Purification, Cloning, and Characterization of the 16S RNA m^{5}C967 Methyltransferase from *Escherichia coli*†

Joan S. Tscherne,‡,§ Kelvin Nurse,‡,|| Paul Popienick,‡,⊥ Hanspeter Michel,# Marek Sochacki,#,@ and James Ofengand*,‡

Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, Miami, Florida 33101, and Hoffman-La Roche, Inc., Nutley, New Jersey 07110

Received August 5, 1998; Revised Manuscript Received December 17, 1998

ABSTRACT: The methyltransferase that forms m^{5}C967 in *Escherichia coli* small subunit ribosomal RNA has been purified, cloned, and characterized. The gene was identified from the N-terminal sequence of the purified enzyme. The gene is a fusion of two open reading frames, *fmu and fmv*, previously believed to be distinct due to a DNA sequencing error. The gene, here named *rsmB*, encodes a 429-amino acid protein that has a number of homologues in prokaryotes, Archaea, and eukaryotes. C-Terminal sequencing of the overexpressed and affinity-purified protein by mass spectrometry methods verified the sequence expected for the gene product. The recombinant protein exhibited the same specificity as the previously described native enzyme; that is, it formed only m^{5}C and only at position 967. C1407, which is also m^{5}C in natural 16S RNA, was not methylated. In vitro, the enzyme only recognized free 16S RNA. 30S ribosomal subunits were not a substrate. There was no requirement for added magnesium, suggesting that extensive secondary or tertiary structure in the RNA substrate may not be a requirement for recognition.

Ribosomal RNA is believed to be the functional heart of the ribosome, containing within itself the structural determinants that define the essentials of ribosome activity (*1*). This is true for the SSU[1] RNA and its decoding function (*2*) as well as for the LSU RNA and its peptidyl transferase activity (*3*, *4*). A major unanswered question in this regard is the role played by the defined set of modified nucleotides that are present in both SSU and LSU RNAs of all ribosomes.

The SSU RNA of *Escherichia coli* contains 10 methylated nucleosides and one Ψ. The Ψ is located at position 516 (*5*), adjacent to the conserved sequence of residues 518−533 which is known to be involved in the decoding process (reviewed in refs *2* and *6*). The synthase which forms this Ψ, and its associated gene, have been identified (*7*). The 10 methylated nucleosides are distributed along the 16S RNA. m^{7}G527 is found within the conserved segment of residues 518−533 mentioned above. m^{2}G966 and m^{5}C967 are located in the loop of a small stem−loop structure which has been implicated in tRNA binding at the P site (*8*). In eukaryotes, these two methylations are replaced by a single highly modified residue, m^{1}acp^{3}U (*9*). m^{2}G1207 occurs in a region believed to be involved in recognition of peptide chain termination codons (*10*). m^{4}Cm1402 and m^{5}C1407 are in the conserved sequence segment of residues 1394−1408, and m^{3}U1498 is in the third conserved sequence region in SSU RNA (residues 1492−1505). These latter two segments are known to interact functionally and are also part of the decoding site (*11*). m^{2}G1516, m$^6_2$A1518, and m$^6_2$A1519 form a tight cluster of methylated bases at the tip of the 3′-terminal stem of 16S RNA. In the proposed tertiary structure for 16S RNA, these 10 seemingly distant modified residues come together forming a compact cage surrounding the location of the anticodon stem−loop structures of A and P site-bound tRNAs (*12*).

The fact that all of the methylated residues occur within functionally significant sites suggests that they should fulfill some specific role. Nevertheless, the nature of that role is unknown. The absence of the two m$^6_2$A residues, caused by a mutation which inactivated the methyltransferase responsible for their biosynthesis, makes the ribosome resistant to the antibiotic kasugamycin, but did not lead to a loss of normal ribosomal function (*13*). The other methylated bases also appear to be nonessential for protein synthesis since ribosomes constructed from totally unmodified 16S RNA are

† Experimental work was performed at the now-defunct Roche Institute of Molecular Biology. Supported by funds from Hoffman-La Roche, Inc. (J.O. and H.M.), and a Markey Foundation grant to the Department of Biochemistry and Molecular Biology, University of Miami School of Medicine (J. O.).

* Corresponding author. E-mail: jofengan@molbio.med.miami.edu. Fax: (305) 243-3955. Postal address: P.O. Box 016129, Miami, FL 33101.

‡ University of Miami School of Medicine.

§ Present address: Department of Science and Technology, Bergen Community College, Paramus, NJ 07652.

|| Present address: RNA Research Group, Department of Molecular Recognition, Smithkline Beecham Pharmaceuticals, 709 Swedeland Rd., King of Prussia, PA 19406.

⊥ Present address: Tarrant County College, Arlington, TX 76018.

Hoffman-La Roche, Inc.

@ Present address: Center for Macromolecular Studies, Polish Academy of Sciences, Sienkiewicza 90-363, Lodz, Poland.

[1] Abbreviations: SSU, small subunit; LSU, large subunit; Ψ, pseudouridine; SAM, *S*-adenosylmethionine; PVDF, poly(vinylidene difluoride); FPLC, fast protein liquid chromatography; IPTG, isopropyl β-D-thiogalactopyranoside; PCR, polymerase chain reaction; DTT, dithiothreitol; ME, mercaptoethanol; PMSF, phenylmethanesulfonyl fluoride; HPLC, high-performance liquid chromatography; Mes, 4-morpholineethanesulfonic acid; ORF, open reading frame.

10.1021/bi981880l CCC: $18.00
Published on Web 01/22/1999

able to carry out all of the partial reactions of in vitro protein synthesis, albeit at ca. half-efficiency (*14*, *15*). There is evidence, however, to suggest that this reduction in efficiency may be due to the lack of one or more of the modified nucleosides (*16*).

The enzymes responsible for the biosynthesis of the methylated nucleosides discussed above are not known, except for the one that makes both m_2^6A residues, and three activities reported previously from this laboratory (*17*−*19*). One of the activities was shown to be specific for formation of m^5C967 and was purified and characterized (*17*−*19*). Only the $m_2^6A1518-1519$-forming enzyme has been cloned (*20*). In this work, we have identified the gene responsible for the m^5C967 methyltransferase by N-terminal sequence analysis of the purified native enzyme, and confirmed the gene identification by overexpression, affinity purification, and characterization of the cloned protein product.

EXPERIMENTAL PROCEDURES

Materials. *S*-Adenosyl[^{3}H]methionine ([^{3}H]SAM) was from Amersham. RNase T1 was from Calbiochem. Wizard DNA Purification Kits and RNasin were from Promega. BA85 cellulose nitrate filters were obtained from Schleicher & Schuell. PVDF membranes were from Millipore Corp. Omega Cells were from Filtron. DEAE Sepharose CL-6B and MonoS FPLC columns were obtained from Pharmacia. Plasmid pCR-Script was from Stratagene. Plasmid pET-15b, the BL21/DE3 and Novablue strains of *E. coli*, and His-Bind resin were from Novagen, Inc. DH5α *E. coli* cells and RNase P1 were from GIBCO-BRL. T4 DNA ligase and restriction enzymes were from New England Biolabs. DNase I was from Worthington. Alumina and IPTG were from Sigma. Deoxyoligonucleotides for the protection experiments were those described previously (*17*). Primers for PCR were prepared as described by Bakin and Ofengand (*21*).

Buffers. Buffer A_x is 20 mM Hepes (pH 7.5 or 8.0 as indicated), 10 mM $Mg(OAc)_2$, *x* mM NH_4Cl, and 2 mM DTT. Buffer B_x is 20 mM Hepes (pH 7.5), 1 mM EDTA, *x* mM NH_4Cl, and 2 mM DTT. Buffer C is 20 mM Hepes (pH 7.8), 20 mM NH_4Cl, 5 mM ME, 0.1 mM EDTA, and 10% glycerol. Buffer D is 20 mM Hepes (pH 8.0), 100 mM NH_4Cl, 5 mM ME, 0.1 mM EDTA, and 6 M urea. Buffer LB is 50 mM Tris-HCl (pH 6.8), 100 mM DTT, 2% SDS, 10% glycerol, and 0.1% bromphenol blue. Binding buffer is 20 mM Hepes (pH 7.9), 0.5 M NaCl, and 5 mM imidazole. Elution buffer is binding buffer with the imidazole concentration raised to 1.0 M.

Purification of the m^5C967 Methyltransferase. The purification procedure was similar to that previously described (*18*, *19*). One hundred grams of *E. coli* MRE600 frozen cell paste (Grain Processing Corp.) grown in M9 medium, harvested in midlog phase, and washed in saline was ground with 200 g of alumina. Three thousand units of DNase I and 100 mL of buffer A_{100} (pH 7.5) containing 0.2 mM PMSF were added, and the mixture was incubated for 15 min at 0 °C. Centrifugation in a Sorvall SS-34 rotor for 30 min at 16 000 rpm yielded the S30. The S30 supernatant was then centrifuged for 4 h at 30 000 rpm in a Beckman Ti45 rotor. The pellet was resuspended in 80 mL of buffer A_{1000} (pH 8.0), underlayed with 40 mL of the same buffer and 37% sucrose, and centrifuged in the same rotor at 45 000 rpm for 16−18 h using slow acceleration. The supernatant of this centrifugation is the high-salt wash. This fraction was adjusted to 35% saturation with $(NH_4)_2SO_4$ and allowed to stir on ice for 4 h. The precipitate was removed by centrifugation, and the supernatant was dialyzed overnight against buffer B_{50}. The dialyzed sample was loaded onto a 5 cm × 43 cm DEAE-Sepharose CL-6B column equilibrated with buffer B_{50} and eluted with a linear gradient of 50 to 500 mM NH_4Cl in buffer B_0. The m^5C methyltransferase activity eluted with a peak at 130 mM NH_4Cl. The peak fractions were pooled, concentrated, and returned to buffer B_{50} by filtration through an Omega cell. The sample was loaded onto a 0.5 cm × 5 cm MonoS column which had been equilibrated with buffer B_{50}. The column was eluted with a linear gradient from 50 to 1000 mM NH_4Cl in buffer B_0. The methyltransferase peak activity eluted at 315 mM NH_4Cl. The peak fractions were pooled and either used directly or made 50% in glycerol for storage at −20 °C.

Cloning, Overexpression, and Affinity Purification of the m^5C Methyltransferase Gene and Gene Product. The putative gene was amplified by PCR. The N-terminal primer extended from position −9 to 22 where A of the initiating ATG codon is 1, with a change at −2 for creation of an *Nde*I site adjacent to the initiating ATG codon. The C-terminal primer, in the reverse orientation, extended from position 1309 to 1339 where the last sense nucleotide is position 1287 (see the Results), and contained mismatches at positions 1331 and 1333 for creation of a *Bam*HI site. After PCR, the amplified product was isolated by agarose gel electrophoresis, extracted using the Wizard DNA Purification Kit, concentrated by membrane filtration (Amicon Microcon 100), inserted into pCR-Script, and used to transform DH5α cells using standard methods. Plasmid DNA was isolated, the gene cut out with *Nde*I and *Bam*HI, and the fragment purified by agarose gel electrophoresis. pET-15b was digested with *Nde*I and *Bam*HI, purified by agarose gel electrophoresis, and incubated with the purified gene insert overnight at 16 °C in a ligation mixture containing 50 mM Tris (pH 7.9), 10 mM $MgCl_2$, 25 μg/mL nuclease-free bovine serum albumin, 10 mM DTT, 1 mM ATP, and T4 DNA ligase. Transformation of Novablue cells was performed by standard methods. Plasmids from two clones containing the insert were transferred into BL21/DE3 cells.

For overexpression, the transformed BL21/DE3 cells were grown in LB (*22*) with 50 μg/mL carbenicillin at 30 °C to an A_{600} of 0.6. IPTG (1 mM) was added to the cells, and incubation was continued at 30 °C to an A_{600} of 1.1−1.6. Cells were recovered by centrifugation and frozen at −70 °C. For analysis of the whole cell contents, cells from 1 mL of culture were suspended in 30 μL of 10 mM Tris (pH 8.0) and an equal amount of 2× buffer LB, heated to 100 °C for 5 min, and then chilled. For preparation of recombinant protein, cells from 100 mL of culture were recovered by centrifugation, washed in 10 mM Hepes (pH 7.5) and 0.17 M NaCl, then sonicated in 6 mL of binding buffer which contained 1 mM PMSF, and centrifuged at 15000*g* to obtain the S15 supernatant and pellet fractions. The pellet was treated in 6 mL of binding buffer containing 6 M urea for 1 h at 0 °C and then centrifuged at 39000*g* to removed insoluble material. The supernatant of this centrifugation is referred to as the S15 pellet (see Figure 2).

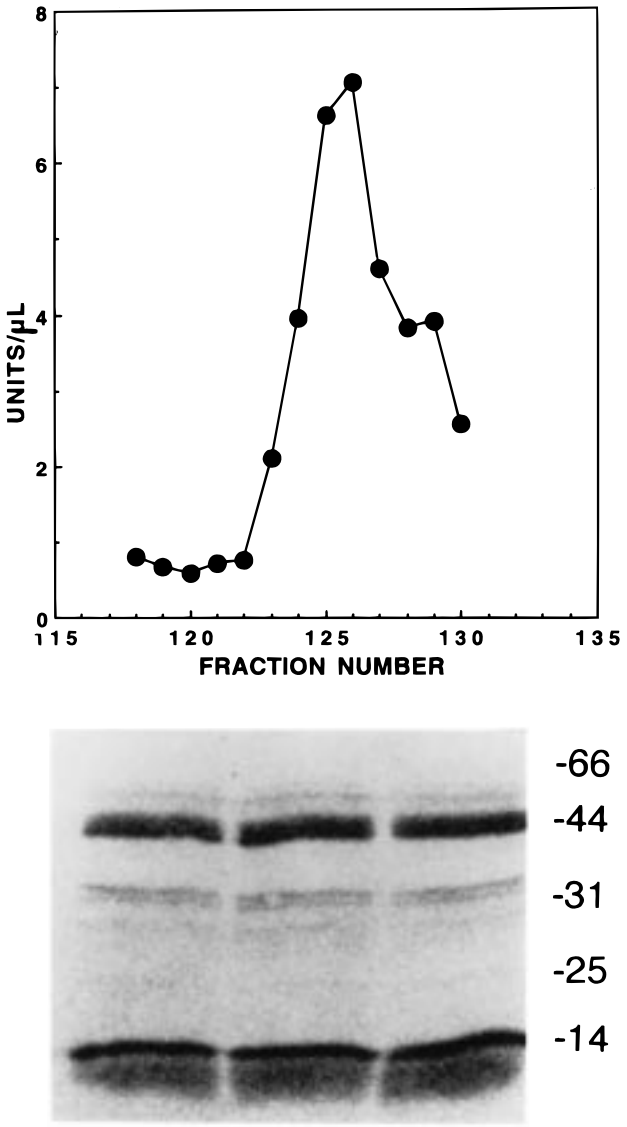

FIGURE 1: Purification of the 16S RNA m^{5}C967 methyltransferase. The top panel shows the final step in the purification of the m^{5}C methyltransferase on an FPLC MonoS column. Chromatography was conducted as described in Experimental Procedures. The bottom panel shows the gel electrophoretic analysis of the purified m^{5}C methyltransferase. Positions of molecular mass standards (Bio-Rad) are shown at the right. The upper heavy band at approximately 44 kDa corresponds to the methyltransferase. The lower heavy band at approximately 14 kDa was shown to be ribosomal protein L9.

The S15 supernatant was applied to a 2.5 mL column of His-Bind resin. Conditions of preparation and operation of the column were as described in the pET System Manual, 4th ed. (Novagen, Inc.). Upon addition of elution buffer, the tagged protein was released. The A_{280}-containing fractions were pooled and dialyzed against buffer C. The S15 pellet from the same culture, solubilized in binding buffer containing 6 M urea as described above, was applied to a 2.5 mL His-Bind column equilibrated in the same buffer. Elution was carried out with elution buffer plus 6 M urea. The pooled A_{280}-containing fractions were dialyzed against buffer D with decreasing concentrations of urea from 6 to 3 M at 1.0 M intervals and then from 3 to 0 M at 0.5 M intervals for 1 h each. Both protein solutions were adjusted to contain 50% glycerol and stored at −20 °C.

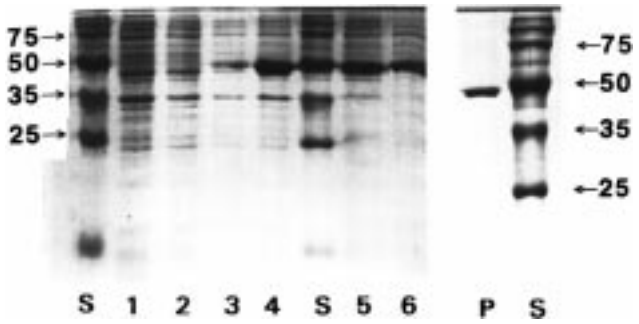

FIGURE 2: Overexpression of the recombinant 16S RNA m^{5}C967 methyltransferase. Cells were transformed with either pET (which lacked the gene insert) or pET/rsmB (which contained the gene for the m^{5}C967 methyltransferase) as described in Experimental Procedures. Cells were sampled before (−) or after (+) induction with IPTG: lane S, molecular mass standards (Novagen); lane 1, pET without IPTG; lane 2, pET with IPTG; lane 3, pET/rsmB without IPTG; lane 4, pET/rsmB with IPTG; lane 5, pET/rsmB S-15 supernatant; lane 6, pET/rsmB S-15 pellet; and lane P, affinity-purified m^{5}C methyltransferase prepared as described in Experimental Procedures.

Methylase Assays. Reaction mixtures contained 100 mM Hepes (pH 7.5), 4 mM $Mg(OAc)_2$ except where otherwise indicated, 200 mM NH_4Cl, either 5 mM DTT or ME, 2 μM [^{3}H]SAM, 500 units/mL RNasin, 100 nM synthetic 16S RNA (prepared as described in ref *7*) or 80 nM 30S ribosomal subunits reconstituted with synthetic 16S RNA (*14*), and enzyme as indicated. Incubation was carried out at 37 °C or as indicated. The reaction was stopped by adding cold 5% trichloroacetic acid. The samples were allowed to precipitate at 0 °C for 10 min and collected on BA85 cellulose nitrate filters. The filters were dissolved in scintillation fluid, and the amount of radioactivity was measured. One unit of activity equals 1 pmol of methyl incorporated per hour under the above conditions.

Nucleoside Analysis. Recombinant m^{5}C methyltransferase was used to radioactively ^{3}H-methylate 16S RNA as described above. The [^{3}H]RNA (7.3 pmol/mL) was digested in a 250 μL incubation mixture (20 mM NaOAc, 0.5 mM $ZnSO_4$, 40 pmol/mL carrier 16S RNA, and 80 μg/mL RNase P1) for 2 h at 37 °C. Then 15.4 μL of 1 M Tris buffer (pH 8.0) and 5.4 units of bacterial alkaline phosphatase were added, and the volume was adjusted to 275 μL and the incubation continued for an additional 2 h at 37 °C. The reaction was stopped by adding acetic acid to a concentration of 27 mM. m^{5}C was added as an internal standard and HPLC conducted as described previously (*17*).

Protection Studies. [^{3}H]Methyl-labeled 16S RNA (0.36 pmol) was hybridized with a 50-fold excess of deoxyoligonucleotide in a 50 μL reaction mixture which also contained 40 mM Mes (pH 6.4), 400 mM NaCl, 9 mM EDTA, and freshly deionized 80% (v/v) formamide. The samples were heat-denatured at 90 °C for 10 min and then cooled at room temperature for 15 min. They were then diluted with 9 volumes of ice-cold RNase buffer [10 mM Tris-HCl (pH 7.5), 300 mM NaCl, and 5 mM EDTA], and RNase T1 (0.2−2.4 Sankyo units) was added. The samples were digested for 30 min at 30 °C. Digestion was stopped by adding 9 volumes of cold 10% TCA. The samples were allowed to precipitate for 10 min on ice and then were filtered on BA85 cellulose nitrate filters; the filters were dissolved, and the amount of radioactivity was counted. Blanks obtained by

digestion with excess RNase T1 plus RNase A were subtracted.

N-Terminal Sequencing. Samples of the purified enzyme were electrophoresed as described below, electroblotted onto a PVDF membrane following standard procedures (*23*), and sequenced. N-Terminal sequencing was carried out as previously described (*24*).

C-Terminal Sequencing by Mass Spectrometry. *(a) Proteolytic Digestion.* For digestion with endoprotease Lys-C (Wako), 2 μg of enzyme was added to 100 μg of protein in 0.3 mL of 100 mM Tris-HCl (pH 9.0), 2 M urea, 0.3 M imidazole, and 0.13 M NaCl and the mixture incubated for 20 h at 37 °C. For the digestion with endoproteinase Glu-C (Sigma), 2 μg of enzyme was added to 100 μg of protein in 0.5 M ammonium bicarbonate, 3 M urea, 0.5 M imidazole, and 0.25 M NaCl and the mixture incubated for 20 h at room temperature. After digestion, the samples were reduced with 10 mM DTT (10 min at 37 °C) and then acidified with acetic acid to a final concentration of 4% (v/v).

(b) Liquid Chromatography−Electrospray Ionization Mass Spectrometry. A reversed phase HPLC system was used for the separation of proteolytic peptides. The digestion mixture was injected onto a 2.1 mm × 22 cm RP300 (Brownlee) reversed phase column. The column was then washed for 20 min with solvent A (0.05% v/v trifluoroacetic acid in water) to remove most of the salts. Peptides were then eluted with an organic solvent gradient of 0 to 60% solvent B (0.045% v/v trifluoroacetic acid in acetonitrile) over the course of 40 min. For liquid chromatography−electrospray ionization (LC−ESI) mass spectrometry (*25*, *26*), the effluent from the reversed phase column was split in a ratio of approximately 5:1. The effluent was introduced into the UV detector at a rate of 125 μL/min, and into the electrospray ionization source of a triple-quadrupole mass spectrometer (TSQ700, Finnigan, San Jose, CA) at a rate of 25 μL/min. Mass spectra were acquired in the positive ionization mode. The scan range was from *m*/*z* 300 to 2000 with a scan time of 6 s. All data were acquired in the centroid mode. Molecular masses were calculated from the mass spectra by using the software provided by the manufacturer. For collision-induced dissociation (CID) mass spectra, a sample aliquot from the corresponding HPLC fraction was injected into a 75 cm long fused silica microcapillary column with an inside diameter of 0.1 mm. The microcapillary was filled at the end with 10 cm of a C-18 reversed phase resin. Peptides were eluted at a rate of ca. 1 μL/min directly into the electrospray ionization source with a 10 min gradient of acetic acid in water (0.5% v/v) to 80% acetonitrile. Conditions for the acquisition of such mass spectra have already been described (*25*).

Protein Determinations. Protein content was assayed by a modified Bradford procedure (Bio-Rad protein assay catalog no. 500-006) using bovine serum albumin as the standard.

Polyacrylamide Gel Electrophoresis. SDS gels were 12% polyacrylamide and contained 0.375 M Tris-HCl (pH 8.8) and 0.1% SDS. The 5% stacking gel contained 0.13 M Tris-HCl (pH 6.8) and 0.1% SDS. Samples were heated at 95 °C for 5 min in buffer LB and then quenched on ice for 5 min before loading. Gels were stained with Coomassie blue. Protein molecular mass (MW) standards were from Bio-Rad and Novagen.

RESULTS

Purification of the 16S RNA m^{5}C967 Methyltransferase. In earlier studies, we found that more than half of the total 16S RNA methylating activity, but only 10% of the protein in an S30 supernatant, was found in the ribosomal high-salt wash fraction (*18*). Consequently, this fraction served as the starting point for the purification. After ammonium sulfate and DEAE column fractionation, the enzyme was further purified by chromatography on a MonoS column (Figure 1, top). The major amount of activity appeared in fractions 125 and 126 with a shoulder at fraction 129. The shoulder could be due to the activity of another rRNA methyltransferase which can also utilize free 16S RNA, or it could be a result of partial proteolysis of the m^{5}C enzyme with retention of some activity. We have not explored this aspect further. The pool of fractions 123−128 yielded two main bands upon polyacrylamide gel electrophoresis (Figure 1, bottom). N-Terminal sequencing of the band at approximately 14 kDa indicated that it was ribosomal protein L9. Since previous results (*18*) had indicated a molecular mass for the enzyme in the range of 35−45 kDa, we suspected that the enzyme activity resided in the 44 kDa band. This expectation was borne out by the experiments described below.

Identification of the Gene. N-Terminal sequencing of the 44 kDa band yielded the sequence MKw/cQRNLRSMAAQAVEQVVEQGQ, where the lowercase letters represent one of two possibilities of a questionable amino acid assignment. A search of the Genbank database available at the time revealed an exact match for 22 of the first 23 amino acids, including the N-terminal methionine with ORF *fmu* (Genbank accession number X77091), at 72.4 min on the *E. coli* chromosome (*27*). However, this ORF encoded a protein with a calculated molecular mass of 26.2 kDa, whereas the polypeptide which generated this N-terminal sequence ran on SDS gels like a 44 kDa protein (Figure 1). Analysis of the *fmu* coding sequence showed that the only way to generate an approximately 45 kDa protein product was by a −1 frameshift or by the insertion of a nucleotide in the published sequence in the sequence window from residues 582−714, counting the A of the initiating ATG codon of *fmu* as 1. This would eliminate the stop codon of *fmu* and result in fusion with the downstream *fmv*, yielding a 48.3 kDa protein. Sequencing of the *E. coli* genome in this region (data not shown) showed the existence of an additional C after C600. This sequence change was subsequently confirmed by Blattner et al. (*28*) and Mazel et al. (*29*).

Cloning of the Gene, Overexpression, and Affinity Purification. The putative gene was cloned into pET-15b by standard methods. As shown in Figure 2, induced cells containing clones with this insert produced large amounts of protein about 44−48 kDa in size (lane 4), whereas no such band was visible in the pET control (lane 2). The overexpressed protein was found in both the S15 supernatant (lane 5) and pellet (lane 6) fractions, although the distribution between these two fractions varied somewhat from preparation to preparation. All experiments were conducted using the soluble supernatant fraction of the enzyme.

Cloning in pET-15b places an N-terminal $(His)_6$ leader on the protein product for affinity purification on Ni^{2+}-containing resin columns. The finding of m^{5}C methyltransferase activity in such an affinity-purified protein would

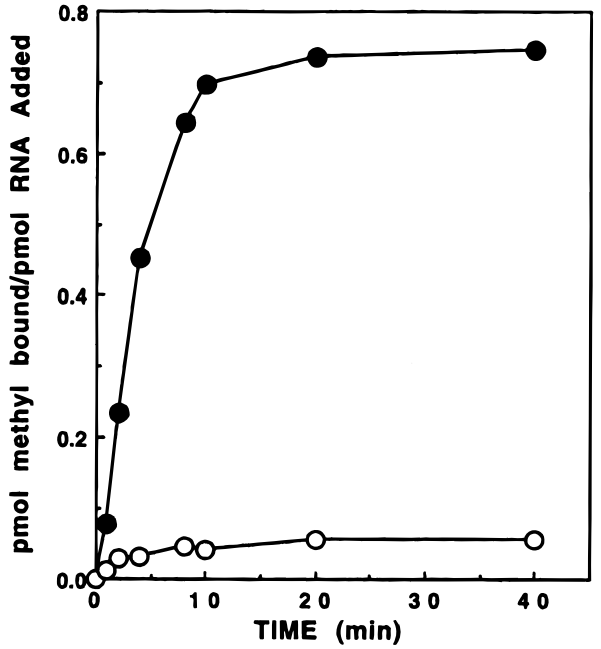

FIGURE 3: Substrate specificity of the recombinant m^{5}C967 methyltransferase. Synthetic 16S RNA (●) and 30S reconstituted particles (○) were prepared, and the extent of methylation was measured as incorporation of ^{3}H into cold TCA precipitable material. The 16S RNA concentration was 100 nM, and 30S particles were present at 80 nM. The Mg^{2+} concentration was 4 mM for methylation of 16S RNA and 2 mM for methylation of the ribosome particles. The enzyme concentration was 0.24 μg/mL.

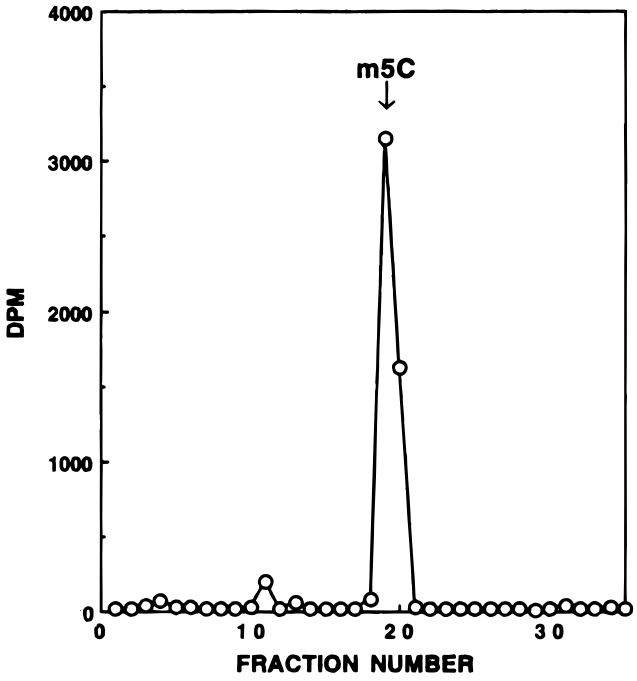

FIGURE 4: Characterization of the methylation product of recombinant m^{5}C967 methyltransferase. [^{3}H]Methyl synthetic 16S RNA was prepared using recombinant m^{5}C967 methyltransferase, digested with ribonuclease P1, and subjected to HPLC for analysis. The position of m^{5}C is shown by the arrow. Its position was determined by the UV absorption of an internal standard (data not shown).

constitute definitive proof that the gene cloned is the gene for the methyltransferase. Figure 2 (lane P) shows that the Ni^{2+} column effectively removed all the contaminating protein from the S15 supernatant fraction, resulting in a single protein band with a molecular mass of 49 kDa. Because the 5′-tag increases the molecular mass by 2179 Da, the expected value for the recombinant protein, 50.5 kDa on the basis of the molecular mass of 48.3 kDa calculated from the ORF, was consistent with that found.

Substrate Specificity of the m^{5}C967 Methyltransferase. Previous work (*17*−*19*) had shown that the enzyme preferred synthetic 16S RNA as a substrate over synthetic 30S ribosomes. Figure 3 shows that this specificity is shared by the recombinant enzyme. The m^{5}C methyltransferase was capable of incorporating 0.75 pmol of methyl groups per picomole of synthetic 16S RNA, while less than 0.1 pmol of methyl groups was incorporated per picomole of synthetic 30S ribosomes under the same conditions. From the data of Figures 3 and 6, a specific activity of 3.1×10^6 units/mg of protein at 37 °C was found, corresponding to a turnover number of 2.6 mol min^{-1} (mol of protein)$^{-1}$. The corresponding value at 30 °C was 1.0×10^6 units/mg of protein (data not shown).

Identification and Localization of the Site of Methylation. Although the cloned enzyme appeared to behave in a fashion completely analogous to that of the native enzyme, conclusive evidence that we had, in fact, cloned the m^{5}C967 methyltransferase required that the nature of the methylation product and the site of methylation be established for the recombinant protein. To this end, [^{3}H]methyl 16S RNA prepared using the recombinant enzyme was digested to nucleosides with ribonuclease P1 and alkaline phosphatase

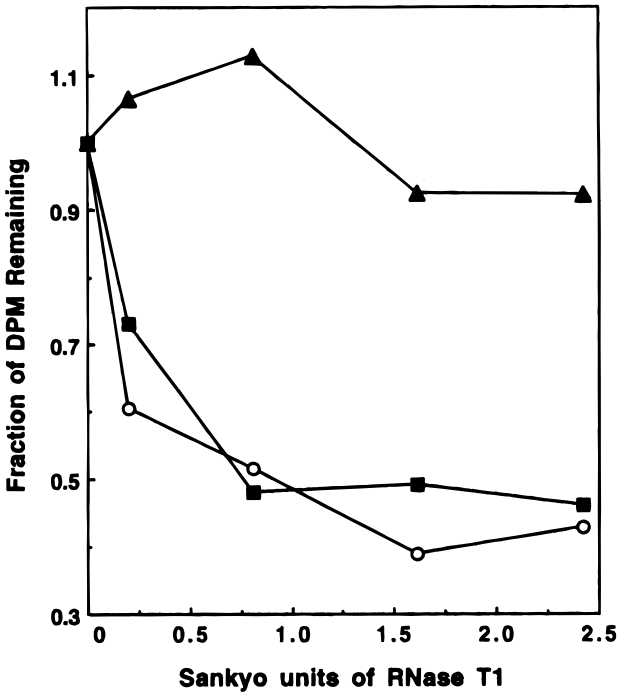

FIGURE 5: Localization of the site of methylation by recombinant m^{5}C967 methyltransferase. [^{3}H]Methyl-labeled 16S RNA (0.36 pmol), prepared using recombinant m^{5}C967 methyltransferase, was hybridized with the indicated oligonucleotide and then digested with RNase T1 (0.2−2.4 Sankyo units) as described in Experimental Procedures: oligomer 958 (▲), oligomer 1398 (■), or no oligomer (○).

and subjected to HPLC analysis (Figure 4). The arrow shows the position of an internal standard of m^{5}C as determined by UV absorption. Ninety-five percent of the radioactivity

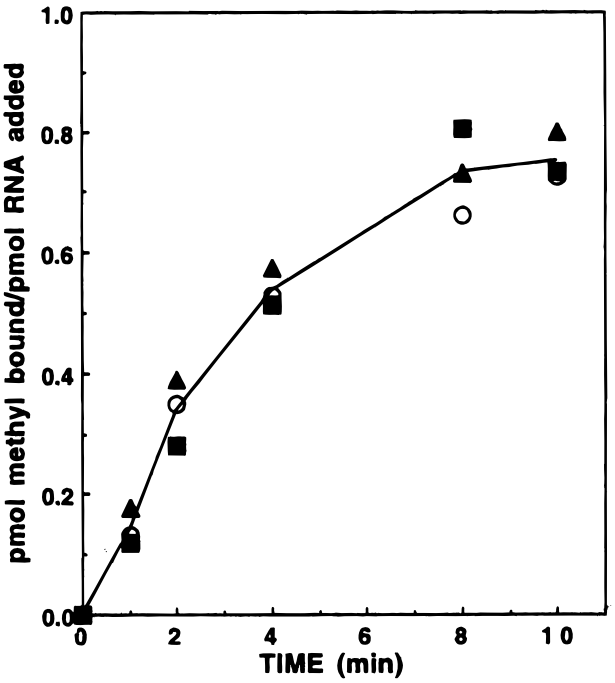

FIGURE 6: Lack of a Mg^{2+} requirement for methylation. Methylation of synthetic 16S RNA was conducted as described in Experimental Procedures except that the incubation mixture contained 4 mM Mg^{2+} (▲), 10 mM Mg^{2+} (○), or 1 mM EDTA (■). The 16S RNA concentration was 100 nM, and the purified enzyme protein concentration was 0.26 μg/mL.

loaded onto the column was found at the m^{5}C position. These results confirm that the recombinant enzyme is an m^{5}C methyltransferase. However, *E. coli* 16S RNA contains two m^{5}C residues, one at position 967 and one at position 1407 (*30*). To determine which one was formed by the recombinant enzyme, hybridization−protection studies were conducted using deoxyoligomers which were complementary to the RNA sequence spanning each of the m^{5}C sites (*17*). Oligomer 958 spanned the region from residue 958 to residue 977, and oligomer 1398 spanned the region from residue 1398 to 1417. As shown in Figure 5, oligomer 958 was capable of protecting the site of the [^{3}H]methyl group from RNase T1 digestion whereas oligomer 1417 was completely ineffective. The inhibition of RNase T1 digestion by oligomer 958 was not due to some general inhibitory effect because the same oligomer preparation failed to inhibit digestion of an RNA labeled by methylation of A1518 and A1519 (*17*). The high baseline value even with excess RNase T1 and no protecting oligomer has been noted previously (*17*, *18*) and is likely due to a partial binding of the RNase T1 digest to the filter matrix. These data, taken together, demonstrate that the recombinant m^{5}C methyltransferase, like the native enzyme (*17*), is specific for C967 in *E. coli* 16S RNA.

Lack of a Mg^{2+} Requirement for Methylation. To test the effect of Mg^{2+} on the rate of methylation, synthetic 16S RNA was methylated using the recombinant enzyme at both 4 and 10 mM Mg^{2+}. These results were compared with the rate of methylation when EDTA was added to a concentration of 1 mM. Since without additions, the Mg^{2+} concentration in the reaction was 0.03 mM, methylation with added 1 mM EDTA means reaction takes place in the absence of Mg^{2+}. Figure 6 shows that neither the rate nor the extent of methylation was affected by these conditions. Although it is not possible to be certain that all the Mg^{2+} had been removed since the RNA was not preincubated with EDTA, it is unlikely that Mg^{2+} is a significant factor for methylation by this enzyme.

Conflicts in Published Sequences of the ORFs fmu and fmv. The fusion of ORFs *fmu* and *fmv* (Swissprot entry P36929) encodes a protein 429 amino acids long with a C-terminal sequence of KLIKK (Figure 7A). This is the C-terminal sequence deduced by Meinnel and Blanquet (*27*) (GenBank accession number X77091) and Schloesser et al. (*31*) (GenBank accession number X52114) from their DNA sequence. It differs from that reported by Blattner et al. (submitted to the EMBL/GenBank/DDBJ databases in January 1997) whose DNA sequence encoded a 386-amino acid protein with a C-terminal sequence of PAD, but is the same as a correction posted in September/October 1997 by this latter group. To determine the sequence definitively, we performed direct C-terminal amino acid sequencing on the overexpressed recombinant protein, since the DNA segment cloned (nt −9 to 1339) was sufficient to include both proposed sequences.

The recombinant protein was digested with endoproteinase Lys-C and endoproteinase Glu-C; digests were analyzed by liquid chromatography−electrospray ionization (LC−ESI) mass spectrometry, and experimental average molecular masses of the eluting proteolytic peptides were recorded online with the mass spectrometer. The identity of the proteolytic peptides was established by comparison of the experimental values with theoretical molecular masses calculated from the known amino acid sequence. Figure 7A shows the amino acid sequence of the longer form. Proteolytic peptides that were identified on the basis of their molecular mass are marked within Figure 7A. Some rather large peptide fragments were produced in certain domains of the protein because the digestion was not complete. However, with both proteolytic enzymes, we did recover peptides from the C-terminus. A Lys-C peptide with a molecular mass of 1742 Da was found. The experimental molecular mass of this peptide is in very good agreement with the theoretical value of the peptide Gln-Asn-Pro-Leu-Gly-Ala-Glu-Glu-Gly-Asp-Gly-Phe-Phe-Tyr-Ala-Lys of 1742.9 Da (LC21 in Figure 7A). Although this peptide is very close to the C-terminus, it did not include the last four amino acids. Digestion of the protein with endoproteinase Glu-C produced a peptide with an average molecular mass of 1387 Da. This experimental molecular mass is in very good agreement with the calculated value of the peptide Gly-Asp-Gly-Phe-Phe-Tyr-Ala-Lys-Leu-Ile-Lys-Lys of 1386.7 Da (V28 in Figure 7A). Because the C-terminus of this peptide consists of two lysines and endoproteinase Glu-C cleaves at the C-terminal side of Glu, we conclude that this sequence indeed provides the C-terminal end of the methyltransferase.

To unambiguously establish the identity of the amino acid sequence, we subjected this Glu-C peptide to collision-induced dissociation analysis. Briefly, doubly charged ions (*m*/*z* 694.5) of the peptide were selected in the first quadrupole of the triple-quadrupole instrument. These ions were then fragmented by subjecting them to collisions with argon gas in the collision chamber. Molecular masses of fragments were then determined in the third quadrupole. The collision-induced dissociation mass spectrum is shown in Figure 7B. The experimental fragmentation pattern is in excellent agreement with the amino acid sequence of the peptide, confirming the identity of this C-terminal peptide.

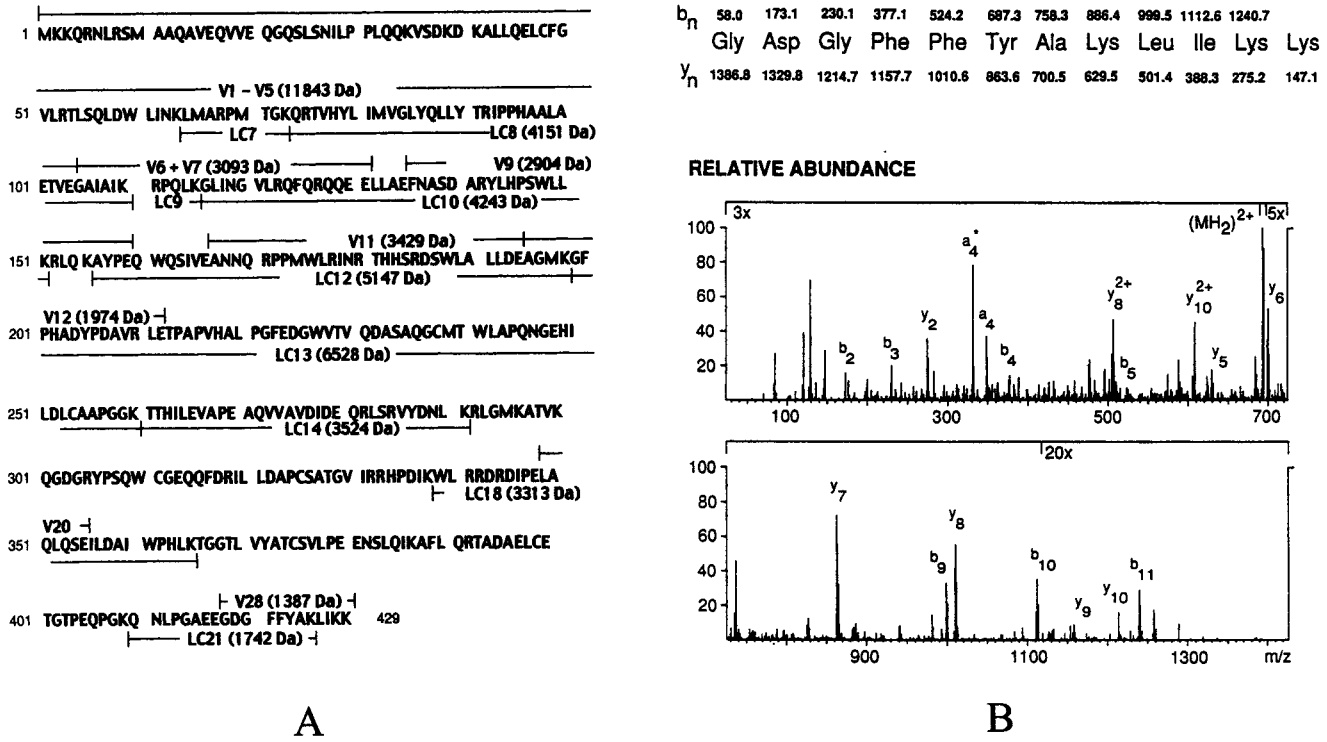

FIGURE 7: C-Terminal sequencing of the recombinant methyltransferase by mass spectrometry. (A) Amino acid sequence of the methyltransferase. Proteolytic peptides which were recovered and identified with LC−ESI mass spectrometry are indicated on top (endoproteinase Glu-C) and below (endoproteinase Lys-C) the amino acid sequence. The number of the peptide is indicated together with the experimental average molecular mass in daltons. (B) Collision-induced dissociation mass spectrum of the C-terminal peptide Gly-Asp-Gly-Phe-Phe-Tyr-Ala-Lys-Leu-Ile-Lys-Lys obtained from a digest with endoproteinase Glu-C. Ions of MH_2^{2+} (*m*/*z* 694.5) were selected for collision analysis. The sequence and the theoretical average molecular masses of possible fragments of type b and y are indicated on the top of the figure.

DISCUSSION

Substrate Specificity and Mg^{2+} Requirements of the Enzyme. The substrate specificity of the cloned enzyme mirrored that of the native one. It recognized protein-free synthetic 16S RNA as a substrate, was inactive with ribosomal particles made from the same synthetic RNA, and, of the two known sites for m^5C, specifically methylated only C967. Although not demonstrated in this work, it had previously been shown (*18*) that the native enzyme did not react with an in vitro transcript of 23S RNA, or with poly(dI-dC), an excellent substrate for DNA m^5C methyltransferases (*32*). It also did not react with isolated 16S or 23S RNAs, that is, with RNAs containing their normal complement of modified bases, showing that no unnatural sites in either rRNA were recognized. Methylation of tRNA or other small RNAs has not been tested so that the possibility of a dual or multiple specificity has not been ruled out. Using native enzyme, the inactivity of ribosomal subunits was shown to be due to the presence of a single ribosomal protein, S19 (*19*). This protein makes close contact with the stem of the stem−loop structure containing C967 as shown by footprinting using both base-specific reagents (*33*) and Fe−EDTA treatment (*34*). It is believed that S19 shuts down C967 methylation by altering the conformation of the stem−loop structure containing C967 so that it is no longer can be recognized by the methyltransferase (*19*). It is reasonable to suppose that the recombinant enzyme also shows these properties, but this has not been tested.

The Mg^{2+} dependence had not been tested previously. The lack of a need for Mg^{2+} was surprising, and implies that primary sequence rather than the secondary or tertiary structure is the most important specificity parameter. However, since in our study the RNA preparation was not preincubated with EDTA to remove all of the Mg^{2+}, it is still possible that the remaining traces of Mg^{2+} were sufficient to confer enough secondary structure on that part of the RNA, for example, on the surrounding stem−loop structure described above, to allow recognition by the methyltransferase. It does seem likely, however, that extensive secondary structure or a tertiary structure is not required.

Gene and Protein Sequence. The gene that encodes the m^5C967 methyltransferase is formed by fusion of two ORFs, *fmu* and *fmv*, by correction of an error in the original sequence of *fmu*. We deduced that there must be such an error, determined the window in which it should have occurred, and found and positioned the missing C residue. Subsequently, the same correction was made by others (*28*, *29*) and the fusion ORF given the name *sun* and listed in the Swissprot database with the accession number P36929. The ORF contains a SAM-binding motif at residues 250−266 (*35*). We now propose that the gene be renamed *rsmB* for (r)ibosomal (s)mall subunit (m)ethyltransferase since the function of the gene product has been identified as an m^5-C967 methyltransferase. B denotes that this is the second gene sequence for an rRNA methyltransferase to be described, the first one being *ksgA*, the gene for the m_2^6A methyltransferase (*20*). We therefore suggest *rsmA* as a suitable alternate name for *ksgA*.

The C-terminal sequence of the overexpressed recombinant methyltransferase was directly determined to be GDGFFY-

AKLIKK by mass spectrometry methods. This direct analysis confirms the sequence deduced from the DNA sequence (*27*, *28*, *31*).

A search of the GenBank database using BLAST 2.0.4 (*36*) identified proteins that were highly significantly similar to the m^5C967 methyltransferase in 13 different organisms. These included other prokaryotes such as *Haemophilus influenzae* and *Bacillus subtilis*, archaea such as *Pyrococcus horikoshii*, *Archaeoglobus fulgidus*, and *Methanococcus jannaschii*, and eukaryotes ranging from yeast to humans. It would appear that the ancestor sequence from which the m^5-C967 methyltransferase was derived is very ancient. It is also interesting that the C-terminal half of RsmB is a remarkable 60% similar to the 184-amino acid central hydrophobic domain of a human proliferation-associated nucleolar protein, P120 (*37*), as is YebU (Swissprot entry P76273), a hypothetical 53.4 kD protein in *E. coli*. YebU is also significantly similar (55% over 137 amino acids) to RsmB. Perhaps YebU is the m^5C1407 methyltransferase.

The protein most similar in sequence to RsmB according to this analysis is from *H. influenzae* (Swissprot entry P44788). This protein is 451 amino acids long compared to 429 for RsmB. They were 49% identical and 65% similar along 416 residues. With such a high level of similarity between the two sequences, it is likely that this gene encodes an equivalent methyltransferase in *H. influenzae* and thus that this organism also has an equivalently located m^5C in its SSU RNA. However, there is no information available about the presence of m^5C at an equivalent site in the SSU RNA of *H. influenzae* or for that matter in other prokaryotes or archaea. In eukaryotes, the equivalent site is occupied by the highly modified m^1acp^3U (*9*). Therefore, if the related eukaryotic gene encodes an m^5C methyltransferase, it must be for another site and/or another RNA. There are no known m^5C residues in the SSU RNA of yeast or humans, but there are two possible sites in LSU RNA (*38*).

Role of Methylation of C967 in the Ribosome. m^5C967 is located next to m^2G966 in a small stem−loop structure which has been implicated in tRNA P site binding both by protection of C967 from modifying agents by bound tRNA (*8*) and by cross-linking of G966 to tRNA at position 32, that is, near the anticodon bases 34−36 (*39*). Replacement of this doubly modified loop in eukaryotes by m^1acp^3U suggests a role for these modifications in tRNA binding, but the proof and mechanism are still to be determined.

As stated in the introductory section, although completely unmodified RNA can make a ribosome which is able to carry out all of its known functions, the efficiency is reduced. The purpose of this work was to identify the gene for the C967 methyltransferase in order to make it possible to delete the gene so as to test the effect of the loss of a single methyl group on ribosome biosynthesis and function. Now that the gene has been identified, such experiments can proceed.

REFERENCES

1. Green, R., and Noller, H. F. (1997) *Annu. Rev. Biochem. 66*, 679−716.
2. Brimacombe, R. (1995) *Eur. J. Biochem. 230*, 365−383.
3. Noller, H. F., Hoffarth, V., and Zimniak, L. (1992) *Science 256*, 1416−1419.
4. Nitta, I., Ueda, T., and Watanabe, K. (1998) *RNA 4*, 257−267.
5. Bakin, A., Kowalak, J. A., McCloskey, J. A., and Ofengand, J. (1994) *Nucleic Acids Res. 22*, 3681−3684.
6. Santer, M., Santer, U., Nurse, K., Bakin, A., Cunningham, P., Zain, M., O'Connell, D., and Ofengand, J. (1993) *Biochemistry 32*, 5539−5547.
7. Wrzesinski, J., Bakin, A., Nurse, K., Lane, B. G., and Ofengand, J. (1995) *Biochemistry 34*, 8904−8913.
8. von Ahsen, U., and Noller, H. F. (1995) *Science 267*, 234−237.
9. Youvan, D. C., and Hearst, J. E. (1981) *Nucleic Acids Res. 9*, 1723−1741.
10. Murgola, E. J., Hijazi, K. A., Göringer, H. U., and Dahlberg, A. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4162−4165.
11. Cunningham, P. R., Nurse, K., Bakin, A., Weitzmann, C. J., Pflumm, M., and Ofengand, J. (1992) *Biochemistry 31*, 12012−12022.
12. Mueller, F., and Brimacombe, R. (1997) *J. Mol. Biol. 271*, 524−544.
13. Van Knippenberg, P. H. (1986) in *Structure, Function, and Genetics of Ribosomes* (Hardesty, B., and Kramer, G., Eds.) pp 412−424, Springer-Verlag, Berlin.
14. Krzyzosiak, W., Denman, R., Nurse, K., Hellman, W., Boublik, M., Gehrke, C. W., Agris, P. F., and Ofengand, J. (1987) *Biochemistry 26*, 2353−2364.
15. Denman, R., Nègre, D., Cunningham, P. R., Nurse, K., Colgan, J., Weitzmann, C., and Ofengand, J. (1989) *Biochemistry 28*, 1012−1019.
16. Cunningham, P. R., Richard, R. B., Weitzmann, C. J., Nurse, K., and Ofengand, J. (1991) *Biochimie 73*, 789−796.
17. Nègre, D., Weitzmann, C., and Ofengand, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4902−4906.
18. Nègre, D., Weitzmann, C., and Ofengand, J. (1990) in *UCLA Symposia on Molecular and Cellular Biology New Series, Volume 128, Nucleic Acid Methylation* (Clawson, G. A., Willis, D. B., Weissbach, A., and Jones, P. A., Eds.) pp 1−17, Wiley-Liss, New York.
19. Weitzmann, C., Tumminia, S. J., Boublik, M., and Ofengand, J. (1991) *Nucleic Acids Res. 19*, 7089−7095.
20. Van Buul, C. P. J. J., and van Knippenberg, P. H. (1985) *Gene 38*, 65−72.
21. Bakin, A., and Ofengand, J. (1993) *Biochemistry 32*, 9754−9762.
22. Zyskind, J. W., and Bernstein, S. I. (1992) *Recombinant DNA Laboratory Manual*, p 187, Academic Press, San Diego, CA.
23. Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035−10038.
24. Denman, R., Weitzmann, C., Cunningham, P. R., Nègre, D., Nurse, K., Colgan, J., Pan, Y.-C., Miedel, M., and Ofengand, J. (1989) *Biochemistry 28*, 1002−1011.
25. Hunt, D. F., Alexander, J. E., McCormack, A. L., Martino, P. A., Michel, H., Shabanowitz, J., Sherman, N., Moseley, M. A., Jorgenson, J. W., and Tomer, K. B. (1991) in *Techniques in Protein Chemistry II* (Villafranca, J. J., Ed.) pp 441−454, Academic Press, New York.
26. Smith, R. D., Loo, J. A., OgorzalekLoo, R. R., Busman, M., and Udseth, H. R. (1991) *Mass Spectrom. Rev. 10*, 359−451.
27. Meinnel, T., and Blanquet, S. (1994) *J. Bacteriol. 176*, 7387−7390.
28. Blattner, F. R., Plunkett, G., III, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science 277*, 1453−1474.
29. Mazel, D., Coic, E., Blanchard, S., Saurin, W., and Marliere, P. (1997) *J. Mol. Biol. 266*, 939−949.
30. Rauè, H. A., Klootwijk, J., and Musters, W. (1988) *Prog. Biophys. Mol. Biol. 51*, 77−129.
31. Schloesser, A., Hamann, A., Bossemeyer, D., Schneider, E., and Bakker, E. P. (1993) *Mol. Microbiol. 9*, 533−544.
32. Pedrali-Noy, G., and Weissbach, A. (1986) *J. Biol. Chem. 261*, 7600−7602.
33. Stern, S., Weiser, B., and Noller, H. F. (1988) *J. Mol. Biol. 204*, 447−481.

34. Powers, T., and Noller, H. F. (1995) *RNA 1*, 194−209.
35. Koonin, E. V., Tatusov, R. L., and Rudd, K. E. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 11921−11925.
36. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res. 25*, 3389−3402.
37. Koonin, E. V. (1994) *Nucleic Acids Res. 22*, 2476−2478.
38. Maden, B. E. H. (1990) *Prog. Nucleic Acid Res. Mol. Biol. 39*, 241−303.
39. Döring, T., Mitchell, P., Osswald, M., Bochkariov, D., and Brimacombe, R. (1994) *EMBO J. 13*, 2677−2685.

BI981880L