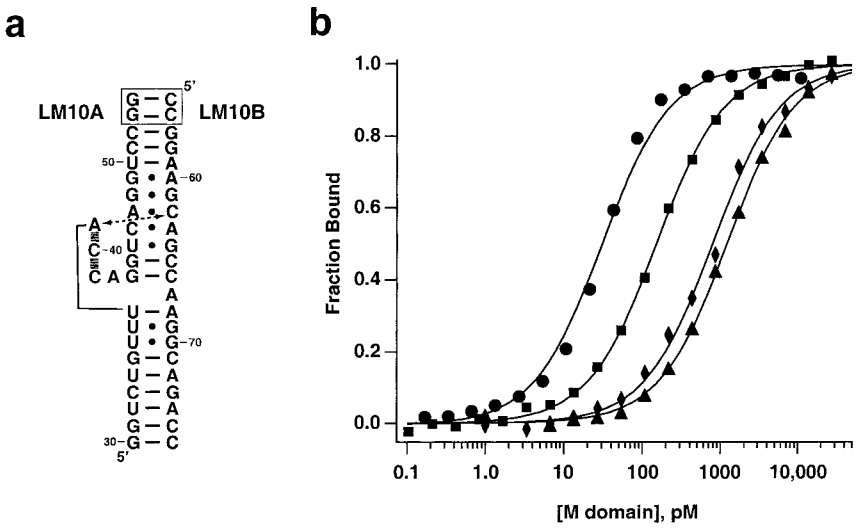

Figure 9. (a) LM10A/B construct used in the functional group mutagenesis studies. The boxed nucleotides replace the wild-type GGAA tetraloop in order to create a two piece RNA. (b) Representative normalized binding curves for wild-type RNA (filled circles) and several mutants (G61(2AP), squares; A60G, diamonds; A47P, triangles).

This pair is not strictly phylogenetically conserved among bacteria, but in many cases is an isosteric C-A^+ pair (Zweib & Larsen, 1997), indicating a preference for a wobble pair at this position. This is supported by the observation that substitution of a C-G and a U-A pair has a $\Delta\Delta G$ of -1.3 kcal/mol and -0.5 kcal/mol, respectively. However, substitution of the isosteric C-A^+ pair yields a $\Delta\Delta G$ of -0.1 kcal/mol, nearly restoring wild-type binding affinity. Adjacent to the G-U wobble is another unusual pair between C46 and A63 that is also not directly recognized by the M domain. While this pair is defined by only a single direct hydrogen bond between the bases, both bases form hydrogen bonds to the ribose-phosphate backbone of the opposite strand (Figure 10(c)). Elimination of one of the hydrogen bonds between A63(N6) and C46(2′OH), however, is the only functional group deletion that stabilizes the protein-RNA interaction. It is interesting to note that in the same type of A-C pair in tRNA(Phe), the 2′-hydroxyl of the cytosine is methylated (Leontis & Westhof, 1999), suggesting this modification may stabilize this unusual pair. Disruption of the single direct hydrogen bond between the two bases by A63-purine or the hydrogen bond between the protein Met383 and A63(2′OH) have only slightly deleterious effects on protein binding ($\Delta\Delta G = -0.3$ and -0.2 kcal/mol, respectively). Conversely, mutations that potentially create Watson-Crick base-pairs at this site or a purine-purine pair have significantly more deleterious effects (Table 2).

At the other side of the symmetric loop, the G49-A60 pair serves a similar role in protein binding (Figure 10(d)). While many permutations of purine-purine pairs support high affinity binding by the M domain, a G-C or U-A pair at this position strongly disrupts protein recognition. This disruption is significantly stronger than elimination of the sole protein contact to this site in 2′deoxy G49, indicating that the presence of a purine-purine pair at this position is important for presenting critical direct recognition elements in the correct structural context to the M domain. Interestingly, while transversion of this pair to an A-G pair supports protein binding fairly well ($\Delta\Delta G = -0.2$ kcal/mol), an A-G pair is never observed in the 4.5 S or 7 S RNA. Since the M domain does not directly recognize these bases, they may play another role in SRP function. This pair is near the putative signal sequence binding site (Batey *et al.*, 2000), suggesting that they could be important for signal recognition or recruitment of the SRP receptor FtsY, which should be sensitive to the bound state of the SRP.

Together, these data present a detailed picture of how each of the conserved non-canonical base-pairs within the symmetric loop promotes M-domain binding. The peripheral pairs U45-G64, C46-A63 and G49-A60 do not directly interact with protein significantly, but are required to present the other nucleotides in the correct structural context. Nucleotides within the asymmetric loop are similarly important; while the C40U does not significantly perturb this interaction, the removal of the cytosine base by hydrazine does interfere with M domain binding. Nucleotides A39, A47, G48 and C62 directly recognize the M domain through a series of highly cooperative sets of interactions, creating an exceptionally tight intermolecular interaction despite the relatively small buried surface at the interface.

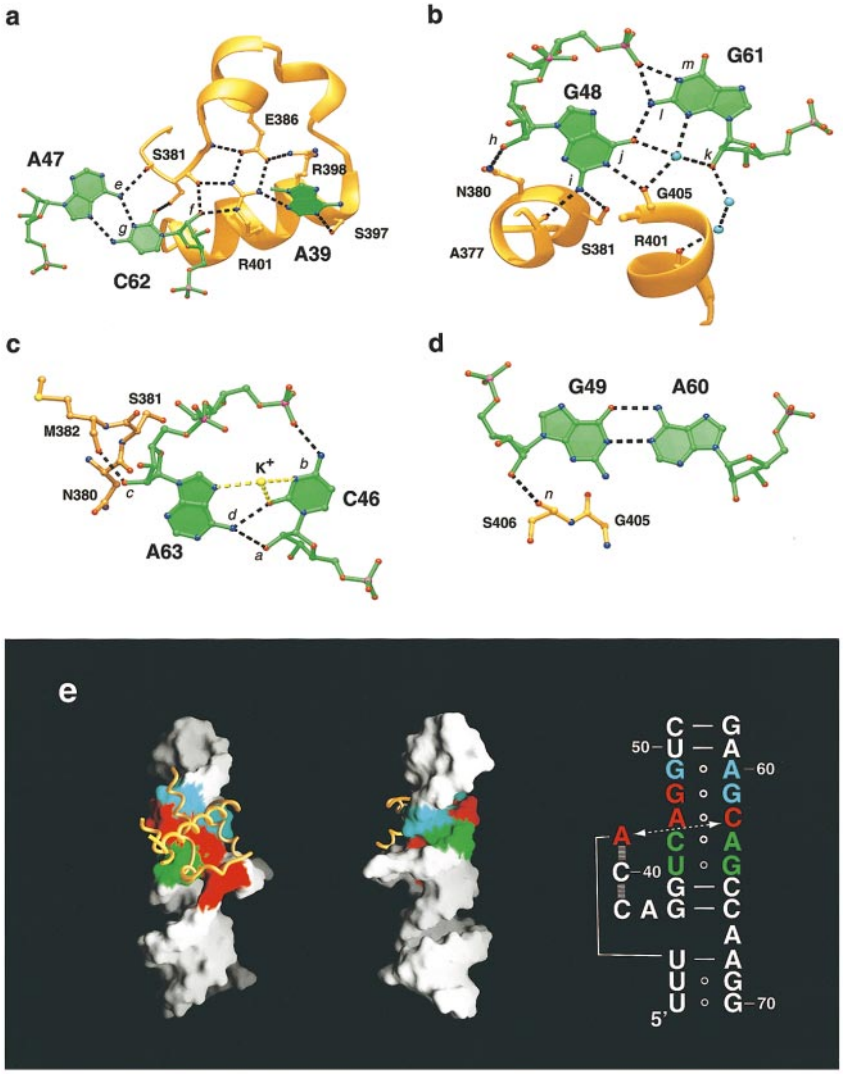

Figure 10. (a)-(d) RNA-protein interactions within the M domain-4.5 S RNA interaction. The light blue spheres in (c) are well ordered solvent molecules at the protein-RNA interface. The italicized small letters correlate to functional group deletions and base substitutions in Table 2. (e) A surface representation of the RNA emphasizing that nucleotides making large energetic contributions (red) are extruded into the minor groove to be presented to the protein (gold ribbon). Flanking nucleotides in the symmetric internal loop (light blue and green) act to place the critical bases in the correct structural context. Strikingly, on the major groove side, these two sets of nucleotides form a contiguous surface despite being separated by the A47-C62 pair. This Figure was made with GRASP (Nicholls *et al.*, 1991).

Materials and Methods

Expression and purification of Ffh-M domain

A vector for the expression of Ffh-M domain was constructed using standard cloning techniques (Sambrook *et al.*, 1989). Two oligonucleotide primers were designed to create *Nco*I and *Bam*HI restriction sites at the 5′ and 3′ end of the gene, respectively, as well as a hexahistidine tag followed by tobacco etch virus (TEV) protease cleavage at the amino terminus of the encoded protein sequence. These primers were used in a polymerase chain reaction with Vent DNA polymerase (New England Biolabs) to amplify the Ffh-M domain from *E. coli* genomic DNA using a standard amplification protocol (Innis *et al.*, 1990). The PCR reaction was subsequently desalted, digested with *Nco*I and *Bam*HI restriction enzymes, and purified on a 1% (w/v) agarose gel. The band corresponding to the insert was excised from the gel and purified using the QIAEX purification kit (Qiagen). The product was ligated into pET15b (Novagen) linearized with *Nco*I and *Bam*HI using T4 DNA ligase and incubated at 16 °C for two hours. The ligation reaction was used to transform *E. coli* DH5α (Gibco-BRL). Transformants were selected on the basis of ampicillan resistance, and the plasmids were sequenced to confirm the correct sequence. The C406S mutation was introduced into the *E. coli* M domain using mutagenic oligonucleotides in a recombinant PCR reaction (Innis *et al.*, 1990).

The Ffh-M domain expression vector was transformed into BL21(DE3)/pLysS (Novagen). Freshly transformed cells were grown in Luria Broth supplemented with 100 μg/ml ampicillan and 67 μg/ml chloramphenicol at 37 °C until the absorbance at 600 nm reached 0.35, protein expression was induced with 1 mM IPTG, and allowed to continue to grow for another four hours. The cells were harvested and lysed by three rounds of freezing in liquid nitrogen and thawing. To the lysate, 10 mM $MgCl_2$ and 100 units of RNase-free DNase were added, and allowed to incubate for 30 minutes at 37 °C to reduce the viscosity, and then centrifuged for 30 minutes at 20,000 *g* to pellet the cellular inclusion bodies, which contain the expressed protein. The supernatant was removed and the pellet was resuspended in 20 ml lysis buffer (50 mM Tris-HCl (pH 8.0), 300 mM NaCl) and 2 ml of 1.0% Triton X-100. The pellet was washed by vigorously vortexing, centrifuged for 30 minutes at

20,000 *g*, and resuspended in Qiagen denaturing buffer A (6 M guanidinium-HCl, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl, pH 8.0). The solution was loaded onto a 10 ml bed volume column containing Superflow Ni^{2+} resin (Qiagen). The column was washed with denaturing buffer B (8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl, pH 8.0) until the absorbance at 280 nm returned to baseline, followed by washing with denaturing buffer C (8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl, pH 6.3), and elution with denaturing buffer E (8 M urea, 0.1 M NaH_2PO_4, 0.01 M Tris-HCl, pH 4.5). The eluate was loaded onto a ProRPC HR 16/10 reverse phase column (Pharmacia) that was equilibrated in ddH_2O/0.65% trifluoroacetic acid. The protein was eluted by running a 30 column volume linear gradient to 50% of buffer B (acetonitrile/0.5% trifluoroacetic acid). Fractions containing Ffh-M were pooled and lyophilized to dryness.

To remove the histidine tag, approximately 50 mg of lyophilized protein was resuspended in 25 ml of a buffer containing 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 5 mM DTT, and 1 mg of recombinant TEV protease and incubated for 24 hours at 30 °C. The cleaved protein was dialyzed into SP buffer A (6 M urea, 10 mM Na-Mes, pH 6.0) and loaded onto a Mono-S column (Pharmacia) equilibrated in the same buffer. The protein was eluted using a 20 column volume linear gradient to 100% SP buffer B (SP buffer A and 1 M NaCl). The appropriate peak was collected, dialyzed into SP buffer C (6 M urea, 10 mM K-Hepes, pH 8.0), reapplied to the Mono-S column equilibrated in the same buffer. Protein was eluted with a 20 column volume linear gradient to 100% SP buffer D (buffer C plus 1 M NaCl), and concentrated in a 10,000 molecular weight cutoff CentriPrep (Amicon). The protein was refolded by exchanging the protein into ddH_2O, further concentrated to 1 mM, and stored at −80 °C.

Construction of plasmids containing RNA genes

For transcription of RNA, plasmids were constructed using a strategy described by Ferré-D'Amaré & Doudna (1996). Briefly, a gene was constructed with hammerhead ribozyme at the 5′ end and a hepatitis delta virus ribozyme at the 3′ end, each of which self-cleave during transcription to yield product RNA with homogeneous termini. The 5′ PCR primer contains a sequence comprising an *Eco*RI restriction site for cloning, a T7 RNA polymerase promoter, and a portion of the hammerhead ribozyme. The 3′ PCR primer includes a *Hin*dIII restriction site for cloning and the terminus of the RZ89 HDV ribozyme (Tanner *et al.*, 1994). In a PCR reaction, these outer primers were coupled with series of internal primers that encoded the 3′ end of the hammerhead ribozyme, the 4.5 S RNA, and the 5′ end of the HDV ribozyme. In a typical PCR reaction, the outer primers were used at a concentration of 500 nM, while the internal primers were below 1 nM.

After amplification, the proper product was cleaved with *Eco*RI and *Hin*dIII and purified on a 2% (w/v) agarose gel. This DNA insert was ligated into pUC19 linearized with *Eco*RI/*Hin*dIII and transformed into competent DH5α *E. coli* cells, and selected on LB/ampicillin plates. Individual transformants were picked and sequenced to verify that the plasmid contained the correct sequence.

Synthesis of RNA

Plasmids containing RNA genes were linearized for use in run-off transcription by cleaving at the unique *Bam*HI site. RNA was transcribed in a reaction containing 40 mM Tris-HCl (pH 8.1), 1 mM spermidine, 5 mM dithiothreitol, 0.01% Triton X-100, 30 mM $MgCl_2$, 8 mM each ribonucleotide 5′ triphosphate, 100 μg/ml linearized plasmid, 0.5 μg/ml inorganic pyrophosphate, and 0.04 mg/ml T7 RNA polymerase. The reaction was incubated for three to four hours at 37 °C, quenched by extracting with an equal volume of phenol equilibrated with Tris-HCl (pH 8.0), and the aqueous layer was precipitated by the addition of 2.5 volumes of 100% ethanol. The RNA was resuspended in formamide stop buffer and purified by electrophoresis on a 12% denaturing polyacrylamide gel in TBE buffer. RNA was visualized by UV shadowing the gel, the correct band excised, and eluted by crushing the gel and incubating in ddH_2O with shaking overnight at 4 °C. The resulting suspension was filtered to remove the gel pieces and exchanged into ddH_2O and concentrated using a pressure concentrator (Amicon) using a 10,000 molecular weight cutoff membrane. The RNA was stored at −20 °C until use.

RNA molecules containing functional group and base substitutions (LM10-A, LM10-B) were synthesized by Dharmacon Research, Inc. (Boulder, CO). The oligonucleotides were deprotected, resuspended in TE buffer, and stored at −20 °C.

MALDI-TOF analysis

For limited proteolysis with endoprotease V8 (B. Mannheim), 38 μM 4.5 S RNA and 20 μM M-domain were incubated in a buffer containing 100 mM sodium phosphate (pH 7.8), and 40 mM NaCl for one hour at 22 °C prior to the addition of a final concentration of 5 ng/μl Protease. Trypsin probing was performed in a reaction containing a buffer of 100 mM ammonium bicarbonate (pH 8.0), 10 mM $MgCl_2$, 2 mM DTT, and 5 ng/μl Protease. Aliquots at time points between five minutes and two hours were taken. One aliquot was quenched by the addition of an equal volume of 2 × SDS-PAGE loading buffer (Sambrook *et al.*, 1989) and subjected to electrophoresis on a 20% PhastGel (Pharmacia). The other aliquot (1 μl) was added to 15 μl matrix solution (a saturated solution of sinapinic acid with 2:1 1% trifluoroacetic acid (aq.:acetonitrile)). One microliter of this solution was spotted onto a sample plate, air dried, and subjected to MALDI-MS on a Voyager-SE linear time-of-flight instrument (PerSeptive Biosystems) (Cohen *et al.*, 1995).

Equilibrium binding assays

The complex between Ffh and Ffh M domain and RNA was observed using a nitrocellulose filter binding assay (Wong & Lohman, 1993). A constant concentration of 5′ end-labeled RNA (0.1 to 1 pM) was incubated with varying concentrations of protein in a buffer containing 20 mM Tris-Hepes (pH 7.5), 200 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01% Igepal C-680, and 0.1 mg/ml tRNA. End-labeled RNA (either the one-piece LM10 RNA or the two-piece LM10A/B RNA) was heated to 90 °C for one minute and cooled on ice for three minutes in TE buffer prior to use. To determine the concentration of active protein in the stock, a stoichiometric titration of the protein against 30 μM 4.5 S RNA was performed. To

ensure that the LM10A/B remains complexed at the low concentrations employed, the annealed RNA was electrophoresed on an 8 % non-denaturing polyacrylamide gel. Reaction volumes of 100 µl were incubated at room temperature for one hour prior to application to filters. Two membranes were used for each experiment: a BA-85 nitrocellulose membrane (Schleicher and Schuell) to retain protein-RNA complexes and a Hybond-N^+ (Amersham) positively charged nylon membrane to retain free RNA. These filters were soaked in wash buffer (20 mM Tris-Hepes (pH 7.5), 200 mM KCl, 10 mM $MgCl_2$) for one hour prior to being placed within a 96-well dot blot apparatus. Aliquots (40 µl) from each reaction were sequentially added to the wells, while vacuum was being applied, followed by washing with 100 µl of wash buffer. Free and bound RNA retained on the filters was quantified by phosphorimaging the filters and analyzing using ImageQuant (Molecular Dynamics). Data were analyzed using non-linear least squares analysis using Igor (Wavemetrics) using the Langmuir isotherm:

$$\Theta = \frac{[\text{protein}]}{[\text{protein}] + K_d} \quad (1)$$

where Θ is the fraction RNA bound and K_d is the apparent equilibrium binding constant. The wild-type LM10A/B complex yields an identical apparent equilibrium binding constant to LM10 RNA, indicating that the two-piece system does not disrupt the ability of M domain to recognize the RNA. Each K_d is the average of at least three independent measurements.

One feature of the nitrocellulose filter binding experiment was that the retention efficiency of mutant complexes decreased as the severity of the mutation increased. Variability of the efficiency of retention has been observed for other protein-nucleic acid interactions monitored by this method and is a recognized pitfall of the technique (Wong & Lohman, 1993). While the wild-type complex typically exhibited 60-80 % of the RNA retained on the nitrocellulose filter at saturation, this decreased to 20-30 % for many of the RNA mutations that have binding affinities between 1-10 nM. Therefore, to accurately measure the equilibrium binding constant for A60C, G48(2AP), G48P, C62U and dC62 mutants, a competition assay was used. The advantage to this method is that the same observed signal, the wild-type RNA bound to M domain, is monitored in every experiment. Binding reactions contained 32 pM body-labeled LM10 RNA, 60 pM FfhM-2(C406S), 20 mM Tris-Hepes (pH 7.5), 200 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 0.01 % Igepal C-680, 0.1 mg/ml tRNA and varying concentrations of mutant LM10A/B complex. The reactions were allowed to incubate for at least 16 hours at room temperature prior to application to filters. Data were analyzed using non-linear least squares analysis using the equation:

$$\Theta = \frac{1}{2T_t}\{K_T + (K_T/K_C)C_t + P_t + T_t - \sqrt{[K_T + (K_T/K_C)C_t + P_t + T_t]^2 - 4T_tP_t}\} \quad (2)$$

in which K_T, K_C, C_t, P_t and T_t are the apparent equilibrium dissociation constants for wild-type and competitor RNA probes, the concentration of the competitor RNA probe, the protein and labeled RNA probe, respectively (Lin & Riggs, 1972). The apparent equilibrium binding constants for several mutants were determined using both direct and competition assays, indicating that these two methods yield consistent results. Thus, the competition assay is valid for measuring apparent equilibrium dissociation constants for select RNA mutants.

RNA alkaline hydrolysis ladder selection

To create a pool of RNA that was uniformly cleaved along the backbone, approximately 50 nmol of 5′- or 3′-end-labeled 4.5 S RNA was resuspended in alkaline hydrolysis buffer (50 mM $NaHCO_3/Na_2CO_3$ (pH 9.2), 1 mM EDTA) and heated to 90 °C for eight minutes. The RNA was immediately precipitated by the addition to 300 mM sodium acetate (pH 5.2), 1 mg/ml tRNA and 2.5 volumes of 100 % ethanol and incubated overnight at −20 °C. The RNA was pelleted by centrifugation, the supernatant removed and air dried and resuspended in 1 × TE buffer. The RNA was resuspended in 1 × binding buffer (50 mM K-Hepes (pH 7.5), 200 mM KCl, 10 mM $MgCl_2$), 0.01 % Igepal C-680 (Sigma Chemical Co.), 0.1 mg/ml tRNA, and 50 pM FfhM-1 and incubated for one hour at room temperature in a total volume of 100 µl. The entire reaction mixture was passed through a nitrocellulose filter (BA85, Schliecher & Schuell) by gentle vacuum and washed twice with 300 µl of 25 mM K-Hepes (pH 7.5). In parallel, a reaction containing no protein was passed through a nitrocellulose filter as a control. To recover the RNA, the filter was soaked in a buffer containing 300 mM NaOAc (pH 5.2), 1 mM EDTA, 1.0 % SDS for one hour at room temperature. The RNA was precipitated by the addition of 2.5 volumes 100 % ethanol and incubated at −20 °C.

Samples of RNA were resuspended in formamide stop buffer and analyzed on an 8 % polyacrylamide/8 M urea denaturing gel. The alkaline hydrolysis reactions were run next to lanes containing end-labeled RNA that was subjected to cleavage by RNase T_1 under denaturing conditions to generate a guanosine sequencing lane.

Nucleotide analog interference analysis

The 4.5 S RNA was transcribed as described above, except that the appropriate amount of a 5′-*O*-(1-thio) nucleoside analog triphosphate was added to the reaction (Ryder *et al.*, 2000). The RNA was bound to FfhM-1, and free and bound RNA were separated by passing the reaction through a nitrocellulose filter. The bound RNA was recovered, precipitated and resuspended in 9 µl ddH_2O, 1 µl 1 mM I_2, and 10 µl formamide stop mix and heated to 90 °C for one minute to cleave phosphorothioate linkages in the RNA for sequencing. Reactions containing selected and unselected RNA were analyzed on an 8 % polyacrylamide/8 M urea denaturing gel.

Crystallization

Prior to crystallization, the appropriate concentrations of RNA and protein were determined using a native gel electrophoresis assay. Reactions containing 30-40 µM of RNA were titrated with increasing quantities of protein to yield 0.8-1.5:1 protein:RNA ratios. 6 µl of each reaction mixture was loaded onto a 6 % native 29:1 polyacrylamide gel in 0.5 × TBE and 2.5 mM $MgCl_2$ and electrophoresed at 5 W, constant wattage for 30 minutes and the RNA visualized by staining with ethidium bromide. A molar ratio of approximately 0.98:1 protein:

RNA as judged from the native polyacrylamide gel was used to form RNP complexes for crystallography.

Crystals were obtained using FfhM-2(C406S) and LM11 RNA. To prepare the complex for crystallization protein and RNA were mixed using individual 1 mM stock solutions. To the protein-RNA solution (approximately 300 µl volume), which would typically be completely aggregated, 1 ml of 8 M urea was added, and the resulting solution dialyzed against one liter of 10 mM K-Hepes (pH 7.5) overnight at 4 °C. The renatured complex was subsequently concentrated in a 10,000 MWCO microconcentrator (Amicon) to approximately 700 µM. Crystals were grown by vapor diffusion using the sitting drop method at 20 °C by mixing 2 µl of the protein-RNA complex and 2 µl of a reservoir solution containing 10 % isopropanol, 50 mM Na-Mes (pH 5.6), 200 mM KCl and 12.5 mM $MgCl_2$. Crystals grew to a maximum dimension of 0.1 mm × 0.1 mm× 0.5 mm within two to three days. Prior to flash-freezing the crystals, they were incubated for no more than two hours in a reservoir solution plus 30 % (v/v) MPD. Selenomethionine labeled complex was crystallized using the same methods, except that 35 mM cyclohexylbutanoyl-*N*-hydroxyethylglucamide (C-HEGA 10, Anatrace) was in both the mother liquor and cryoprotecting solutions.

Data collection and refinement

Data from three crystals were merged together to produce the data set used to solve and refine the structure. Raw images were processed with DENZO (Otwinowski & Minor, 1997) and the intensities were scaled and merged with SCALEPACK (Otwinowski & Minor, 1997). The structure of the SRP complex was determined by molecular replacement using the program AMoRe (Navaza, 1994). The search model used was derived from the previously reported SRP structure (PDB 1dul) (Batey *et al.*, 2000) by excluding the solvent and metal ions. Data in the 15.0-4.0 Å resolution range and an integration radius of 20.0 Å were used in the calculations of the rotation and translation functions. The correct solution had a correlation coefficient of 0.696 and an *R*-factor of 31.5 %. Structure refinement was started with the data set of 40.0-1.8 Å resolution using the program CNS (Brünger *et al.*, 1998) in which bulk solvent and anisotropic *B* corrections were applied. 5 % of the total reflections were excluded for free *R*-factor (R_{free}) calculations to monitor the refinement progress. The initial free and the conventional *R*-factors were 33.0 and 29.8 %, respectively. Alternate cycles of model building and several rounds of refinement led to building of the residues 1, 10-12, 21, 22 and 84 into the model. The refinement was proceeded further by extending the resolution stepwise from 1.8 to 1.52 Å. Probable water molecules (a total of 196) making hydrogen bonds with either the protein or RNA atoms were located in difference Fourier maps contoured at 3σ level. When the R_{free} of the model had dropped to 23.1 %, the structure was subjected to restrained anisotropic temperature factor refinement by using SHELX-97 (Sheldrick & Schneider, 1997). Distance, planarity, and chiral volume and anti-bumping restraints were applied from the onset of refinement. Updating the .res file option was used to pickup further water molecules into the density. A total of 287 water molecules were added to the model in stages. Water molecules were rejected when the corresponding peaks were not found in the difference Fourier map. Magnesium ions were judged by refinement of their thermal *B* factors, by their octahedral coordination geometry as well as the distance (2.1 to 2.2 Å) of their inner-sphere coordinated water molecules. Potassium ions were identified by their peak height in the electron density (>5 SDs above mean density level in each case) as well as inspection of the nature of the functional groups forming inner sphere coordinations. To verify the identities of these peaks as monovalent cations, crystals of the FfhM-2(C406S):LM11 RNA complex were soaked with thallium acetate and anomalous difference maps calculated from the X-ray diffraction data (Basu *et al.*, 1998). In the RNA, all sugar puckers are C3′-*endo* except for A39, G49, A54, G57, G58, G64, A68, A72, and C76 which are $^{3}_{2}$T, and A42 which is $^{2}_{1}$T. The validity of anisotropic *B*-value refinement at the given resolution was supported by the R_{free} statistics, which dropped from 23.5 to 19.9 %. The auxiliary program SHELXPRO was used for preparation of input parameters for the refinement, for map calculation and for model analysis. Model building and real-space refinement were carried out using the graphics program O (Jones *et al.*, 1991). In addition, the programs PROCHECK and WHATIF were used for structure evaluation. The final results of the refinement are listed in Table 1.

Atomic coordinates

Coordinates of the refined SRP core have been deposited in the RCSB PDB (accession number 1HQ1).

Acknowledgments

We thank L. Lucast and M. Talavera for technical assistance, A. Ferré-D'Amaré, J. Ippolito, and J. Murray for useful discussions, and P. Adams, L. Doherty, J. Kieft and J. Murray for critical comments on the manuscript. R.T.B. was funded through a postdoctoral fellowship from the Jane Coffin Childs Memorial Fund.

References

Abe, Y., Shodai, T., Muto, T., Mihara, K., Torii, H., Nishikawa, S.-I., Endo, T. & Kohda, D. (2000). Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. *Cell,* **100**, 551-560.

Agalarov, S. C., Sridhar Prasad, G., Funke, P. M., Stout, C. D. & Williamson, J. R. (2000). Structure of the S15,S6,S18-rRNA complex: assembly of the 30 S ribosome central domain. *Science,* **288**, 107-113.

Althoff, S., Selinger, D. & Wise, J. A. (1994). Molecular evolution of SRP cycle components: functional implications. *Nucl. Acids Res.* **22**, 1933-1947.

Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science,* **289**, 905-920.

Basu, S., Rambo, R. P., Strauss-Soukup, J., Cate, J. H., Ferré-D'Amaré, A. R., Strobel, S. A. & Doudna, J. A. (1998). A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nature Struct. Biol.* **5**, 986-992.

Batey, R. T. & Williamson, J. R. (1996). Interaction of the *Bacillus stearothermophilus* ribosomal protein S15 with 16 S rRNA: I. Defining the minimal RNA site. *J. Mol. Biol.* **261**, 536-549.

Batey, R. T., Rambo, R. P., Lucast, L., Rha, B. & Doudna, J. A. (2000). Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science,* **287**, 1232-1239.

Bernstein, H. D., Zopf, D., Freymann, D. M. & Walter, P. (1993). Functional substitution of the signal recognition particle 54-kDa subunit by its *Escherichia coli* homolog. *Proc. Natl Acad. Sci. USA,* **90**, 5229-5233.

Brown, J. W., Nolan, J. M., Haas, E. S., Rubio, M. A. T., Major, F. & Pace, N. R. (1996). Comparative analysis of ribonuclease P RNA using gene sequences from natural microbial populations reveals tertiary structural elements. *Proc. Natl Acad. Sci. USA,* **93**, 3001-3006.

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. & Warren, G. L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

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

Clemons, W. M., Jr, Gowda, K., Black, S. D., Zwieb, C. & Ramakrishnan, V. (1999). Crystal structure of the conserved subdomain of human protein SRP54 M at 2.1 Å resolution: evidence for the mechanism of signal peptide binding. *J. Mol. Biol.* **292**, 697-705.

Cohen, S. L., Ferré-D'Amaré, A. R., Burley, S. K. & Chait, B. T. (1995). Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. *Protein Sci.* **4**, 1088-1099.

Conn, G. L., Draper, D. E., Lattman, E. E. & Gittis, A. G. (1999). Crystal structure of a conserved ribosomal protein-RNA complex. *Science,* **284**, 1171-1174.

Conte, L. L., Chothia, C. & Janin, J. (1999). The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **285**, 2177-2198.

Conway, L. & Wickens, M. (1989). Modification interference analysis of reactions using RNA substrates. *Methods Enzymol.* **180**, 369-379.

Costa, M. & Michel, F. (1995). Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.* **14**, 1276-1285.

Costa, M. & Michel, F. (1997). Rules for RNA recognition of GNRA tetraloops deduced by *in vitro* selection: comparison with *in vivo* evolution. *EMBO J.* **16**, 3289-3302.

de Gier, J.-W. L., Mansournia, P., Valent, Q. A., Phillips, G. J., Luirink, J. & von Heijne, G. (1996). Assembly of a cytoplasmic membrane protein in *Echerichia coli* is dependent on the signal recognition particle. *FEBS Letters,* **399**, 307-309.

Doherty, E. A. , Batey, R. T., Masquida, B. & Doudna, J. A. (2001). A universal mode of helix packing in RNA. *Nature Struct. Biol.* **In the press**.

Draper, D. E. (1999). Themes in RNA-protein recognition. *J. Mol. Biol.* **293**, 255-270.

Ferré-D'Amaré, A. R. & Doudna, J. A. (1996). Use of *cis*- and *trans*-ribozymes to remove 5′ and 3′ heterogeneities from milligrams of in vitro transcribed RNA. *Nucl. Acids Res.* **24**, 977-978.

Ferré-D'Amaré, A. R. & Doudna, J. A. (2000). Crystallization and structure determination of a hepatitis delta virus ribozyme: use of the RNA-binding protein U1A as a crystallization module. *J. Mol. Biol.* **295**, 541-556.

Ferré-D'Amaré, A. R., Zhou, K. & Doudna, J. A. (1998a). Crystal structure of a hepatitis delta virus ribozyme. *Nature,* **395**, 567-574.

Ferré-D'Amaré, A. R., Zhou, K. & Doudna, J. A. (1998b). A general module for RNA crystallization. *J. Mol. Biol.* **279**, 621-631.

Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. & Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. *Nature,* **314**, 235-238.

Frankel, A. D. (2000). Fitting peptides into the RNA world. *Curr. Opin. Struct. Biol.* **10**, 332-340.

Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J. (1990). *PCR Protocols: A Guide to Methods and Applications*, Academic Press, Inc., New York.

Jancarik, J. & Kim, S. H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Crystallog.* **24**, 409-411.

Joachimiak, A. & Sigler, P. B. (1991). Crystallization of protein-DNA complexes. *Methods Enzymol.* **208**, 82-99.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A,* **47**, 110-119.

Jordan, S. R., Whitcombe, T. V., Berg, J. M. & Pabo, C. O. (1985). Systematic variation in DNA length yields highly ordered repressor-operator cocrystals. *Science,* **230**, 1383-1385.

Jovine, L., Hainzl, T., Oubridge, C., Scott, W. G., Li, J., Sixma, T. K., Wonacott, A., Skarzynski, T. & Nagai, K. (2000). Crystal structure of the ffh and EF-G binding sites in the conserved domain IV of *Escherichia coli* 4.5 S RNA. *Structure Fold. Des.* **8**, 527-540.

Keenan, R. J., Freymann, D. M., Walter, P. & Stroud, R. M. (1998). Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell,* **94**, 181-191.

Kleywegt, G. J. & Jones, T. A. (1994). A super position. *ESF/CCP4 Newsletter,* **31**, 9-14.

Lentzen, G., Moine, H., Ehresmann, C., Ehresmann, B. & Wintermeyer, W. (1996). Structure of 4.5 S RNA in the signal recognition particle of *Escherichia coli* as studied by enzymatic and chemical probing. *RNA,* **2**, 244-253.

Leontis, N. B. & Westhof, E. (1999). Recurrent RNA motifs: analysis at the basepair level. In *RNA Biochemistry and Biotechnology* (Barciszewski, J. & Clark, B. F. C., eds), pp. 45-61, Kluwer Academic Publishers, London.

Lin, S.-Y. & Riggs, A. D. (1972). *lac* repressor binding to non-operator DNA: detailed studies and a comparison of equilibrium and rate competition methods. *J. Mol. Biol.* **72**, 671-690.

Michel, F. & Westhof, E. (1990). Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J. Mol. Biol.* **216**, 585-610.

Miller, J. D., Wilhelm, H., Gierasch, L., Gilmore, R. & Walter, P. (1993). GTP binding and hydrolysis by the signal recognition particle during initiation of protein translocation. *Nature,* **366**, 351-354.

Miller, J. D., Bernstein, H. D. & Walter, P. (1994). Interaction of the *E. coli* Ffh/4.5 S ribonucleoprotein and

FtsY mimics that of mammalian signal recognition particle and its receptor. *Nature,* **367**, 657-659.
Murthy, V. L., Srinivasan, R., Draper, D. E. & Rose, G. D. (1999). A complete conformational map for RNA. *J. Mol. Biol.* **291**, 313-327.
Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A,* **50**, 157-163.
Newitt, J. A. & Bernstein, H. D. (1997). The N-domain of the signal recognition particle 54-kDa subunit promotes efficient signal sequence binding. *Eur. J. Biochem.* **245**, 720-729.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins: Struct. Funct. Genet.* **11**, 281-296.
Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326.
Pan, J. D. T. & Woodson, S. (1997). Folding of RNA involves parallel pathways. *J. Mol. Biol.* **273**, 7-13.
Peluso, P., Herschlag, D., Nock, S., Freymann, D. M., Johnson, A. E. & Walter, P. (2000). Role of 4.5 S RNA in assembly of the bacterial signal recognition particle with its receptor. *Science,* **288**, 1640-1643.
Pley, H., Flaherty, K. & McKay, D. (1994). Model for an RNA tertiary interaction from the structure of an intermolecular complex between a GAAA tetraloop and an RNA helix. *Nature,* **372**, 111-113.
Poritz, M. A., Strub, K. & Walter, P. (1988). Human SRP RNA and *E. coli* 4.5 S RNA contain a highly homologous structural domain. *Cell,* **55**, 4-6.
Poritz, M. A., Bernstein, H. D., Strub, K., Zopf, D., Wilhelm, H. & Walter, P. (1990). An *E. coli* ribonucleoprotein containing 4.5 S RNA resembles mammalian signal recognition particle. *Science,* **250**, 1111-1117.
Powers, T. & Noller, H. F. (1995). Hydroxyl radical footprinting of ribosomal proteins on 16 S rRNA. *RNA,* **1**, 194-209.
Price, S. R., Ito, N., Oubridge, C., Avis, J. M. & Nagai, K. (1995). Crystallization of RNA-protein complexes. I. Methods for the large-scale preparation of RNA suitable for crystallographic studies. *J. Mol. Biol.* **249**, 398-408.
Price, S. R., Oubridge, C., Varani, G. & Nagai, K. (1998). Preparation of RNA:protein complexes for X-ray crystallography and NMR. In *RNA:Protein Interactions: A Practical Approach* (Smith, C. W. J., ed.), pp. 37-74, Oxford University Press, Oxford.
Query, C. C., Bentley, R. C. & Keene, J. D. (1989). A specific 31-nucleotide domain of U1 RNA directly interacts with the 70 K small nuclear ribonucleoprotein component. *Mol. Cell Biol.* **9**, 4872-4881.
Ribes, V., Römisch, K., Giner, A., Dobberstein, B. & Tollervey, D. (1990). *E. coli* 4.5 S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell,* **63**, 591-600.
Rook, M. S., Treiber, D. K. & Williamson, J. R. (1998). Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape. *J. Mol. Biol.* **281**, 609-620.
Ryder, S. P., Ortoleva-Donnelly, L., Kosek, A. B. & Strobel, S. A. (2000). Chemical probing of RNA by nucleotide analog interference mapping. *Methods Enzymol.* **317**, 92-109.
Saenger, W. (1984). *Principles of Nucleic Acid Structure. Springer Advanced Texts in Chemistry* (Cantor, C. S., ed.), Springer-Verlag, New York.
Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schatz, G. & Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science,* **271**, 1519-1526.
Schmitz, U., Freymann, D. M., James, T. L., Keenan, R. J., Vinayak, R. & Walter, P. (1996). NMR studies of the most conserved RNA domain of the mammalian signal recognition particle (SRP). *RNA,* **2**, 1213-1227.
Schmitz, U., Behrens, S., Freymann, D. M., Keenan, R. J., Lukavsky, P., Walter, P. & James, T. L. (1999). Structure of the phylogenetically most conserved domain of SRP RNA. *RNA,* **5**, 1419-1429.
Scott, W. G., Finch, J. T., Grenfell, F., Fogg, J., Smith, T., Gait, M. J. & Klug, A. (1995). Rapid crystallization of chemically synthesized hammerhead RNAs using a double screening procedure. *J. Mol. Biol.* **250**, 327-332.
Serganov, A. A., Masquida, B., Westhof, E., Cachia, C., Portier, C., Garber, M., Ehresmann, B. & Ehresmann, C. (1996). The 16 S rRNA binding site of *Thermus thermophilus* ribosomal protein S15: comparison with *Escherichia coli* S15, minimum site and structure. *RNA,* **2**, 1124-1138.
Sheldrick, G. M. & Schneider, T. R. (1997). SHELXL: high-resolution refinement. *Methods Enzymol.* **277**, 319-343.
Silverman, S. K. & Cech, T. R. (1999). Energetics and cooperativity of tertiary hydrogen bonds in RNA structure. *Biochemistry,* **38**, 8691-8702.
Struck, J. C. R., Toschka, H. Y., Specht, T. & Erdmann, V. A. (1988). Common structural features between eukaryotic 7 SL RNAs, eubacterial 4.5 S RNA and scRNA and archaebacterial 7 S RNA. *Nucl. Acids Res.* **16**, 7740.
Sundberg, E. J. & Mariuzza, R. A. (2000). Luxury accommodations: the expanding role of structural plasticity in protein-protein interactions. *Structure Fold. Des.* **8**, R137-R142.
Szewczak, A. A. & Cech, T. C. (1997). An RNA internal loop acts as a hinge to facilitate ribozyme folding and catalysis. *RNA,* **3**, 838-849.
Tanner, N. K., Schaff, S., Thill, G., Petit-Koskas, E., Crain-Denoyelle, A. & Westhof, E. (1994). A three-dimensional model of hepatitis delta virus ribozyme based on biochemical and mutational analyses. *Curr. Biol.* **4**, 488-498.
Turner, D. H., Sugimoto, N., Kierzek, R. & Dreiker, S. D. (1987). Free energy increments for hydrogen bonds in nucleic acid base-pairs. *J. Am. Chem. Soc.* **109**, 3783-3785.
Walter, P. (1995). Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Harvey Lect.* **91**, 115-131.
Westhof, E. & Fritsch, V. (2000). RNA folding: beyond Watson-Crick pairs. *Structure Fold. Des.* **8**, R55-R65.
Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W. & Ramakrishnan, V. (1999). A detailed view of a ribosomal active site: the structure of the L11-RNA complex. *Cell,* **97**, 491-502.
Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr, Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T. & Ramakrishnan, V. (2000). Structure of the 30 S ribosomal subunit. *Nature,* **407**, 327-339.

Wong, I. & Lohman, T. M. (1993). A double-filter method for nitrocellulose-filter binding: application to protein-nucleic acid interactions. *Proc. Natl Acad. Sci. USA,* **90**, 5428-5432.

Wood, H., Luirink, J. & Tollervey, D. (1992). Evolutionarily conserved nucleotides within the *E. coli* 4.5 S RNA are required for association with P48 *in vitro* and for optimal functional *in vivo*. *Nucl. Acids Res.* **20**, 5919-5925.

Zheng, N. & Gierasch, L. M. (1997). Domain interactions in *E. coli* SRP: stablization of the M domain by RNA is required for effective signal sequence modulation of NG domain. *Mol. Cell,* **1**, 79-87.

Zopf, D., Bernstein, H. D., Johnson, A. E. & Walter, P. (1990). The methionine-rich domain of the 54 kD protein subunit of the signal recognition particle contains an RNA binding site and can be cross-linked to a signal sequence. *EMBO J.* **9**, 4511-4517.

Zweib, C. & Larsen, N. (1997). The Signal Recognition Particle database (SRPDB). *Nucl. Acids Res.* **25**, 107-108.

Edited by D. Draper

(Received 27 October 2000; received in revised form 13 December 2000; accepted 15 January 2001)