Structure of tRNA pseudouridine synthase TruB and its RNA complex: RNA recognition through a combination of rigid docking and induced fit

Hu Pan, Sanjay Agarwalla, Demetri T. Moustakas, Janet Finer-Moore, and Robert M. Stroud*

Department of Biophysics and Biochemistry, University of California, San Francisco, CA 94143

Contributed by Robert M. Stroud, August 29, 2003

RNA pseudouridine synthase, TruB, catalyzes pseudouridine formation at U55 in tRNA. This posttranscriptional modification is almost universally conserved and occurs in the T arm of most tRNAs. We determined the crystal structure of *Escherichia coli* TruB apo enzyme, as well as the structure of *Thermotoga maritima* TruB in complex with RNA. Comparison of the RNA-free and -bound forms of TruB reveals that this enzyme undergoes significant conformational changes on binding to its substrate. These conformational changes include the ordering of the ''thumb loop,'' which binds right into the RNA hairpin loop, and a 10° hinge movement of the C-terminal domain. Along with the result of docking experiments performed on apo TruB, we conclude that TruB recognizes its RNA substrate through a combination of rigid docking and induced fit, with TruB first rigidly binding to its target and then maximizing the interaction by induced fit.

RNA modification is a posttranscriptional process whereby certain nucleotides are altered after their initial incorporation into an RNA chain. Pseudouridine (Ψ) is the most abundantly found modification in RNA (1). It is found in most RNAs, including transfer, ribosomal, transfer-messenger, small nuclear, and small nucleolar guide RNAs. Many Ψ residues are highly conserved and appear to be confined to the functionally important part of RNA. For example, Ψs are clustered within the peptidyl transferase center of the ribosome (2), are conserved within regions of small nuclear RNAs that are involved in RNA–RNA interactions (3), and have been implicated in spliceosome assembly (4).

The most obvious structural effect of Ψ formation is the creation of a new hydrogen bond donor N–H, located where C5 used to be. It has been shown that pseudouridylation has the effect of enhancing local RNA stacking in both single-stranded and duplex regions, resulting in increased conformational stability (5, 6). Certain genetic mutants lacking specific Ψ residues in tRNA or rRNA exhibit difficulties in translation, display slow growth rates, and fail to compete effectively with wild-type strains in mixed culture (7–10). All of the evidence indicates that Ψs play an important and critical role in RNA-mediated cellular processes. The precise role of this modification, however, remains unclear.

Ψ synthases catalyze the isomerization of U to Ψ. The general mechanism for these enzymes requires a nucleophilic attack on C6 of the uracil ring in the target U by a universally conserved Asp residue, which leads to the breakage of the glycosidic bond, followed by a rotation of the uracil ring and reattachment of the C5 atom of the uracil to C1′ of the ribose (Fig. 1*A*) (11). In prokaryotes, pseudouridylation is mediated by a set of enzymes that are site- or region-specific; each of these enzymes specifies the formation of just one or sometimes several Ψs in RNA. Although the reaction catalyzed by each of these enzymes is the same, the substrate specificity varies from simple stem–loop structures to larger and more complex RNA.

TruB catalyzes Ψ55 formation in tRNAs. This modification is almost universally conserved among different species and tRNAs (8, 12). It has been shown that cells lacking TruB are unable to compete with wild-type cells effectively (8). In another study, the TruB-null mutant exhibited a defect in survival during rapid transfer from 37° to 50°C; this indicated that TruB-effected Ψ55 formation in tRNA contributes to thermal stress tolerance in *Escherichia coli*, possibly by optimizing the stability of the tRNA at high temperatures (13).

The crystal structure of a complex between *E. coli* TruB and a fragment of tRNA, solved by Hoang and Ferre-D'Amare (14), provided a glimpse into the molecular details of how the enzyme binds the T loop of tRNA and positions the target U in the active site. However, structures of the unbound protein and of protein–RNA intermediate complexes are required to define the mechanism by which TruB assembles into a catalytically competent complex with its cognate tRNA. We report the crystal structures of *E. coli* TruB in the absence of RNA, and of *Thermotoga maritima* TruB (*tm*TruB) in complex with a 17-base stem–loop of RNA, an analogue of a tRNA T arm. Comparison of the apo and RNA-bound forms of TruB provides insight into the structural basis for RNA recognition and specificity.

Materials and Methods

Protein Expression and Purification. The *E. coli truB* clone in pET-15b was a gift from J. Ofengand (University of Miami, Miami) (12). The *Nde*I–*Bam*HI insert containing the *truB* gene from this construct was subcloned into *Nde*I–*Bam*HI sites of pET-11a (Novagen) to obtain a construct without the $(His)_6$ tag and was used for protein expression. The protein was expressed in *E. coli* BL21(DE3) cells and was purified by successive chromatographic steps on DEAE-Sepharose, hydroxyapatite, heparin-Sepharose, and Q-Sepharose columns. *T. maritima truB* gene was PCR-amplified from its genomic DNA and cloned into *Nde*I–*Not*I sites of plasmid pET-28b (Novagen). The resulting construct had an N-terminal $(His)_6$ tag. Protein was expressed in *E. coli* BL21(DE3) cells and was purified by metal affinity chromatography using Talon resin (Clontech). The $(His)_6$ tag was cleaved by using thrombin and the protein was further purified on a hydroxyapatite column.

Crystallization and Data Collection. *E. coli* TruB protein was concentrated to 3 mg/ml in 10 mM Tris (pH 7.5)/2 mM EDTA/2 mM DTT. Crystals were grown by the vapor-phase diffusion method by mixing with equal volumes of protein and a well solution containing 8–10% (wt/vol) polyethylene glycol (PEG) 4,000, 0.2 mM magnesium sulfate, and 100 mM sodium cacodylate (pH 6.5). Selenomethionine-TruB crystals were obtained in similar conditions with 6% of PEG 4,000.

RNA oligonucleotide was obtained from Dharmacon. *tm*TruB was concentrated to 2.5 mg/ml in 10 mM Tris (pH 7.5)/2 mM

Abbreviations: Ψ, pseudouridine; *tm*TruB, *Thermotoga maritima* TruB.

Data deposition: The atomic coordinates of *Escherichia coli* TruB and the *tm*TruB–RNA complex have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1R3E and 1R3F).

*To whom correspondence should be addressed at: S412C University of California–Genentech Hall, 600 16th Street, San Francisco, CA 94143-2240. E-mail: stroud@msg.ucsf.edu.

www.pnas.org/cgi/doi/10.1073/pnas.2135585100

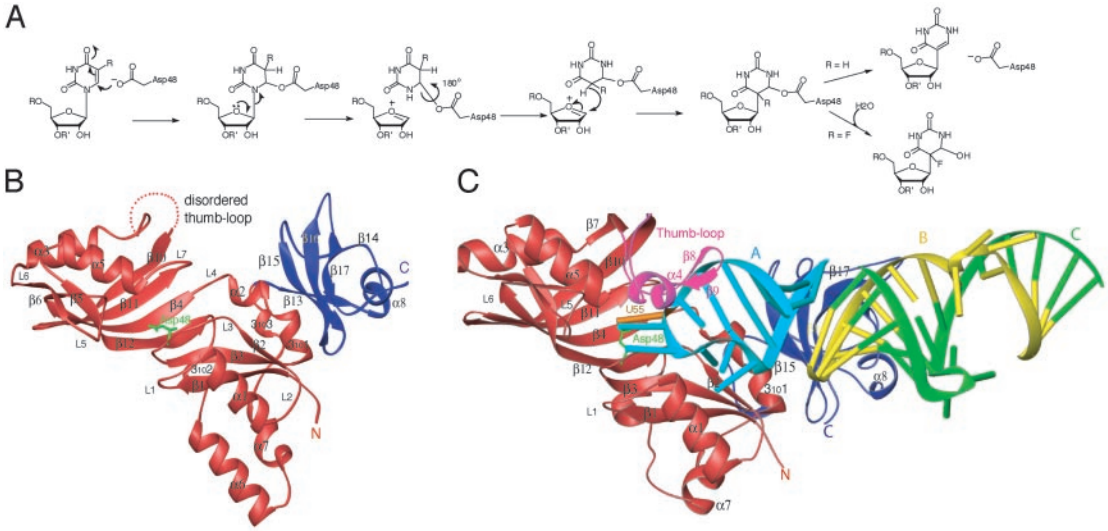

Fig. 1. Mechanism of Ψ synthases and overall structure of TruB with and without RNA bound. (*A*) Mechanism proposed for Ψ synthases (25). R represents H or F. If R = F, a stable covalent intermediate is expected. However, in the crystal structure, a hydrolyzed product analogue was observed. (*B*) Overall structure of apo TruB. The catalytic domain is shown in red, and the C-terminal domain is shown in blue. Residues 124–152 are disordered and indicated by a dotted line. Catalytic residue Asp-48 is shown in green. (*C*) Overall structure of the *tm*TruB–RNA complex. The color scheme for the protein part is the same as in *B*. The thumb loop is in magenta. The three pieces of RNA are as follows: molecule A in cyan, molecule B in yellow, and molecule C in green. U55 is shown in orange. These and other similar illustrations were generated by using RIBBONS (26).

EDTA/2 mM DTT and mixed with the RNA at a molar ratio of 1:1.6. Crystals were grown by vapor-phase diffusion. The protein–RNA mixture was incubated at room temperature for 30 min and then mixed with an equal volume of reservoir solution [1.9 M ammonium sulfate/5 mM $MgCl_2$/100 mM sodium cacodylate (pH 6.5)]. Selenomethionine–*tm*TruB–RNA cocrystals were obtained in similar conditions with 1.6 M ammonium sulfate.

Apo TruB crystals were frozen in mother liquor with 10% glycerol added. Three-wavelength anomalous dispersion data were collected to 1.85 Å at 100 K from a single selenomethionine TruB crystal at the Advanced Light Source (Beamline 5.0.2) by using an ADSC charge-coupled device detector. Apo TruB crystals grew in space group $P2_1$ with cell dimensions a = 48.03 Å, b = 71.01 Å, c = 48.67 Å, and β = 117.87°, n = 1. Three-wavelength anomalous dispersion data were also collected from a selenomethionine–*tm*TruB–RNA complex crystal to 2.1 Å at 100 K at the Advanced Light Source (Beamline 8.3.1). Crystals were frozen in mother liquor with 15% glycerol added. The crystals belonged to space group $C2$ with cell dimensions a = 98.53 Å, b = 159.36 Å, c = 44.53 Å, and β = 97.81°, n = 1. Diffraction intensities were integrated and reduced by using the program DENZO and were scaled by using SCALEPACK (15). A summary of the crystallographic data is shown in Table 1.

Multiwavelength Anomalous Dispersion Phasing, Model Building, and Refinement. Anomalous difference Patterson maps were calculated for the selenomethionine data of *E. coli* apo TruB, from which a heavy-atom positional search gave solutions for five of the six selenium atoms. Heavy-atom positional refinement and subsequent phase calculations were conducted by using CNS (16). Phase improvement by means of iterative density modification resulted in well defined 1.85-Å electron density maps. Subsequent refinement was carried out by using CNS, resulting in an R factor of 22.8% (R_{free} = 25.7%). The final model includes residues 8–123, 153–314, and 176 water molecules. A 29-amino acid segment containing residues 124–152 was not observed in the electron density. An analysis of the geometry shows that all parameters are well within the expected values at this resolution (Table 1).

Although both apo TruB and *E. coli* TruB–RNA complex structures were used as search models for a molecular replacement solution for the *tm*TruB–RNA complex structure, neither yielded a satisfactory solution. Therefore, multiwavelength anomalous dispersion data were collected to obtain experimental phases. All four selenium atoms were located by CNS. An interpretable electron density map was obtained after solvent flattening and phase extension to 2.1 Å against native data. The model was built and refined in CNS to an R factor of 22.3% (R_{free} = 27.3%). The final model includes all of the protein residues 10–315 (the *E. coli* TruB numbering is used), three RNA molecules, with 17 bases in each molecule, and 187 water molecules. Analysis of the geometry shows that all parameters are well within expected values at this resolution (Table 1).

Molecular Docking. DOCK 4.0.1 (17) was used to generate and evaluate a large number of possible orientations of RNA in the active site of apo TruB. The crystal structure of the apo TruB monomer and the 17-base RNA stem–loop from the structure of yeast $tRNA^{Phe}$ (Protein Data Bank ID code 1EHZ) (18) were used as a starting point for the docking. The apo TruB structure was prepared for docking by using SYBYL 6.5 (Tripos Associates, St. Louis). The active site of TruB was characterized by creating a solvent-accessible molecular surface over all residues within 10 Å of Asp-48 by using a molecular surface program, DMS. Next the DOCK accessory SPHGEN (17) was used to generate a set of 460 spheres over the surface. To remove biologically irrelevant structures, the set of 10,000 orientations was filtered to keep only those structures that satisfied a 10-Å distance constraint between the C6 of U55 and the Oδ of Asp-48. The 460 remaining structures were then ranked according to their contact score, and the top 100 structures were selected for visual inspection.

Results and Discussion

The Overall Structure of Apo TruB. Apo TruB has an α/β fold that is common to many ribosomal proteins as well as RNA-binding protein domains (19). The structure of TruB consists of two distinct domains, a large catalytic domain and a small C-terminal domain (Fig. 1*B*). The universally conserved Asp-48 is located at the edge of the RNA-binding cleft. The overall fold of the apo structure is similar to that of the *E. coli* TruB–RNA complex (14) but with one striking difference: a 29-aa segment (residues

Table 1. Statistics for the crystallographic data of apo TruB and *tm*TruB–RNA complex

	Apo TruB			*tm*TruB–RNA complex			
	Peak	Inflection	Remote	Native	Peak	Inflection	Remote
Data collection							
λ, Å	0.97926	0.97942	0.9610	0.9730	0.9796	0.9800	1.1271
Resolution, Å	1.85	1.85	1.85	2.1	2.9	2.9	2.8
Mosaicity, °	0.381	0.386	0.385	0.55	0.85	0.90	0.80
No. of reflections	426,177	428,770	426,647	525,236	632,468	630,898	592,567
No. of unique reflections	24,780	24,925	24,806	38,642	15,120	15,127	16,613
Completeness, % (last shell)	100 (100)	100 (100)	100 (100)	98.4 (97.5)	100 (100)	100 (100)	100 (100)
$I/\sigma(I)$ (last shell)	16 (5.1)	20.6 (3.9)	23.7 (4.9)	58.8 (45.3)	49.4 (13.0)	54.4 (14.3)	70.0 (20.7)
R_{merge}, % (last shell)	9.5 (45.3)	7.1 (57.9)	6.9 (47.0)	7.6 (73.1)	12.7 (66.2)	11.3 (61.7)	10.9 (71.5)
Multiwave anomalous dispersion phasing							
R_{cullis} (acentric/centric)		0.51/0.44			0.82/0.66		
Phasing power (acentric/centric)		2.4/2.7			1.0/1.1		
Figure of merit		0.65			0.40		
Figure of merit dm		0.97			0.93		
Refinement and model quality							
No. of reflections		46,517			34,732		
Resolution, Å		40–1.85			40–2.10		
No. of protein atoms		2,158			2,469		
No. of RNA atoms		0			1,068		
No. of water molecules		176			293		
$R_{crystal}$, %		22.8			22.3		
R_{free}, %		25.7			27.3		
rms deviation bond lengths, Å		0.006			0.0065		
rms deviation bond angles, °		1.29			1.36		
Ramachandran plot, %							
Most favored		92.4			83.6		
Favored		7.6			13.4		
Allowed		0			3.0		
Disallowed		0			0		

124–152) of the apoprotein corresponding to the thumb-loop region is disordered in the crystals and was not observed in the electron density map.

The Overall Structure of the *tm*TruB–RNA Complex. A 17-base stem–loop RNA with a sequence corresponding to the T arm (nucleotides 49–65) of yeast $tRNA^{Phe}$ was used as a substrate of TruB for crystallization. This 17-base T arm was shown to contain all of the essential features for RNA recognition and has a k_{cat}/K_m value almost identical to that of the intact tRNA (20). In an attempt to obtain a covalent TruB–RNA complex, the target U55 was modified to a 5-fluorouridine.

The overall fold of *tm*TruB is in general agreement with that of the *E. coli* TruB–RNA complex with a Cα rms difference of 1.48 Å (Fig. 1*C*). Although the sequence identity between the two species is ≈30%, the core regions of the two structures are almost identical, with many conserved residues in the RNA-binding cleft aligned in the same positions. There are two major differences between the structures of *tm*TruB and *E. coli* TruB. First, *tm*TruB lacks the helix corresponding to α6 of *E. coli* TruB, resulting in a more compact fold, which is 12 Å shorter in one dimension. Second, on the C-terminal domain of *tm*TruB, a loop between β15 and β16 is much longer (nine residues longer) and partially disordered compared with that of *E. coli* TruB.

There are three RNA molecules per protein–RNA complex. RNA molecule A has a stem–loop structure and binds into the catalytic cleft of TruB as expected. RNA molecules B and C form a 14-bp duplex with a three-base bulge in each strand and stack end-to-end against the RNA molecule A (Fig. 1*C* and 2*A*). This duplex mimics the acceptor stem of the intact tRNA. Similarly, the 22-base RNA fragment in the crystal structure of the *E. coli* TruB complex had dimerized by base-pairing the overhanging 5′ G residue of one RNA with the same residue of a symmetry mate, resulting in the extension of the RNA helix (14). The end of the RNA duplex that is not stacked on RNA molecule A interacts with a number of residues on α3 of a symmetry-related neighboring protein molecule. Phe-113′ forms stacking interactions with G65 of the RNA molecule C, the side chain of Arg-109′ hydrogen bonds to O2 of C49, and Glu-106′ hydrogen bonds to N4 of C49 of the RNA molecule B. In fact, these interactions are the only crystal contacts between the layers along the *b* axis, and the whole RNA stem–loop duplex structure serves as a set of stitches linking protein molecules together.

Based on secondary structure prediction of RNA and the NMR structure of the T arm of yeast $tRNA^{Phe}$ (21), the 17-base RNA used for crystallization forms a single stable stem–loop species in solution. Self-complementary small RNAs usually crystallize in the form of RNA duplexes, stacked end to end (22). However, hairpin forms are usually observed in the protein–RNA complex crystals. It is rather unique that, in our structure, two completely different RNA folds (hairpin and duplex) of an RNA molecule are present in the same crystal form. The crystal-packing forces probably drove the conversion of the stem–loop into the duplex conformation seen in the crystal structure. This special RNA molecule arrangement also illustrates the significant influence of crystal-packing forces on nucleic acids and the inherent flexibility of RNA oligomers.

Conformational Change on RNA Binding. Comparison of the structures of the *tm*TruB–RNA and *E. coli* TruB–RNA complexes with that of *E. coli* apo TruB reveals significant conformational changes induced by the RNA substrate. The most prominent change is that a 29-aa segment (residues 124–152) of the protein corresponding to the thumb-loop region, which is totally disor-

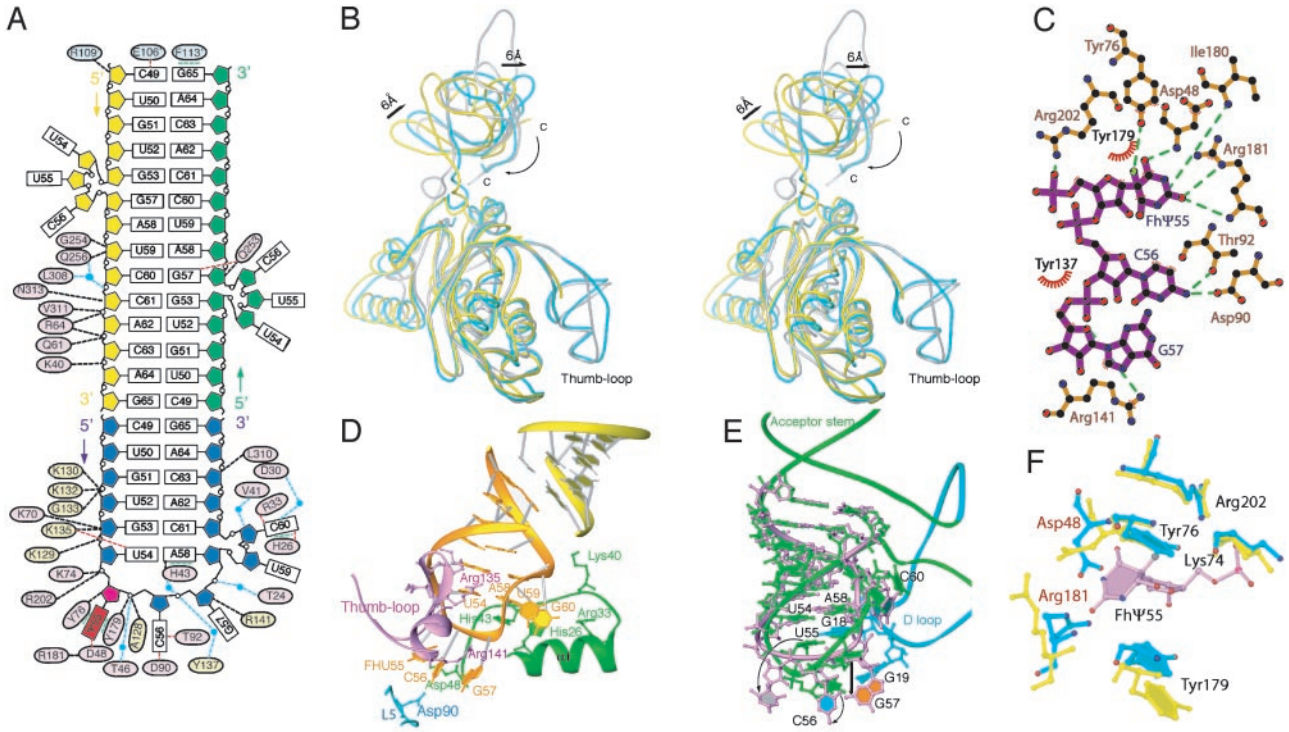

Fig. 2. Protein–RNA interactions and conformational changes on RNA binding. (*A*) Schematic representation of protein–RNA interactions. Ψ55 is highlighted with red, and the backbones of the RNA molecules A, B, and C are shown in blue, yellow, and green, respectively. Water-mediated contacts are shown in dashed blue lines, directed interactions to the RNA bases are shown in dashed red lines, interactions to the RNA backbones are shown in dashed black lines, and stacking interactions are shown in dashed green lines. (*B*) Superposition of *E. coli* apo TruB (yellow), *tm*TruB–RNA complex (gray), and *E. coli* TruB–RNA complex (cyan) (14). (*C*) Interactions between the three flipped-out bases and the protein residues. Hydrogen bonds are shown as green dashed lines, and hydrophobic interactions are shown as red ''eyelashes.'' The figure was generated by using the program LIGPLOT (27). (*D*) Detailed view of the specific protein–RNA interactions. (*E*) Superposition of the T arm bound to TruB (lavender) with corresponding residues from the structure of $tRNA^{Phe}$ (acceptor stem shown in green, D loop shown in blue). (*F*) Superposition of the catalytic residues of apo TruB (yellow) and the corresponding residues from *tm*TruB–RNA complex (blue).

dered in the apoprotein, becomes ordered on RNA binding. It is possible that the thumb loop is unstructured in the RNA-free state, or it oscillates between free and bound forms, and the RNA binding shifts the equilibrium to the bound form. The ordered thumb loop forms two short β-strands (β8 and β9), and a short α-helix (α4) protrudes into the major groove of the RNA loop and makes direct contacts with the RNA bases on the loop through hydrogen bonds and hydrophobic interactions (Fig. 2 *A* and *D*). The sequence alignment among the TruB family (data not shown) shows that this region is highly conserved, indicating that the thumb-loop region plays an important role in RNA recognition and substrate specificity.

Another major difference between the RNA-free and RNA-bound forms of TruB lies in the C-terminal domain. The C-terminal domain moves as a rigid body toward RNA on complexation. This movement results in the maximum displacement of 6 Å at the tip of the domain. The superposition of the two RNA complex structures together with the apo structure of *E. coli* TruB is shown in Fig. 2*B*. The 17-base RNA stem–loop in the *tm*TruB cocrystal structure was not long enough to make contact with the C-terminal domain; instead, this domain makes contacts with the extended RNA helix formed by the RNA duplex. The backbone atoms of the stem–loop and the first five base pairs of the duplex in our complex structure overlap well with the T stem–loop and the equivalent part of the acceptor stem of the yeast $tRNA^{Phe}$ (18) with an rms deviation of 1.2 Å. Therefore, the stem–loop and RNA duplex together mimic the intact tRNA and show how tRNA would bind to TruB. Residues Gln-253, Gly-254, and Gln-256, located on the linker of the two domains (the loop between α7 and β14), make hydrogen bonds with the RNA backbone. The C-terminal loop swings down ≈16 Å and hydrogen bonds to the RNA through residues Leu-308, Leu-310, Val-311, and Asn-313 (Fig. 2 *A* and *B*). The interactions between the C-terminal domain and the RNA are nonspecific, because all of the hydrogen bonds are made through the RNA backbone.

Another notable conformational change that occurs on RNA binding is the movement of loop L5 closer to the active site, where $O^{\delta 1}$ of the highly conserved residue Asp-90 accepts a hydrogen bond from N4 of one of the flipped-out bases, C56 (Fig. 2*C*). Taken together, the structural differences between the RNA-free and -bound forms of TruB suggest that binding of RNA substrate induces large conformational changes of the protein to ensure proper binding of a substrate.

RNA Recognition. The interactions between *tm*TruB and RNA are extensive; ≈3,900 $Å^2$ of surface area is buried as a result of protein–RNA interactions. Protein–RNA interactions are schematically represented in Fig. 2*A*. A major contribution to these interactions comes from the thumb-loop region, which anchors the RNA loop into the active site cleft, resulting in an RNA loop that is almost completely enclosed by the protein. A number of highly conserved residues, Ala-128, Lys-129, Lys-130, Gly-132, Arg-135, and Arg-141, from this thumb-loop region form extensive direct hydrogen bonds with residues U54, C56, and G57 of the RNA loop and residues U52 and G53 of the RNA stem (Fig. 2 *A* and *D*). This thumb loop, along with loops L3 and L7, helps clamp the RNA loop into the active site. Five residues from the α1–β2 region also make a significant number of contacts with five of the seven RNA loop residues (Fig. 2 *A* and *D*).

In contrast to the protein, the RNA retains its shape on

binding compared with the equivalent portion of the free $tRNA^{Phe}$ structure, except that the target base, U55, as well as its adjacent bases C56 and G57, flip out of the RNA loop and are involved in numerous specific interactions with the protein rather than other RNA bases. The superposition of the 17-base RNA in the *tm*TruB–RNA cocrystal structure and the equivalent portion of the free $tRNA^{Phe}$ structure is shown in Fig. 2*E*. In the intact tRNA, Ψ55 makes a hydrogen bond with G18 of the D loop, C56 base-pairs with G19, and G57 is sandwiched between G18 and G19. U59 and C60 also form a number of hydrogen bonds with the sugar-phosphate backbone of the D loop. On forming the protein–RNA complex, these T-loop and D-loop interactions must be disrupted to flip out U55, C56, and G57 for recognition and catalysis. In the *tm*TruB–RNA cocrystal structure, U55 is positioned in the active site and makes extensive base-stacking and hydrogen-bond interactions with many conserved residues of TruB (details in the next section). The two adjacent bases, C56 and G57, are stabilized by only protein–RNA interactions. C56 forms direct hydrogen bonds with Asp-90 and Thr-92 from loop L5 and Tyr-179 from α5. G57 forms hydrogen bonds with Arg-141 on the thumb loop and Thr-46 on β3 (Fig. 2*C*).

The loop in the T arm of bacterial tRNAs has the consensus sequence U54-U55-C56-A/G57-A58-X59-C/U60 (23). Kinetic studies with T-arm mutants have also determined the favored nucleotides for TruB recognition at each position in the T loop (20, 24). The interactions observed between the RNA and protein in the cocrystal structure rationalize the requirement of this consensus-loop sequence for recognition by TruB. U54 and A58 are two conserved bases in all tRNAs and form a reverse-Hoogsteen base pair that stacks over the G53·C61 base pair. TruB cannot catalyze Ψ formation in tRNA mutants with any substitution at either the U54 or A58 position (20, 24). In the crystal structure, a hydrogen bond is observed between O4 of U54 and the NH1 of Arg-135 on the thumb loop. A58 stacks with a completely conserved His-43. Two hydrogen bonds are also formed between His-43 and A58 (ND1···N6 and O···N6). tRNA mutants with A or G substitution at position 56 are not substrates for TruB, whereas a C56U mutant has a 40-fold lower k_{cat} than the wild-type substrate (20). In our structure, N4 of C56 is a hydrogen bond donor to the conserved Asp-90 and a hydrogen bond acceptor to Thr-92. A uracil substitution at this position would disrupt the hydrogen bonds with Asp-90. This hydrogen bond appears to be important for the proper orientation for the substrate RNA in the active site, and lower k_{cat} observed with the C56U mutant is presumably due to the distortion of the bound conformation of the neighboring target U55. At the positions of G57 and U59, mutations are very well tolerated with no more than 3-fold lower k_{cat}/K_m for any mutant (20). In the structure, hydrogen bonds are observed between the protein residues and the RNA backbone at the positions of G57 and U59 but not with the bases; substitution with any other three bases at these two positions can be tolerated without steric clashes. Finally, mutations at C60 can also be well tolerated. Arg-33 forms a hydrogen bond with O2 of C60 in the structure. It can be replaced by a uracil without disturbing the hydrogen bond or introducing steric clashes with its surrounding protein residues. Furthermore, the T stem portion of the T arm does not contain any conserved residue. Protein residues make hydrogen bonds with the backbone of the T stem but do not interact with the bases. This observation suggests that identities of individual bases in the T stem are not important for recognition by TruB, and that the T stem is required mainly to maintain a stable stem–loop structure. Previous kinetic studies with T-arm mutants corroborate these findings (20, 23). In summary, Ψ55 synthase recognizes both tRNA tertiary structure and specific sequences surrounding the nucleotide to be modified. Specific recognition of the consensus of the loop portion of RNA appears to be essential, whereas TruB does not have any requirement for a specific sequence in the stem region of RNA.

The Conformational Changes of the Conserved Residues in the RNA-Binding Cleft. The active site cleft is lined with many positively charged or hydrophobic residues, and many of these residues are highly conserved. A close inspection of these conserved residues in the RNA-binding cleft of the RNA-free and -bound structures of *E. coli* TruB and *tm*TruB reveals important differences in this region. The universally conserved catalytic residue, Asp-48, has two different conformations in RNA-free and -bound forms (Fig. 2*F*). In the apo form, Asp-48 accepts a single hydrogen bond from Arg-181, which is a highly conserved residue in all Ψ synthase families, and points outward from where U55 binds, leaving a relatively large cavity in the active site. In the RNA-bound form, Asp-48 forms a salt bridge and two hydrogen bonds with Arg-181 and points toward the U55. Thus, this RNA substrate-induced conformational switch of Asp-48 and Arg-181 is important for bringing the catalytic groups into proper alignment for catalysis and closing up the active site cavity. There are also other conformational differences between the residues lining the active site cleft in the free protein and RNA complexes. Many of the side chains of the active-site residues move slightly toward the RNA bases in the RNA-bound form (Fig. 2*F*). U55 is sandwiched between the aromatic rings of Tyr-179 and Tyr-76. Interestingly, Tyr-179 is positioned further from the U55 binding site in the apo structure, leaving a relatively big U55-binding pocket. Once the U55 flips into this cavity, Tyr-179 moves toward the RNA base and clamps the uracil ring of U55 into a productive orientation for catalysis.

Two possible mechanisms have been suggested for the catalysis of Ψ formation (24). In one mechanism the attack of the aspartate nucleophile on the C6 of U was proposed, whereas another mechanism proposed the attack of aspartate on the C1′ of the ribose sugar. Subsequent studies on pseudouridine synthase I (TruA) favored the first of these mechanisms (Fig. 1*A*) (25). U55 in the RNA substrate used in our study was modified

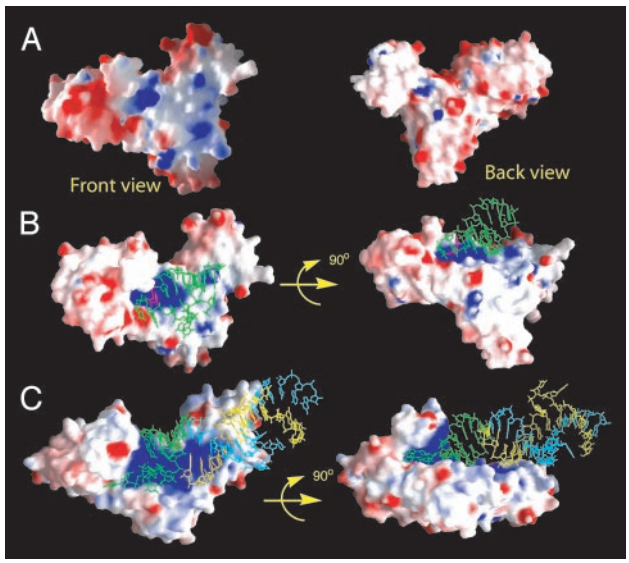

Fig. 3. Electrostatic surface potential of TruB. The protein surfaces are colored by their electrostatic potentials, from red (−10 *kT*) to blue (+10 *kT*). (*A*) The front and back surfaces of apo TruB. (*B*) The front surface of apo TruB with 17-base RNA (shown in green with U55 in magenta) docked into the active site and the view 90° away. (*C*) The front surface of *tm*TruB with bound RNA substrate and the view 90° away. The RNA molecules A, B, and C are shown in green, yellow, and blue, respectively. The figure was prepared by using GRASP (28).

to a 5-fluorouridine, a mechanism-based inhibitor, to capture the intermediate TruB–RNA covalent complex. However, our structure shows that there is no covalent bond between the carboxylate of Asp-48 and C6 of the base. The base is rotated by about 180° along the glycosidic bond such that the C6 lies on the opposite side, away from Asp-48. In our structure, RNA in the active site is found to contain a product analogue, 5-fluoro-6-hydroxy-Ψ, at position 55. This result is identical to that obtained in the *E. coli* TruB–RNA complex structure (14). This analogue possibly is the result of the hydrolysis of the ester bond between the carboxylate and the uracil ring. The presence of the hydroxyl group on the C6 confirms that the catalysis of Ψ formation proceeds by the attack on C6 of the base and not C1′ of the sugar.

TruB Accesses Its Substrate by Rigid Docking. The surface electrostatic potential of apo TruB shows a striking feature: there is a large, positively charged surface stretching from the RNA-binding cleft to the C-terminal domain (Fig. 3*A*). In contrast, on the other side of the molecule, the surface is mostly negatively charged. Such a large, positively charged surface provides an attractive potential for RNA binding. Moreover, the structure of TruB bound to RNA reveals that this enzyme recognizes the preformed three-dimensional structure of the T stem–loop primarily through shape complementarity. Therefore, TruB appears to first recognize its RNA substrate by a combination of electrostatic interactions and surface complementarity, which is enhanced by the hydrophobic interactions.

To test the hypothesis that TruB accesses its substrate initially through charge and shape complementarity, a docking experiment was carried out on the apo TruB structure with a 17-base stem–loop from the structure of yeast $tRNA^{Phe}$. Only one side of the TruB molecular surface, which contains the RNA-binding cleft, was selected for docking. A distance restraint of 10 Å between the catalytic residue Asp-48 and the uracil base of Ψ55 was applied to remove biologically irrelevant structures. The docked structure with the most favorable score is almost identical to the TruB–RNA complex structure (Fig. 3 *B* and *C*). This finding suggests that in the early stage of RNA recognition, structure-dependent and charge-dependent rigid docking plays an important role in presenting the RNA in a defined context for optimal association.

Roles of Induced Fit. Our structures indicate that one of the roles of the thumb loop is to provide a binding surface that can fit tightly into the RNA loop, thus optimizing the RNA–protein interaction. The RNA-free and RNA-bound structures effectively illustrate the ordering of the thumb loop to achieve RNA-binding specificity. In the absence of RNA, the thumb loop is disordered, leaving the RNA-binding cleft exposed to solvent. On binding to the RNA substrate, the thumb loop becomes ordered and protrudes into the RNA loop, thus closing the open active site and locking the RNA into the RNA-binding cleft. The highly conserved residues on the thumb loop form a hydrogen-bond network with the RNA loop. The C-terminal domain also moves toward RNA to facilitate appropriate adjustment of the catalytic domain to its substrate and further stabilize the RNA–protein interactions. The conformational changes and ordering of the thumb loop induced by the RNA substrate may be instrumental in bringing the target U into the active site of TruB with proper alignment for catalysis.

Induced fit is an important feature of RNA and DNA recognition and is widely observed in RNA- and DNA-binding proteins. The induced-fit mechanism offers important advantages in cellular regulation, as the inherent flexibility of the proteins allows local or global structures to be modified in response to multiple cellular partners and ensures fine control over binding affinity. Indeed, flexibility in TruB appears to be essential for recognition of multiple tRNAs and for tight binding.

General Mechanism of RNA Recognition. Taken together, the structures of apo and RNA-bound TruB suggest that the enzyme binds its substrate in two steps, first rigidly docking to its cognate RNA, then maximizing interactions through induced fit. The initial recognition of tRNA substrate would be achieved by rigid docking through highly complementary charges and shapes of the two counterparts. After the correct substrate is docked into the active site of TruB in a correct orientation, the thumb loop refolds and closes up the active site by a snug "fit" of this part of the protein into the RNA loop. The RNA major groove is in contact with active-site residues on the thumb loop that assist in positioning the U55 for the pseudouridinylation reaction. Once the enzyme and RNA are in close contact, more specific recognition and adjustments of the catalytic residues occur. His-43 pushes into the RNA loop, displacing G18, which was base-pairing with U55, thereby liberating U55. Tyr-179 and Tyr-76 close around the flipped-out U55, stabilizing the optimum orientation of this base for catalysis. The catalytic residue Asp-48, coordinated by Arg-181, is brought close to C6 of U55, ready for nucleophilic attack. A network of direct or water-mediated hydrogen bonds is formed in this step. Finally, the thumb loop must reopen to release the RNA. There is no cofactor needed for reaction; how and what triggers the release of the RNA require new experimental data. This report shows that Ψ synthase utilizes an induced-fit mechanism for substrate recognition. The structures confirm the essential role of the highly conserved thumb loop. These results provide insight into the molecular details of how Ψ synthase recognizes its substrate and modifies the target U base.

We thank the staff of Beamlines 5.0.2 and 8.3.1 at the Advanced Light Source for assistance with data collection, W. Chen and G. Kuznetsov for technical assistance, and P. J. Greene and J. C.-H. Chen for critical reading of the manuscript. This research was supported by National Institutes of Health Grant GM51232 (to R.M.S.).

1. Ofengand, J., Bakin, A., Wrzesinski, J., Nurse, K. & Lane, B. G. (1995) *Biochem. Cell Biol.* **73,** 915–924.
2. Bakin, A., Lane, B. G. & Ofengand, J. (1994) *Biochemistry* **33,** 13475–13483.
3. Gu, J., Patton, J. R., Shimba, S. & Reddy, R. (1996) *RNA* **2,** 909–918.
4. Yu, Y. T., Shu, M. D. & Steitz, J. A. (1998) *EMBO J.* **17,** 5783–5795.
5. Davis, D. R. (1995) *Nucleic Acids Res.* **23,** 5020–5026.
6. Newby, M. I. & Greenbaum, N. L. (2002) *Nat. Struct. Biol.* **9,** 958–965.
7. Charette, M. & Gray, M. W. (2000) *IUBMB Life* **49,** 341–351.
8. Gutgsell, N., Englund, N., Niu, L., Kaya, Y., Lane, B. G. & Ofengand, J. (2000) *RNA* **6,** 1870–1881.
9. Gutgsell, N. S., Del Campo, M. D., Raychaudhuri, S. & Ofengand, J. (2001) *RNA* **7,** 990–998.
10. Urbonavicius, J., Durand, J. M. & Bjork, G. R. (2002) *J. Bacteriol.* **184,** 5348–5357.
11. Huang, L., Pookanjanatavip, M., Gu, X. & Santi, D. V. (1998) *Biochemistry* **37,** 344–351.
12. Nurse, K., Wrzesinski, J., Bakin, A., Lane, B. G. & Ofengand, J. (1995) *RNA* **1,** 102–112.
13. Kinghorn, S. M., O'Byrne, C. P., Booth, I. R. & Stansfield, I. (2002) *Microbiology* **148,** 3511–3520.
14. Hoang, C. & Ferre-D'Amare, A. R. (2001) *Cell* **107,** 929–939.
15. Otwinoswski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
16. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54,** 905–921.
17. Ewing, T. J., Makino, S., Skillman, A. G. & Kuntz, I. D. (2001) *J. Comput. Aided Mol. Des.* **15,** 411–428.
18. Shi, H. & Moore, P. B. (2000) *RNA* **6,** 1091–1105.
19. Varani, G. & Nagai, K. (1998) *Annu. Rev. Biophys. Biomol. Struct.* **27,** 407–445.
20. Gu, X., Yu, M., Ivanetich, K. M. & Santi, D. V. (1998) *Biochemistry* **37,** 339–343.
21. Koshlap, K. M., Guenther, R., Sochacka, E., Malkiewicz, A. & Agris, P. F. (1999) *Biochemistry* **38,** 8647–8656.
22. Holbrook, S. R. & Kim, S. H. (1997) *Biopolymers* **44,** 3–21.
23. Becker, H. F., Motorin, Y., Sissler, M., Florentz, C. & Grosjean, H. (1997) *J. Mol. Biol.* **274,** 505–518.
24. Mueller, E. G. (2002) *Nat. Struct. Biol.* **9,** 320–322.
25. Gu, X., Liu, Y. & Santi, D. V. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 14270–14275.
26. Carson, M. (1997) *Methods Enzymol.* **277,** 493–505.
27. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995) *Protein Eng.* **8,** 127–134.
28. Nicholls, A., Bharadwaj, R. & Honig, B. (1993) *Biophys. J.* **64,** A116–A125.