Crystal Structure of the *Thermus thermophilus* 16 S rRNA Methyltransferase RsmC in Complex with Cofactor and Substrate Guanosine*[^5]

Hasan Demirci, Steven T. Gregory, Albert E. Dahlberg, and Gerwald Jogl[^1]

From the Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island 02912

---

Post-transcriptional modification is a ubiquitous feature of ribosomal RNA in all kingdoms of life. Modified nucleotides are generally clustered in functionally important regions of the ribosome, but the functional contribution to protein synthesis is not well understood. Here we describe high resolution crystal structures for the N2-guanine methyltransferase RsmC that modifies residue G1207 in 16 S rRNA near the decoding site of the 30 S ribosomal subunit. RsmC is a class I S-adenosyl-L-methionine-dependent methyltransferase composed of two methyltransferase domains. However, only one S-adenosyl-L-methionine molecule and one substrate molecule, guanosine, bind in the ternary complex. The N-terminal domain does not bind any cofactor. Two structures with bound S-adenosyl-L-methionine and S-adenosyl-L-homocysteine confirm that the cofactor binding mode is highly similar to other class I methyltransferases. Secondary structure elements of the N-terminal domain contribute to cofactor-binding interactions and restrict access to the cofactor-binding site. The orientation of guanosine in the active site reveals that G1207 has to disengage from its Watson-Crick base pairing interaction with C1051 in the 16 S RNA and flip out into the active site prior to its modification. Inspection of the 30 S crystal structure indicates that access to G1207 by RsmC is incompatible with the native subunit structure, consistent with previous suggestions that this enzyme recognizes a subunit assembly intermediate.

---

Post-transcriptional modification is a ubiquitous feature of rRNA in all kingdoms of life. The position of many modified nucleotides is conserved and generally located in functionally important regions in the ribosome (1), although the functions have been identified for few of the modifications. There are three basic types of rRNA modifications as follows: pseudouridylation, base methylation, and ribose methylation. N2-Methylation of guanine is the second most abundant modification in rRNA after pseudouridine formation (reviewed in Ref. 2). Complete modification maps of bacterial rRNA have been constructed for only two species, *Escherichia coli* and *Thermus thermophilus* (3). Given the divergent conditions for growth of these organisms, it was unexpected that they differ little in their 16 S rRNA modifications. *E. coli* 16 S rRNA has 11 modifications, whereas *T. thermophilus* has 14, with 9 modifications being identical between the two species. A similar situation applies to the 23 S rRNA (4). Thus, it is predicted that there will be significant similarity in the enzymology of modification of these two organisms.

A number of enzymes responsible for rRNA modification in *E. coli* have been identified and characterized, and several structures have been solved. However, little is known about how these enzymes recognize their substrates or, in many instances, what those substrates actually are. Some methyltransferases, such as KsgA, recognize subunit assembly intermediates (5), whereas others, such as RsmG, appear to methylate subunits once assembly has been completed (6–8). Only the structures of the 5-methyluridine methyltransferases RlmD and TrmA are known in complex with their respective RNA substrates (9, 10).

RsmC (EC 2.1.1.52) is an AdoMet[^2]-dependent methyltransferase responsible for N2-methylation of G1207 in 16 S rRNA and was first described in 1999 (11). AdoMet-dependent methyltransferases are divided into five structurally different classes (12, 13). RsmC belongs to the largest class I family, which is characterized by a central seven-stranded β-sheet that is flanked by three helices on each side and is structurally similar to Rossmann-fold domains. Class I methyltransferases modify a wide variety of substrates, including nucleic acids and proteins. Substrate specificity is achieved by additional substrate recognition domains in many of these enzymes (14).

The target of RsmC, G1207 (*E. coli* numbering), is situated in helix 34 of 16 S rRNA (15), a component of the ribosome long known to be involved in codon recognition, and forms a Watson-Crick base pair with C1051. Substitutions at C1054 or C1200 act as suppressors of nonsense mutations (16, 17), whereas a number of substitutions at or near these positions have been shown to influence decoding accuracy (18). These data can now be rationalized using crystal structures of the 30 S

---

[^1]: To whom correspondence should be addressed: Dept. of Molecular Biology, Cell Biology and Biochemistry, Brown University, Box G-E129, Providence, RI 02912. Tel.: 401-863-6123; Fax: 401-863-6114; E-mail: Gerwald_Jogl@brown.edu.

[^2]: The abbreviations used are: AdoMet, S-adenosyl-L-methionine; AdoHcy, S-adenosyl-L-homocysteine; r.m.s.d., root mean square deviation.
subunit from *T. thermophilus* in which C1054 can be seen to participate directly in codon recognition through interactions with the codon-anticodon complex (19). Thus, like a number of methylated nucleotides of 16S rRNA, G1207 is in a position to potentially influence the fidelity of codon recognition. Substitutions at G1207 produce dominant-lethal phenotypes (20), although this may result more from effects on local RNA structure than on the methylation status of this base because rsmC is dispensable (21). A role for RsmC in 30S subunit assembly has been suggested based on the observation that the enzyme most efficiently methylates 30S subunits at low magnesium concentrations, although the physiological relevance of this observation is far from clear (11).

The structure of the apo-form of *E. coli* RsmC was recently solved (21) revealing that the enzyme is composed of two domains with nearly identical folds, presumably the result of a gene duplication event. It was not clear from this structure precisely how the cofactor AdoMet or the substrate guanosine would dock into the active site. To gain insights into the substrate recognition mechanism, we determined high resolution crystal structures of RsmC from the extreme thermophile *T. thermophilus* in a ternary complex with AdoMet and guanosine bound in the active site as well as in complexes with AdoMet or AdoHcy. The orientation of the guanine base in the active site clearly shows that the substrate G1207 needs to disengage from its Watson-Crick interaction with C1051 in the context of the 16S rRNA structure. In addition, the base orientation is guided by hydrophobic stacking interactions with protein residues in a manner similar to class I N^6^-adenine DNA methyltransferases, and the modified nitrogen atom is located in a similar position in the active site. However, the direction of base insertion into the active site is dissimilar, and the mechanism of overall substrate recognition is unlikely to be related between the two enzymes. Structural comparison with RsmC from *E. coli* reveals a conserved positively charged surface region in the noncatalytic domain that is not detected by sequence alignment but may function in rRNA binding.

**EXPERIMENTAL PROCEDURES**

Cloning—The full-length *rsmC* gene (GenBank™ accession numberYP_143799) from *T. thermophilus* HB8 genomic DNA was amplified by PCR using a forward primer containing an NdeI restriction site (GGAATTCATATGAGCCTGAC-GCG-... was amplified by PCR using a forward primer containing an expression construct contained a C-terminal hexahistidine tag. Heat-denatured *E. coli* proteins were separated by centrifugation at 15,000 rpm at 4°C for 30 min. The soluble proteins were heat-treated at 65°C for 30 min to precipitate *E. coli* proteins. Heat-denatured *E. coli* proteins were separated by centrifugation at 15,000 rpm at 4°C for 30 min. The soluble C-terminally hexahistidine-tagged *T. thermophilus* RsmC was further purified by affinity chromatography with nickel-nitrilotriacetic acid resin (Qiagen). Untagged proteins were removed with a buffer containing 20 mM Tris-HCl (pH 8.5) and 250 mM NaCl and 10 mM imidazole (pH 8.5). Recombinant RsmC was then eluted with the same buffer containing 150 mM imidazole, and eluted protein was immediately mixed with 5% v/v glycerol. RsmC was unstable and aggregated in the absence of glycerol. The protein was further purified with anion exchange chromatography (QFF) (GE Healthcare) at pH 8.5, using a linear gradient of 10 mM to 1 mM NaCl concentration. RsmC fractions were pooled and concentrated and applied to a size exclusion chromatography S200 column (GE Healthcare) at pH 8.5 and 200 mM NaCl. The purified RsmC was concentrated to 24 mg/ml in a buffer containing 20 mM Tris (pH 8.5), 200 mM NaCl, and 5% v/v glycerol for crystallization trials. The C-terminal hexahistidine tag was not removed for crystallization. For the production of selenomethionyl proteins, the expression construct was transformed into B834(DE3) cells (Novagen). The bacterial growth was carried out in defined LeMaster medium (22), and the protein was purified using the same protocol as for the wild-type protein. To form the RsmC-AdoHcy complex, purified RsmC was mixed with 2 mM AdoHcy, incubated at 50°C for 10 min, and slowly cooled down to room temperature. To form the RsmC-AdoMet-guanosine complex, purified RsmC was mixed with 2 mM chloride salt of AdoMet and 1 mM guanosine hydrate and incubated at 50°C for 10 min before performing crystallization experiments.

Crystallization—Crystals of RsmC complexed with AdoMet and AdoHcy were obtained with the microbatch technique under oil at 4°C. 1 μl of protein solution was mixed with the reservoir solution, which contained 0.1 M HEPES (pH 7.5) and 1.5 mM lithium sulfate monohydrate. Initial crystals grew over the course of 3–5 days with maximum dimensions of 0.6 × 0.6 × 0.6 mm. Crystals of RsmC in complex with AdoMet and guanosine hydrate were obtained with the microbatch technique under oil at 4°C. 1 μl of protein solution was mixed with the reservoir solution, which contained 0.17 mM lithium sulfate monohydrate, 0.085 M Tris-HCl (pH 8.5), 25.5% w/v PEG4000, and 15% glycerol. Crystals grew over the course of 3–5 days with maximum dimensions of 0.6 × 0.6 × 0.6 mm. Crystals of the RsmC-AdoMet-guanosine ternary complex were flash-frozen by plunging into liquid nitrogen directly from their mother liquor, whereas RsmC-AdoMet and RsmC-AdoHcy crystals were cryo-protected by rapid soaking into a solution containing mother liquor supplemented with the addition of 30% v/v glycerol before freezing.

Data Collection—X-ray diffraction data for all crystals were collected on a MAR CCD detector at the X4C beamline of the National Synchrotron Light Source in Brookhaven. For the initial structure determination, a selenomethionyl single wave-
Structure of the RsmC Methyltransferase

### Results

- The overall structure of RsmC consists of two structurally related domains (residues 3–175 and residues 187–373) connected by a short linker region (Fig. 1 and Fig. 2A), probably the result of a gene duplication event followed by domain subfunctionalization (21). Both domains form a canonical Rossmann-like fold with minor variations (Fig. 2A). The N-terminal domain consists of a six-stranded β-sheet that is

---

**Table 1**

| Data collection | RsmC1 AdoMet + guanosine | RsmC2 AdoHcy | RsmC3 AdoMet | RsmC4 anomalous |
|----------------|---------------------------|--------------|--------------|-----------------|
| Space group    | P2₁2₁2₁                   | P2₁2₁2₁      | P2₁2₁2₁      | P2₁2₁2₁         |
| Cell dimensions| a, b, c (Å)                | a, b, c (Å)  | a, b, c (Å)  | a, b, c (Å)     |
|                | 50.9, 86.2, 95.1           | 51.0, 88.0, 94.9 | 50.7, 88.4, 95.4 | 51.1, 86.5, 95.1 |
| Resolution (Å) | 30 to 1.55 (1.61 to 1.55) | 30 to 1.55 (1.61 to 1.55) | 30 to 1.58 (1.64 to 1.58) | 30 to 1.9 (1.97 to 1.90) |
| Rwork/Rfree   | 0.066 (0.418)              | 0.059 (0.357) | 0.051 (0.398) | 0.035 (0.198)   |
| I/σI           | 23.8 (2.2)                 | 24.7 (2.6)   | 27.3 (3.0)   | 50.2 (8.5)      |
| Completeness (%)| 97.8 (91.6)              | 99.1 (96.5)  | 98.6 (92.4)  | 100 (100)       |
| Redundancy     | 5.7 (3.1)                  | 5.5 (4.6)    | 5.9 (4.2)    | 6.4 (5.5)       |

---

**Structure Determination and Refinement**—The locations of three selenium atoms (out of four expected selenium atoms) were determined with the program Selenolyte (24) based on the anomalous differences in the single wavelength anomalous data set. Reflection phases to 1.9 Å were calculated with Selenolyte. An initial model was built with Resolve (25), and ARP/wARP (26, 27) was used for subsequent model building. The atomic coordinates from the first model were then used as initial models for refinement against all three complex data sets. All models were checked and completed with Coot (28). Crystallographic refinement was performed with the program Refmac (29) from the CCP4 package (30). Final models of all three structures contain residues 3–372 in chain A. Residues of the C-terminal hexahistidine tag were disordered and not included in the final models. The stereochemical quality of the models was assessed with Procheck (31). The Ramachandran statistics (most favored/additionally allowed/generously allowed/disallowed) are 94.5/4.9/0.6/0.0% for RsmC1, 94.2/5.2/0.6/0.0% for RsmC2, and 94.5/4.9/0.6/0.0% for RsmC3. The refinement statistics are summarized in Table 1. Figures were generated using Pymol (32) and Jmol (33). Sequence alignments were generated with ClustalW2 (34) and Staccato (35).

---

**Atomic Coordinates**—The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 3DMH, 3DMG, and 3DMF for RsmC1, RsmC2, and RsmC3, respectively.

---

**Results**

The structure of RsmC (375 amino acids) was solved by single-wavelength anomalous dispersion methods from seleno-methionine-labeled protein in space group P2₁2₁2₁, initially to 1.9 Å resolution (data set RsmC4). We report three high resolution structures (up to 1.55 Å) for the enzyme in complex with guanosine and AdoMet, in complex with AdoHcy, and in complex with AdoMet (referred to as RsmC1, RsmC2, and RsmC3, respectively). The crystallographic R/Rfree factors are 0.19/0.23, 0.18/0.22, and 0.18/0.22 for the data sets, respectively. The structures are highly similar with an overall root mean square deviation of 0.19 Å between RsmC1 and RsmC2. There is one molecule in the asymmetric unit. The majority of residues (94.5/94.2/94.5%) are in the most favored region of the Ramachandran plot, and there are no residues in the disallowed region. In all three structures, electron density was well defined except for the first two and last three residues and for residues 309–314 in a loop region near the active site (discussed below). The data collection and refinement statistics are given in Table 1.

The overall structure of RsmC consists of two structurally related domains (residues 3–175 and residues 187–373) connected by a short linker region (Fig. 1 and Fig. 2A), probably the result of a gene duplication event followed by domain subfunctionalization (21). Both domains form a canonical Rossmann-like methyltransferase fold with minor variations (Fig. 2A). The N-terminal domain consists of a six-stranded β-sheet that is
flanked on both sides by six helices of varying length. The C-terminal domain consists of a seven-stranded \( \beta \)-sheet flanked by five helices. Both domains contain an additional \( \beta \)-hairpin that is also observed in other methyltransferases and forms a highly conserved hydrophobic interface with the core domain. There is one AdoMet and one guanosine molecule bound to the C-terminal catalytic domain, whereas the N-terminal domain contains a bound sulfate ion (from the precipitant used for crystallization). The first helix a1 is the only unique secondary structure element observed. It connects to the side of the C-terminal domain and contributes to the cofactor-binding site. The two domains in RsmC, including the \( \beta \)-hairpin, align with a least squares r.m.s.d. of 1.7 Å (Fig. 2C). Although the secondary structure elements generally align quite well, many small variations are observed especially in the loop regions flanking the cofactor- and substrate-binding sites of the catalytic domain.

Cofactor Binding in the Active Site of RsmC—In contrast to other methyltransferases, recombinant RsmC retains AdoMet tightly bound throughout purification and crys-
Structure of the RsmC Methyltransferase

Guanosine Binding in the RsmC Active Site—The position of the guanosine molecule in the active site is mainly determined by a hydrophobic stacking interaction with the Phe-308 side chain and two hydrogen bonds of Asn-305 and Asp-213 with the guanine molecule in the active site. Residues in the N-terminal domain are colored red and yellow. The C-terminal domain and AdoMet bound in the catalytic domain are shown in gray for reference.

The comparison of the cofactor-binding site in the catalytic domain with the equivalent region in the N-terminal domain shows the key differences rendering the N-terminal domain noncatalytic (Fig. 3D). Hydrophobic interactions of Val-318 and Glu-262 with the adenine base have been replaced with hydrogen-bonding Tyr-117 and Thr-74. More significantly, Arg-28 and Arg-113 cooperate to bind the sulfate ion in the active site, engaging in a hydrogen bond to the guanine N7 atom (2.9 Å), and its amino group nitrogen atom is located at a distance of 3.5 Å from the methionine carbonylate oxygen. AdoMet and AdoHcy are bound in an identical manner. Phe-207 is the only residue that is shifted in the two structures, with the phenyl ring translating laterally by 0.6 Å (a comparison of AdoMet- and AdoHcy-binding is given in supplemental Fig. S2A).

The presence of the noncatalytic N-terminal domain restricts the accessibility of the cofactor-binding site significantly. Helix α1 and the loop region connecting βα2 with α1 effectively extend the adenine-binding pocket and create a deep cofactor-binding site that is likely only accessible in the absence of substrate rRNA. Both AdoMet and AdoHcy are bound in an identical manner. Phe-207 is the only residue that is shifted in the two structures, with the phenyl ring translating laterally by 0.6 Å (a comparison of AdoMet- and AdoHcy-binding is given in supplemental Fig. S2A).

AdoMet could be displaced only during crystallization in the presence of an excess of AdoHcy as observed in data set RsmC2. This in part may reflect the thermostable nature of this enzyme.

As in other class I methyltransferases, AdoMet is bound in a canonical conformation above the β-sheet and close to the conserved GXGXXG methyltransferase signature motif (residues 241–245 between strand βc1 and helix αc3, see Fig. 1 and Fig. 3, A and B). The adenine ring is located in a mainly hydrophobic pocket that is lined by residues Phe-322 and Val-318 on one side and the main chain and C-β atom of Asp-263 on the other side. The presence of the charged aspartate in this position is unusual when compared with other methyltransferases; however, its participation in a hydrogen bond to the main-chain nitrogen of Arg-76 and the Ser-75 side chain might further stabilize the interdomain coordination in the thermophilic enzyme. The NH atom of the adenine group is coordinated by a hydrogen bond to Asp-288, which also interacts with the Tyr-8 hydroxyl group from the N-terminal helix αa1. Both ribose hydroxyl groups form hydrogen bonds to Glu-262. This represents another variation in this structure, as position 262 in methyltransferases is typically occupied by an aspartate residue (as in RsmC from E. coli; a more comprehensive sequence alignment of RsmC homologs is given in supplemental Fig. S1). The Ser-216 side chain coordinates the methionine carboxyl group of the cofactor, and the methionine amino group interacts with the main-chain carbonyl of Gly-241 and with Asp-239 via solvent water molecules. Asn-305, which is also a conserved signature residue in the class I methyltransferase cofactor-binding site, engages in a hydrogen bond to the guanine N7 atom (2.9 Å), and its amino group nitrogen atom is located at a distance of 3.5 Å from the methionine carbonylate oxygen. The hydrophobic Phe-207 approaches the AdoMet sulfur atom and might contribute to the enzymatic function by destabilizing the cofactor. The presence of the noncatalytic N-terminal domain restricts the accessibility of the cofactor-binding site significantly. Helix α1 and the loop region connecting βα2 with α1 effectively extend the adenine-binding pocket and create a deep cofactor-binding site that is likely only accessible in the absence of substrate rRNA. Both AdoMet and AdoHcy are bound in an identical manner. Phe-207 is the only residue that is shifted in the two structures, with the phenyl ring translating laterally by 0.6 Å (a comparison of AdoMet- and AdoHcy-binding is given in supplemental Fig. S2A).

The comparison of the cofactor-binding site in the catalytic domain with the equivalent region in the N-terminal domain shows the key differences rendering the N-terminal domain noncatalytic (Fig. 3D). Hydrophobic interactions of Val-318 and Glu-262 with the adenine base have been replaced with hydrogen-bonding Tyr-117 and Thr-74. More significantly, Arg-28, Arg-113, Asn-53, and His-35 extend into the cofactor-binding pocket precluding any AdoMet binding in the N-terminal domain. Instead, the role of this domain may be to engage in substrate-rRNA interactions. Consistent with this notion, Arg-28 and Arg-113 cooperate to bind the sulfate ion in the N-terminal domain and by analogy might function in coordinating a phosphate group from the substrate rRNA.
the base. The two hydroxyl groups of the ribose moiety are oriented toward the protein surface and are not engaged in hydrogen bonds with ordered solvent water molecules. Comparison of the guanosine-bound RsmC1 structure with the substrate-free RsmC2 structure shows that the Asn-341 side chain occupies the position of the guanosine N7 atom and that the Asn-341 side chain rotates by about 120° upon guanosine binding. In addition, the side chains of Asp-213, Phe-308, and Tyr-365 reorient slightly when guanosine is bound (supplemental Fig. S2).

To investigate the possibility that adenosine might also bind in the active site of RsmC, we performed the same experiment with adenosine instead of guanosine. However, there was no indication of adenosine binding in a 2.2-Å diffraction data set from crystals obtained under the same conditions as the guanosine-bound crystals (data not shown).

DISCUSSION

Comparison with Structurally Related Methyltransferases—A data base search with the SSM algorithm (36) showed that the structure most closely related to T. thermophilus RsmC is E. coli RsmC (21) followed by a structure of unknown function (Protein Data Bank code 1DUS) from Methanococcus jannaschii, which has been proposed to represent the RsmC homolog from this organism (37). RsmC from E. coli is the only structure with the same two-domain architecture, whereas other structurally related methyltransferases share structural similarity for only the catalytic domain. RsmC from T. thermophilus and E. coli share only 26% identical residues with each other in the structure-based sequence alignment (Fig. 1), but the two structures can be superimposed with an r.m.s.d. of 1.6 Å for 292 Cα atoms. The RsmC C-terminal domain by itself also shows strong similarity to the catalytic domains of protein methyltransferases such as the ribosomal protein L11 trimethyltransferase PrmA (38). The r.m.s.d. of the catalytic domain with the T. thermophilus PrmA is 1.5 Å for 147 Cα atoms.

It is interesting to note that the thermophilic RsmC represents a complete domain duplication, whereas the mesophilic RsmC has lost the leading N-terminal β-hairpin (and the single N-terminal helix; Fig. 4A). These additional secondary structure elements contribute to the cofactor-binding site and create a better-defined binding pocket in our RsmC structure. The RsmC structure from E. coli was reported in the absence of cofactor or substrate suggesting that cofactor binding is significantly weaker in the absence of the N-terminal extension. Conversely, the higher cofactor-binding affinity of the Thermus enzyme is clearly reflected in our observation that the enzyme always co-purifies with bound cofactor. In addition, we found that extended dialysis of the cofactor-bound enzyme (up to 3
days in large volumes at 4 °C did not suffice to remove the cofactor. The remaining interacting residues in the cofactor-binding site and the major determinants for guanosine placement are well conserved between both enzymes (Fig. 4B).

A significant difference between both structures is seen in the loop region surrounding Phe-207. The side chain of this residue approaches the AdoMet reactive methyl group, and the hydrophobic environment in the active site might contribute to catalysis. Phe-207 is conserved in the sequence alignment, but the location of the corresponding Phe-176 in E. coli is further removed from the AdoMet-binding site. In addition, the Phe-176 side chain was not modeled in the E. coli structure for comparison. A similar structural flexibility of this loop has also been observed in the PrmA trimethyltransferase (discussed below). The side chain conformations for residues equivalent to Asp-213 and Phe-308 in the guanosine-binding region are different in the E. coli structure in the absence of guanosine suggesting a structural rearrangement upon substrate binding. Tyr-365 is replaced with Phe-193 in PrmA. Replacement of Tyr-365 with Trp-247 effectively closes off the guanosine-binding site in PrmA. Phe-308 is replaced with Tyr-193 in PrmA that is oriented in the opposite direction and interacts with the AdoMet base as a functional replacement of Tyr-8 in RsmC (the N-terminal β-hairpin is not present in PrmA).

The position of Phe-207 in RsmC is equivalent to Phe-99 in PrmA. The loop region surrounding Phe-99 contains several glycine residues and is disordered in several structures in the absence of cofactor. A highly similar loop and phenylalanine placement are also observed in the PrmA protein methyltransferase (39). These combined observations led us to propose a functional significance for the observed flexibility and the induced fit of this loop upon cofactor binding. In combination with the significant difference of this loop region in the apo-RsmC structure from E. coli, it seems likely that a similar induced fit mechanism is present in these enzymes. Further studies will be required to determine the functional significance of these observations.

**Implications for rRNA Binding**—The RsmC target site, G1207, is located in helix 34 of 16 S rRNA and base pairs with C1051 (Fig. 5A). The orientation of guanosine in our RsmC-AdoMet-guanosine ternary complex structure shows that G1207 has to disengage from its Watson-Crick pairing with C1051 and flip out of the helix to insert into the RsmC active site. Such base flipping in RNA-enzyme complexes has been observed previously. The TruB pseudouridine synthase-RNA complex revealed that three bases flip from the tRNA analog T loop to insert U55 into the active site for pseudouridylation.

---

**Figure 5. Implications for substrate binding.** A. Schematic representation of the 16 S rRNA. The substrate G1207, the base-pairing C1051, and the functionally important C1054 are shown as sticks. The position of the close-up figure in the rRNA structure is indicated in the lower left figure. The positions of helices 34 and 18 are indicated. B. Comparison of substrate base orientation between RsmC (left) and TaqI (right). The cofactor (cofactor homolog in TaqI) is shown at the bottom and the substrate base at the top of each figure. C and D, Surface representation colored by electrostatic charge for RsmC from *T. thermophilus* and *E. coli*. AdoMet is shown as yellow sticks and guanosine as salmon sticks. AdoMet and guanosine from the Thermus structure are shown with the *E. coli* structure for comparison.
A number of aminoacyl-tRNA synthetase-tRNA co-crystal structures indicate that these enzymes differentiate among tRNAs by flipping out and monitoring the anticodon bases (reviewed in Ref. 41). The flipping out of G1207 is unique in that this base engages in Watson-Crick pairing in the context of the ribosome.

Nucleotide flipping has also been observed for DNA methyltransferases such as the C^r^-cytosine methyltransferase M.Hhal (42) or the N^r^-adenine methyltransferase TaqI (43, 44). A functional similarity between RsmC and N^r^-adenine methyltransferases has been proposed based on sequence analysis and the conservation of the N^105NPY^108 signature motif (H^105NPY^108 in TaqI, Protein Data Bank code 1G38 (45)). Phe-308 provides the base stacking interaction for the substrate guanine in RsmC. The structural comparison of this region with TaqI shows that Tyr-108 provides the same hydrophobic interaction for the substrate base (Fig. 5B). However, the tyrosine hydroxyl group engages in a hydrogen bond with Asn-105 that repositions the tyrosine side chain and results in a substantially different orientation of the substrate base. Thus, simple replacement of Phe with Tyr results in a significantly different base insertion direction and contributes to the differing catalytic activity between these two enzyme structures.

**Implications for Ribosome Binding**—Inspection of the 30 S subunit crystal structure (46) indicates that access to G1207 by RsmC would be occluded in the native structure by helix 18 (Fig. 5A) and would require either rotation of the 30 S head around the neck region or a substantial backward motion toward the solvent side of the subunit. Movement of the head relative to the rest of the 30 S subunit occurs with the open-closed transition during codon recognition (47) suggesting that the capacity of this structure to engage in such motions is an intrinsic feature. The observation that methylation of 30 S subunits by RsmC is more efficient at low magnesium concentration (0.9 mM) than at high magnesium concentration (6 mM) has been taken to imply that a subunit assembly intermediate is the actual substrate (11). Alternatively, low magnesium conditions may allow greater freedom of motion of the head. What would facilitate such motion in vivo is unclear, although restrictions of head motion might be established by one or more late binding proteins, such as S2, which has binding sites on both the head and the body of the 30 S subunit (46). Other methyltransferases, such as KsgA, modify the ribosome prior to completion of assembly, and negative determinants for modification are the incorporation of one or more ribosomal proteins (5). In this context, it is interesting to note that the noncatalytic domains of both RsmC structures place a positively charged surface area close to the active site that suggests a functional contribution of the duplicated N-terminal domain for substrate RNA binding (Fig. 5, C and D). Although this observation is consistent with the functional requirements for these enzymes, the residues contributing to the charged surface patch are not conserved between RsmCs from different organisms (supplemental Fig. S1).

Despite the presence of RsmC orthologs among bacteria and the proximity of G1207 to C1054 and the site of codon recognition, the base identity at position 1207 of bacterial 16 S rRNAs is somewhat variable. The 1051:1207 base pair is found almost entirely as G-C and C-G at similar frequencies (15), raising the obvious question as to the function of RsmC in species having a C at position 1207. Substitutions at G1207 produce dominant-lethal phenotypes in *E. coli* (20) suggesting that Watson-Crick pairing of G1207 and C1051 is essential for maintaining a conformation of C1054 suitable for participation in codon recognition. The local structure of this helix is quite remarkable as it contains several bases splayed out of the helix as well as two sharp reversals of backbone trajectory that are established by multiple noncanonical base pairings and several base-triple interactions. One possible role for the interaction of RsmC with helix 34 may be to aid in pre-organization of this highly unorthodox structure and prevent misfolding during 30 S subunit assembly.

**Acknowledgments**—We thank Ashwin Cadambi, Megha Katti, Sising He, and Holly Careskey for their assistance with crystallization experiments; John Schwanow and Randy Abramowitz for access to the X4C beamline at the National Synchrotron Light Source; and Hua Li for help with data collection at the synchrotron.

**REFERENCES**

1. Decatur, W. A., and Fournier, M. J. (2002) Trends Biochem. Sci. 27, 344–351
2. Sergiev, P. V., Bogdanov, A. A., and Dontsova, O. A. (2007) Nucleic Acids Res. 35, 2295–2301
3. Guymon, R., Pomerantz, S. C., Crain, P. F., and McCloskey, J. A. (2006) Biochemistry 45, 4888–4899
4. Mengel-Jorgensen, I., Jensen, S. S., Rasmussen, A., Poehlsgaard, J., Iversen, J. J. L., and Kirpekar, F. (2006) *J. Biol. Chem.* 281, 22108–22117
5. Xu, Z., O’Farrell, H. C., Rife, J. P., and Culver, G. M. (2006) *Nat. Struct. Mol. Biol.* 15, 534–536
6. Okamoto, S., Tamaru, A., Nakajima, C., Nishimura, K., Tanaka, Y., Tokuyama, S., Suzuki, Y., and Ochi, K. (2007) *Mol. Microbiol.* 63, 1096–1106
7. Nishimura, K., Johansen, S. K., Inaoka, T., Hosaka, T., Tokuyama, S., Tahara, Y., Okamoto, S., Kawamura, F., Douthwaite, S., and Ochi, K. (2007) *J. Bacteriol.* 189, 6068–6073
8. Nishimura, K., Inaoka, T., Tokuyama, S., Okamoto, S., and Ochi, K. (2007) *J. Bacteriol.* 189, 3876–3883
9. Lee, T. T., Agarwalla, S., and Stroud, R. M. (2005) *Cell* 120, 599–611
10. Alian, A., Lee, T. T., Griner, S. L., Stroud, R. M., and Finer-Moore, J. (2008) *Proc. Natl. Acad. Sci. U. S. A.* 105, 6876–6881
11. Tischerer, J. S., Nurse, K., Popieniek, P., and Ofengand, J. (1999) *J. Biol. Chem.* 274, 924–929
12. Kozbial, P. Z., and Mushgian, A. R. (2005) *BMC Struct. Biol.* 5
13. Schubert, H. L., Blumenthal, R. M., and Cheng, X. D. (2003) *Trends Biochem. Sci.* 28, 329–335
14. Martin, J. L., and McMillan, F. M. (2002) *Curr. Opin. Struct. Biol.* 12, 783–793
15. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D’Souza, L. M., Da, Y. S., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pand, N., Shang, Z. D., Yu, N., and Gutell, R. R. (2002) *BMC Bioinformatics* 3
16. Goring, H. U., Hijazi, K. A., Murgola, E. J., and Dahlberg, A. E. (1991) *Proc. Natl. Acad. Sci. U. S. A.* 88, 6603–6607
17. Gregory, S. T., and Dahlberg, A. E. (1995) *Nucleic Acids Res.* 23, 4234–4238
18. Moine, H., and Dahlberg, A. E. (1994) *J. Mol. Biol.* 243, 402–412
19. Ogle, J. M., Brodersen, D. E., Clemons, W. M., Tarry, M. J., Carter, A. P., and Ramakrishnan, V. (2001) *Science* 292, 897–902
20. Jemiolo, D. K., Taurence, J. S., and Giese, S. (1991) *Nucleic Acids Res.* 19, 4259–4265
21. Sunita, S., Purta, E., Durawa, M., Tkaczuk, K. L., Swathi, J., Bujnicki, J. M., and Sivaraman, J. (2007) *Nucleic Acids Res.* 35, 4264–4274
22. Hendrickson, W. A., Horton, J. R., and Lemaster, D. M. (1990) *EMBO J.* 9,
