

---

# Identification and characterization of RsmE, the founding member of a new RNA base methyltransferase family

---

GEORGETA N. BASTUREA, KENNETH E. RUDD, and MURRAY P. DEUTSCHER

Department of Biochemistry and Molecular Biology, University of Miami Miller School of Medicine, Miami, Florida 33101-6129, USA

## ABSTRACT

A variety of RNA methyltransferases act during ribosomal RNA maturation to modify nucleotides in a site-specific manner. However, of the 10 base-methylated nucleotides present in the small ribosomal subunit of *Escherichia coli*, only three enzymes responsible for modification of four bases are known. Here, we show that the protein encoded by *yggJ*, a member of the uncharacterized DUF558 protein family of predicted alpha/beta (trefoil) knot methyltransferases is responsible for methylation at U1498 in 16S rRNA. The gene is well-conserved across bacteria and plants, and likely performs the same function in other organisms. A *yggJ* deletion strain lacks the methyl group at U1498 as well as the specific methyltransferase activity. Moreover, purified recombinant YggJ specifically methylates m<sup>3</sup>U1498 in vitro. The deletion strain was unaffected in exponential growth in rich or minimal media at multiple temperatures, but it was defective when grown in competition with isogenic wild-type cells. Based on these data, we conclude that *yggJ* is the founding member of a family of RNA base methyltransferases, and propose that it be renamed *rsmE*.

**Keywords:** *E. coli*; methyltransferase; protein family; RNA modification

## INTRODUCTION

RNA methyltransferases (MTase) are enzymes that catalyze the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to an acceptor RNA. A variety of these enzymes act during ribosomal RNA maturation to modify nucleotides in a site-specific manner. Ten methylated nucleotides are present in the small ribosomal subunit of *Escherichia coli*, all distributed in important functional regions, such as the tRNA binding sites, the decoding region, and the mRNA binding area (Ofengand and Del Campo 2004). Their specific function is not well understood, but they are thought to be involved in maintaining and stabilizing ribosome three-dimensional structure, resulting in optimization of overall ribosome activity (Decatur and Fournier 2002). Addition of a methyl group presumably leads to slight changes in the local environment that influences rRNA folding and its interactions with specific proteins,

tRNAs, and mRNA, affecting the efficiency of protein synthesis.

The 3' domain of the small ribosomal subunit contains two highly conserved regions that are part of the decoding site: nucleotides 1394–1408 and 1492–1505. The only methylated nucleotide in the second region, U1498, seems to be important for the protein synthetic process. It has been suggested, based on experimental results with in vitro reconstituted ribosomes that lacked modified nucleotides (Ringquist et al. 1993), that this nucleotide might be involved in aminoacyl-tRNA selection. In addition, a possible role for m<sup>3</sup>U1498 in intersubunit communication was suggested based on its position at the top of helix 44, a region that in the crystal structure of the 70S particle (Yusupov et al. 2001) interacts with conserved regions in 23S rRNA (Decatur and Fournier 2002).

Relatively little is known about the MTases that produce methyl modifications in 16S rRNA. To date, only three genes coding for such enzymes have been identified: *ksgA* encodes RsmA (van Buul and van Knippenberg 1985), which is responsible for four methylations at residues A1518 and A1519, and *rsmB* (Tscherne et al. 1999a) and *rsmC* (Tscherne et al. 1999b), which encode two enzymes that produce m<sup>5</sup>C967 and m<sup>2</sup>G1207, respectively. In addition, almost nothing is known about how the removal of a specific

---

**Reprint requests to:** Murray P. Deutscher, University of Miami Miller School of Medicine, Department of Biochemistry and Molecular Biology, P.O. Box 016129, Miami, FL 33101-6129, USA; e-mail: mdeusch@med.miami.edu; fax: (305) 243-3955.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.2283106>.

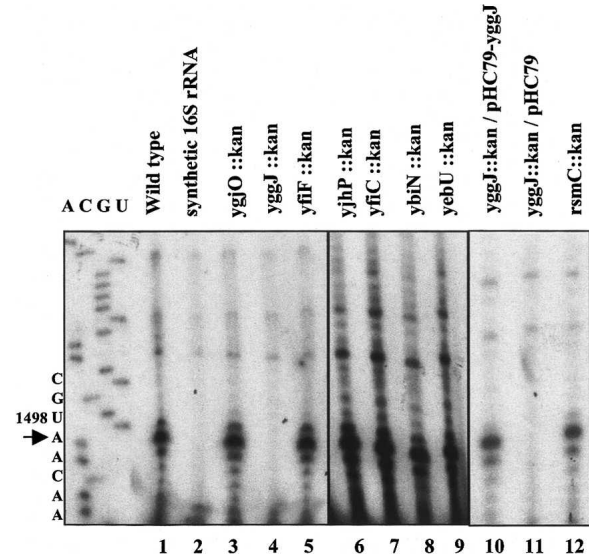
MTase affects cells. It is known that loss of the four methyl groups added by RsmA has little effect on ribosome function (van Knippenberg 1986); however, the effect of deletion of the other two MTase genes has not been determined.

Here, we describe the identification of the MTase gene responsible for the  $m^3U1498$  modification in *Escherichia coli*. The gene was identified by analysis of several putative MTase deletion strains and was confirmed by gene complementation, overexpression, and in vitro characterization of the purified enzyme. Most importantly, the gene belongs to a previously uncharacterized alpha/beta knot MTase family, and is the first  $m^3U$  MTase gene to be identified. The enzyme has high specificity for U1498 and is not involved in methylation of the only other  $m^3U$  site in the *E. coli* ribosome, position 1915 in 23S rRNA. Based on these studies, we propose that the *yggJ* gene be renamed *rsmE*.

## RESULTS

### Deletion of *yggJ* gene is associated with lack of methylation at U1498

Many genes have been predicted to be MTases, but few have been confirmed, and little is known about their actual biological roles. In order to identify the gene responsible for methylation of U1498 in the small ribosomal subunit of *E. coli*, we selected for analysis several genes predicted to encode RNA MTases whose function have not been determined. These include *yjhP* (Katz et al. 2003), *yfiC* (Katz et al. 2003), *ybiN* (Bujnicki 2000; Bujnicki and Rychlewski 2002; Katz et al. 2003), *yebU* (Reid et al. 1999; Katz et al. 2003), *ygiO* (Bujnicki 2000; Bujnicki and Rychlewski 2002), *yfiF* (Koonin and Rudd 1993) and *yggJ* (Forouhar et al. 2003). Each of the genes was replaced with a kanamycin resistance cassette using the method of Datsenko and Wanner (2000). Deletions were verified by the acquisition of kanamycin resistance and by the size of the PCR product generated upon amplification from the N- and C-terminal ends of the gene (data not shown). Total ribosomal RNA extracted from wild-type and mutant cells grown to exponential phase (Ofengand et al. 2001) was tested for the presence of a methylated U1498 by reverse transcription primer extension. In this method, reverse transcriptase stops at a site 3' to any base methylated at positions usually involved in Watson-Crick base pairing ( $m^3U$ ,  $m^2G$ , or  $m^2A$ ), due to reduced H bonding. As shown in Figure 1, use of this procedure with a primer specific to the region spanning nucleotides 1506–1525 in 16S rRNA revealed that for wild type (lane 1), *rsmC* deletion (lane 12), and most of the putative MTase gene deletion strains (lanes 3, 5–9), reverse transcriptase stopped at position A1499, just one base 3' to  $m^3U1498$ . As a control, no stop was obtained when in vitro synthesized, unmethylated 16S rRNA was used as the template for primer extension (lane 2). Of most interest was the absence of a stop for strain *yggJ::kan*

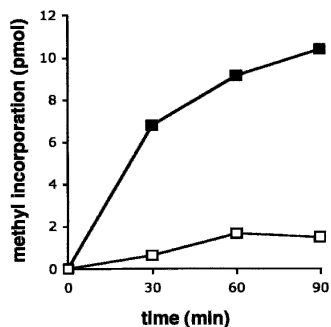


**FIGURE 1.** Primer extension analysis of RNA extracted from putative RNA methyltransferase deletion strains. Total RNA was annealed to a [ $^{32}P$ ]-labeled primer complementary to residues 1506–1525 in 16S rRNA, prior to extension by M-MLV-RT, as described in Materials and Methods. Products of the reaction were separated on 8% acrylamide/7 M urea denaturing gels and visualized by autoradiography. A stop at position A1499, noted by an arrow, indicates the presence of a methyl group at U1498.

(lane 4) indicating that rRNA in this strain lacks the methyl group at position 1498.

### S100 extracts from the *yggJ* deletion strain lack the $m^3U1498$ specific methyltransferase activity

Earlier work showed that S100 extracts prepared from wild-type cells contain a  $m^3U1498$  specific MTase activity that modifies synthetic 30S particles reconstituted from in vitro synthesized RNA and ribosomal proteins (Negre et al. 1989). We can detect the same activity in wild-type extracts by incorporation of tritiated methyl groups from [ $^3H$ ]-SAM into 30S particles prepared from the mutant strain, which lacks this modification (Fig. 2). In contrast, extracts prepared from the *yggJ* deletion strain have ninefold lower activity on the same substrate. The small amount of methyl groups incorporated by the mutant S100 extract is likely due to incorporation at other sites in the mutant 30S particles. For example, it is known that methylation at A1518 and A1519 are late events since 27S particles without any  $m^6_2A$  have been isolated from *E. coli* (Hayes et al. 1971). It is possible that these events are slowed even further by the absence of the  $m^3U1498$  methylase resulting in an increase of incompletely modified ribosomes in the 30S subunit preparations. Methylation at U1498 was also detected by reverse transcription upon treatment of synthetic 30S particles with wild-type S100 extract, but not with the *yggJ* mutant extract (data not shown), confirming the position of methyl incorporation.



**FIGURE 2.**  $m^3U1498$  methyltransferase activity of wild-type and mutant S100 extracts. Reactions were carried out as described in Materials and Methods, with mutant 30S particles purified from the *yggJ* deletion strain as substrate. Reaction mixtures contained 500 nM 30S subunit, 100  $\mu$ M [ $^3H$ ]-SAM, and 320  $\mu$ g of extract protein. The amount of incorporated methyl groups was monitored by acid precipitation and scintillation counting. (Filled squares) Wild-type extract; (open squares) *yggJ* mutant extract. One experiment representative of three repeats is shown.

### Methyl group at U1498 is recovered by gene complementation

To confirm that the *yggJ* gene is responsible for U1498 methylation, gene complementation was performed. The *yggJ* gene region, including 200 bp upstream of the putative ATG start codon (<http://ecogene.org>), was amplified from genomic DNA by PCR and cloned into the low copy number vector pHC79 (Hohn and Collins 1980), such that the gene would be expressed from its own promoter. MG1655 *yggJ::kan* was transformed with either the empty pHC79 vector or the *yggJ*-containing plasmid, grown to exponential phase, and tested for the presence of U1498 methylation. Figure 1 shows that U1498 is methylated when the mutant strain is transformed with the plasmid containing the cloned gene (lane 10), but there is no stop at A1499 with RNA extracted from cells transformed with the empty plasmid (lane 11). These data indicate that synthesis of the methyl group on U1498 requires the presence of *yggJ*.

### Purified protein encoded by *yggJ* has $m^3U1498$ specific methyltransferase activity

To examine the *in vitro* activity of the *yggJ* gene product, the gene was subcloned into pET28a. The recombinant DNA obtained codes for an N-terminal His-tag fusion protein, which allowed easy purification by affinity chromatography. Figure 3 shows that a protein product of ~29 kDa was induced in the presence of IPTG, as expected from the DNA sequence. The protein was purified to homogeneity by metal affinity chromatography (Fig. 3) and was assayed for its MTase activity (Fig. 4). The purified enzyme was able to incorporate [ $^3H$ ] methyl groups from [ $^3H$ ]-SAM into 30S particles purified from the mutant strain but not into 30S particles purified from wild type.

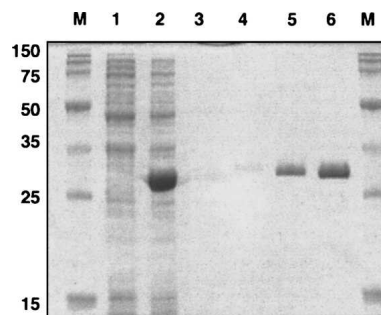
Approximately one methyl group could be incorporated per mutant 30S subunit at high levels of enzyme (data not shown). Taken together, these results demonstrate that *yggJ* encodes the MTase specific for  $m^3U1498$ . Based on this, we propose that the *yggJ* gene be renamed *rsmE*, as the protein product is the fourth ribosomal small subunit methyltransferase to be conclusively identified (Ofengand and Del Campo 2004).

### RsmE is not involved in methylation at the only other $m^3U$ site in ribosomal RNA

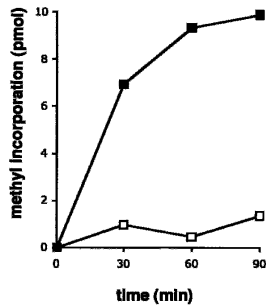
One additional  $m^3U$  is present in *E. coli*. This methylated nucleotide is  $m^3\psi1915$  in 23S rRNA. The uridine at this position is isomerized into a pseudouridine by RluD, but methylation at this position is independent of RluD (Raychaudhuri et al. 1998). It should be noted that pseudouridine does not lead to a stop in the primer extension with this procedure. We examined whether RsmE might have a dual specificity and also be involved in methylation of this residue. A reverse transcriptase specific stop at position 1916 was obtained when RNAs extracted from either wild-type or *rsmE::kan* strains were analyzed, while no stop was obtained in the case of synthetic 23S rRNA (data not shown). These results indicate that the methyl group at U1915 is not eliminated by deletion of *rsmE* and therefore that RsmE is not involved in methylation at this second site. Consequently, the  $m^3U$  methylase encoded by *rsmE* is highly specific for U1498 in 16S rRNA.

### Effect of *rsmE* on cell growth

Deletion of the *rsmE* gene leads to the absence of both the gene product and the U1498 methyl group. To determine if there is any physiological effect caused by this loss, we measured growth rates of wild-type and mutant cells



**FIGURE 3.** Overexpression and purification of recombinant  $m^3U1498$  methyltransferase. Protein samples were analyzed by SDS-PAGE. Cells transformed with pET28a plasmid containing the gene for  $m^3U1498$  methyltransferase were sampled before (lane 1) and after (lane 2) induction with 1 mM IPTG. Recombinant protein was affinity-purified, as described in Materials and Methods. (Lanes 3–6) 0.2, 0.4, 2.0, and 4  $\mu$ g, respectively, purified His-RsmE. (Lane M) Molecular mass standards with mass values on the left.



**FIGURE 4.** Methyl acceptor activity of wild-type and mutant 30S subunits. 30S particles (1  $\mu$ M) purified from either wild-type (open squares) or the *yggJ* mutant strain (filled squares) were methylated with [ $^3$ H]-SAM (100  $\mu$ M) and purified His-tagged RsmE (165 nM). The amount of incorporated methyl groups was monitored as in Figure 2. One experiment representative of three repeats is shown.

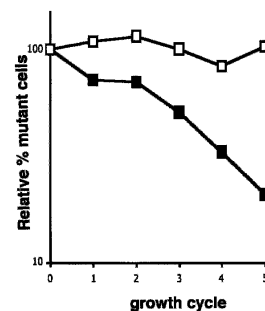
under a variety of conditions. No significant difference between the wild-type and the RsmE deletion strain was observed when cells were grown in either LB or M9 glucose media at 25°C, 37°C, or 42°C (data not shown). To determine whether any subtle differences in growth could be observed, mutant cells were grown in competition with wild-type cells using a previously described method (Gutgsell et al. 2000). In this procedure, equal amounts of wild-type and mutant cells are grown together over the entire range of growth cycle, and the mutant cells remaining in the culture at the end of each cycle are determined by resistance to kanamycin. As shown in Figure 5, there was an exponential decrease in the percent of mutant cells as the number of growth cycles increased. A different MTase deletion strain (MG1655 *rsmC::kan*) containing the same kanamycin cassette served as a control and showed essentially no decrease with the growth cycle. These data indicate that the effect observed was due to the absence of RsmE. We also introduced plasmid pHC79 carrying wild-type *rsmE* into the deletion–insertion strain and tested its growth in competition with wild-type cells carrying an empty plasmid. The presence of wild-type *rsmE* was sufficient to completely restore the ability of the mutant cells to grow in competition with wild-type cells (data not shown).

### RsmE belongs to a distinct protein family of previously unknown function

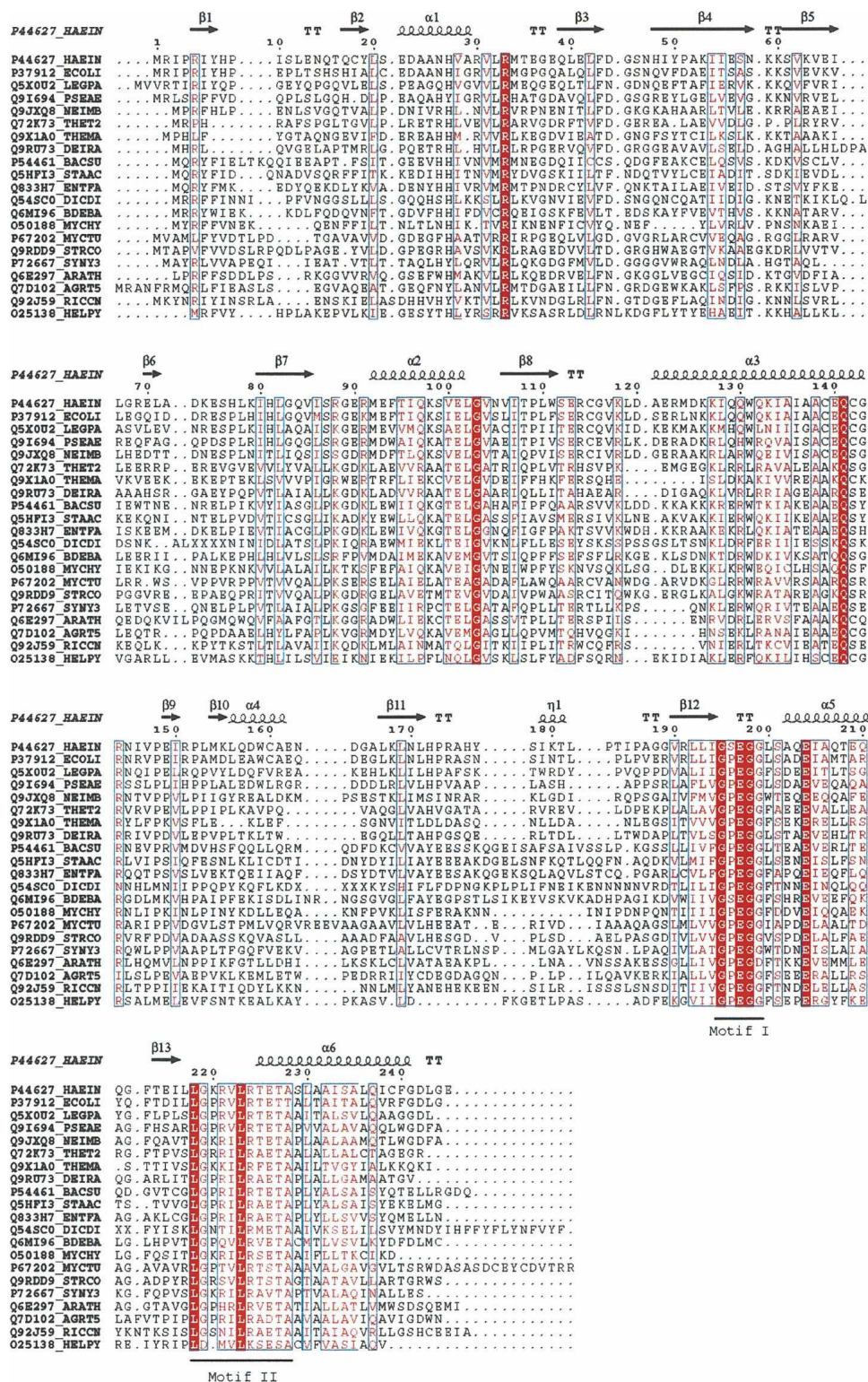
A Blastp search against the complete UniProtKB database (<http://www.ebi.ac.uk/uniprot/>) using the RsmE protein sequence as query and a threshold E value of 0.001 identified 184 bacterial and plant proteins with highly significant similarity to the m<sup>3</sup>U1498 MTase. Since the only plant homolog found in the Blastp search was from *Arabidopsis thaliana*, we searched the NCBI EST database with RsmE and found homologs in many plant species (data not shown). Twenty-one protein sequences, representative of

different groups of bacteria and plants, were selected for further analysis. The sequences, together with crystal structure data from *Haemophilus influenzae*, YggJ (Protein Data Bank [PDB] entry 1NXZ), *Bacillus subtilis*, YqeU (PDB entry 1VHK), *Thermotoga maritima* Tm1380 (PDB entry 1Z85), and *Thermus thermophilus* Tt1573 (PDB entry 1V6Z), were used in a structure-based multiple sequence alignment on the T-COFFEE server (<http://igs-server.cnrs-mrs.fr/Tcoffee>; Poirot et al. 2004). The alignment was visualized with ESript 2.2 (<http://esript.ibcp.fr/ESript/cgi-bin/ESript.cgi>; Gouet et al. 1999) and is presented along with the corresponding secondary structure elements defined by the crystal structure of the *H. influenzae* protein (Forouhar et al. 2003; Fig. 6). Analysis of the generated alignment revealed highly conserved residues that localize to regions suggested to be important for cofactor and substrate binding, as well as for dimerization. However, the conserved motifs differ from ones characteristic of the other known trefoil knot MTase families (Gustafsson et al. 1996; Anatharaman et al. 2002).

The most conserved motif is Motif I ([XGP/SEGG/D], where X is a hydrophobic residue), and spans region 193–198 (*E. coli* numbering). Pro and Ser are both present across bacterial species, while the last Gly in the motif can be replaced by Asp, primarily in plants. For example, a Gly to Asp change is found in *A. thaliana*, but it can also be accepted in prokaryotes since an Asp is present in *Poryphyromonas gingivalis* (accession no. Q7MT38). A second, slightly less conserved region spans residues 217–227. This region contains multiple residues with conservative changes, as well as highly conserved Leu residues at positions 217 and 223. Gln142, Arg144, and Glu203 are also well conserved. Based on the crystal structure, the first two localize to the surface of helix 3, while Glu203 is located on helix 5. They are all situated in close proximity to the



**FIGURE 5.** Growth competition of *rsmE* mutant cells and wild-type cells. MG1655 *rsmE::kan* (filled squares) and MG1655 *rsmC::kan* (open squares) were each mixed with equal amounts of wild-type cells and grown as described (Gutgsell et al. 2000). Cultures were diluted  $\sim 10^6$  prior to plating on LB plates to determine total cell number or on LB/kanamycin plates to determine the number of kanamycin-resistant cells. The percent of Kan<sup>r</sup> cells was determined at each cycle. The data were normalized to 100% at cycle 0 and plotted on a logarithmic scale.



**FIGURE 6.** Multiple sequence alignment of RsmE family protein sequences. The structure-based alignment was generated with T-COFFEE (<http://igs-server.cnrs-mrs.fr/Tcoffee>) and visualized with ESpript2.2 (<http://esript.ibcp.fr/ESript/cgi-bin/ESript.cgi>). Secondary structure elements shown were derived from the crystal structure of *H. influenzae* HI0303 (Yggj) protein, PDB entry 1NXZ. Manual modifications of the obtained alignment were performed as follows: (1) Fifty-two N-terminal residues from the *A. thaliana* sequence were removed; (2) an alternative start site for the *Deinococcus radiodurans* was used based on the lack of conservation of the first 11 residues, following the GenBank suggested start site and the lack of a potential ribosome binding site upstream of this start codon; (3) three sequence inserts of 37, 10, and 9 residues were removed from the *Dictyostelium discoideum* and are represented by four or two X signs starting at positions 75, 163, and 211, respectively (*H. influenzae* numbering). The SwissProt organism codes used are AGRT5, *Agrobacterium tumefaciens*; ARATH, *Arabidopsis thaliana*; BACSU, *Bacillus subtilis*; BDEBA, *Bdellovibrio bacteriovorus*; DEIRA, *D. radiodurans*; DDCDI, *D. discoideum*; ECOLI, *Escherichia coli*; ENTFA, *Enterococcus faecalis*; HAEIN, *Haemophilus influenzae*; HELPY, *Helicobacter pylori*; LEGPA, *Legionella pneumophila*; MYCHY, *Mycoplasma hyopneumoniae*; MYCTU, *Mycobacterium tuberculosis*; NEIMB, *Neisseria meningitidis*; PSEAE, *Pseudomonas aeruginosa*; RICCN, *Rickettsia conorii*; SYNY3, *Synechocystis* sp.; STAAC, *Staphylococcus aureus*; STRCO, *Streptomyces coelicolor*; THEMA, *Thermotoga maritima*; THET2, *Thermus thermophilus*.

SAM binding site and therefore might be involved in catalysis. However, the specific roles of these residues remain to be investigated. Other residues in the protein also may be involved in substrate recognition or in catalysis. For example, Arg33 is very well conserved and has been proposed, based on surface charge distribution, to be part of the RNA binding site (Forouhar et al. 2003).

All the other proteins in the RsmE family are of unknown function and constitute a previously uncharacterized protein family designated UPF0088 by SwissProt (<http://www.ebi.ac.uk/swissprot/>), DUF558 (accession no. PF04452) by Pfam (<http://www.sanger.ac.uk/Software/Pfam/>), and IPR006700 by InterPro (<http://www.ebi.ac.uk/interpro/>). Together with five other families, DUF558 belongs to the alpha/beta knot superfamily. Among these, only two families, grouped in the class of the SPOUT (SpoU-TrmD) MTases, have an attributed function. One, the SpoU methylase family, includes tRNA- and rRNA-specific ribose MTases, while the second, the tRNA m<sup>1</sup>G methylase family, includes the TrmD family of MTases. The other three families contain proteins of unknown function (<http://www.sanger.ac.uk/Software/Pfam/>). Based on our data and the protein sequence analysis, we suggest that members of the DUF558 family, possibly all orthologs of RsmE, are rRNA m<sup>3</sup>U MTases. We do not exclude the possibility that some of the orthologs can be methylases specific to other bases, but we do not believe that they can be 2'-O-ribose MTases, as was originally suggested for *yggJ* by Forouhar et al. (2003).

## DISCUSSION

RNA methylation is a process mediated by MTases that are site- or region-specific. The ability of these enzymes to accommodate various RNA substrates is mirrored in their structural diversity. All of their currently known structures reveal the presence of diverse substrate recognition domains. This is in dramatic contrast to their conserved three-dimensional catalytic domain, which is shared by members of most MTase families (Schubert et al. 2003). These structural features explain their capacity for performing the same catalytic reaction despite their lack of overall sequence similarity. Five MTase structural classes have been distinguished to date, but most of the identified RNA MTases belong to class I, characterized by the presence of a common, conserved Rossmann fold SAM binding domain (Schubert et al. 2003). Much less conservation is noticed at the sequence level, where only a few conserved motifs are present, most of them being part of the SAM binding region (Fauman et al. 1999). Greater sequence similarity can be found in specific subfamilies, such as the m<sup>2</sup>G MTases, m<sup>5</sup>C MTases, m<sup>6</sup>A MTases, m<sup>1</sup>G MTases, m<sup>5</sup>U MTases, or MTases of the RrmJ/fibrilysin family.

During the past few years, other SAM-dependent MTases were identified and were found to have a completely different SAM binding fold, an alpha/beta knot structure

(Schubert et al. 2003). Two of the six families of this class, SpoU and TrmD (SPOUT members), include ribose Gm and Cm (Renalier et al. 2005) or m<sup>1</sup>G MTases, respectively. Several additional uncharacterized *E. coli* genes, including *yggJ*, also were predicted to encode MTases with alpha/beta knot structures (Koonin and Rudd 1993; Forouhar et al. 2003). However, enzymes for all the ribose Gm and m<sup>1</sup>G sites in *E. coli* are already known, so these uncharacterized genes presumably code for MTases with other specificities.

The crystal structures of RsmE family proteins from *H. influenzae*, YggJ (PDB entry 1NXZ), *B. subtilis*, YqeU (PDB entry 1VHK), *T. maritima* Tm1380 (PDB entry 1Z85), and *T. thermophilus* Tt1573 (PDB entry 1V6Z), have been determined. They revealed dimeric proteins with a C-terminal domain that also contains the alpha/beta knot fold. Based on these findings it was suggested that *yggJ* codes for an RNA 2'-O-ribose MTase (Forouhar et al. 2003). However, contrary to this prediction, it is shown here that YggJ actually is an m<sup>3</sup>U-specific MTase, highly specific for position 1498 in 16S rRNA.

The alpha/beta knot fold defines a superfamily of predicted SAM MTases, with six known families. Two of these families (SPOUT members) were previously shown to contain bona fide MTases. Bioinformatic analysis (Fig. 6) indicates that YggJ is distinct from the SPOUT families and has its own specific motifs. The gene is very well conserved across bacteria, but family homologs are also found in plants. Each organism has a single YggJ family member, with no indication of horizontal gene transmission, suggesting they all may be isofunctional orthologs. Also, all the encoded proteins belong to DUF558, one of the four alpha/beta knot-containing predicted MTase families whose function has not yet been defined. Based on this study, we now propose that DUF558 is an m<sup>3</sup>U MTase family, and that its founding member, YggJ, be renamed RsmE. This conclusion is further supported by the gene's high degree of conservation across bacteria and plants, as illustrated by its phylogenetic tree (Fig. 7).

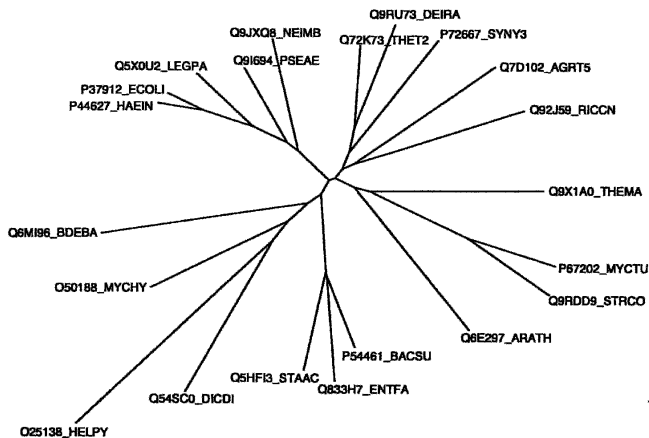
## MATERIALS AND METHODS

### Materials

[<sup>3</sup>H]-SAM was purchased from Amersham Pharmacia Biotech. [ $\alpha$ -<sup>32</sup>P]ATP was obtained from Perkin-Elmer. T4 DNA ligase, T4 polynucleotide kinase, and restriction enzymes were from New England Biolabs. DNase I was from Worthington. Taq DNA polymerase and PfuTurbo DNA polymerase were from Promega and Stratagene, respectively. Moloney murine leukemia virus reverse transcriptase (M-MLV-RT) was from Invitrogen.

### Buffers

Buffer A contained 20 mM HEPES (pH 7.5), 60 mM NH<sub>4</sub>Cl, 10.5 mM Mg (OAc)<sub>2</sub>, 0.5 mM EDTA and 10 mM 2-mercaptoetha-



**FIGURE 7.** Phylogenetic tree analysis of the RsmE family. The tree was generated using ClustalX1.81 based on the T-COFFEE alignment and visualized with PhyloDendron (<http://iubio.bio.indiana.edu/treeapp/>).

no; buffer B was the same as buffer A, but with 1 M  $\text{NH}_4\text{Cl}$  added; buffer C contained 20 mM HEPES (pH 7.5), 60 mM  $\text{NH}_4\text{Cl}$ , and 10 mM 2-mercaptoethanol; buffer D contained 50 mM  $\text{NaPO}_4$  (pH 7.0), 300 mM NaCl; buffer E was the same as buffer C with the addition of 120 mM  $\text{NH}_4\text{Cl}$  and 10% glycerol.

### Generation of knockout strains

MG1655 deletion strains were obtained by the gene replacement method of Datsenko and Wanner (2000). Deletion alleles containing the N- and C-terminal fragments of the deleted genes linked to a kanamycin cassette were constructed by PCR amplification from plasmid pKD13, using Taq DNA polymerase and appropriate primers. The wild-type genes were then replaced with the PCR amplified alleles as described (Datsenko and Wanner 2000). Selection on kanamycin plates and colony PCR using primers specific to regions flanking the candidate genes were used to identify the deletion strains.

### Primer extension analysis

Total cellular RNA was extracted as described (Ofengand et al. 2001), and annealed to a [ $^{32}\text{P}$ ]5'-end labeled primer specific to region 1506–1525 of 16S rRNA. The primer was extended by M-MLV-RT according to the manufacturer's protocol. Primer extension products were precipitated with ethanol; dissolved in DNA loading buffer containing 89% formamide, 4% TE buffer (10 mM Tris-HCl at pH 8.0, 1 mM EDTA), 0.12% bromophenol blue, and 0.17% xylene cyanol; separated on 8% acrylamide/7 M urea sequencing gels; and visualized by exposure to X-ray film.

### Gene complementation

The *yggJ* gene region containing 200 bp upstream of the putative ATG start codon predicted in EcoGene (<http://www.ecogene.org>) was amplified from genomic DNA by PCR, using PfuTurbo DNA polymerase. The N-terminal primer was 5'-TCTGTCGACT

TAGCCCAAATCGCCA-3', which introduced an EcoRI restriction site. The C-terminal primer was, in reverse orientation, 5'-TTTGAATTCATGCGTATCCCCCGCA-3', and introduced a Sall restriction site. The amplified fragment was separated by low-melt agarose gel electrophoresis, phenol extracted, ethanol precipitated, and digested with the appropriate restriction enzymes. The digested fragment was ligated with T4 DNA ligase into pHC79 plasmid (Hohn and Collins 1980), digested with the same restriction enzymes, and transformed into JM109 cells. Plasmids containing the correct size insert were extracted and transformed into MG1655 *yggJ*::kan. The same strain was also transformed with the empty pHC79 vector as a control. Transformed cells were grown in LB medium containing 50  $\mu\text{g}/\text{mL}$  kanamycin and 100  $\mu\text{g}/\text{mL}$  carbenicillin at 37°C to logarithmic phase prior to further analysis.

### Preparation of S100 extracts and purification of 30S ribosomal particles

S100 extracts and 30S ribosomal particles were prepared from the RNase I-deficient strains MG1655 I::tet (Gutgsell et al. 2005) and MG1655 I::tet *yggJ*::kan. The MG1655 I::tet *yggJ*::kan was obtained by a phage P1 mediated transduction between MG1655 I::tet and MG1655 *yggJ*::kan. The two strains were grown to logarithmic phase ( $A_{600} \sim 0.6\text{--}0.8$ ), washed with Tris buffer saline (TBS) containing 10 mM Mg (OAc) $_2$ , and frozen at  $-80^\circ\text{C}$ . Pellets were allowed to thaw in 2 mL of buffer A per gram of wet cell pellet and passed through a French press at 18,000 psi. Cell debris was removed by centrifugation at 16,000 rpm for 30 min in a SS-34 Sorvall rotor. The supernatant fraction was treated with 2 units of DNase I on ice for 15 min and centrifuged in a Beckman 60Ti rotor at 32,000 rpm for 4 h to obtain ribosome free S100 extracts. The S100 fraction was dialyzed against buffer A without  $\text{Mg}^{2+}$  and EDTA. The wild-type and mutant ribosome pellets were washed with buffer B and centrifuged through an 18% sucrose cushion for 18 h at 44,000 rpm in a Beckman 60Ti rotor. The ribosomal pellet was resuspended in buffer B containing only 2 mM  $\text{Mg}^{2+}$  to allow disassembly into 50S and 30S particles. The two subunits were separated by centrifugation on a 10%–30% sucrose gradient for 17 h at 21,000 rpm in a Beckman SW28 rotor. Fractions containing the 30S particles were collected, concentrated by centrifugation through an Amicon Ultra 30K molecular cutoff filter, quickly frozen in liquid nitrogen, and stored at  $-80^\circ\text{C}$ .

### Cloning and overexpression of the $\text{m}^3\text{U}$ methyltransferase gene

The *yggJ* gene was amplified by PCR of genomic DNA, using PfuTurbo DNA polymerase. The N-terminal primer was 5'-TTTGCTAGCATGCGTATCCCCCGCA-3', which introduced a NheI restriction site. The C-terminal primer, in reverse orientation, was 5'-TCTCTCGAGTTAGCCCAAATCGCCA-3' and introduced an XhoI restriction site. The amplified product was purified as described above and digested with NheI and XhoI. The pET28a vector (Novagen), also digested with NheI and XhoI, and the gene product were ligated with T4 DNA ligase and transformed into JM109 cells. This procedure resulted in isolation of plasmids containing the gene insert fused with an N-terminal His-tag cod-

ing region. DNA sequencing of the gene-containing plasmid showed that the correct gene had been amplified. For overexpression, BL21 (DE3) cells were transformed with isolated plasmids and grown in LB medium containing 50  $\mu\text{g}/\text{mL}$  kanamycin at 37°C to an  $A_{600}$  of 0.6. Isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) at 1 mM was added to the culture, and the incubation continued for 2 h. Cells were recovered by centrifugation, washed with TBS, and stored at  $-80^\circ\text{C}$ .

### Affinity purification

His-tagged YggJ was purified by cobalt-immobilized metal affinity chromatography on BD Talon resin, according to the manufacturer's instructions. Briefly, cell pellets were thawed on ice using 2 mL buffer B per 25 mL cell culture and passed through a French press at 18,000 psi. Cell debris was removed by centrifugation at 16,000 rpm for 20 min in a SS-34 Sorvall rotor. The supernatant fraction was treated with 2 units of DNase I on ice for 15 min. His-tagged protein in the supernatant fraction was allowed to bind to the Talon resin pre-equilibrated with buffer D, while rocking on a platform at 4°C for 1 h. Unbound protein was removed by washing the resin three times with buffer D. The resin was then transferred to a 2-mL gravity flow column and washed with five column volumes of buffer D. Protein was eluted with 5 mL of buffer D containing 150 mM imidazole, collected in 500- $\mu\text{L}$  fractions, and then analyzed by SDS-PAGE. Fractions containing pure protein were pooled, dialyzed against buffer E, quickly frozen in liquid nitrogen, and stored at  $-80^\circ\text{C}$ . Protein concentration was determined from the  $A_{280}$  absorption value and by the method of Bradford (1976). The absorption coefficient ( $\epsilon$ ; = 13,940  $\text{M}^{-1}\text{cm}^{-1}$ ) was determined based on the number of tryptophan and tyrosine residues according to the method of Gill and von Hippel (1989).

### Methylase assays

Reaction mixtures contained 20 mM HEPES (pH 7), 100 mM  $\text{NH}_4\text{Cl}$ , 5 mM DTT, 100  $\mu\text{M}$  SAM, 1  $\mu\text{Ci}$  [ $^3\text{H}$ ]-SAM, 10 mM  $\text{Mg}(\text{OAc})_2$ , substrate and protein as indicated. Reaction mixtures (50  $\mu\text{L}$ ) were incubated at 37°C and the reaction product was precipitated with 2.5 mL ice-cold 10% TCA/2% casamino-acids solution. Fifty micrograms of salmon sperm DNA was added per reaction to facilitate precipitation. Samples were incubated on ice for 10 min prior to collection on microglass fiber filters (VWR). Filters were washed with 24 mL 2.5%TCA/2% casamino acids, followed by 5 mL 1:1 95% ethanol/diethyl ether and allowed to dry. Radioactivity was determined by scintillation counting.

### Cell growth rate and growth competition

To determine the exponential growth rate, overnight cultures of the *E. coli* strains were diluted to an  $A_{600} \sim 0.01$  in 50 mL of medium in 500-mL flasks and incubated at different temperatures with shaking at 250 rpm. Measurements of  $A_{600}$  were determined from samples taken at equal time intervals. Doubling times were determined from a plot of the log  $A_{600}$  versus time. Growth competition assays were performed as described (Gutgsell et al. 2000).

### ACKNOWLEDGMENTS

This work was initiated under the direction of Professor James Ofengand (deceased December 6, 2004). This work was supported by grant GM58879 from the National Institutes of Health.

Received November 4, 2005; accepted December 1, 2005.

### REFERENCES

- Anatharaman, V., Koonin, E.V., and Aravid, L. 2002. SPOUT: A class of methyltransferases that includes spoU and trmD RNA methylase superfamilies and novel superfamilies of predicted prokaryotic RNA methyltransferases. *J. Mol. Microbiol. Biotechnol.* **4**: 71–75.
- Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**: 248.
- Bujnicki, J.M. 2000. Phylogenomic analysis of 16S rRNA:(guanine-N2) methyltransferases suggests new family members and reveals highly conserved motifs and a domain structure similar to other nucleic acid amino-methyltransferases. *FASEB J.* **14**: 2365–2368.
- Bujnicki, J.M. and Rychlewski, L. 2002. RNA:(guanine-N2) methyltransferases RsmC/RsmE and their homologs revisited—Bioinformatic analysis and prediction of the active site based on the uncharacterized MJ0882 protein structure. *BMC Bioinformatics* **3**: 10.
- Datsenko, K.A. and Wanner, B.L. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci.* **97**: 6640–6645.
- Decatur, W.A. and Fournier, M.J. 2002. rRNA modifications and ribosome function. *Trends Biochem. Sci.* **27**: 344–351.
- Fauman, E.B., Blumenthal, R.M., and Cheng, X. 1999. Structure and evolution of AdoMet-dependent methyltransferases. In *S-adenosylmethionine-dependent methyltransferases: Structures and functions* (eds. X. Cheng and R.M. Blumenthal), pp. 1–38. World Scientific, River Edge, NJ.
- Forouhar, F., Shen, J., Xiao, R., Acton, T.B., Montelione, G.T., and Tong, L. 2003. Functional assignment based on structural analysis: Crystal structure of the yggJ protein (HI0303) of *Haemophilus influenzae* reveals an RNA methyltransferase with a deep trefoil knot. *Proteins* **53**: 329–332.
- Gill, S.C. and von Hippel, P.H. 1989. Calculation of protein extinction coefficients from amino-acids sequence data. *Anal. Biochem.* **182**: 319–326.
- Gouet, P., Courcelle, E., Stuart, D.I., and Metz, F. 1999. ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305–308.
- Gustafsson, C., Reid, R., Greene, P.J., and Santi, D.V. 1996. Identification of new RNA modifying enzymes by iterative genome search using known modifying enzymes as probes. *Nucleic Acids Res.* **24**: 3756–3762.
- Gutgsell, N., Englund, N., Niu, L., Kaya, Y., Lane, B.G., and Ofengand, J. 2000. Deletion of the *Escherichia coli* pseudouridine synthase gene *truB* blocks formation of pseudouridine 55 in tRNA in vivo, does not affect exponential growth, but confers a strong selective disadvantage in competition with wild-type cells. *RNA* **6**: 1870–1881.
- Gutgsell, N., Deutscher, M.P., Ofengand, J. 2005. The pseudouridine synthase RluD is required for normal ribosome assembly and function in *Escherichia coli*. *RNA* **11**: 1141–1152.
- Hayes, F., Hayes, D., Fellner, P., and Ehresmann, C. 1971. Additional nucleotide sequences in precursor 16S ribosomal RNA from *E. coli*. *Nat. New Biol.* **232**: 54–55.
- Hohn, B. and Collins, J. 1980. A small cosmid for efficient cloning of large DNA fragments. *Gene* **11**: 291–298.
- Katz, J.E., Dlakic, M., and Clarke, S. 2003. Automated identification of putative methyltransferases from genomic open reading frames. *Mol. Cell Proteomics* **2**: 525–540.



- Koonin, E.V. and Rudd, K.E. 1993. SpoU protein of *Escherichia coli* belongs to a new family of putative rRNA methylases. *Nucleic Acids Res.* **21**: 5519.
- Negre, D., Weitzmann, C., and Ofengand, J. 1989. In vitro methylation of *Escherichia coli* 16S ribosomal RNA and 30S ribosomes. *Proc. Natl. Acad. Sci.* **86**: 4902–4906.
- Ofengand, J. and Del Campo, M. 2004. Modified nucleosides in *Escherichia coli* ribosomal RNA. In *EcoSal—Escherichia coli and salmonella: Cellular and molecular biology* (ed. R. Curtiss) ASM Press, Washington, DC. <http://www.ecosal.org>.
- Ofengand, J., Del Campo, M., and Kaya, Y. 2001. Mapping pseudouridines in RNA molecules. *Methods* **25**: 365–373.
- Poirot, O., Suhre, K., Abergel, C., O'Toole, E., and Notredame, C. 2004. 3DCoffee@igs: A web server for combining sequences and structures into a multiple sequence alignment. *Nucleic Acids Res.* **32**: W37–W40.
- Raychaudhuri, S., Conrad, J., Hall, B.G., and Ofengand, J. 1998. A pseudouridine synthase required for the formation of two universally conserved pseudouridines in ribosomal RNA is essential for normal growth of *Escherichia coli*. *RNA* **4**: 1407–1417.
- Reid, P., Greene, P.J., and Santi, D.V. 1999. Exposition of a family of RNA m<sup>5</sup>C methyltransferases from searching genomic and proteomic sequences. *Nucleic Acids Res.* **27**: 3138–3145.
- Renalier, M.H., Joseph, N., Gaspin, C., Threbault, P., and Mougin, A. 2005. The Cm56 tRNA modification in archaea is catalyzed either by a specific 2'-O-methylase, or a C/D sRNP. *RNA* **11**: 1051–1063.
- Ringquist, S., Cunningham, P., Weitzmann, C., Formenoy, L., Pleij, C., Ofengand, J., and Gold, L. 1993. Translation initiation complex formation with 30 S ribosomal particles mutated at conserved positions in the 3'-minor domain of 16 S RNA. *J. Mol. Biol.* **234**: 14–27.
- Schubert, H.L., Blumenthal, R.M., and Cheng, X. 2003. Many paths to methyltransfer: A chronicle of convergence. *Trends Biochem. Sci.* **28**: 329–335.
- Tscherne, J.S., Nurse, K., Popienick, P., Michel, H., Sochacki, M., and Ofengand, J. 1999a. Purification, cloning, and characterization of the 16S RNA m<sup>3</sup>C967 methyltransferase from *Escherichia coli*. *Biochemistry* **38**: 1884–1892.
- Tscherne, J.S., Nurse, K., Popienick, P., and Ofengand, J. 1999b. Purification, cloning, and characterization of the 16 S RNA m<sup>2</sup>G1207 methyltransferase from *Escherichia coli*. *J. Biol. Chem.* **274**: 924–929.
- van Buul, C.P. and van Knippenberg, P.H. 1985. Nucleotide sequence of the ksgA gene of *Escherichia coli*: Comparison of methyltransferases effecting dimethylation of adenosine in ribosomal RNA. *Gene* **38**: 65–72.
- van Knippenberg, P.H. 1986. Structural and functional aspects of the N<sup>6</sup>,N<sup>6</sup>-dimethyladenosines in 16S ribosomal RNA. In *Structure, function, and genetics of ribosomes* (eds. B. Hardesty and G. Kramer)pp. 412–424. Springer Verlag, Berlin.
- Yusupov, M.M., Yusupova, A., Baucom, K., Earnest, T.N., Cate, J.H., and Noller, H.F. 2001. Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**: 883–896.