Published online 30 December 2011 *Nucleic Acids Research, 2012, Vol. 40, No. 9* ***4071–4085***
doi:10.1093/nar/gkr1287

Base methylations in the double-stranded RNA by a fused methyltransferase bearing unwinding activity

Satoshi Kimura, Yoshiho Ikeuchi, Kei Kitahara, Yuriko Sakaguchi, Takeo Suzuki and Tsutomu Suzuki*

Department of Chemistry and Biotechnology, Graduate School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

Received February 24, 2011; Revised November 28, 2011; Accepted December 14, 2011

ABSTRACT

Modifications of rRNAs are clustered in functional regions of the ribosome. In Helix 74 of *Escherichia coli* 23S rRNA, guanosines at positions 2069 and 2445 are modified to 7-methylguanosine(m^7G) and N^2-methylguanosine(m^2G), respectively. We searched for the gene responsible for m^7G2069 formation, and identified *rlmL,* which encodes the methyltransferase for m^2G2445, as responsible for the biogenesis of m^7G2069. *In vitro* methylation of rRNA revealed that *rlmL* encodes a fused methyltransferase responsible for forming both m^7G2069 and m^2G2445. We renamed the gene *rlmKL*. The N-terminal RlmL activity for m^2G2445 formation was significantly enhanced by the C-terminal RlmK. Moreover, RlmKL had an unwinding activity of Helix 74, facilitating cooperative methylations of m^7G2069 and m^2G2445 during biogenesis of 50S subunit. In fact, we observed that RlmKL was involved in the efficient assembly of 50S subunit in a mutant strain lacking an RNA helicase *deaD*.

INTRODUCTION

The ribosome is a ribonucleoprotein complex that translates genetic information on mRNA into the corresponding amino acid sequence to generate a protein. The ribosome consists of small and large subunits, each of which is composed of ribosomal RNAs (rRNAs) and ribosomal proteins. The rRNAs form the basic structure of the ribosome, and play central roles in the fundamental processes of protein biosynthesis (1–6). In all domains of life, rRNAs contain residues with various chemical modifications that are enzymatically introduced after transcription. These modifications are primarily base methylation, ribose methylation and pseudouridylation (7,8). In prokaryotes, rRNA modifications are introduced by site-specific RNA-modifying enzymes, whereas in eukaryotes, most rRNA modifications, namely 2′-*O*-methylations and pseudouridylations, are introduced by a series of small nucleolar RNA–protein complexes (snoRNPs) composed of a protein methyltransferase or pseudouridylase and a specific snoRNA as a guide RNA to determine the precise position for modification (9). rRNA modifications are clustered near the active sites of the ribosome such as the decoding center, subunit interfaces, transfer RNA (tRNA) or mRNA binding sites and the peptidyl-transferase center (10,11), implying that these modifications are required for proper function of the rRNAs. However, many rRNA modifications are dispensable for cell viability (12). In fact, the lack of a single rRNA modification confers only a mild defect in cell growth and ribosomal function. Hence, the exact function of each modification is unknown. However, collective disruption of several modifications causes a severe defect in ribosomal function and cell growth (13,14). Therefore, each rRNA modification plays a specific role in fine-tuning ribosomal activity and/or biogenesis (8).

The *Escherichia coli* ribosome contains 17 species of modified nucleotides at 36 positions in the rRNAs: 7 species are found at 11 positions in 16S rRNA and 15 species at 25 positions in 23S rRNA (8). The modifications are 10 pseudouridines, 24 methylated nucleosides, 1 dihydrouridine and 1 5-hydroxycytidine (15). Although all of the pseudouridine synthase genes have been identified in *E. coli*, several methyltransferases still remain to be identified. In the decoding center of the small subunit of the ribosome, rRNA modifications play roles in maintaining the translational fidelity by suppressing frameshifts and read-through activity (16–18), and ensuring accurate translational initiation (19,20). In the large 50S subunit of the ribosome, rRNA-modifying enzymes such as RrmJ/FtsJ and RluD are required for efficient assembly of the subunit (21–24), indicating that the enzymes themselves

*To whom correspondence should be addressed. Tel: +81 3 5841 8752; Fax: +81 3 5841 0550; Email: ts@chembio.t.u-tokyo.ac.jp

act as assembly factors for efficient biogenesis of the ribosome (25). In addition, some rRNA modifications are involved in modulating susceptibility to certain antibiotics (26,27).

To understand the functional and physiological aspects of rRNA modifications, identifying the genes and enzymes responsible for each modification is essential. We developed a method of genome-wide screening called ribonucleome analysis to identify genes responsible for RNA modifications using a reverse genetic approach combined with mass spectrometry (MS) (28). This approach has successfully identified many genes responsible for tRNA modifications among the uncharacterized genes in *E. coli* and *Saccharomyces cerevisiae* (29–33). Recently, we used this method to identify two methyltransferases, encoded by *rsmH* and *rsmI*, which are responsible for biogenesis of N^4, 2′-*O*-dimethylcytidine (m^4Cm) at position 1402 in the *E. coli* 16S rRNA (20).

In Helix 74, in domain V of the 23S rRNA, G2069 and G2445 are modified to form 7-methylguanosine (m^7G) and N^2-methylguanosine (m^2G), respectively (Figure 1A and B). The methyltransferase RlmL, encoded by *rlmL*/*ycbY*, catalyzes *S*-adenosylmethionine (Ado-Met)-dependent m^2G2445 formation (34). Although deletion of *rlmL*/*ycbY* resulted in a slight growth reduction phenotype, the functional and physiological role of m^2G2445 remains unclear (34). The methyltransferase mediating the biogenesis of m^7G2069 has not yet been identified. In this study, we employed a genome-wide screen of uncharacterized genes in *E. coli* using the ribonucleome analysis to search for the gene responsible for m^7G2069 formation. We happened to identify *rlmL*/*ycbY* as responsible for the biogenesis of m^7G2069. In fact, *rlmL*/*ycbY* encodes a fused methyltransferase with dual active sites responsible for forming both m^7G2069 and m^2G2445 (Figure 1C). Thus, *rlmL*/*ycbY* has been renamed *rlmKL*. Genetic and biochemical characterization of this unique enzyme has revealed that cooperative methylation of Helix 74 by RlmKL plays a key role in the efficient assembly of the 50S subunit.

MATERIALS AND METHODS

RNA mass spectrometry

Each rRNA fraction (200 fmol) was digested at 37°C for 30 min in a 10 µl reaction mixture containing 10 mM ammonium acetate (pH 5.3) and 5 U/µl RNase T1 (Epicentre), or 10 mM ammonium acetate (pH 7.7) and 1 ng/µl RNase A (Ambion). Subsequently, an equal volume of 0.1 M triethylamine-acetate (TEAA) (pH 7.0) was added to the reaction mixture for LC/MS. Analysis of RNA fragments by capillary liquid chromatography (LC) coupled with nano electrospray (ESI) LC/MS was carried out using a tandem quadrapole time-of-flight (QqTOF) mass spectrometer (QSTAR XL, Applied Biosystems) and a linear ion trap-orbitrap hybrid mass spectrometer (LTQ Orbitrap XL, Thermo Fisher Scientific). Both systems are equipped with a nano electrospray source and a splitless nano HPLC system (KYA Technologies), as described previously (20). Conditions and solvent systems for capillary LC have been described previously (28). All procedures for the ribonucleome analysis were conducted as described previously (20).

Sucrose density gradient centrifugation

Subunit profiling by sucrose density gradient (SDG) centrifugation was carried out as previously described (35). Briefly, *E. coli* cells were grown in 250 ml of Luria-Bertani (LB) medium in a 500-ml flask with vigorous shaking at 25°C. Cells were harvested from 50 ml culture by centrifugation when the cell density reached an A_{600} of 0.5. The cell pellet was resuspended in 1 ml of cold buffer [20 mM HEPES–KOH (pH 7.6), 0.5 mM $Mg(OAc)_2$, 100 mM NH_4Cl, 6 mM β-mercaptoethanol]. A cell lysate was prepared by the lysozyme-freeze–thaw method, as described (35) and cleared by centrifugation at 15 000 rpm for 15 min at 4°C. The concentration of total RNA in the lysate was estimated by measuring A_{260}. In total, 10 U of A_{260} of the lysate were layered on top of a 10–40% SDG prepared in cold buffer and then separated by ultracentrifugation in a Beckman SW-28 Rotor at 20 000 rpm for 14 h at 4°C. Fractions were collected from the gradient using a Piston Gradient Fractionator (BIOCOMP) and the position of the ribosomal subunits was monitored by A_{260} using a ultra violet (UV) monitor (ATTO AC-5200).

In vitro methylation assay

For *in vitro* reconstitution of m^7G2069 and m^2G2445 formation, we used 23S rRNA or a series of rRNA fragments as substrates. rRNA fragments and transcripts 1–9 except for transcript 6 were transcribed *in vitro* by T7 RNA polymerase (see Supplementary Data, 'Materials and Methods' section). The transcript 6 was chemically synthesized (Sigma genosys).

A reaction mixture (50 µl) consisting of 200 mM NH_4OAc, 40 mM Tris–HCl (pH 7.5), 3 mM $MgCl_2$, 6 mM β-mercaptoethanol, 1 mM Ado-Met, 0.2 µM 23S rRNA (or rRNA fragments) and 0.1 µM recombinant RlmKL [or RlmK(CTD), or RlmL(NTD)] was incubated at 37°C for 2 h. Substrate RNAs were recovered from aliquots of the reaction mixture by phenol–chloroform extraction and ethanol precipitation. RNA was digested by RNase T1 or RNase A and analyzed by LC/MS. To examine the time course of methylation (Figure 4C and D and Supplementary Figures S2 and S4), a 100-µl reaction mixture was prepared and 10-µl aliquots were taken at each time point and mixed with phenol–chloroform to stop the reaction. For preparing the m^7G2069-containing domain V RNA, 50 pmol domain V RNA was incubated at 37°C with 200 pmol RlmK (CTD) in 100 µl reaction mixture for 2 h. Phenol–chloroform-extracted RNA was passed through NAP-5 column (GE healthcare) to remove Ado-Met (Ado-Hcy), and recovered by ethanol precipitation. In RNase T1 digested samples, we detected a 16-mer fragment bearing m^7G2069 (Um^7GA ACCUUUACUAUAGp), or a 14-mer fragment without methylation (AACCUUUACUAUAGp). In RNase A digested samples, a tetramer fragment with (m^7GAACp) or without (GAACp) m^7G2069 and a hexamer fragment

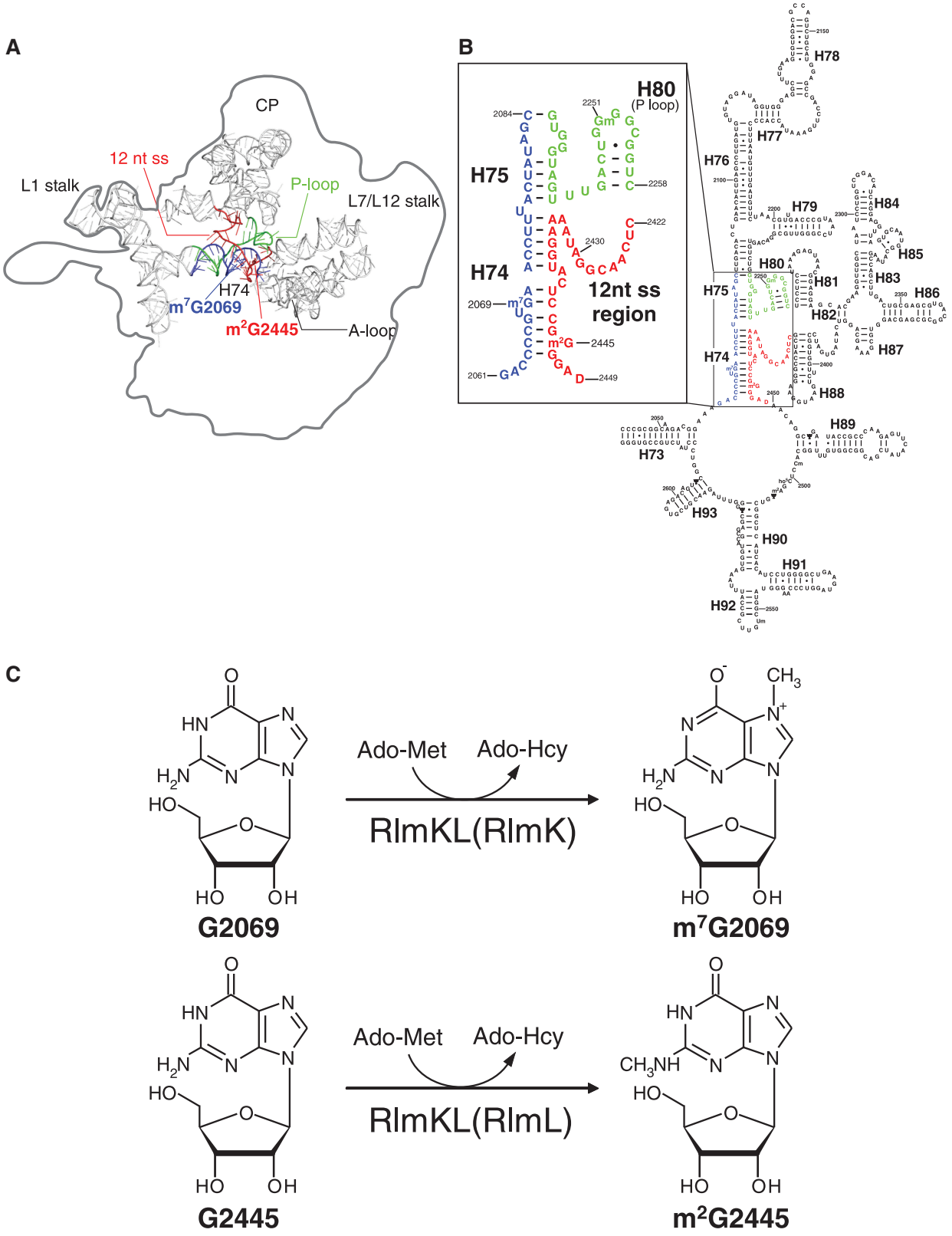

Figure 1. Tertiary and secondary structures of the *E. coli* 23S rRNA domain V and enzymatic formation of 7-methylguanosine (m^7G2069) and N^2-methylguanosine (m^2G2445). (**A**) The tertiary structure of domain V in *E. coli* 23S rRNA (PDB ID 2aw4). The strands G2061-C2084, G2235-C2258, C2422-D2449 are colored blue, green and red, respectively. Coordinates for Helix 93 are not displayed because it lies in front of Helix 74. The silhouette of 50S subunit was showed in gray line. (**B**) The secondary structure of domain V in the *E. coli* 23S rRNA with modified nucleotides, 7-methylguanosine (m^7G), N^2-methylguanosine (m^2G), 2′-*O*-methylguanosine (Gm), dihydrouridine (D), pseudouridine (Ψ), 2′-*O*-methylcytidine (Cm), 5-hydroxycytidine (ho^5C), N^2-methyladenosine (m^2A) and 2′-*O*-methyluridine (Um). The numbers of the RNA helices and nucleotides are indicated. Watson–Crick-type base pairs are shown by bars, and wobble base pairs by black dots. Helix 74 and surrounding regions are boxed on the left. The color code for each strand is the same as (A). (**C**) Enzymatic formation of m^7G2069 and m^2G2445 is catalyzed by RlmK and RlmL, respectively, using Ado-Met as a methyl-group donor. Ado-Met is converted to Ado-Hcy (*S*-adenosylhomocysteine) by the reaction.

with (Gm^2GGGAUp) or without (GGGGAUp) m^2G2445 were detected. The methylated and unmethylated fragments were quantified by summing the areas of the mass chromatograms (±0.1 tolerance) of all detected ions. The frequency of methylation was calculated from the ratio of the areas of the mass chromatograms for the methylated and unmethylated fragments. For accurate quantification of both methylations, we generated a standard curve for the area ratio of the chromatograms against the ratio of the methylated fragments (Supplementary Figure S5). Using these curves, we could determine the exact values for methylation.

For the mutational study of transcript 3, methylation assay was performed as followed. The reaction mixtures (20 µl) consisting of the same buffer, except for the 1 µM of substrate RNA and 48 µM of ^{14}C-labeled Ado-Met (MP Bio Inc.), were incubated at 37°C for 30 min. Reaction was stopped and the RNA substrate was extracted by the addition of 50 µl phenol–chloroform. The half-volume of the extracted RNA was spotted onto a Whatman 3 mm filter disc. The discs were washed three times with 5% trichloroacetic acid (Nacalai Tesque) and dried with ethanol. Total radioactivity for both m^7G2069 and m^2G2445 methylations was then measured by liquid scintillation counting. The remaining half of RNA was treated by chloroform and precipitated with ethanol. The recovered RNA was digested with MazF (Takara) in the reaction mixture (5 µl) consisting of 1 × MazF digestion buffer (Takara), 10 U MazF and the RNA at 37°C for 1 h. A digested RNA was mixed with 5 µl of 2 × loading solution (9 M urea, 0.025 % bromophenol blue and 0.025% xylene cyanol) and electrophoresed by 10% polyacrylamide gel with 7 M urea. The gel was stained by Toluidine blue O and dried *in vacuo*. Radioactivity was visualized and quantified using a FLA7000 system (Fujifilm). There were mainly three bands, which were the intact 123-mer substrate undigested by MazF, m^7G2069-containing 27-mer fragment and m^2G2445-containing 96-mer fragment (Supplementary Figure S7). The ratio of m^7G2069 and m^2G2445 was determined from the signal intensities of both fragments. Each methylation was determined from the ratio of both methylations and the total sum of the methylations measured by the liquid scintillation counting.

RNA unwinding assay

The unwinding assay was basically carried out according to the literatures for RNA helicases (36,37). 5′-^{32}P-labeled transcript 8 and non-labeled transcript 9 were annealed to prepare the duplex substrate for RlmKL. In this experiment, one guanosine residue was added to 5′-end of transcript 8 for efficient *in vitro* transcription. Annealing was performed in 8 µl aliquot containing 200 mM NH_4OAc, 40 mM Tris–HCl (pH 7.6), 3 mM $MgCl_2$, 6 mM β-mercaptoethanol, 2 pmol transcript 8, 2 pmol transcript 9 and 5′-^{32}P-labeled transcript 8 (220 000 cpm) as a tracer. The mixture was heated at 90°C for 1 min and stood at room temperature for 20 min. After 3 min pre-incubation at 30°C, 0.8 µl of the mixture was mixed with 19.2 µl solution containing 200 mM NH_4OAc, 70 mM Tris–HCl (pH 7.6), 3 mM $MgCl_2$, 60 mM KCl, 6 mM β-mercaptoethanol, 0.6 mM dithiothreitol (DTT) 1 mM Ado-Met, 4 pmol non-labeled transcript 8 and 10 µg of RlmKL (100 pmol) or acetylated bovine serum albumin (BSA) (B2518 Sigma-Aldrich). The reaction mixture was incubated at 30°C. At each period of time, 5 µl aliquot was taken and mixed with 1.67 µl of 4 × quench solution containing 120 mM ethylenediaminetetra acetic acid, 2.4% sodium dodecyl sulfate (SDS), 40% glycerol and bromphenol blue (BPB) and left on ice. The quenched mixtures were analyzed by 15% native polyacrylamide gel electrophoresis (PAGE) (29:1) at 4°C. The gel was exposed to the imaging plate to visualize and quantify the duplex and the unwound single strand by FLA7000 imaging analyzer system (Fujifilm). The plot was fitted to one phase decay equation [dsRNA unwound] = $[dsRNA]_0 \cdot(1-\exp(-k\cdot\text{time}/[dsRNA]_0))$ and the half-life was calculated, using GraphPad PRISM 5 (Graph Pad software).

Other materials and methods are described in Supplementary Data.

RESULTS

Identification of *rlmL*/*ycbY* as the enzyme mediating m^7G2069 formation

To identify gene products responsible for rRNA modifications in *E. coli*, we employed a reverse genetic approach combined with RNA MS (ribonucleome analysis) for genome-wide screening in *E. coli* (20,28). Total RNA, consisting primarily of rRNAs and mRNAs, was prepared from 94 genomic-deletion strains of *E. coli* (38), each of which lacks about 20 kb (~20 genes) (see Supplementary Data). Thus, this analysis covered almost 2000 genes, ~40% of the total 4762 genes of *E. coli*. Total RNA was digested into fragments by base-specific ribonucleases (RNase T1 or RNase A) and the digests were analyzed by capillary LC coupled with nano ESI LC/MS. In these complex mixtures of RNA fragments, we can detect RNA fragments containing modified nucleoside(s) as multiply-charged negative ions by monitoring the specific mass-to-charge ratio (m/z) value. Using both RNase T1 and RNase A, we can detect all modifications to be analyzed. In wild-type *E. coli*, we clearly detected the triple-charged negative ion (m/z 1706) of a 16-mer fragment including m^7G2069 (Um^7GAACCUUUACUA UAGp; MW 5120) produced by RNase T1 digestion, and the double-charged negative ion (m/z 1024) of a hexamer fragment including m^2G2445 (Gm^2GGGADp; MW 2049) produced by RNase A digestion (Figures 1B and 2A). Each fragment was analyzed by collision-induced dissociation (CID) by MS/MS to confirm the exact position of the modification (Figure 2B). When we analyzed RNA fragments from the deletion strain OCR55 (*ΔpncB-rmf*), the 16-mer fragment bearing m^7G2069 was absent (Figure 2A). If 7-methylation of G does not occur at position 2069, an unmodified 14-mer fragment (AACCU UUACUAUAGp; m/z 1485, MW 4454) should be detected. In fact, we clearly detected this 14-mer fragment from the OCR55 strain (data not shown), confirming the absence of m^7G2069. These data indicated that

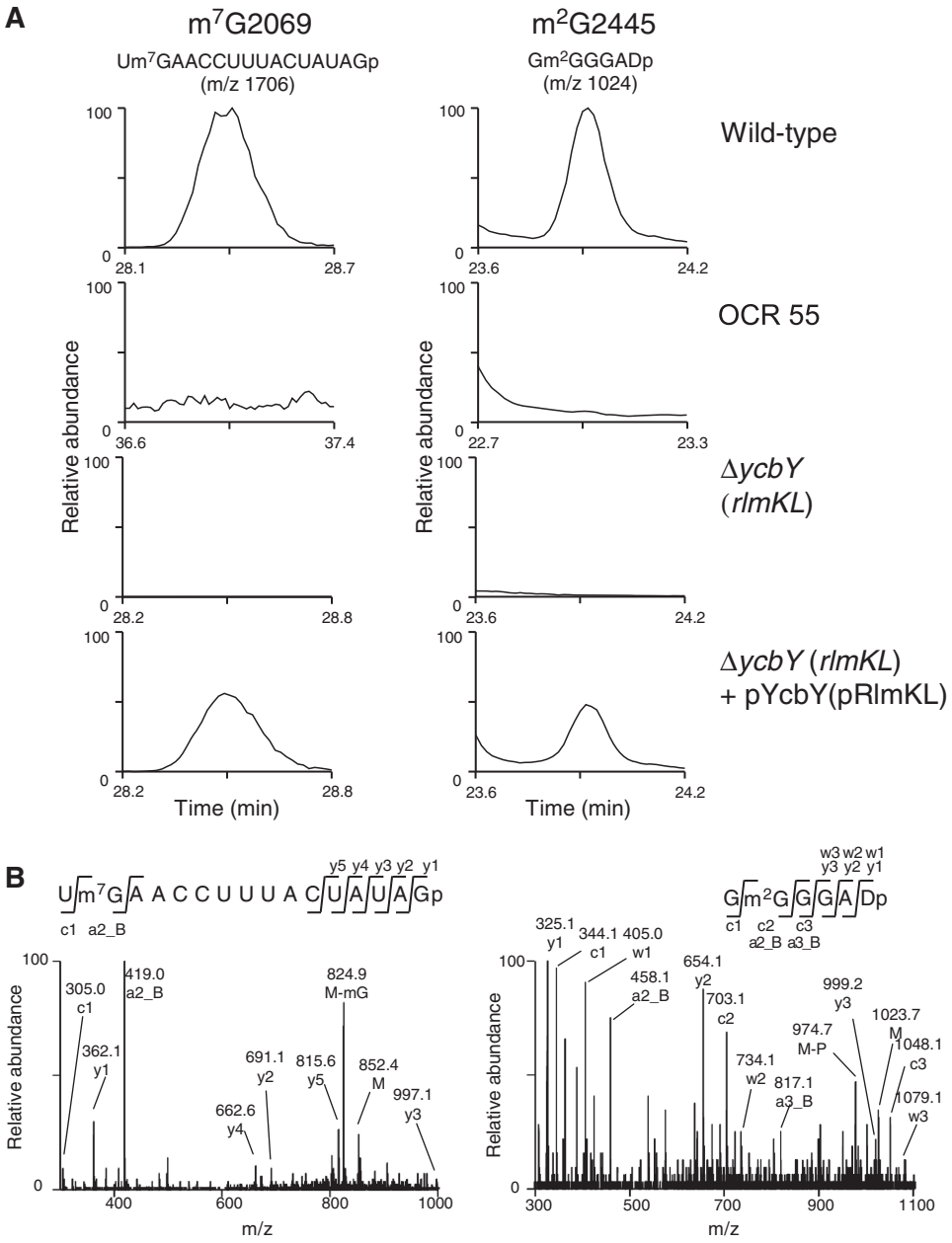

Figure 2. *ycbY* (*rlmKL*) is responsible for both m^7G2069 and m^2G2445 formation. (**A**) LC/MS analyses of RNase T1 (left panels) and RNase A (right panels) digests of rRNAs from the wild-type (Top row of panels), genome-deletion strain OCR55 (Second row), *ΔycbY* (*rlmKL*) strain (Third row) and *ΔycbY* (*rlmKL*) strain carrying pRlmKL (Bottom row). Left and right panels show mass chromatograms detecting the triply-charged ions of the 16-mer fragment bearing m^7G2069 (*m*/*z* 1706) and the doubly charged ions of the hexamer fragment bearing m^2G2445 (*m*/*z* 1024), respectively. (**B**) Collision-induced dissociation (CID) spectra of RNA fragments bearing m^7G2069 (left panel) and m^2G2445 (right panel). The six-charged ion (*m*/*z* 852.4) of the 16-mer fragment carrying m^7G2069 and doubly charged ion (*m*/*z* 1023.7) of the hexamer fragment carrying m^2G2445 were used as parent ions for CID. The sequences were confirmed by assignment of the product ions. The nomenclatures for product ions of nucleic acids are as suggested in the literature (41).

a gene responsible for m^7G2069 formation resides in this deleted region. Since *rlmL*/*ycbY* is also localized in this region, no hexamer fragment bearing m^2G2445 was observed (Figure 2A). To identify the gene responsible for m^7G2069 formation, we analyzed 23 single gene knockout strains (KO collections), each of which lacks one gene in the deleted region of OCR55. The *ΔrlmL*/*ycbY* strain was found to lack the 16-mer fragment bearing m^7G2069 (Figure 2A). To our surprise, *rlmL*/*ycbY* gene encodes a known methyltransferase, RlmL, which is responsible for m^2G2445 formation (34) and in fact, no hexamer fragment bearing m^2G2445 was detected in this strain (Figure 2A). When a plasmid encoding *rlmL*/*ycbY* was introduced into the *ΔrlmL*/*ycbY* strain, both m^7G2069 and m^2G2445 were detected in the total RNA (Figure 2A). These data clearly demonstrated that *rlmL*/*ycbY* is responsible not only for m^2G2445 formation, but also for m^7G2069 formation.

Therefore, we renamed the gene *rlmKL* according to the preferred nomenclature (39,40).

RlmKL is a fused methyltransferase consisting of an N-terminal RlmL domain and a C-terminal RlmK domain

Escherichia coli RlmKL is a fused methyltransferase consisting of 702 amino acid residues, containing a thiouridine synthases, methylases and PSUSs (THUMP) domain and two methyltransferase domains (Figure 3A). According to the cluster of orthologous group (COG) database (42), the N-terminal domain (NTD) and C-terminal domain (CTD) of RlmKL have been annotated as COG0116 and COG1092, respectively. Each domain has an Ado-Met-dependent methyltransferase motif (Figure 3A). The *rlmKL* gene is conserved in γ-proteobacteria (Supplementary Figure S1). However, in β-proteobacteria, *Neisseria meningitidis* bears COG0116 and COG1092 as separate proteins encoded at different genomic loci (Supplementary Figure S1). Since *rlmKL* appears to be responsible for two different methyl modifications, we hypothesized that each methyltransferase domain catalyzes the formation of one modification, either m^7G2069 or m^2G2445. To examine this possibility, we carried out complementation tests for the two methyl modifications in the *ΔrlmKL* strain by introducing various *rlmKL* mutants. We constructed *rlmKL* mutants in which conserved amino acid residues were substituted with Ala, including D195A, N309A, N397A, R530A, D568A and D597A. As a positive control, introduction of plasmid-encoded *rlmKL* into the *ΔrlmKL* strain restored both methyl modifications (Figure 3B). Subsequently, each of the *rlmKL* mutants was introduced into the *ΔrlmKL* strain, and formation of the methyl modifications was examined. The N309A mutation in the NTD impaired m^2G2445 formation, but rescued m^7G2069 formation (Figure 3B). In contrast, the D568A mutation in the CTD did not rescue m^7G2069 formation, but m^2G2445 was efficiently formed in this mutant (Figure 3B). These observations suggested that the NTD and CTD of RlmKL encode an RlmL for m^2G2445 formation and an RlmK for m^7G2069 formation, respectively. Asn397 is located in the linker region between the NTD and CTD. Since the N397A mutant exhibited no m^2G2445 formation but rescued m^7G2069 formation, Asn397 should be considered part of RlmL. The other mutations, including D195A, R530A and D597A, did not affect methylation (data not shown).

We next examined complementation of methyl modifications by the separate N-terminal domain [RlmL(NTD)] and C-terminal domain [RlmK(CTD)] constructs. Given the C-terminal end of the *N. meningitidis* RlmL homolog (Supplementary Figure S1) and that Asn397 should be part of RlmL, Met1 to Arg398 of RlmKL was used for RlmL(NTD), while Asn376 to Ala702 was cloned for RlmK(CTD). The RlmL(NTD) construct rescued m^2G2445 formation, and the RlmK(CTD) construct restored m^7G2069 formation (Figure 3B). These data indicated that each methyltransferase domain could complement each methylation defect when supplied on a multicopy plasmid. Moreover, m^7G2069 formation did not require pre-existing m^2G2445 and vice versa: both methylations occur independently.

Neisseria meningitidis has separate protein homologs for RlmL and RlmK. We next examined whether these homologs have the same activities as the *E. coli* RlmKL. We cloned the *N. meningitidis rlmL* (NMB0455) and *rlmK* (NMB1367) homologs into multicopy plasmids, and introduced each construct into the *E. coli ΔrlmKL* strain. As shown in Figure 3C, the *N. meningitidis rlmL* and *rlmK* homologs rescued m^2G2445 and m^7G2069 formation, respectively, demonstrating that the *N. meningitidis* RlmL and RlmK proteins have activities corresponding to the NTD and CTD of *rlmKL*.

In vitro reconstitution of m^7G2069 and m^2G2445 catalyzed by RlmKL

Methylation of m^7G2069 and m^2G2445 was reconstituted *in vitro* using recombinant RlmKL. The *E. coli* RlmKL and the RlmK(CTD) and RlmL(NTD) constructs were expressed recombinantly and purified to homogeneity (Figure 4A). As substrates, we isolated 50S ribosomal subunits and naked 23S rRNA lacking both m^7G2069 and m^2G2445 from the *ΔrlmKL* strain. The recombinant proteins were incubated with the 50S subunit and the naked 23S rRNA in the presence or absence of Ado-Met as a methyl donor and the rRNAs were then extracted and analyzed by LC/MS. No methylation was formed when the 50S subunit was used as the substrate for RlmKL (Figure 4B). When the naked 23S rRNA was used as a substrate, both m^7G2069 and m^2G2445 were successfully reconstituted by RlmKL in the presence of Ado-Met (Figure 4B). These results clearly showed that RlmKL uses the naked 23S rRNA as a substrate, indicating that methyl modifications of Helix 74 by RlmKL are likely to be introduced at an early step in 50S assembly in the cell. Next, m^7G2069 and m^2G2445 formation by RlmK(CTD) and RlmL(NTD), respectively, was reconstituted. As expected, m^7G2069 was generated by RlmK(CTD), whereas m^2G2445 was generated by RlmL(NTD) (Figure 4B). These data demonstrated that RlmKL is a fused methyltransferase, with each of the domains, N-terminal RlmL and C-terminal RlmK, acting independently as a methyltransferase. To facilitate the *in vitro* experiments, we examined the 23S rRNA domain V, which was transcribed *in vitro*. In fact, both methylations were efficiently introduced to the domain V transcript.

RlmK supports the N^2-methylation of G2445 catalyzed by RlmL

To investigate the functional advantages of a fused methyltransferase, we examined the cooperativity of methylations by the two domains. We employed the domain V transcript as a substrate to compare the activities of m^2G2445 formation catalyzed by RlmL(NTD) and RlmKL (Figure 4C). The initial velocity of m^2G2445 formation by RlmKL was much higher than that by RlmL(NTD). Whether the RlmL(NTD)-catalyzed m^2G2445 formation was affected by the addition of RlmK(CTD) *in trans* was also examined (Figure 4C). The initial velocity of m^2G2445 formation was increased

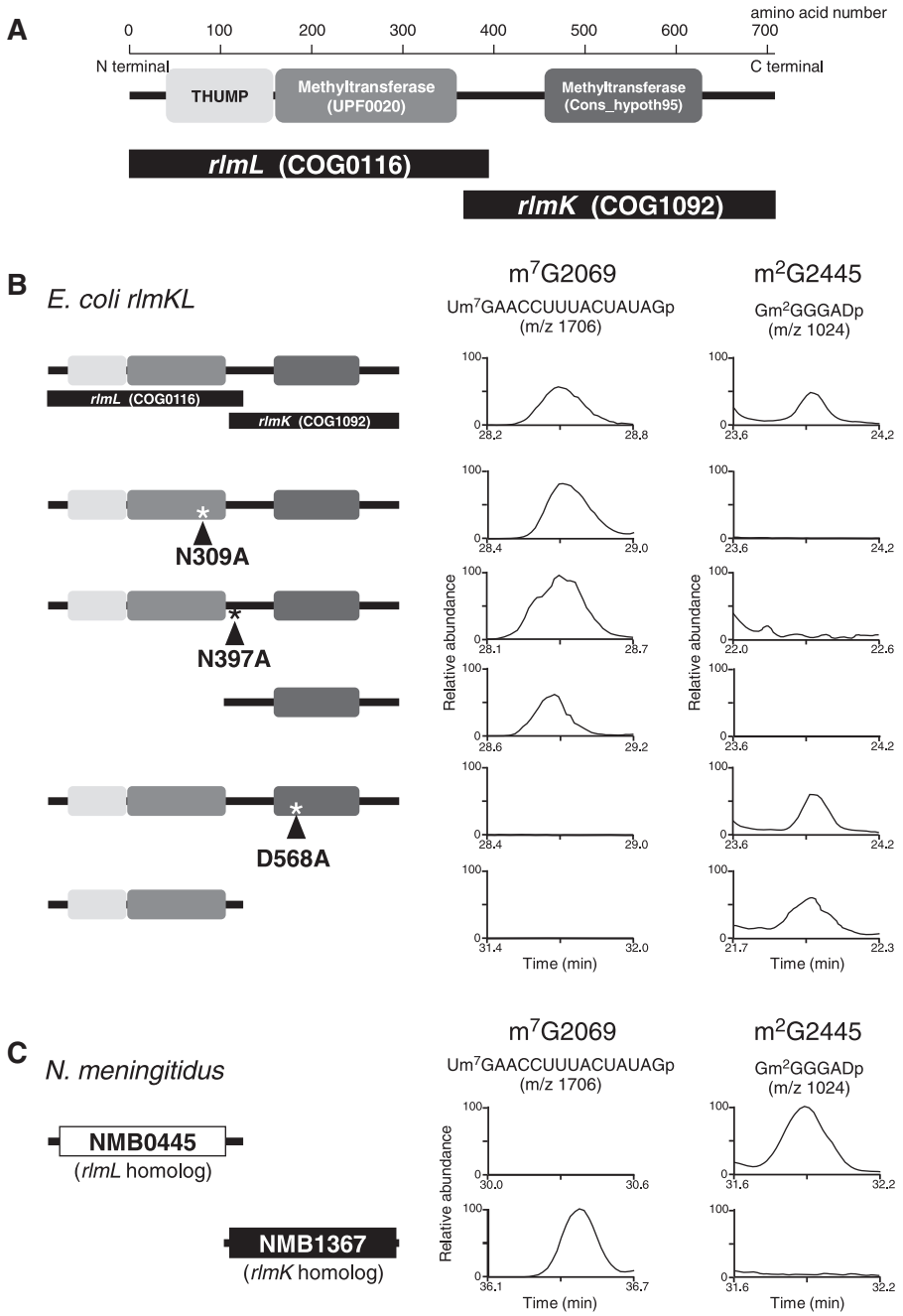

Figure 3. Complementation study of m^7G2069 and m^2G2445 formation by plasmid-encoded RlmKL and mutants. (**A**) Domain structure of *E. coli* RlmKL. RlmKL consists of two methyltransferase domains, COG0116 and COG1092. We designated the N-terminal domain (COG0116) as RlmL and the C-terminal domain (COG1092) as RlmK. Each domain has an Ado-Met-dependent methyltransferase motif, UPF0020 in the RlmL domain and Cons_hypoth 95 in the RlmK domain. A THUMP domain was found in the N-terminal region of the RlmL domain. (**B**) Complementation of the *ΔrlmKL* strain by introducing plasmid-encoded RlmKL or RlmKL mutants. On the left hand side, the domain structures and the mutation positions of RlmKL are shown. On the right hand side, mass chromatograms detecting the 16-mer fragment carrying m^7G2069 (*m*/*z* 1706) and the hexamer fragment carrying m^2G2445 (*m*/*z* 1024), respectively, are shown for the *ΔrlmKL* strain transformed with pRlmKL (top panels), pRlmKL N309A (second panels), pRlmKL N397A (third panels), pRlmK (fourth panels), pRlmKL D568A (fifth panels) and pRlmL (bottom panels). (**C**) Complementation of the *E. coli ΔrlmKL* strain by introducing *N. meningtidis* RlmL (NMB0445) or RlmK (NMB1367). Mass chromatograms on the right panels are the same as in (B).

depending on the concentration of RlmK(CTD), showing that the addition of RlmK(CTD) elevated the m^2G2445-forming activity of RlmL(NTD). In the presence of 0.8 μM RlmK(CTD), 0.1 μM RlmL(NTD) exhibited an m^2G2445-forming activity level similar to 0.1 μM RlmKL (Figure 4C). It was speculated that m^2G2445 formation by RlmL(NTD) might be enhanced by the presence of m^7G2069 generated by RlmK(CTD) in

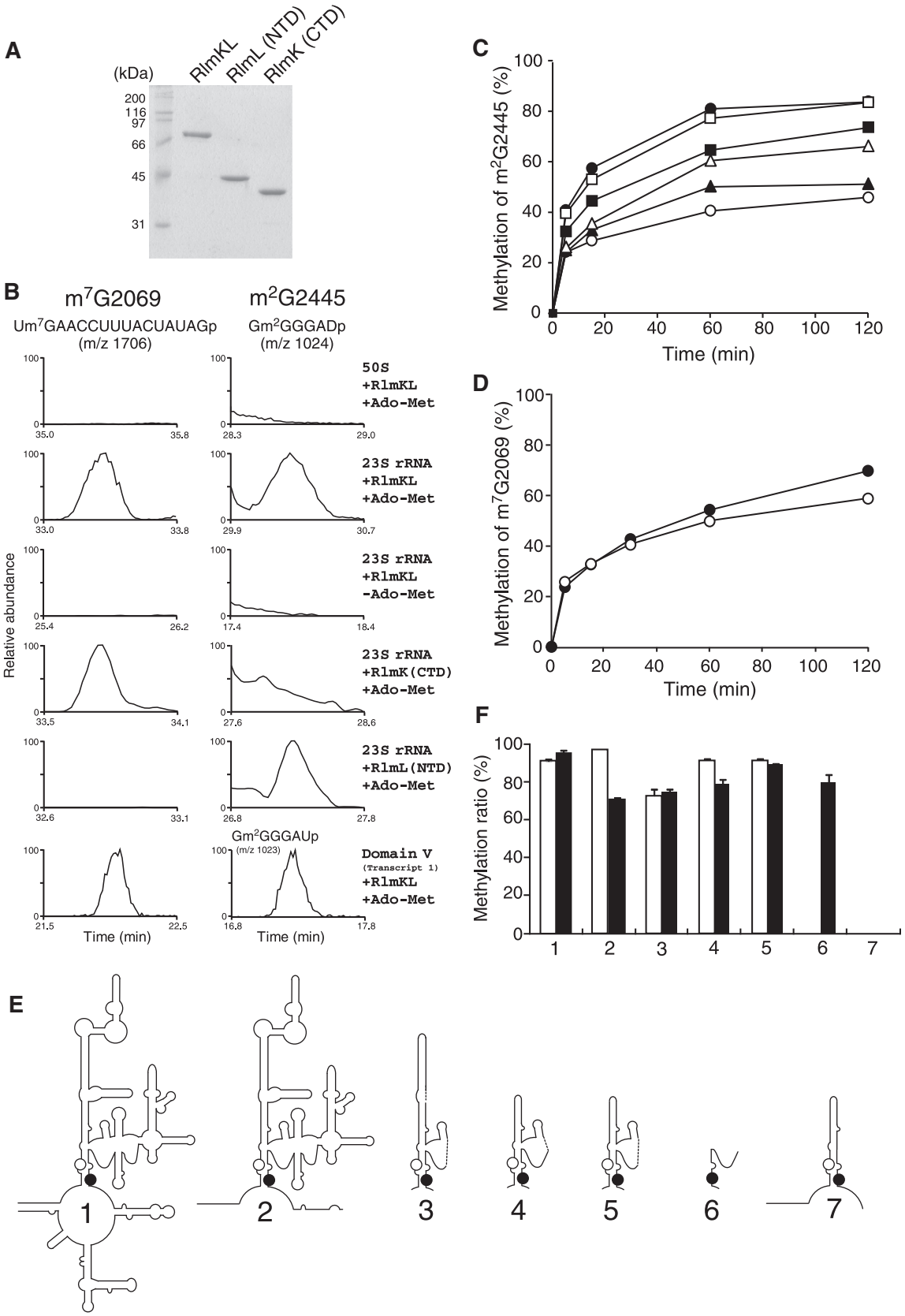

Figure 4. *In vitro* reconstitution of m^7G2069 and m^2G2445 formation by RlmKL. (**A**) SDS–PAGE analysis of the purified recombinant RlmKL, RlmL(NTD) and RlmK(CTD) stained with coomassie Brilliant Blue R-250. (**B**) *In vitro* reconstitution of m^7G2069 and m^2G2445 by RlmKL, RlmL(NTD) or RlmK(CTD). Left and right panels show mass chromatograms detecting the triple-charged ions of the 16-mer fragment bearing m^7G2069 (*m/z* 1706) and the double-charged ions of the hexamer fragment bearing m^2G2445 (*m/z* 1024), respectively. Conditions for *in vitro*

(continued)

substrate RNA. To examine this possibility, we prepared the domain V transcript with m^7G2069 (82% was modified). However, the presence of m^7G2069 did not affect the m^2G2445-forming activity (Supplementary Figure S2). In contrast, the initial velocities of m^7G2069 formation by RlmKL and RlmK were very similar (Figure 4D). In these experiments, we clearly demonstrated that the fusion of these two methyltransferases enhances the efficiency of m^2G2445 formation.

Minimization of the substrate recognized by RlmKL

To investigate substrate recognition by RlmKL, we minimized the substrate RNA by preparing six truncated RNAs, based on domain V, by *in vitro* transcription and examined the transcripts for methylation by RlmKL (Figure 4E and F and Supplementary Figure S3). By the end point measurement, over 90% of the domain V transcript (transcript 1) was methylated at both positions (Figure 4E and F). The methylation of transcript 2 indicated that the lower parts of domain V, including Helices 73, 89 and 90–93, are not necessary for either methylation (Figure 4E and F). Even when the upper part of domain V was truncated, as in transcripts 3, 4 and 5 (Supplementary Figure S3) m^7G2069 and m^2G2445 were efficiently formed (Figure 4E and F). These results indicated that removal of the upper part (Helices 76–79), as well as the right part (Helices 81–88) with connecting residues C2258 and C2422 (transcript 4) or connecting residues C2258 and A2426 (transcripts 3 and 5) resulted in little effect on either methylation (Figure 4E and F). In contrast, transcript 7 (Figure 4E and Supplementary Figure S3), which lacks Helix 80 and the 12 nt single-strand (ss) region (Figure 1B), was not methylated at either position (Figure 4F). This result indicated that Helix 80 and the 12 nt ss region are critical sites necessary for m^2G2445 and m^7G2069 formation. Additionally, in the 29-mer single-stranded transcript 6 which consists of residues C2422 to A2450 (Figure 4E and Supplementary Figure S3), m^2G2445 was efficiently formed (Figure 4F). This result indicated that duplex formation of H74 is not required for the m^2G2445 formation.

As the single-stranded RNA (transcript 6) was a good substrate for RlmKL, we examined a substrate specificity of RlmL(NTD) using transcripts 4 and 6 as substrates (Supplementary Figure S4). RlmL(NTD) catalyzed the m^2G2445 formation on transcript 6, but did not efficiently methylate transcript 4 which forms Helix 74 (Figure 4E). This result strongly suggests that RlmL prefers the single-stranded RNA substrate for the m^2G2445 formation. Taken together with the observation that RlmK domain has a function to support the m^2G2445 formation by RlmL(NTD), we speculated that RlmK domain unwinds the duplex structure of Helix 74, thereby facilitating the efficient N^2-methylation of G2445 catalyzed by the RlmL domain.

Unwinding activity of RlmKL

To examine the unwinding activity of RlmKL, we performed the unwinding assay used for RNA helicases (36,37). The 5′-^{32}P-labeled transcript 8 and non-labeled transcript 9 were annealed to prepare the duplex substrate for RlmKL (Figure 5A). The duplex and single strands were resolved by native PAGE to monitor the unwinding of the duplex with the passage of reaction time (Figure 5B). In the presence of RlmKL, 71% of the duplex substrate was unwound in the first 3 min, and 92% was unwound in 30 min (Figure 5B). As a control, in the presence of BSA, 29% of the duplex substrate was spontaneously converted to the single strand, and 66% was unwound in 30 min (Figure 5B). In the presence of RlmKL, the half-life ($t_{1/2}$) of the duplex substrate was 2.5 min, whereas in the presence of BSA, $t_{1/2}$ was 6.0 min (Figure 5C). We also examined the unwinding activity of RlmKL in the absence of Ado-Met, and found no change in the activity (data not shown). These results suggested that RlmKL has an activity to unwind Helix 74 during substrate recognition and methylation.

Substrate recognition by RlmKL

To identify the sites required for both methylations, we constructed a series of variants based on transcript 3 and measured the methylation capacity of each variant. The secondary structure of transcript 3, the detailed constructs of the variants and their methylation activities are shown in Figure 6A, B and C, respectively. The efficiency of methylation was measured as described in 'Materials and Methods' section (Supplementary Figure S7). The base flipping of Helix 80 reduced m^7G2069 formation. In the P-loop, three variants, G2250C, G2251C and G2252C, lacked m^7G2069 formation, and G2253C reduced m^7G2069 formation, whereas the C2254A variant showed normal activity. All mutants in P-loop and H80 have efficient m^2G2445 forming capacities. In the 12 nt ss region, three variants, C2427U, G2429A and A2430U showed reduced m^7G2069 forming activity, whereas G2428C and G2429A lacked m^2G2445 formation and C2427U and A2430U and 2431-2433 mutation reduced the m^2G2445 formation. In Helix 74, decreased

Figure 4. Continued
methylation are indicated in the right hand side. The hexamer fragment with m^2G2445 produced by RNase A digestion of domain V is $Gm^2GGGAUp$ (*m*/*z* 1023, MW 2047). (**C**) Time-course of m^2G2445 formation in domain V (transcript 1) catalyzed by 0.1 μM RlmL(NTD) (open circle), 0.1 μM RlmL(NTD) with 0.1 μM RlmK(CTD) (closed triangle), 0.1 μM RlmL(NTD) with 0.2 μM RlmK(CTD) (open triangle), 0.1 μM RlmL(NTD) with 0.4 μM RlmK(CTD) (closed square), 0.1 μM RlmL(NTD) with 0.8 μM RlmK(CTD) (open square) and 0.1 μM RlmKL (closed circle). (**D**) Time-course of m^7G2069 formation in domain V (transcript 1) catalyzed by 0.1 μM RlmK(CTD) (open circle) and 0.1 μM RlmKL (closed circle). (**E**) Truncated RNA substrates. Open and closed circles in the secondary structures indicate the positions of G2069 and G2445, respectively. Detailed constructs for transcripts 3–7 are shown in Supplementary Figure S3. (**F**) Truncation of domain V (transcript 1) to identify the minimum substrate. The ratio of methylation (%) was calculated from the area of mass chromatograms for the methylated and unmethylated fragments at the end point of the reaction (see Materials and Methods). White and black bars represent the efficiency of m^7G2069 and m^2G2445 formation, respectively.

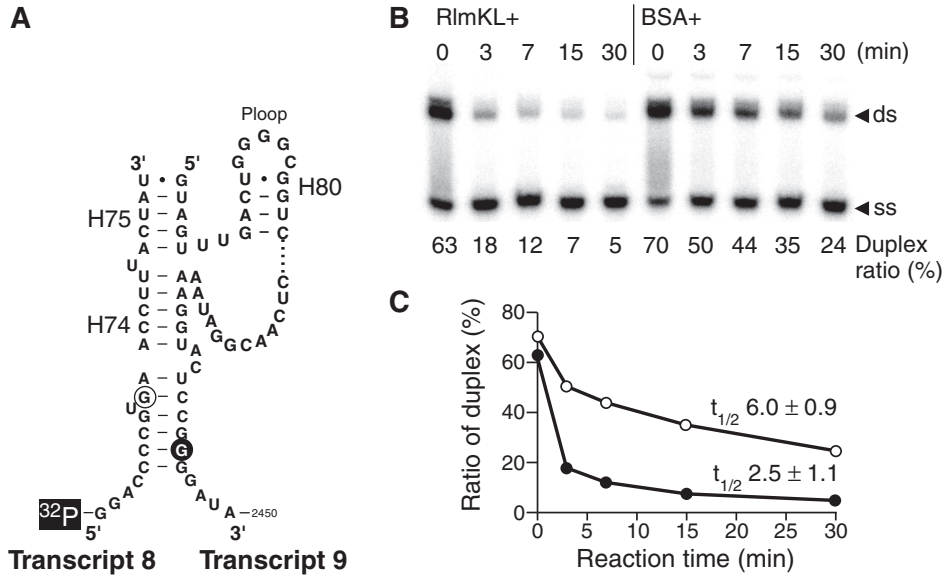

Figure 5. Unwinding activity of RlmKL. (**A**) The RNA substrate used for the unwinding assay. 5′-end of transcript 8 was labeled by ^{32}P. (**B**) The unwinding assay by RlmKL. The duplex and single strands were separated by native PAGE and the radio activity of transcript 8 was visualized and quantified by the imaging analyzer. Duplex ratio (%) described at the bottom of the panel stands for the remaining duplex proportion in total radio activity of transcript 8. The left 5 lanes and the right 5 lanes show time course monitoring of the unwinding activity in the presence of RlmKL (RlmKL$^+$) and BSA (BSA$^+$), respectively. (**C**) Duplex ratio (%) was plotted against reaction time (min). Open and closed circles indicate BSA$^+$ and RlmKL$^+$, respectively. The half-lives ($t_{1/2}$) for both conditions were calculated by non-linear curve fitting and shown as average ±SD; values of three independent experiments.

m^{2}G2445 formation was detected in the base-flipping variant C2072G/G2437C. In addition, A2070G/U2441C, G2067C/C2443G and C2066G/G2444C mutants lacked the m^{2}G2445 formation. The five base-flipping variants, A2071U/U2438A, A2070G/U2441, G2067C/C2443G, C2066G/G2444C, C2064G/G2446C and the base substitution mutant U2068A showed no m^{7}G2069 forming activity. The decreased m^{7}G2069 formation was observed in C2065G/G2445C mutant. No significant change in the methylation of the A2439G and C2440G variants of the internal loop of Helix 74 were observed.

In summary, the sites required for m^{7}G2069 formation and m^{2}G2445 formation are illustrated in Figure 6D. In Helix 80 and the P-loop, G2250, G2251, G2252 and G2253 are critical only for 7-methylation of G2069, and the stem of H80 is also involved in m^{7}G2069 formation. In the 12 nt ss region, C2427, G2429 and A2430 are recognized for both methylations, and especially G2428 and G2429 were critical for N^2-methylation of G2445 and additionally, the residues 2431–2433 are recognized for m^{2}G2445 formation. In Helix 74, C2064/G2446, C2065/G2445, C2066/G2444, G2067/C2443, U2068, A2070/U2441 and A2071/U2438 were essential for 7-methylation of G2069. For m^{2}G formation, duplex formation was not necessary as described above. Therefore, in the 3′ single strand of Helix 74, U2441, C2443 and G2444 were required for N^2-methylation of G2445. The relative orientations of these bases are shown in the crystal structure of the 50S subunit (Figure 6E). These results clearly demonstrated that different sets of residues were involved, respectively, in 7-methylation of G2069 and N^2-methylation of G2445, suggesting that the N-terminal RlmL and C-terminal RlmK domains in RlmKL recognize distinct residues in the substrate.

rlmKL is involved in efficient 50S assembly

The protein product of *ycbY* has been reported to co-sediment with the early assembly intermediate of the 50S ribosomal subunit in *E. coli* (43). In addition, YcbY interacts with the DeaD and SrmB proteins, which are DEAD-box RNA helicases involved in 50S subunit formation (44–46). These results suggested that RlmKL may be involved in an early stage of the biogenesis of the 50S subunit. However, disruption of the *rlmKL* gene showed almost no growth phenotype even when cells were cultured at low temperature (Figure 7A), and no significant change in ribosome profile was detected (Figure 7B). Since YcbY interacts with DeaD and SrmB in *E. coli* (45), we explored the synthetic growth phenotype of the *ΔrlmKL* strain in which RNA helicases, including *deaD*, *srmB*, *rhlE* and *dbpA*, were knocked out. As shown in Figure 7A, we observed a synthetic growth phenotype with *deaD*, in particular at low temperature. When cultured at 37°C, the doubling time of the *ΔrlmKL*/*ΔdeaD* strain (28.4 min) was delayed by 4.5 min relative to the *ΔrlmKL* strain (23.9 min) and 2.9 min relative to the *ΔdeaD* strain (25.5 min). At 28°C, the *ΔrlmKL*/*ΔdeaD* strain exhibited a growth defect resulting in a doubling time of 99.0 min, a 50 min delay relative to the *ΔrlmKL* strain (49.0 min) and a 21.7 min delay relative to the *ΔdeaD* strain (77.3 min). This result suggested that

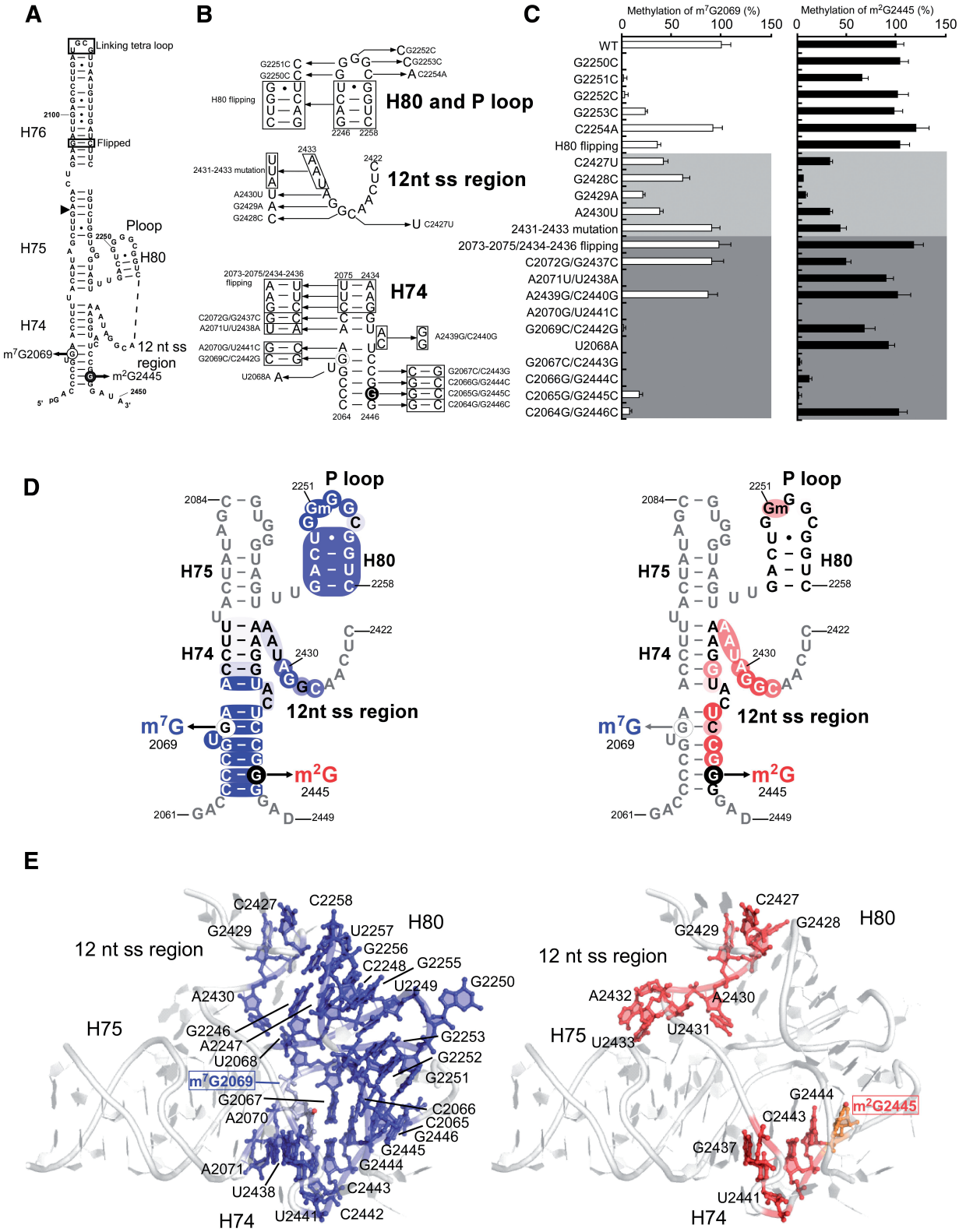

Figure 6. Substrate recognition of RlmKL. (**A**) The secondary structure of transcript 3, which contains base flipping mutation C2096G/G2193C to be cut by MazF at the single position. Closed arrowhead show the cleavage position by MazF. (**B**) Design of mutants based on the transcript 3. Arrows and boxes indicate the substitutions. (**C**) Methylation activities of the transcript 3 variants. The left and right graphs show the end point of m^7G2069 and m^2G2445 formation, respectively. The relative activities of methylation (%) were calculated and normalized by wild-type sequence. (**D**) The recognized residues for m^7G2069 and m^2G2445 formation in transcript 3. The residues whose mutations reduced m^7G2069 or m^2G2445 formation to <50% are depicted as white letters. In the left panel, the residues that are depicted in blue background indicates the m^7G2069 forming activity. In the right panel, the residues that are depicted in red background indicates the m^2G2445 formation activity. Color gradation of blue and red background for each residue represents magnitude of importance for the methylation. (**E**) Crystal structure of Helix 74 and surrounding region in the *E. coli* 50S subunit. Coordinates were obtained from PBD (2aw4) (3). We added the methyl groups for m^7G2069 and m^2G2445, shown in red, respectively. The recognized residues for m^7G2069 and m^2G2445 formation are shown in blue and red, respectively.

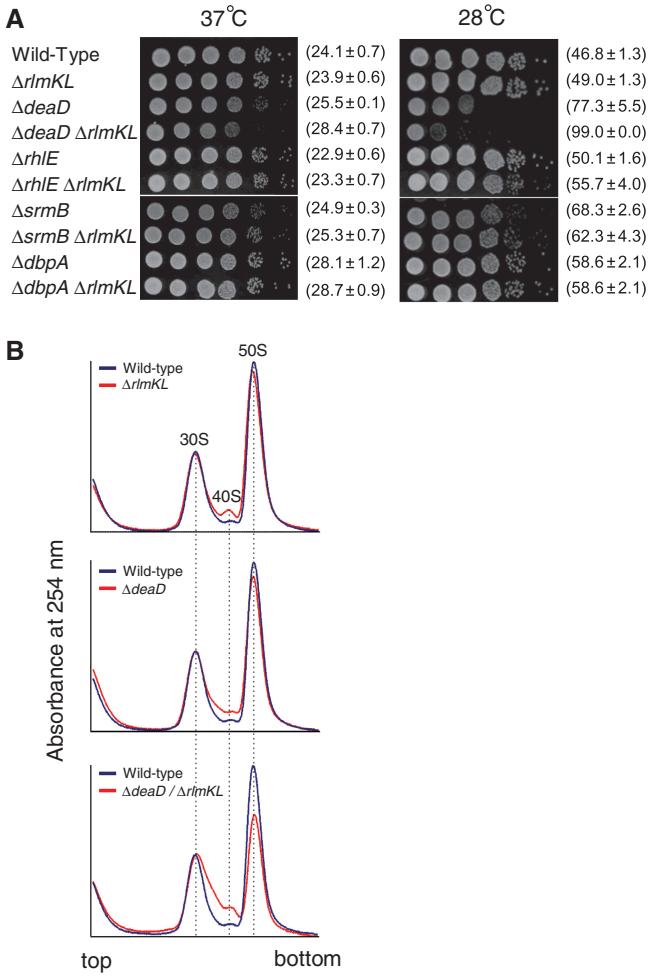

Figure 7. Synthetic growth phenotype of *ΔrlmKL* and ribosomal subunit profiling. (**A**) Exploring the synthetic growth phenotype of *ΔrlmKL* in the presence of the disruption of RNA helicases, including *deaD*, *rhlE*, *srmB* and *dbpA*. Each strain was serially diluted (1:10 dilutions), spotted onto LB plates and incubated at 37°C or 28°C for 11 and 22 h, respectively. The doubling time (in minutes) is shown to the right of each panel. (**B**) SDG profiling of ribosomal subunits in cell lysates in the presence of 0.5 mM Mg^{2+} for the wild-type (blue line), *ΔrlmKL* (red line), *ΔdeaD* (red line) and *ΔrlmKL*/*ΔdeaD* (red line) strains. The amount of each subunit was quantified by UV absorbance at 254 nm. The peak positions for the 30S and 50S subunits are indicated. The peak height was normalized to the 30S peak.

RlmKL and the DeaD RNA helicase are cooperatively involved in efficient assembly of the 50S subunit. Therefore, the ribosome profile of the *ΔrlmKL*/*ΔdeaD* strain was analyzed by sucrose density gradient centrifugation. In the *ΔdeaD* strain (Figure 7B), a slight decrease in the 50S subunit compared with the wild-type strain, and an accumulation of a 40S fraction that contains an intermediate of the 50S subunit, were detected. In the *ΔrlmKL*/*ΔdeaD* strain (Figure 7B), an apparent decrease in the 50S subunit compared with the wild-type and *ΔrlmKL* strains and an accumulation of an intermediate 30–40S fraction were observed. These results indicated that deletion of *rlmKL* in the *ΔdeaD* background caused a considerable defect in 50S assembly.

DISCUSSION

We identified RlmKL as a unique methyltransferase responsible for the formation of both m^2G2445 and m^7G2069 in 23S rRNA. Genetic and biochemical studies revealed that RlmKL is composed of tandem methyltransferase domains, an N-terminal RlmL and C-terminal RlmK. Each domain is a distinct individual methyltransferase responsible for one of the methylations. RlmKL is a novel type of fused methyltransferase bearing dual active sites responsible for different methyl modifications at distinct positions.

In general, the conjugation of enzymes involved in the same pathway has the advantage of facilitating efficient and rapid catalysis of sequential reactions. There are some examples in metabolic pathways for the biogenesis of amino acids and fatty acids (47). Another example of conjugated enzymes can be found in non-ribosomal peptide synthesis in bacteria (48). Among RNA-modifying enzymes, we reported previously that a series of enzymes responsible for wybutosine (yW) synthesis are fused in some cases (32,49).

Another example of a fused methyltransferase for rRNAs is RsmC, the methyltransferase responsible for N^2-methylation of G1207 in 16S rRNA (50). RsmC consists of two methyltransferase domains. The CTD has a *bona fide* methyltransferase that catalyzes m^2G1207 formation using Ado-Met. The NTD has homology to Ado-Met-dependent methyltransferases but has no methyltransferase activity due to mutations in the conserved residues for Ado-Met binding. However, the NTD is required for the stability of the CTD and plays a supportive role in substrate recognition by the CTD (50). In RlmKL, conjugation of RlmL and RlmK is advantageous for facilitating the m^2G2445 formation. In addition, the m^2G2445 formation catalyzed by the RlmL(NTD) can be enhanced by the supportive substrate recognition by the RlmK(CTD), similar to the function of the non-catalytic methyltransferase domain in RsmC.

COG0116 and COG1092 are widely distributed in bacteria and archaea. Due to their sequence similarity, COG0116 and COG1092 can be aligned and placed in the same phylogenetic tree (Supplementary Figure S6). The non-rooted neighbor-joining (NJ) tree shown in Supplementary Figure S6 reveals that COG0116 and COG1092 are clearly separate from each other with 100% bootstrap probability. In this analysis, we found that both COGs were composed of at least two distinct classes of enzymes. COG0116 contains RlmL, as well as Trm14 which is a N^2-methyltransferase for archaeal tRNAs (51), while COG1092 contains RlmK, as well as RlmI, which is responsible for m^5C1962 formation in 23S rRNA (52). The strong similarity between these RNA methyltransferases makes classification of the enzymes complicated. By examining the phylogenetic clusters containing RlmL and RlmK, we can confirm that RlmKL is only present in γ-proteobacteria, and that the separated RlmK and RlmL are found in the genus *Neisseria*, suggesting that the conjugation of these two enzymes should have occurred in the common ancestor of γ-proteobacteria.

According to the phylogenetic analysis of COG1092 (Supplementary Figure S6), RlmK is a paralog of RlmI. Structural analysis revealed that RlmI has three domains, a Rossman-fold domain, an EEHEE domain, and a PUA domain (53). RlmK also has two first domains, the Rossman-fold and EEHEE domains. The Rossman-fold is found in a typical methyltransferase domain containing the Ado-Met binding motif. The EEHEE domain is an RNA recognition module composed of β-sheets and an α-helix, and is found in rRNA methyltransferases including RlmI and RlmD (54). RlmD is the methyltransferase responsible for m^3U1939 formation. According to the crystal structure of the ternary complex consisting of the partial RNA substrate, RlmD and Ado-Met (54), the EEHEE domain binds to the single-stranded RNA 5′ of the target uridine, U1939. Therefore, the EEHEE domain is thought to be a single-stranded RNA binding module that recognizes the sequence immediately 5′ of the target site. Together with our observation that RlmK recognizes residues C2064, C2065 C2066, G2067 and U2068, which are residues immediately 5′ of the target G2069, it is possible that the EEHEE domain of RlmK recognizes the 5′ strand adjacent to G2069 in Helix 74.

RlmL has a THUMP domain (55) in its N-terminal region. The THUMP domain is frequently found in various tRNA-modifying enzymes including ThiI (56), CDAT8 (57), Tan1p (58), Pus10 (59) and PAB1281 (60). The THUMP domain is an RNA-binding domain, and, according to a structural study of ThiI, the THUMP domain together with the ancillary N-terminal domain configures the surface for RNA binding (56). Deleting these two domains of ThiI resulted in the complete loss of tRNA binding (61). These results indicate that the THUMP domain of RlmL may be involved in recognizing G2428 and G2429 in the 12 nt ss region which are apart from the methylated residue G2445.

Based on this study, we propose a catalytic mechanism for the cooperative modifications of Helix 74 (Figure 8). Our biochemical data has shown that RlmK recognizes Helix 80, the conserved bases in the P-loop, 12 nt ss region and Helix74, whereas RlmL recognizes critical residues in the 12 nt ss region and the 3′ strand of Helix 74 (Figure 6D). As shown in the crystal structure of the 50S subunit (Figure 6E), the residues required for m^2G2445 formation are interrupted by the residues essential for m^7G2069 formation. Thus, both domains of RlmKL cannot bind Helix 74 at the same time in the folded structure of the 23S rRNA. In addition, RlmL(NTD) preferentially recognizes and methylates the single-stranded substrate rather than the duplex substrate, suggesting that Helix 74 should be unwound to serve as the preferred substrate for RlmL(NTD) during dimethyl-formation. In fact, we could directly observe an unwinding activity of RlmKL with the duplex substrate formed by transcripts 8 and 9 (Figure 5C), strongly supporting the proposed mechanism of cooperative methylation by RlmKL. As discussed above, RlmK has an EEHEE domain that may recognize the 5′ residues neighboring G2069 in Helix 74. Since several residues in the 5′ strand of Helix 74 are essential for m^7G2069 formation, the EEHEE domain may recognize these bases and unwind Helix 74, releasing the 3′ ss for RlmL(NTD) recognition.

We observed that the *ΔrlmKL/ΔdeaD* strain exhibited a synthetic growth phenotype, especially at low temperatures and had defects in 50S subunit assembly (Figure 7). This result indicated that RlmKL and the DeaD RNA helicase work redundantly in ribosome assembly. As the DeaD RNA helicase possess an ATP-dependent RNA chaperone activity that modulates the local structure of rRNA, RlmKL may act as an assembly factor. RlmKL does not recognize the assembled 50S subunit but uses the naked 23S rRNA as a substrate. In addition, RlmKL co-sediments with the assembly intermediate of the 50S subunit (43), suggesting that RlmKL directly interacts with the 23S rRNA at an early step in 50S assembly. Since RlmKL unwinds Helix 74 during dimethyl formation, the local structural alterations in domain V, induced by RlmKL recognition, may be involved in the correct folding of the rRNA after transcription, domain–domain interactions, and/or hierarchical association of ribosomal proteins in the early steps of 50S assembly. Further study will be required to elucidate the precise role of RlmKL as an assembly factor in ribosome biogenesis.

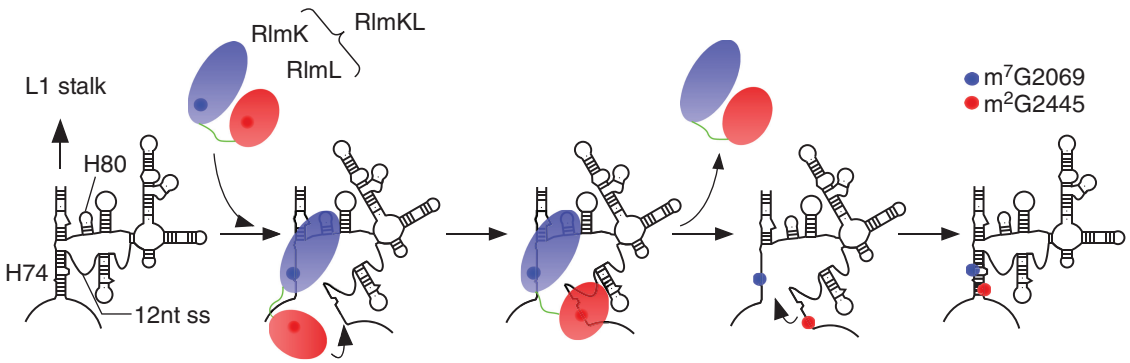

Figure 8. Schematic depiction of a catalytic mechanism for the cooperative modifications of Helix 74. After transcription of domain V in the 23S rRNA, the C-terminal RlmK domain of RlmKL unwinds Helix 74 upon recognizing Helix 80 and Helix 74, and then methylates G2069. The single-stranded 3′ side of Helix 74 is then recognized by the N-terminal RlmL domain, and G2445 is methylated. Finally, Helix 74 reforms a duplex, and domain V is restructured.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1, Supplementary Figures 1–7, Supplementary Methods and Supplementary References [20,38,62–69].

ACKNOWLEDGEMENTS

We are grateful to the Suzuki lab members for many fruitful discussions and technical advice. We also thank J. Kato (Tokyo Metropolitan University) and NBRP-*E. coli* at National Institute of Genetics (NIG) for providing materials.

FUNDING

Grants-in-aid for scientific research on priority areas from the Ministry of Education, Science, Sports and Culture of Japan (to T.S.); JSPS Fellowship for Japanese Junior Scientists (to S.K.); the New Energy and Industrial Technology Development Organization (NEDO) (grant to T.S.). Funding for the open access charge: Japan Ministry of Education, Science, Sports and Culture.

Conflict of interest statement. None declared.

REFERENCES

1. Ban,N., Nissen,P., Hansen,J., Moore,P.B. and Steitz,T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science*, **289**, 905–920.
2. Wimberly,B.T., Brodersen,D.E., Clemons,W.M. Jr, Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
3. Schuwirth,B.S., Borovinskaya,M.A., Hau,C.W., Zhang,W., Vila-Sanjurjo,A., Holton,J.M. and Cate,J.H. (2005) Structures of the bacterial ribosome at 3.5 A resolution. *Science*, **310**, 827–834.
4. Selmer,M., Dunham,C.M., Murphy,F.V.t., Weixlbaumer,A., Petry,S., Kelley,A.C., Weir,J.R. and Ramakrishnan,V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science*, **313**, 1935–1942.
5. Korostelev,A., Trakhanov,S., Laurberg,M. and Noller,H.F. (2006) Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell*, **126**, 1065–1077.
6. Schluenzen,F., Tocilj,A., Zarivach,R., Harms,J., Gluehmann,M., Janell,D., Bashan,A., Bartels,H., Agmon,I., Franceschi,F. *et al.* (2000) Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*, **102**, 615–623.
7. Chow,C.S., Lamichhane,T.N. and Mahto,S.K. (2007) Expanding the nucleotide repertoire of the ribosome with post-transcriptional modifications. *ACS Chem. Biol.*, **2**, 610–619.
8. Lapeyre,B. (2005) Conserved ribosomal RNA modification and their putative roles in ribosome biogenesis and translation. In: Grosjean,H. (ed.), *Fine-Tuning of RNA Functions by Modification and Editing, Topics in Current Genetics*, Vol. 12. Springer, NY, pp. 263–284.
9. Yu,Y., Terns,M.R. and Terns,P.M. (2005) Mechanisms and functions of RNA-guided RNA modification. In: Grosjean,H. (ed.), *Fine-Tuning of RNA Functions by Modification and Editing, Topics in Current Genetics*, Vol. 12. Springer, Berlin Heidelberg.
10. Brimacombe,R., Mitchell,P., Osswald,M., Stade,K. and Bochkariov,D. (1993) Clustering of modified nucleotides at the functional center of bacterial ribosomal RNA. *FASEB J.*, **7**, 161–167.
11. Decatur,W.A. and Fournier,M.J. (2002) rRNA modifications and ribosome function. *Trends Biochem. Sci.*, **27**, 344–351.
12. Purta,E., O'Connor,M., Bujnicki,J.M. and Douthwaite,S. (2009) YgdE is the 2′-O-ribose methyltransferase RlmM specific for nucleotide C2498 in bacterial 23S rRNA. *Mol. Microbiol.*, **72**, 1147–1158.
13. Liang,X.H., Liu,Q. and Fournier,M.J. (2007) rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol. Cell*, **28**, 965–977.
14. King,T.H., Liu,B., McCully,R.R. and Fournier,M.J. (2003) Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol. Cell*, **11**, 425–435.
15. Havelund,J.F., Giessing,A.M., Hansen,T., Rasmussen,A., Scott,L.G. and Kirpekar,F. (2011) Identification of 5-hydroxycytidine at position 2501 concludes characterization of modified nucleotides in E. coli 23S rRNA. *J. Mol. Biol.*, **411**, 529–536.
16. Liang,X.H., Liu,Q. and Fournier,M.J. (2009) Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA*, **15**, 1716–1728.
17. Baudin-Baillieu,A., Fabret,C., Liang,X.H., Piekna-Przybylska,D., Fournier,M.J. and Rousset,J.P. (2009) Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res.*, **37**, 7665–7677.
18. Ejby,M., Sorensen,M.A. and Pedersen,S. (2007) Pseudouridylation of helix 69 of 23S rRNA is necessary for an effective translation termination. *Proc. Natl Acad. Sci. USA*, **104**, 19410–19415.
19. Das,G., Thotala,D.K., Kapoor,S., Karunanithi,S., Thakur,S.S., Singh,N.S. and Varshney,U. (2008) Role of 16S ribosomal RNA methylations in translation initiation in Escherichia coli. *EMBO J.*, **27**, 840–851.
20. Kimura,S. and Suzuki,T. (2010) Fine-tuning of the ribosomal decoding center by conserved methyl-modifications in the Escherichia coli 16S rRNA. *Nucleic Acids Res.*, **38**, 1341–1352.
21. Gutgsell,N.S., Deutscher,M.P. and Ofengand,J. (2005) The pseudouridine synthase RluD is required for normal ribosome assembly and function in *Escherichia coli*. *RNA*, **11**, 1141–1152.
22. Bugl,H., Fauman,E.B., Staker,B.L., Zheng,F., Kushner,S.R., Saper,M.A., Bardwell,J.C. and Jakob,U. (2000) RNA methylation under heat shock control. *Mol. Cell*, **6**, 349–360.
23. Caldas,T., Binet,E., Bouloc,P., Costa,A., Desgres,J. and Richarme,G. (2000) The FtsJ/RrmJ heat shock protein of *Escherichia coli* is a 23S ribosomal RNA methyltransferase. *J. Biol. Chem.*, **275**, 16414–16419.
24. 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.
25. Wilson,D.N. and Nierhaus,K.H. (2007) The weird and wonderful world of bacterial ribosome regulation. *Crit. Rev. Biochem. Mol. Biol.*, **42**, 187–219.
26. Poehlsgaard,J. and Douthwaite,S. (2005) The bacterial ribosome as a target for antibiotics. *Nat. Rev. Microbiol.*, **3**, 870–881.
27. Toh,S.M. and Mankin,A.S. (2008) An indigenous posttranscriptional modification in the ribosomal peptidyl transferase center confers resistance to an array of protein synthesis inhibitors. *J. Mol. Biol.*, **380**, 593–597.
28. Suzuki,T., Ikeuchi,Y., Noma,A., Suzuki,T. and Sakaguchi,Y. (2007) Mass spectrometric identification and characterization of RNA-modifying enzymes. *Methods Enzymol.*, **425**, 211–229.
29. Ikeuchi,Y., Shigi,N., Kato,J., Nishimura,A. and Suzuki,T. (2006) Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thiouridine biosynthesis at tRNA wobble positions. *Mol. Cell*, **21**, 97–108.
30. Ikeuchi,Y., Kitahara,K. and Suzuki,T. (2008) The RNA acetyltransferase driven by ATP hydrolysis synthesizes N4-acetylcytidine of tRNA anticodon. *EMBO J.*, **27**, 2194–2203.
31. Soma,A., Ikeuchi,Y., Kanemasa,S., Kobayashi,K., Ogasawara,N., Ote,T., Kato,J., Watanabe,K., Sekine,Y. and Suzuki,T. (2003) An RNA-modifying enzyme that governs both the codon and

amino acid specificities of isoleucine tRNA. *Mol. Cell*, **12**, 689–698.
32. Noma,A., Kirino,Y., Ikeuchi,Y. and Suzuki,T. (2006) Biosynthesis of wybutosine, a hyper-modified nucleoside in eukaryotic phenylalanine tRNA. *EMBO J.*, **25**, 2142–2154.
33. Noma,A., Sakaguchi,Y. and Suzuki,T. (2009) Mechanistic characterization of the sulfur-relay system for eukaryotic 2-thiouridine biogenesis at tRNA wobble positions. *Nucleic Acids Res.*, **37**, 1335–1352.
34. Lesnyak,D.V., Sergiev,P.V., Bogdanov,A.A. and Dontsova,O.A. (2006) Identification of Escherichia coli m2G methyltransferases: I. the ycbY gene encodes a methyltransferase specific for G2445 of the 23 S rRNA. *J. Mol. Biol.*, **364**, 20–25.
35. Kitahara,K. and Suzuki,T. (2009) The ordered transcription of RNA domains is not essential for ribosome biogenesis in *Escherichia coli*. *Mol. Cell*, **34**, 760–766.
36. Diges,C.M. and Uhlenbeck,O.C. (2005) *Escherichia coli* DbpA is a 3′ –> 5′ RNA helicase. *Biochemistry*, **44**, 7903–7911.
37. Bizebard,T., Ferlenghi,I., Iost,I. and Dreyfus,M. (2004) Studies on three E. coli DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. *Biochemistry*, **43**, 7857–7866.
38. Hashimoto,M., Ichimura,T., Mizoguchi,H., Tanaka,K., Fujimitsu,K., Keyamura,K., Ote,T., Yamakawa,T., Yamazaki,Y., Mori,H. *et al.* (2005) Cell size and nucleoid organization of engineered *Escherichia coli* cells with a reduced genome. *Mol. Microbiol.*, **55**, 137–149.
39. Andersen,N.M. and Douthwaite,S. (2006) YebU is a m5C methyltransferase specific for 16 S rRNA nucleotide 1407. *J. Mol. Biol.*, **359**, 777–786.
40. Ofengand,J. and Del Campo,M. (2004) Modified Nucleosides of Escherichia coli Ribosomal RNA. In: Curtiss,R. (ed.), *Escherichia coli and Salmonella: Cellular and Molecular Biology*. ASM Press, Washington, DC.
41. Mcluckey,S.A., Vanberkel,G.J. and Glish,G.L. (1992) Tandem mass-spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectr.*, **3**, 60–70.
42. Tatusov,R.L., Koonin,E.V. and Lipman,D.J. (1997) A genomic perspective on protein families. *Science*, **278**, 631–637.
43. Jiang,M., Sullivan,S.M., Walker,A.K., Strahler,J.R., Andrews,P.C. and Maddock,J.R. (2007) Identification of novel Escherichia coli ribosome-associated proteins using isobaric tags and multidimensional protein identification techniques. *J. Bacteriol.*, **189**, 3434–3444.
44. Charollais,J., Dreyfus,M. and Iost,I. (2004) CsdA, a cold-shock RNA helicase from *Escherichia coli*, is involved in the biogenesis of 50S ribosomal subunit. *Nucleic Acids Res.*, **32**, 2751–2759.
45. Butland,G., Krogan,N.J., Xu,J., Yang,W.H., Aoki,H., Li,J.S., Krogan,N., Menendez,J., Cagney,G., Kiani,G.C. *et al.* (2007) Investigating the in vivo activity of the DeaD protein using protein-protein interactions and the translational activity of structured chloramphenicol acetyltransferase mRNAs. *J. Cell Biochem.*, **100**, 642–652.
46. Charollais,J., Pflieger,D., Vinh,J., Dreyfus,M. and Iost,I. (2003) The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli. *Mol. Microbiol.*, **48**, 1253–1265.
47. Srere,P.A. (1987) Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem.*, **56**, 89–124.
48. Stein,T. and Vater,J. (1996) Amino acid activation and polymerization at modular multienzymes in nonribosomal peptide biosynthesis. *Amino acids*, **10**, 201–227.
49. Noma,A., Ishitani,R., Kato,M., Nagao,A., Nureki,O. and Suzuki,T. (2010) Expanding role of the Jumonji C domain as an RNA hydroxylase. *J. Biol. Chem.*, **285**, 34503–34507.
50. Sunita,S., Purta,E., Durawa,M., Tkaczuk,K.L., Swaathi,J., Bujnicki,J.M. and Sivaraman,J. (2007) Functional specialization of domains tandemly duplicated within 16S rRNA methyltransferase RsmC. *Nucleic Acids Res.*, **35**, 4264–4274.
51. Menezes,S., Gaston,K.W., Krivos,K.L., Apolinario,E.E., Reich,N.O., Sowers,K.R., Limbach,P.A. and Perona,J.J. (2011) Formation of m2G6 in Methanocaldococcus jannaschii tRNA catalyzed by the novel methyltransferase Trm14. *Nucleic Acids Res.*, **39**, 7641–7655.
52. Purta,E., O'Connor,M., Bujnicki,J.M. and Douthwaite,S. (2008) YccW is the m5C methyltransferase specific for 23S rRNA nucleotide 1962. *J. Mol. Biol.*, **383**, 641–651.
53. Sunita,S., Tkaczuk,K.L., Purta,E., Kasprzak,J.M., Douthwaite,S., Bujnicki,J.M. and Sivaraman,J. (2008) Crystal structure of the Escherichia coli 23S rRNA:m5C methyltransferase RlmI (YccW) reveals evolutionary links between RNA modification enzymes. *J. Mol. Biol.*, **383**, 652–666.
54. Lee,T.T., Agarwalla,S. and Stroud,R.M. (2005) A unique RNA Fold in the RumA-RNA-cofactor ternary complex contributes to substrate selectivity and enzymatic function. *Cell*, **120**, 599–611.
55. Aravind,L. and Koonin,E.V. (2001) THUMP–a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases and RNA methylases. *Trends Biochem. Sci.*, **26**, 215–217.
56. Waterman,D.G., Ortiz-Lombardia,M., Fogg,M.J., Koonin,E.V. and Antson,A.A. (2006) Crystal structure of Bacillus anthracis ThiI, a tRNA-modifying enzyme containing the predicted RNA-binding THUMP domain. *J. Mol. Biol.*, **356**, 97–110.
57. Randau,L., Stanley,B.J., Kohlway,A., Mechta,S., Xiong,Y. and Soll,D. (2009) A cytidine deaminase edits C to U in transfer RNAs in Archaea. *Science*, **324**, 657–659.
58. Johansson,M.J. and Bystrom,A.S. (2004) The Saccharomyces cerevisiae TAN1 gene is required for N4-acetylcytidine formation in tRNA. *RNA*, **10**, 712–719.
59. McCleverty,C.J., Hornsby,M., Spraggon,G. and Kreusch,A. (2007) Crystal structure of human Pus10, a novel pseudouridine synthase. *J. Mol. Biol.*, **373**, 1243–1254.
60. Armengaud,J., Urbonavicius,J., Fernandez,B., Chaussinand,G., Bujnicki,J.M. and Grosjean,H. (2004) N2-methylation of guanosine at position 10 in tRNA is catalyzed by a THUMP domain-containing, S-adenosylmethionine-dependent methyltransferase, conserved in Archaea and Eukaryota. *J. Biol. Chem.*, **279**, 37142–37152.
61. Tanaka,Y., Yamagata,S., Kitago,Y., Yamada,Y., Chimnaronk,S., Yao,M. and Tanaka,I. (2009) Deduced RNA binding mechanism of ThiI based on structural and binding analyses of a minimal RNA ligand. *RNA*, **15**, 1498–1506.
62. Spedding,G. (1990) *Isolation and Analysis of Ribosomes from Prokaryotes, Eukaryotes, and Organelles, Ribosomes and Protein Synthesis, A Practical Approach*. Oxford University Press, New York.
63. Hanada,T., Suzuki,T., Yokogawa,T., Takemoto-Hori,C., Sprinzl,M. and Watanabe,K. (2001) Translation ability of mitochondrial tRNAsSer with unusual secondary structures in an in vitro translation system of bovine mitochondria. *Genes Cells*, **6**, 1019–1030.
64. Sampson,J.R. and Uhlenbeck,O.C. (1988) Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. *Proc. Natl Acad. Sci. USA*, **85**, 1033–1037.
65. Sato,N.S., Hirabayashi,N., Agmon,I., Yonath,A. and Suzuki,T. (2006) Comprehensive genetic selection revealed essential bases in the peptidyl-transferase center. *Proc. Natl Acad. Sci. USA*, **103**, 15386–15391.
66. Kitagawa,M., Ara,T., Arifuzzaman,M., Ioka-Nakamichi,T., Inamoto,E., Toyonaga,H. and Mori,H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA Res.*, **12**, 291–299.
67. Lesnyak,D.V., Osipiuk,J., Skarina,T., Sergiev,P.V., Bogdanov,A.A., Edwards,A., Savchenko,A., Joachimiak,A. and Dontsova,O.A. (2007) Methyltransferase that modifies guanine 966 of the 16 S rRNA: functional identification and tertiary structure. *J. Biol. Chem.*, **282**, 5880–5887.
68. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L. and Mori,H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006 0008.
69. 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. USA*, **97**, 6640–6645.