ABSTRACT

In bacterial 16S rRNAs, methylated nucleosides are clustered within the decoding center, and these nucleoside modifications are thought to modulate translational fidelity. The N\(^4\), 2'-O-dimethylcytidine (m\(^4\)Cm) at position 1402 of the *Escherichia coli* 16S rRNA directly interacts with the P-site codon of the mRNA. The biogenesis and function of this modification remain unclear. We have identified two previously uncharacterized genes in *E. coli* that are required for m\(^4\)Cm formation. *mraW* (renamed *rsmH*) and *yraL* (renamed *rsmI*) encode methyltransferases responsible for the N\(^4\) and 2'-O-methylations of C1402, respectively. Recombinant RsmH and RsmI proteins employed the 30S subunit (not the 16S rRNA) as a substrate to reconstitute m\(^4\)Cm\(^{1402}\) in the presence of S-adenosylmethionine (Ado-Met) as the methyl donor, suggesting that m\(^4\)Cm\(^{1402}\) is formed at a late step during 30S assembly in the cell. A luciferase reporter assay indicated that the lack of N\(^4\) methylation of C1402 increased the efficiency of non-AUG initiation and decreased the rate of UGA read-through. These results suggest that m\(^4\)Cm\(^{1402}\) plays a role in fine-tuning the shape and function of the P-site, thus increasing decoding fidelity.

INTRODUCTION

The ribosome is the translator of genetic information on mRNAs that results in protein synthesis. The crystal structures of ribosomal subunits and their functional complexes with various substrates have helped us to elucidate the mechanisms of protein synthesis. The small (30S) subunit of the ribosome, containing the 16S RNA, plays a critical role in deciphering the codons on mRNAs (1–4). The neck region of the 30S subunit provides a path for mRNA along with the three tRNA binding sites, the A-, P- and E-sites. At the P-site, precise codon–anticodon pairing occurs between the mRNA and a peptidyl-tRNA during the elongation cycle to maintain the reading frame (5–7). In the initiation step, the AUG codon at the P-site is recognized by formyl-methionyl (fMet)-tRNA\(^{fMet}\) and this recognition is assisted by initiation factors (8,9). This step is kinetically controlled by the formation and stability of the 30S initiation complex (30S IC), which is followed by the formation of the 70S initiation complex (70S IC) (10,11). The flexibility of the 16S rRNA during these conformational rearrangements is believed to affect the fidelity of initiation (12), suggesting that the local structure of the P-site affects the fidelity of AUG codon selection. The P-site in the small subunit is formed by several conserved residues (positions 1400–1405 and 1496–1502) in the top part of helix 44 in the 16S rRNA (Figure 1A) (7). Certain point mutations in the P-site of *Escherichia coli* 16S rRNA resulted in growth reduction (13,14). In particular, a single base deletion of C1400 resulted in a lethal phenotype (15). In an in vitro study, a mutant ribosome in which G1401–C1501 was flipped to C1401–G1501 showed full tRNA-binding activity at the P-site and mostly normal peptide-bond formation; however, formation of the first peptide bond was largely inhibited in the mutant ribosome (16). A U1498G mutation also affected the formation of the first peptide bond (17,18). These data suggest that certain conserved residues in the P-site of the 16S rRNA are required for the initiation of translation by the ribosome.

Ribosomal RNAs are decorated with various post-transcriptional modifications (19). Seven species of RNA modifications are found at 11 positions in the 16S rRNA, and some are clustered in the decoding center (20). They are required for the fine-tuning of local RNA structure (21), RNA–RNA interactions (21), 30S subunit assembly (22,23) and antibiotic resistance (24). Recently, Umesh and his colleagues showed that...
methyl-modifications in the 16S rRNA are required for the stringent selection of the initiator tRNA (25), suggesting that the translation initiation step is affected and modulated by certain methyl-modifications. However, it is still unclear which modifications are dominantly associated with the translation initiation. At the P-site of the 16S rRNA there is a unique dimethyl modification, N\(^4\), 2'-O-dimethylcytidine (m\(^4\)Cm) at position 1402 (Figure 1A and B), which was discovered in 1966 (26). C1402 is universally conserved (27), and the dimethyl modification is thought to be conserved in bacteria (28–30). A C to U replacement at position 1402 resulted in a slow growth phenotype (13), and m\(^4\)Cm1402 is thought to be important for ribosome function. However, the enzymes and biogenesis for this modification have not been characterized. In the crystal structure of the \(E. coli\) 30S subunit complexed with mRNA and tRNA\(^{i\text{Met}}\), the codon-anticodon helix at the P-site interacts with the major groove on the top part of helix 44 (5–7). The backbone of the P-site codon is recognized by m\(^4\)Cm1402 through hydrogen bonds between the N\(^4\) of m\(^4\)Cm1402 and a phosphate between the 2nd and 3rd letters of the P-site codon (Figure 5B) (5–7). These observations suggest that m\(^4\)Cm1402 is involved in the fine-tuning of the local structure of the P-site, and the correct recognition of the initiation codon.

Figure 1. The secondary structure of the decoding center in the \(E. coli\) 16S rRNA and the enzymatic formation of m\(^4\)Cm1402. (A) The secondary structure of the \(E. coli\) 16S rRNA is shown on the left. Helix 44 and surrounding regions are shown on the right, with helix and base numbers indicated. A number of modified nucleosides including 5-methylcytidine (m\(^5\)C), N\(^4\), 2'-O-dimethylcytidine (m\(^4\)Cm) and 3-methyluridine (m\(^3\)U) are also shown. The bases annotated as A-site (blank circle), P-site (shaded with light gray) and E-site (shaded with dark gray) are indicated (4,6). (B) Enzymatic formation of m\(^4\)Cm1402. N\(^4\)-methylation and 2'-O-methylation of C1402 are catalyzed by RsmH and RsmI, respectively, using Ado-Met as a methyl group donor. After the reaction, Ado-Met is converted to Ado-Hcy.
To investigate the function and biogenesis of m^7Cm1402, it is necessary to identify the genes that encode enzymes responsible for m^7Cm1402 formation. We have performed a genome-wide screen for genes responsible for RNA modifications using a reverse genetics approach combined with mass spectrometry (ribonucleome analysis) (31). By this approach we have identified many genes responsible for tRNA modifications (32–34). For this study, we modified the approach in order to screen for genes responsible for rRNA modifications. We successfully identified two methyltransferases responsible for m^7Cm1402 formation in the E. coli 16S rRNA (Figure 1B), and further characterized the biogenesis and function of m^7Cm1402.

MATERIALS AND METHODS

Strains and mediums

A series of single deletion strains each bearing a kanamycin resistance marker (Km^r) was obtained from the Genetic Stock Research Center, National Institute of Genetics, Japan (35). A series of knockout strains with transposon insertions was kindly provided by Dr. T. Miki (36). The E. coli K-12 strain BW25113 (lac^F rrrBiT44 ΔlacZ_W116 hsdR514 ΔaraBAD_A133 ΔrhaBAD_LD728) was used for the ‘one-step inactivation of chromosomal genes’ procedure (37). The following primers were used to amplify the chloramphenicol acetyltransferase gene (Cm^r) with 40-nt extensions at both ends that were homologous to sequences of the yraL gene: ΔyrαL-F (Cm^r) (5’-atcgtgaatctgattatcaaccaacctctgttgtaagggcataatcttgaagtggtt–3’) and ΔyrαL-R (Cm^r) (5’-ttgcc ccaaggcgccttttgctagcctttctgtgcacacgaatagataagcc ggc-3’). The pBT vector (Stratagene) was used as a template for Cm^r. The amplified sequence was used to construct BW25113 ΔyrαL::Cm. To construct Δrrm2A/Δrrm2L, Δrrm2L::Cm was transferred to Δrrm2A (Km^r) by P1-transduction. The E. coli strains were grown in 1 ml of Luria–Bertani (LB) medium in 96-well plates at 37°C overnight. Doubling time was determined using OD_{600} measurements in a plate reader (Molecular Device, Inc.) every 15 min.

Ribonucleome analysis

The mutant E. coli strains were grown in 1 ml of LB medium in 96-well plates at 37°C overnight. Total RNA was prepared from the harvested cells using an RNA preparation kit (RNeasy, Qiagen) and the rRNAs were isolated by electrophoresis in a 4% polyacrylamide gel containing 7 M urea. About 100 fmol of rRNAs were digested with RNase T1 or RNase A, and the digests were analyzed by capillary liquid chromatography (LC) nano electrospray ionization/mass spectrometry as described previously (31,38) with the following modifications. The concentration of RNase T1 was adjusted to 5 U/μl. We employed a linear ion trap-orbitrap hybrid mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific) equipped with a custom-made nanospray ion source, a Nanovolume Valve (Valco Instruments), and a splitless nano HPLC system (DiNa, KYA Technologies). Each digest was mixed with TEAA and loaded onto a nano-LC trap column (C18, 0.5 × 1.0 mm, desalted, and then concentrated with 0.1 M TEAA (pH 7.0). RNA fragments were eluted from the trap column and directly injected into a C18 capillary column (H/Q Sil; 3 μm C18, 100-A pore size; 0.1 × 100 mm, KYA Technologies). The solvent system consisted of 0.4 M 1,1,1,3,3, 3-hexafluoro-2-propanol (HFIP; pH 7.0, adjusted with triethylamine; solvent A) and 0.4 M HFIP in 50% methanol (solvent B), and the samples were chromatographed at a flow rate of 300 nl/min using a linear gradient of 10–90% solvent B over 35 min. The chromatographic eluent was sprayed from a sprayer tip attached to the capillary column. The ionization voltage was set to −1.9 kV and ions were scanned in the negative polarity mode. For nucleoside analysis, RNAs were digested into nucleosides and analyzed by LC/MS using ion-trap mass spectrometry as described previously (31).

Preparation of rRNA fragments

Escherichia coli ribosomes were obtained basically as described previously (39,40). rRNAs were prepared from the ribosomes by phenol/chloroform extraction and ethanol precipitation (41). The 49-mer fragment covering positions 1378–1426 was carved out from the 16S rRNA using a complementary oligonucleotides procedure (42). The 20 A_{260} unit rRNA was mixed with 4 nmol of deoxyoligonucleotide (5’-caaccaacctccagttgtagcgggtgtgaacaagcgggaacg-3’), complementary to positions 1378–1426 in a buffer consisted of 50 mM Tris–HCl (pH 7.6) and 115 mM KCl. The mixture was incubated at 90°C for 5 min then cooled to 37°C. One-ninth volume of 500 mM NH4OAc (pH 5.3), 0.5 μg RNase A and 500 U RNase T1 was added, and the mixture was incubated at 37°C for 1 h to digest the non-protected regions of the rRNA. The protected DNA/RNA heteroduplex was recovered by phenol/chloroform extraction and ethanol precipitation, and then purified by the electrophoresis in a 15% polyacrylamide gel containing 7 M Urea.

Expression and purification of recombinant RsmH and RsmI

The strains carrying pCA24N plasmids for expression of RsmH and RsmI were obtained from the ASKA clone collection [NBRP (NIG, Japan): E. coli] (43). The recombinant RsmH and RsmI proteins, with N-terminal 6 × His-tags, were expressed in soluble form by induction with 0.1 mM IPTG, and purified by using the AKTA chromatography system with the His-trap chelating HP column (Amersham Biosciences). The pooled proteins were dialyzed against a buffer consisting of 50 mM Tris–HCl (pH 7.5), 1 mM dithiothreitol and 100 mM KCl. To remove endogenous Ado-Met, the recombinant proteins were incubated with 2 mM S-adenosylhomocysteine (Ado-Hcy) in a buffer consisted of 50 mM Tris–HCl (pH 8.0), 10 mM MgCl2, 100 mM NH4Cl for 1 h at 37°C, and then dialyzed (44). The protein concentration was determined using the Bio-Rad protein assay kit.
with bovine serum albumin as the standard. Glycerol was added to the protein solution to a final concentration of 30%.

In vitro reconstitution of m\(^4\)Cm1402

The 30S subunit and the tightly-coupled 70S ribosomes were isolated from the \(\Delta rsmH/\Delta rsmI\) strain by sucrose density gradient centrifugation as described previously (45). The 16S rRNA was prepared from the 30S fraction using Isogen (Nippon gene). In vitro methylations were performed at 37°C for 1 hour in 100-μl reaction mixtures containing 100 mM NH\(_4\)Cl, 20 mM HEPES-KOH (pH 7.6), 10 mM Mg(OAc)\(_2\), 6 mM β-mercaptoethanol, 1 mM Ado-Met, 2.5 pmol 16S rRNA (or 5 pmol of 30S subunit or tightly-coupled 70S) and 10 pmol of recombinant proteins (RsmH and/or RsmI). rRNA was recovered from an aliquot of the reaction mixture using Isogen (Nippon Gene). This was digested with RNase \(T_1\) and then subjected to LC/MS analysis as described earlier.

Luciferase reporter constructs and assays

The ORFs of Rluc and Fluc were PCR-amplified from pJD375, which was kindly provided by Dr. Jonathan Dinman (Univ. of Maryland). A series of dual luciferase reporter constructs was made for this study. The sequences and names of DNA primers for vector construction were shown in Supplementary Table S1. The Firefly and Renilla luciferase fusion ORF was amplified from pJD375 using the primers LucNcoI-f and LucHindIII-r. The PCR product was cloned into the NcoI and HindIII sites of pQE60 (Qiagen), and resulting plasmid was named pJD375 using the primers LucNcoI-f and LucHindIII-r. The PCR product was cloned into the NcoI and HindIII sites of pQE60 (Qiagen), and resulting plasmid was named pQE60-FRLuc. For the series of reporter vectors designed to examine non-AUG initiation, the UAA stop codon was carved out from the 16S rRNA (Figure 2A). The window sequences were designed based on constructs corresponding start codons: AUG, AUU, UUG and GUG.

RESULTS

Identification of \(rsmH\) and \(rsmI\), which are responsible for m\(^4\)Cm1402 formation

We employed a reverse genetics approach combined with MS (ribonuclease analysis) to identify the genes responsible for m\(^4\)Cm1402 formation in \(E.\ coli\). To specifically detect m\(^4\)Cm at position 1402 in the 16S rRNA, we set up conditions LC/MS to analyze complex mixtures of rRNA fragments digested with base-specific ribonucleases (RNases). Total RNA was prepared from a wild-type strain of \(E.\ coli\) and digested with RNase \(T_1\), and then subjected to LC/MS. In the complex mixture of RNA fragments we clearly detected the doubly charged negative ion (m/z 652.1) in the tetramer fragment containing m\(^4\)Cm (m\(^4\)CCCGp, MW 1306.2), which originated from the 16S rRNA (Figure 2A).

Since m\(^4\)Cm is a dimethyl modification, we picked up 24 uncharacterized genes in \(E.\ coli\) that were annotated as methyltransferases in the Clusters of Orthologous Groups (COG) database (48), as candidates for enzymes involved in m\(^4\)Cm1402 formation. Total RNAs were prepared from \(E.\ coli\) strains, each knocked out at one of the 24 candidate genes, and subjected them to the LC/MS analysis described earlier. As a result we identified two strains, \(\Delta mraW\) and \(\Delta yraL\), which both produced mono-methylated tetratmers (m\(^3\)CCCGp or CmCCGp, MW 1292.2) instead of the dimethylated tetramer produced by wild-type \(E.\ coli\) (Figure 2A). Therefore, both genes encode methyltransferases for \(N^4\)-methylation or \(2\'-O\)-methylation of C1402.

For a more detailed analysis, we used complementary oligodeoxynucleotides (42) to isolate an RNA fragment of 49 mer containing C1402 (positions 1378–1426), which was carved out from the 16S rRNA. We used LC/MS to analyze this fragment from wild-type and mutant \(E.\ coli\) strains. In the fragment isolated from wild-type \(E.\ coli\) we observed the dimethylated tetramer (m\(^3\)CmCCGp, MW 1306.2) produced by RNase \(T_1\) digestion and the dimethylated trimer (Gm\(^4\)CmCp, MW 1001.2) produced by RNase A digestion (Figure 2B). Each fragment was analyzed by collision-induced dissociation to confirm the exact positions of the dimethylation (Figure 2C). In the
Figure 2. Identification of rsmH and rsmI. (A) LC/MS analysis of the total RNA from E. coli wild-type (left panels), ΔmraW (rsmH) (middle panels) and ΔyraL (rsmI) (right panels), digested by RNase T1. The top and second panels show mass chromatograms detecting the tetramer fragments containing m⁴Cm1402 (m/z 652.1, top panels), or Cm1402 or m⁴C1402 (m/z 645.1, second panels). (B) LC/MS analysis of the 49-mer fragment (1378–1426) in the 16S rRNA from E. coli wild-type (left panels), ΔmraW (rsmH) (middle panels) and ΔyraL (rsmI) (right panels), after digested with RNase T1 (upper panels) and RNase A (lower panels). Top panels: mass chromatograms detecting the tetramer fragments containing m⁴Cm1402 (m/z 652.1). Second panels: detection of the tetramers containing Cm1402 or m⁴C1402 (m/z 645.1). Third panels: detection of trimer fragments containing m⁴Cm1402 (m/z 1000.2). Fourth panels: detection of trimer fragments containing Cm1402 (m/z 986.2). Although the dimer Gm⁴Cp must have been produced in this treatment, we could not detect it because of chemical noise in the early eluents. (C) Collision-induced dissociation (CID) spectra of the tetramer (left) and trimer (right) fragments bearing m⁴Cm1402. The doubly charged ion (m/z 652.1) of the tetramer and singly charged ion (m/z 1000.2) of the trimer were employed as parent ions for CID. The sequence of each fragment with modifications was confirmed by assignment of the product ions. The nomenclatures for product ions of nucleic acids are the same as those used in the literature (64). The regions with 10-fold magnification are indicated with ‘×10’. (D) Nucleoside analysis of the 49-mer fragment from wild-type (left panels), ΔmraW (rsmH) (middle panels) and ΔyraL (rsmI) (right panels). The top panels show the UV trace at 254 nm, the second and bottom panels show base peak presentations of the mass chromatograms to detect dimethylcytidine (m/z 272) and monomethylcytidine (m/z 258), respectively.
ΔmraW strain, a mono-methylated tetramer (m₄CmCCGp or CmCCGp, MW 1292.2) was observed in the RNase T₁ digest and a mono-methylated trimer (GmCCp, MW 987.2) was observed in the RNase A digest. If mraW encodes a methyltransferase responsible for the 2′-O-methylation of m₄Cm1402, the mono-methylated dimer (Gm₄Cp, MW 682.1) would be produced by RNase A digestion. Therefore, this result demonstrated that mraW encodes a methyltransferase responsible for the N⁰-methylation of m₄Cm1402. In the ΔylaL strain, a mono-methylated tetramer (m₄CCCGGp or CmCCGp, MW 1292.2) was observed in the RNase T₁ digest, whereas the mono-methylated trimer (GmCCp, MW 987.2) was not detected in the RNase A digest. Although the dimer Gm₄Cp must have been produced in this treatment, we could not detect it because of chemical noise in the early eluents. The data suggested that yraL encodes a methyltransferase responsible for the 2′-O-methylation of m₄Cm1402.

To confirm these results we carried out nucleoside analyses of the 49-mer fragment from the 16S rRNA. In wild-type E. coli, we observed a positively charged proton-adduct of m₄Cm (m/z 272) (Figure 2D). In the ΔmraW strain, no m₄Cm could be found; instead Cm (m/z 258) appeared (Figure 2D). In ΔylaL strain, no m₄Cm could be found, and we observed m⁴Cm (m/z 258) as an overlapping peak with that of m⁴C1407. This approximately doubled the peak intensity of the mono-methylated cytidines in the mass chromatogram (Figure 2D). Taken together, our results indicate that mraW and yraL encode methyltransferases responsible for the transfer of the N⁰-methyl and 2′-O-methyl groups, respectively, onto C1402. We therefore re-named mraW as rsmH, and yraL as rsmI, according to the established rules of nomenclature (49,50) (Figure 1B).

### Growth properties of knockout strains of rsmH and/or rsmI

To evaluate the cellular functions of m₄Cm1402 in E. coli, we measured the doubling times of single and double knockout strains of rsmH and rsmI (Table 1). When cultured at 37°C, the ΔrsmH and ΔrsmI strains showed 15% and 12% increases in doubling times, respectively, compared with wild-type E. coli. The ΔrsmH/ΔrsmI double knockout strain showed a 29% increase in doubling time. The knockout strains showed similar increases in doubling times when cultured at 33°C and 42°C. These data suggested that a clear growth phenotype was seen in the double knockout strain, whereas each of the single knockout strains showed similar but milder phenotypes. Thus, the defect in m₄Cm modification resulted in a slight but significant change in cellular growth properties.

### In vitro reconstitution of m₄Cm1402

Next we tried to reconstitute m₄Cm1402 in vitro using recombinant RsmH and RsmL. N-terminal His-tagged recombinant RsmH and RsmI were expressed and purified by Ni affinity chromatography (Figure 3A). We employed naked 16S rRNA, 30S subunits, and tightly-coupled 70S ribosomes without C1402 methylation, which were prepared from the ΔrsmH/ΔrsmI strain, as the substrates for the in vitro reconstitution. Recombinant proteins were incubated with the substrates in the presence of S-adenosylmethionine (Ado-Met) as a methyl group donor, and then the rRNAs were isolated for LC/MS analysis. As shown in Figure 3B, when the 30S subunit was used as substrate, m₄C1402 and Cm1402 were formed by RsmH and RsmI, respectively, although the level of N⁰-methylation was quite low under these conditions. In addition, m₄Cm1402 was successfully reconstituted when RsmH and RsmI were incubated together, but only in the presence of Ado-Met. Neither type of methylation was observed when the naked 16S RNA or tightly-coupled 70S ribosomes were used as substrates (Figure 3B). These results confirmed that RsmH and RsmI are AdoMet-dependent methyltransferases responsible for the methylation of C1402. Our results also revealed that RsmH and RsmI recognize the 30S subunit as a substrate, but not the naked 16S RNA or the tightly-coupled 70S ribosome.

### Dimethylation of C1402 is involved in the efficiency of non-AUG initiation and fidelity of decoding

Since m₄Cm1402 directly interacts with the phosphate backbone of the P-site codon, the dimethyl modification may modulate P-site functions. We constructed a dual luciferase reporter system to evaluate the fidelity of initiation, reading frame maintenance, and stop codon read-through by ribosomes lacking the dimethyl-modification of C1402. First, to evaluate the efficiency of translation initiation at non-AUG codons, we employed reporter constructs encoding Renilla luciferase (Rluc) starting with the AUG codon, and Firefly luciferase (Fluc) starting with various codons (AUG, AUU, UUG, GUG and AAA) on a single transcript (Figure 4A). The GUG, UUG and AUU codons have been shown to serve as initiation codons in E. coli, and normalizing the ratio to that of wild-type E. coli (relative F/R value). As shown in Figure 4A, the ΔrsmH

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>Doubling time (min)</th>
<th>ΔrsmH</th>
<th>ΔrsmI</th>
<th>ΔrsmH/ΔrsmI</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>30.5 ± 0.4</td>
<td>34.5 ± 1.9</td>
<td>31.4 ± 0.2</td>
<td>36.5 ± 0.5</td>
</tr>
<tr>
<td>37</td>
<td>20.4 ± 0.3</td>
<td>23.5 ± 0.8</td>
<td>22.9 ± 0.6</td>
<td>26.3 ± 2.5</td>
</tr>
<tr>
<td>42</td>
<td>20.9 ± 0.8</td>
<td>22.7 ± 3.1</td>
<td>22.9 ± 0.3</td>
<td>26.8 ± 0.4</td>
</tr>
</tbody>
</table>

The doubling times shown here are the means and standard deviations of three independent experiments.
strain, lacking the N^4-methylation of m^4Cm1402, showed a 1.8-fold increase in the rate of AUU initiation (relative to AUG initiation) compared with those in wild-type and ΔrsmI strains. This result suggests that the N^4-methylation of m^4Cm1402 negatively regulates the efficiency of initiation at the AUU codon. In addition, ΔrsmH/ΔrsmI showed a 1.8-fold increase in relative AUU initiation, a 20% decrease in relative UUG initiation and a 10% decrease in relative GUG initiation compared with the wild type. These data indicate that both the 2′-O-methylation and N^4-methylation of C1402 are needed for efficient initiation at the UUG and GUG codons.

Next, to measure the efficiency of stop codon read-through and reading frame maintenance in the absence of C1402 methylation, we employed fusion constructs containing the Rluc and Fluc open reading frames separated only by short windows containing a stop codon (UGA or UAG), or a frameshift site (+1 or –1) (Figure 4B). The translational efficiency of each construct was assessed by measuring the F/R values. As shown in Figure 4B, the ΔrsmH strain showed a decrease of 50% in UGA read-through. In contrast, the ΔrsmI strain showed a 1.3-fold increase in UGA read-through efficiency. However, the relative F/R value for the UGA construct in the ΔrsmH/ΔrsmI strain was similar to that of the ΔrsmH strain, indicating that rsmH plays a dominant role in the efficiency of UGA read-through. In addition, we observed that the ΔrsmI strain showed a 1.3-fold increase in +1 frameshift and a 1.2-fold increase in –1 frameshift. The ΔrsmH strain showed a 20% decrease in UAG read-through, a 10%
and reading frame maintenance in *E. coli*.

The reporter genes *Rluc* (Renilla) and *Fluc* (Firefly), starting with various codons (AUG, AUU, UUG, GUG or AAA), are tandemly arranged in the same transcript. The efficiency of translation initiation was assessed by dividing the chemiluminescence of *Fluc* by that of *Rluc* (the F/R value). For each initiation codon, the relative F/R values for the knockout strains were normalized to that of the wild-type strain. SD: Shine–Dalgarno sequence. 

**Figure 4.** Non-AUG initiation, stop codon read-through and reading frame maintenance in the absence of *rsmH* and/or *rsmI*. (A) Efficiency of translation initiation at non-AUG codons in *E. coli* wild-type (WT) (white), *Arsml* (gray), *ArsmH* (stripe) and *Arsml/ArsmH* (black). The reporter constructs with various initiation codons are shown at the top. The reporter genes *Rluc* (Renilla), starting with AUG codon, and *Fluc* (Firefly), corresponding to *E. coli* C1402 is found in *Saccharomyces cerevisiae*, humans and *Arabidopsis thaliana*. It is known that the C (UGAUGA) and D (CUA) box (C/D box) snoRNA guides 2′-O-methylation at that site in eukaryotic cytoplasmic rRNA (52). These observations suggest that 2′-O-methylation at this site is conserved not only in bacteria but also in eukaryotes.

During the biogenesis of ribosomes, rRNA modifications are introduced at different stages of subunit formation to assist in the efficient assembly of ribosomal particles (53). Some modifications occur on naked rRNAs during transcription, while others are introduced at early or later stages of assembly. In an *in vitro* reconstitution of m4Cm1402, both RsmH and RsmI used the assembled 30S particle as a substrate for methylation, while neither enzyme recognized the naked 16S rRNA or the tightly coupled 70S ribosome. Considering that m4Cm1402 is localized at the P-site, RsmH and RsmI might access the 30S subunit from the inter-subunit side and enter the P-site to specifically recognize C1402. In this model, an association between the 30S and 30S subunits would disturb the interactions of both enzymes. This would explain why neither enzyme used the tightly coupled 70S ribosome as a substrate. Since the efficiency of N4-methylation by RsmH was quite low when the 30S subunit was used as a substrate, we speculate that RsmH (and RsmI) employs the inter-subunit interface of the 30S subunit as an actual substrate at the late stage of ribosomal assembly in the cell.

In the crystal structure of 70S ribosome complexed with mRNA, and tRNAs, the P-site in helix 44 of the 16S
The m^4Cm^1402 is shown as a set of yellow spheres, and its methyl groups are depicted as red spheres. The coordinates of the methyl groups of m^4Cm^1402 are calculated based on a previously published structure (6). (B) The tertiary structure of the P-site in the 30S subunit, including the AUG codon of the mRNA, the CAU anticodon of tRNA^Met and m^3G^1349, C^1400, m^4Cm^1402, m^4U^1498 and A^1500 in the 16S rRNA. The same color code was used as described for (A). Mg^2+ ions are depicted as magenta spheres. The methyl group of m^3U^1498 is also depicted as a 3-methyl group of m^3U^1498; C^1400, m^4Cm^1402, m^4U^1498 and A^1500 in the 16S rRNA. The coordinates of the methyl groups of m^4Cm^1402 are shown as red spheres. The hydrogen bonds formed by m^4Cm^1402 are shown as red dash lines.

Figure 5. Structural insights into the P-site function of m^4Cm^1402. (A) Crystal structure of the 30S subunit with the P-site tRNA^Met and mRNA (5). The 16S rRNA, the ribosomal proteins, the anticodon stem–loop of the P-site tRNA, and the mRNA are depicted in light blue, green, blue, and orange, respectively. The m^4Cm^1402 is shown as a set of yellow spheres, and its methyl groups are depicted as red spheres. The coordinates of the methyl groups of m^4Cm^1402 are calculated based on a previously published structure (6). (B) The tertiary structure of the P-site in the 30S subunit, including the AUG codon of the mRNA, the CAU anticodon of tRNA^Met and m^3G^1349, C^1400, m^4Cm^1402, m^4U^1498 and A^1500 in the 16S rRNA. The same color code was used as described for (A). Mg^2+ ions are depicted as magenta spheres. The methyl group of m^4U^1498 is also depicted as a 3-methyl group of m^3U^1498; C^1400, m^4Cm^1402, m^4U^1498 and A^1500 in the 16S rRNA. The coordinates of the methyl groups of m^4Cm^1402 are shown as red spheres. The hydrogen bonds formed by m^4Cm^1402 are shown as red dash lines.

rRNA is shaped mainly by five conserved bases: m^4Cm^1402, C^1403, m^4U^1498, A^1499 and A^1500 (Figures 1A and 5). m^4Cm^1402 confers the following interactions: the N^4 nitrogen of the amino group forms a hydrogen bond with the oxygen atom of the phosphate backbone between the 2nd and 3rd letters of the P-site codon, the N^3 nitrogen forms a hydrogen bond with the N^6 amino group of A^1500, and the base stacks on the cytosine ring of C^1403. In addition, the N^4-methyl group of m^4Cm^1402 is close to the N^2-methyl group of m^4U^1498; this distance is calculated to be 3-4 A (Figure 5B). Since the van der Waals diameter of a methyl group is 2.0 A, these two methyl groups could interact via a van der Waals attraction. In addition, the N^4-methyl group donates an electron to the base, to strengthen the stacking force to C^1403 and the hydrogen bond between the N^3 of m^4Cm^1402 and the N^6 of A^1500. It is known that 2'-O-methylation stabilizes the sugar pucker as a C3'-endo conformation (54,55). In fact, as seen in the crystal structure, m^4Cm^1402 adopts the C3'-endo form. If it adopted the C2'-endo form, m^4Cm^1402 could no longer associate with the P-site mRNA in the current structure. Taken together, the N^4-methylation by RsmH and the 2'-O-methylation by RsmI are involved in stabilizing the local structure and interactions of the P-site to accommodate the codon–anticodon helix.

Using the dual luciferase reporter assay, we showed that the deletion of rsmH or rsmI affects the efficiency of non-AUG initiation and the fidelity of translation. A relative increase in initiation at the AUU codon was induced by a lack of N^4-methylation. It is supposed that the formation of the 30S IC with a non-AUG initiation codon is kinetically excluded due to the unstable binding of initiator fMet-tRNA^Met and the lower rates of IF3 dissociation and 50S docking (10,11). As discussed earlier, the lack of N^4-methylation at C1402 is thought to modulate the local structure of the P-site, and might allow the site to accept the non-canonical codon–anticodon helix. When the AUU codon is recognized by the initiator tRNA at the P-site, the 3rd letter of AUU codon is mismatched with the 1st letter of the anticodon. Considering that the N^4 nitrogen of m^4Cm^1402 forms a hydrogen bond with the 5' phosphate of the 3rd letter of the AUU codon, the lack of N^4-methylation might directly influence this interaction to accommodate the AUU codon, leading to efficient formation of the 30S IC and then the 70S IC. In E. coli, AUU is used as the initiation codon in the infC gene, which encodes IF3. The expression level of IF3 is controlled by the efficacy of AUU initiation (56,57). Since IF3 is involved in the selection and formation of the 30S IC by accelerating the influx of initiator tRNA (10), it is thought that a low concentration of IF3 enhances the efficiency of AUU initiation and translation of infC, thus maintaining the cellular concentration of IF3. Based on our results, N^4-methylation of m^4Cm^1402 by RsmH also plays a role in regulating AUU initiation. Furthermore, the dimethylmodification of C1402 contributes slightly to the efficacy of translation initiation at GUG or UUG. When GUG or UUG is recognized by the initiator tRNA, the 1st letter of each codon is recognized by the 3rd letter of the anticodon as a G-U or U-U wobble pair, a situation that is different from the case in AUU initiation. The dimethyl-modification of C1402 might stabilize this non-canonical pairing at the P-site. Therefore we suggest that the dimethyl-modification of C1402 plays an important role in P-site function, especially in start codon selection.

The N^3-methyl group of m^4U^1498 is located in the vicinity of the N^4-methyl group of m^4Cm^1402 at the P-site. The gene rsmE was shown to encode a
methyltransferase for m\textsuperscript{3}U1498 (58). No growth phenotype could be observed when a Δrsme\textsuperscript{E} strain was grown alone, but it was defective when grown in competition with wild-type E. coli (58). In addition, when 16S rRNA methylations were low, the rsme\textsuperscript{E} deletion apparently decreased the stringency of initiator tRNA selection (25). Taken together with our results, these observations indicate that methyl-modifications clustered at the P-site collectively work to maintain accurate translation initiation.

The dual luciferase assays also revealed that the relative rate of UGA read-through activity was increased in the absence of A\textsuperscript{4}-methylation by RsmH. This result was unexpected because read-through efficiency is determined at the A-site (59,60). However, it is known that some genetic mutations at the P-site influence read-through activity (61). In addition, a recent report indicated that a mismatch introduced in the codon-anticodon pairing at the P-site elevated the affinity of the release factor for the A-site (62), suggesting that the geometry of the P-site codon–anticodon helix affects termination codon recognition at the A-site. It can be speculated that the lack of A\textsuperscript{4} methylation of C1402 influences the configuration of the P-site codon–anticodon helix, thus modulating the activity of RF2 and the UGA codon specificity at the A-site.

In this study, we showed that m\textsuperscript{4}Cm1402 in the 16S rRNA modulates the accuracy of P-site function. In yeast, it has been reported that loss of rRNA modifications in the decoding center collectively affected biogenesis and function of the small subunit (63). It is possible that m\textsuperscript{4}Cm1402 is also involved in other ribosomal functions and in the biogenesis of the 30S subunit. Further investigations are needed to clarify these issues.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We are grateful to Drs J. Dinman, T. Miki and H. Mori for providing materials. We thank the Suzuki lab members including Drs Takeo Suzuki, Y. Ikeuchi, K. Kitahara, T. Yokoyama, and Y. Sakaguchi for technical support and for many fruitful discussions.

FUNDING
Grants-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports, and Culture of Japan (MEXT) (to T.S.); the Global COE Program for Chemistry Innovation (to S.K.); and the New Energy and Industrial Technology Development Organization (NEDO) (to T.S.). Funding for open access charge: MEXT.

Conflict of interest statement. None declared.

REFERENCES

