Published online 17 December 2015

Nucleic Acids Research, 2016, Vol. 44, No. 2 ***509–523***
doi: 10.1093/nar/gkv1470

NAR Breakthrough Article

Biogenesis and growth phase-dependent alteration of 5-methoxycarbonylmethoxyuridine in tRNA anticodons

Yusuke Sakai, Kenjyo Miyauchi, Satoshi Kimura 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 November 09, 2015; Revised November 30, 2015; Accepted December 02, 2015

ABSTRACT

Post-transcriptional modifications at the anticodon first (wobble) position of tRNA play critical roles in precise decoding of genetic codes. 5-carboxymethoxyuridine (cmo^5U) and its methyl ester derivative 5-methoxycarbonylmethoxyuridine ($mcmo^5U$) are modified nucleosides found at the anticodon wobble position in several tRNAs from Gram-negative bacteria. cmo^5U and $mcmo^5U$ facilitate non-Watson–Crick base pairing with guanosine and pyrimidines at the third positions of codons, thereby expanding decoding capabilities. By mass spectrometric analyses of individual tRNAs and a shotgun approach of total RNA from *Escherichia coli*, we identified $mcmo^5U$ as a major modification in $tRNA^{Ala1}$, $tRNA^{Ser1}$, $tRNA^{Pro3}$ and $tRNA^{Thr4}$; by contrast, cmo^5U was present primarily in $tRNA^{Leu3}$ and $tRNA^{Val1}$. In addition, we discovered 5-methoxycarbonylmethoxy-2′-*O*-methyluridine ($mcmo^5Um$) as a novel but minor modification in $tRNA^{Ser1}$. Terminal methylation frequency of $mcmo^5U$ in $tRNA^{Pro3}$ was low (≈30%) in the early log phase of cell growth, gradually increased as growth proceeded and reached nearly 100% in late log and stationary phases. We identified CmoM (previously known as SmtA), an AdoMet-dependent methyltransferase that methylates cmo^5U to form $mcmo^5U$. A luciferase reporter assay based on a +1 frameshift construct revealed that terminal methylation of $mcmo^5U$ contributes to the decoding ability of $tRNA^{Ala1}$.

INTRODUCTION

RNA molecules are frequently modified post-transcriptionally on their nucleobase and ribose moieties. These modifications carry qualitative information embedded in RNA molecules associated with various biological processes. To date, more than 100 species of modifications have been identified in various RNAs from all domains of life (1), the majority of which are found in tRNAs. tRNA modifications play critical roles in decoding properties and stabilization of tertiary structure (2,3). In particular, a wide variety of modifications occur at the first (wobble) position of the anticodon. These modifications play pivotal roles in modulating codon recognition and ensuring accurate translation of the genetic code (4).

In the classical wobble hypothesis proposed by Crick (5), uridine (U34) at the wobble position pairs with A and G at the third letter of the codon. In the decoding system of Mycoplasma species and mitochondria (6–9), however, U34 can recognize any of the four bases in a family box due to its conformational flexibility, a phenomenon termed 'four-way wobbling' (4), although efficiency of this type of wobbling strongly depends on the second and third letters of codons. To restrict decoding capability, tRNAs responsible for two codon sets ending in a purine (NNR) often contain 5-methyl-(2-thio)uridine derivatives [$xm^5(s^2)U$] at the wobble position (3,4), including 5-carboxymethylaminomethyl-(2-thio)uridine [$cmnm^5(s^2)U$] and 5-methylaminomethyl-(2-thio)uridine [$mnm^5(s^2)U$] in bacterial tRNAs, 5-methoxycarbonylmethyl-(2-thio)uridine [$mcm^5(s^2)U$] and its derivatives in eukaryotic cytoplasmic tRNAs and 5-taurinomethyl-(2-thio)uridine [$\tau m^5(s^2)U$] in mitochondrial tRNAs. Due to the conformational rigidity of xm^5s^2U modifications, which are largely fixed in the C3′-*endo* ribose pucker conformation (10), xm^5s^2U prefers to base-pair with A and G, thus preventing misreading of near-cognate

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

codons ending in pyrimidine (NNY) (2,11). In addition, C5-substituent of xm^{5}U modification plays a critical role in stabilizing U-G wobble pairing at the A-site of ribosome (12,13).

By contrast, to expand decoding capacity in most bacterial species, 5-hydroxyuridine derivatives (xo^{5}U) are present at the wobble position of tRNAs responsible for recognizing family boxes. 5-carboxymethoxyuridine (cmo^{5}U, also called uridine-5-oxy acetic acid) (Figure 1A) is present in tRNAs from Gram-negative bacteria including *Escherichia coli* and *Salmonella enterica* (3,4,14). cmo^{5}U was first reported as a minor nucleoside at the wobble position of *E. coli* tRNA^{Val1} (15,16). The chemical structure of cmo^{5}U was determined in 1970 (17). Subsequently, cmo^{5}U was also found in *E. coli* tRNA^{Ser1} (18), tRNA^{Ala1} (19) and tRNA^{Leu3} (20). In addition, cmo^{5}U is present in tRNA^{Pro3} from *S. enterica* (serovar Typhimurium)(21). The presence of the methyl ester derivative of cmo^{5}U, 5-methoxycarbonylmethoxyuridine (mcmo^{5}U, also called uridine-5-oxy acetic acid methyl ester) (Figure 1A) was predicted by a series of *in vitro* studies (22–24). In Gram-positive bacteria, 5-methoxyuridine (mo^{5}U) is present in tRNA^{Thr} from *Bacillus subtilis* (25). Collectively, these findings suggested that cmo^{5}U or mcmo^{5}U are present at the anticodon wobble position in *E. coli* tRNA^{Val1}, tRNA^{Ala1}, tRNA^{Ser1}, tRNA^{Thr4}, tRNA^{Pro3} and tRNA^{Leu3} (14,20). However, the exact frequency of these modifications in each tRNA has not been determined.

cmo^{5}U and mcmo^{5}U enable non-Watson–Crick base pairing with guanosines and pyrimidines at the third positions of codons, thereby expanding decoding capability. *In vitro* studies revealed that cmo^{5}U at wobble position allows tRNAs to recognize codons ending in G or U (18,26–29). In addition, tRNAs with cmo^{5}U are able to read codons ending in C in the family boxes for Ala, Pro and Val *in vivo* (30–32). Solution structure of cmo^{5}U nucleoside analyzed by NMR prefers to adopt the C2′-*endo* ribose pucker conformation, providing a mechanistic insight into base pairing between cmo^{5}U and pyrimidines (10). Such base pairing at the ribosomal A-site was visualized in a crystal structure of the 30S ribosome in complex with an anticodon stem-loop (ASL) bearing cmo^{5}U and its cognate codon (33). cmo^{5}U forms an intramolecular hydrogen bond that pre-structures the anticodon loop, enabling cmo^{5}U to pair with all four bases at the wobble position in the mRNA. Unexpectedly, cmo^{5}U adopts the C3′-*endo* form and pairs with G at the third letter of codon by the standard Watson–Crick geometry, rather than classical U-G wobble geometry, indicating that the enol form of the uracil base is involved in this base pairing interaction (33).

A pathway consisting of multiple enzymatic reactions has been proposed for biogenesis of cmo^{5}U and mcmo^{5}U (see Figure 7). Björk and colleagues made a primary contribution to characterizing this pathway (3,14). cmo^{5}U biogenesis requires chorismate, an end product of the shikimate pathway, which is involved in biosynthesis of aromatic amino acids and vitamins (34,35). First, U34 is hydroxylated to form 5-hydroxyuridine (ho^{5}U) via an unknown pathway. One carbon atom of the 5-carboxymethoxy-group in cmo^{5}U is derived from methionine, indicating the involvement of AdoMet methyltransferase in this step (35). Two genes, *cmoA* and *cmoB*, both of which have AdoMet binding motifs, are responsible for formation of cmo^{5}U (31). Subsequent structural analyses revealed that CmoA contains a novel derivative of AdoMet, *S*-adenosyl-*S*-carboxymethyl-L-homocysteine (SCM-SAH or Cx-SAM) (36,37). Kim and Almo's group successfully reconstituted cmo^{5}U formation in tRNA *in vitro* using recombinant CmoA and CmoB in the presence of AdoMet and prephenate (a derivative of chorismate) (37). In this process, CmoA first synthesizes SCM-SAH from AdoMet and prephenate. Next, CmoB employs SCM-SAH as a substrate to transfer its carboxymethyl-group to the 5-hydroxy group of ho^{5}U, thereby generating cmo^{5}U on tRNA (see Figure 8). In some tRNAs, cmo^{5}U is further methylated to form mcmo^{5}U. Biochemical studies in *S. enterica* cell lysate revealed that cmo^{5}U in tRNA^{Ser1} and tRNA^{Ala1} is further methylated by an unidentified AdoMet-dependent methyltransferase, whereas cmo^{5}U in tRNA^{Val1} is not (22). Initially, the cmo^{5}U methyltransferase was predicted to be encoded by *supK* (23); however, *supK* was subsequently identified as *prfB* which encodes release factor 2 (RF2) (38). CmoA was also speculated to use its second function to methylate cmo^{5}U to generate mcmo^{5}U (31). However, it is now clear that CmoA is responsible for catalyzing SCM-SAH formation (37). Thus, the gene responsible for the cmo^{5}U methyltransferase remains to be identified.

In this study, we isolated individual tRNAs from *E. coli* and analyzed the modification status of each tRNA by mass spectrometry. We found that mcmo^{5}U was present as a major modification in tRNA^{Ala1}, tRNA^{Ser1}, tRNA^{Pro3} and tRNA^{Thr4}, whereas cmo^{5}U is primarily present in tRNA^{Leu3} and tRNA^{Val1}. In addition, we discovered 5-methoxycarbonylmethoxy-2′-*O*-methyluridine (mcmo^{5}Um) (Figure 1A) as a novel derivative of mcmo^{5}U in tRNA^{Ser1}. Frequency of terminal methylation of mcmo^{5}U in tRNA^{Pro3} was dependent on growth phase. Moreover, we identified the cmo^{5}U methyltransferase, which we named CmoM, that methylates cmo^{5}U to form mcmo^{5}U in the presence of AdoMet. This terminal methylation of mcmo^{5}U contributes to the decoding ability of tRNA^{Ala1}.

MATERIALS AND METHODS

Strains and plasmid construction

The *E. coli* deletion strains Δ*cmoB*::Km^r, Δ*trmL*::Km^r, Δ*cmoM* (Δ*smtA*)::Km^r, and their parental strain BW25113, were obtained from the National BioResource Project (NBRP), National Institute of Genetics, Japan (39). All strains were cultured in LB media at 37°C.

To generate a vector for expression of recombinant CmoM with an N-terminal His6-tag, *cmoM* was polymerase chain reaction (PCR)-amplified from the BW25113 genome and inserted into the *Nde*I/*Not*I site of pET-28a (Novagen) to yield pET-cmoM-N-His6. To complement Δ*cmoM*, *cmoM* with its 5′ flanking region (including the promoter) was PCR-amplified and cloned into the low-copy plasmid pMW118 (Nippon Gene) to yield pMW−cmoM (psmtA). Point mutations were introduced in pMW-cmoM by QuikChange^{TM} site-directed mutagenesis (Agilent Technologies). For dual-luciferase reporters, the fusion gene

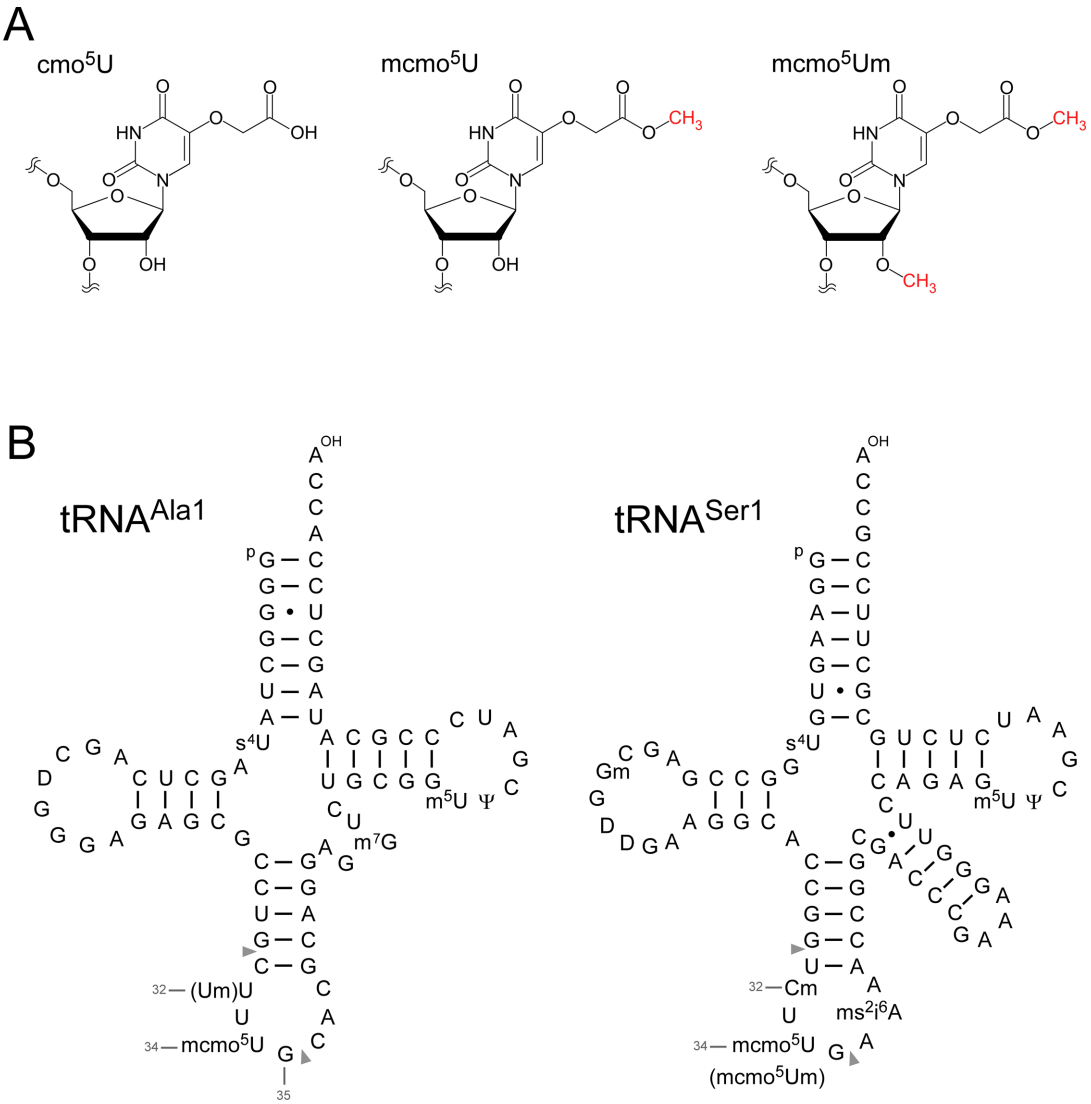

Figure 1. 5-carboxymethoxyuridine (cmo^{5}U) and *E. coli* tRNAs. (**A**) Chemical structures of 5-carboxymethoxyuridine (cmo^{5}U, left), 5-methoxycarbonylmethoxyuridine (mcmo^{5}U, center) and 5-methoxycarbonylmethoxy-2′-*O*-methyluridine (mcmo^{5}Um, right). (**B**) Secondary structures of *E. coli* tRNA^{Ala1} (left) and tRNA^{Ser1} (right) with post-transcriptional modifications: 4-thiouridine (s^{4}U), 2′-*O*-methylguanosine (Gm), dihydrouridine (D), 2′-*O*-methylcytidine (Cm), 2′-*O*-methyluridine (Um), 5-methoxycarbonylmethoxyuridine (mcmo^{5}U), 5-methoxycarbonylmethoxy-2′-*O*-methyluridine (mcmo^{5}Um), 2-methylthio-*N*^{6}-isopentenyladenosine (ms^{2}i^{6}A), 7-methylguanosine (m^{7}G), 5-methyluridine (m^{5}U) and pseudouridine (Ψ). The position numbers of the residues (gray letters) are displayed according to the nucleotide numbering system (64). Pairs of gray triangles indicate the positions of cleavage by RNase T$_1$ that generate RNA fragments containing the wobble positions. The sequence of *E. coli* tRNA^{Ala1} is the same as that of tRNA^{Ala1B} (65).

of *Renilla* and firefly luciferases was PCR-amplified from pQE-Luc(+1) (40) using primers containing *Nco*I and *Xho*I sites, and then inserted into the corresponding site of pBAD/*Myc*-His (Invitrogen) to yield pBAD-RFLuc. Subsequently, the RF2 recoding site (32,41), including an SD-like sequence and GCG as a test codon, was introduced by PCR into the linker region between the two luciferases to yield pBAD-RFLucGCG. Two variants in which the test codon was replaced by UCG (pBAD-RFLucUCG) and GG (pBAD-RFLucGG) were generated by QuikChange™ site-directed mutagenesis. All constructs used in this study were verified by Sanger sequencing. The primers used in this study are listed in Supplementary Table S1.

RNA extraction and tRNA isolation

Total RNA from each *E. coli* strain was extracted by phenol in acidic condition (42). Individual tRNAs were isolated by reciprocal circulating chromatography, as described previously (42,43). The 5′-terminal ethylcarbamate amino-modified DNA probes used in this method are listed in Supplementary Table S1.

Mass spectrometry of tRNA modifications

For RNA fragment analysis, isolated tRNA (1.25 pmol) was digested at 37°C for 1 h in 12.5 μl of a solution containing 20 mM NH_4OAc (pH 5.3) and 125 U RNase T$_1$. The digested RNA was mixed with 12.5 μl of 0.1 M triethylamine-acetate

(pH 7.0), and 10 μl of the digest was analyzed by capillary liquid chromatography (LC) coupled to nano electrospray (ESI)/mass spectrometry (MS) on a linear ion trap-Orbitrap hybrid mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific) as described (42,44).

Total nucleosides were analyzed basically as described (44,45). Isolated $tRNA^{Ser1}$ (110 pmol) was digested at 37°C for 3 h in 15 μl of a solution containing 20 mM trimethylamine-HCl (TMA-HCl) (pH 7.0), 0.05 U of nuclease P1 and 0.1 U of BAP. The digested RNA (100 pmol) was adjusted as 50 μl of 90% acetonitrile and subjected to HILIC/ESI-MS using Q Exactive™ Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific) equipped with a Dionex UltiMate™ 3000 LC System (Thermo Fisher Scientific) using a ZIC-cHILIC column (3 μm, 2.1 × 150 mm, Merck Millipore). For ribonucleome analysis, a reverse genetic approach combined with RNA-MS (44), nucleosides of total RNAs obtained from knockout strains were subjected to RPC-ESI/MS using LCQ Advantage ion-trap mass spectrometer (Thermo Fisher Scientific) equipped with an HP1100 liquid chromatography system (Agilent Technologies) using an Inertsil ODS-3 column (2.1 × 250 mm, GL science).

Shotgun analysis of tRNA fragments by RNA-MS

Total RNA of each strain cultured in 2–20 ml of LB medium was extracted with TRIzol™ (Life Technologies). Then, 50–250 μg of total RNA was dissolved in 800 μl of 3 M NH_4OAc (pH 5.3), mixed with 640 μl (0.8 vol) of isopropanol at room temperature and centrifuged at 15 000 rpm for 10 min to precipitate long RNAs including rRNAs. The supernatant was collected and precipitated with ethanol. Depletion of rRNA was verified by denaturing PAGE. The resultant small RNA fraction (50 ng) was digested for 1 h at 37°C with RNase T_1 (125 U) in 20 mM NH_4OAc (pH 5.3), and the digested RNA (20 ng) was subjected to capillary LC coupled to nanoESI/MS as described above.

Mutation study of CmoM by genetic complementation

The *E. coli* Δ*cmoM* strain was transformed with psmtA, its mutant derivatives or an empty vector (pMW118), and the functional importance of the mutations was assessed by monitoring restoration of $mcmo^5U$ in the transformants. Transformants were cultured overnight at 37°C in 2 ml of LB medium containing 100 μg/ml ampicillin. Total RNA was subjected to the shotgun analysis as described above to detect the anticodon-containing fragment of $tRNA^{Pro3}$. Modification frequency was determined by calculating the intensity ratio of mass chromatograms between UmUcmo^{5}UGp (m/z 683.567, z = 2) and UmUmcmo^{5}UGp (m/z 690.575, z = 2).

Preparation of recombinant protein

E. coli BL21 (DE3) transformed with pET-cmoM-N-His6 was cultured at 37°C to an OD_{600} of ≈0.7, supplemented with 0.1 mM IPTG and cultured at 37°C for an additional 4 h. Cells were harvested and disrupted by sonication in lysis buffer consisting of 50 mM HEPES-KOH (pH 7.5), 100 mM KCl, 10 mM $MgCl_2$, 1 mM dithiothreitol and 0.2 mM PMSF. Purification of recombinant protein was performed basically as described (42,46) using a HisTrap column (GE Healthcare) with a linear gradient of imidazole (25–500 mM). The purified protein was dialyzed in lysis buffer, added with glycerol to f.c. 30% and stored at –20°C.

In vitro reconstitution of $mcmo^5U$

$tRNA^{Ser1}$, a substrate for CmoM, was isolated from the Δ*cmoM* strain. *In vitro* methylation was performed for 1 h at 37°C in a 10 μl reaction mixture containing 50 mM HEPES-KOH (pH 7.5), 100 mM KCl, 10 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 1 mM AdoMet, 10 pmol of $tRNA^{Ser1}$ and 1 pmol of recombinant CmoM. After the reaction, tRNA was extracted with acidic phenol and chloroform, followed by ethanol precipitation. The tRNAs were digested with RNase T_1 and subjected to LC/MS as described above.

Luciferase reporter assay

The luciferase reporter assay was performed essentially as described (40). *E. coli* wild-type, Δ*cmoM* and Δ*cmoB* strains were transformed with the series of dual-luciferase reporters described above. Each transformant was precultivated at 37°C in 2 ml LB medium containing 100 μg/ml ampicillin overnight. The preculture (1%) was inoculated to 2 ml of LB medium containing 100 μg/ml ampicillin and 100 μM arabinose to induce expression of the reporter. When the OD_{600} reached 0.3–0.7, 1 ml aliguot was centrifuged, and the pelleted cells were resuspended in 200 μl of lysis buffer [10 mM HEPES-KOH (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$, 7 mM 2-mercaptoethanol, 400 μg/ml lysozyme]. Cell lysates were prepared by the freeze-thaw method (47) and cleared by centrifugation. The reporter assay was carried out on 5 μl of lysate using GloMax™ 96 Microplate Luminometer (Promega) with the Dual-Luciferase™ Reporter Assay System (Promega). The Fluc luminescence signal was normalized against the Rluc luminescence signal.

RESULTS

Modification status of individual tRNAs analyzed by mass spectrometry

Because the methyl ester of $mcmo^5U$ is easily hydrolyzed during preparation and handling of tRNAs (22), little information is available regarding the presence of $mcmo^5U$ in individual tRNAs. To determine which tRNA species contain this modification, we employed reciprocal circulating chromatography (RCC) (43) to isolate six tRNA species ($tRNA^{Ala1}$, $tRNA^{Leu3}$, $tRNA^{Pro3}$, $tRNA^{Ser1}$, $tRNA^{Thr4}$ and $tRNA^{Val1}$) (Figure 1B) predicted to contain cmo^5U or $mcmo^5U$ from *E. coli* cells harvested at stationary phase. Each individual tRNA was digested with RNase T_1 and subjected to capillary liquid chromatography (LC)/nanoelectrospray ionization mass spectrometry (MS) to precisely analyze its post-transcriptional modifications (44). The RNA fragment containing the wobble modification was detected as multiply-charged negative

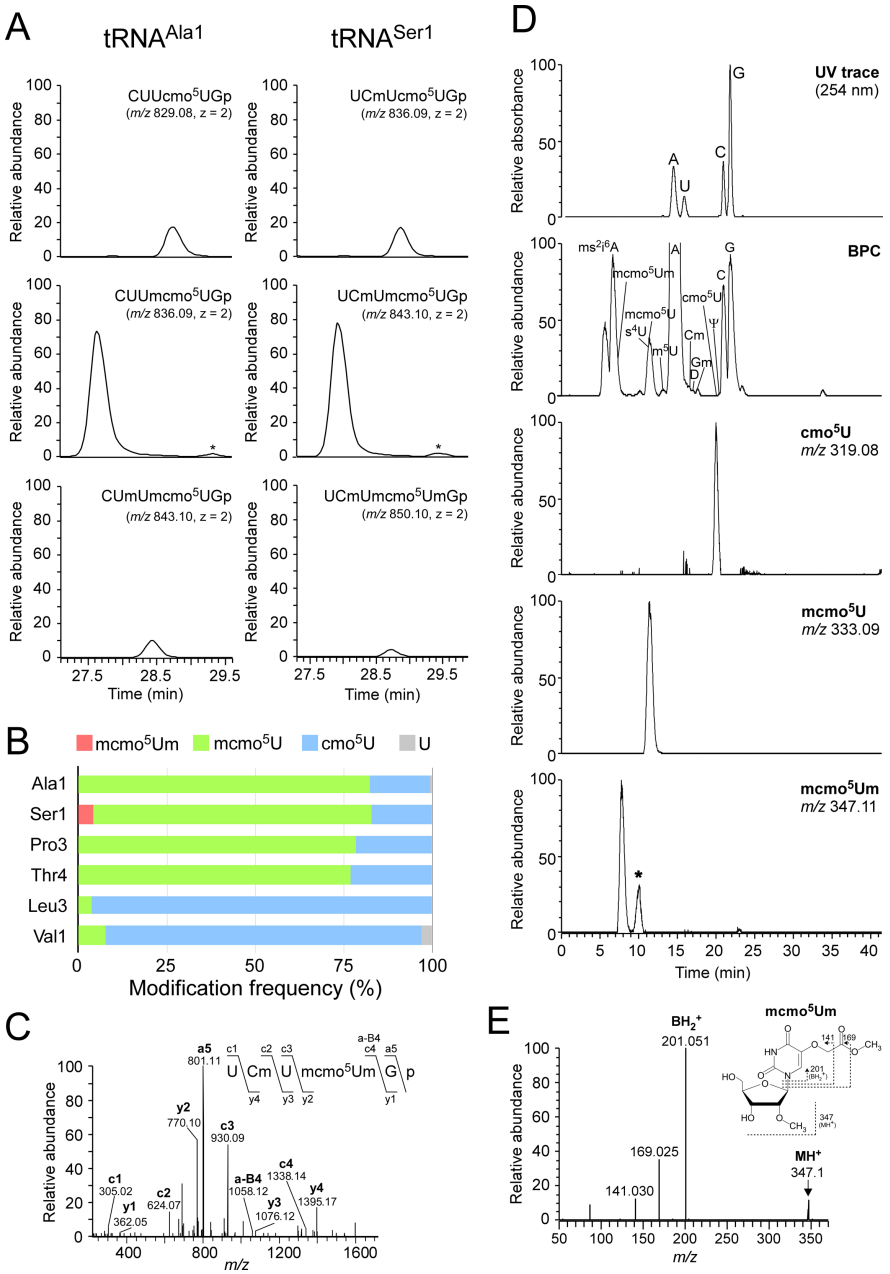

Figure 2. Mass spectrometric analysis of individual tRNAs isolated from stationary-phase *E. coli*. (**A**) Mass chromatograms of RNase T_1-digested fragments containing cmo^{5}U and its derivatives from tRNA^{Ala1} (left panels) and tRNA^{Ser1} (right panels) isolated from stationary-phase *E. coli*. Top and middle panels: extracted-ion chromatograms (XIC) for doubly-charged negative ions of cmo^{5}U34-containing fragments (CUUcmo^{5}UGp of tRNA^{Ala1}, MW 1660.18, *m/z* 829.08; UCmUcmo^{5}UGp of tRNA^{Ser1}, MW 1674.19, *m/z* 836.09) and mcmo^{5}U34-containing fragments (CUUmcmo^{5}UGp of tRNA^{Ala1}, MW 1674.19, *m/z* 836.09; UCmUmcmo^{5}UGp of tRNA^{Ser1}, MW 1688.21, *m/z* 843.10), respectively. Bottom left and bottom right panels: XICs for doubly-charged negative ions of Um32/mcmo^{5}U34-containing fragment of tRNA^{Ala1} (CUmUmcmo^{5}UGp, MW 1688.21, *m/z* 843.10) and mcmo^{5}Um34-containing fragment of tRNA^{Ser1} (UCmUmcmo^{5}UmGp, MW 1702.22, *m/z* 850.10), respectively. The peaks marked with asterisks represent Um32/cmo^{5}U-containing fragment (CUmUcmo^{5}UGp, MW 1674.19, *m/z* 836.09) in tRNA^{Ala1} and cmo^{5}Um-containing fragment (UCmUcmo^{5}UmGp, MW 1688.21, *m/z* 843.10) in tRNA^{Ser1}. The RNA fragments containing unmodified C32 are also present in tRNA^{Ser1}, but they are not described here due to high frequency of Cm32 in tRNA^{Ser1} isolated from stationary-phase *E. coli*. (**B**) Modification frequencies of cmo^{5}U and its derivatives at position 34 in six species of tRNAs isolated from stationary-phase *E. coli*. Relative composition of each modification was calculated from the peak area ratio of mass chromatograms of RNase T_1-digested fragments containing mcmo^{5}Um34 (red), mcmo^{5}U34 (green), cmo^{5}U34 (blue) or U (gray). (**C**) Collision-induced dissociation (CID) spectrum of a fragment of *E. coli* tRNA^{Ser1} containing mcmo^{5}Um34. The doubly-charged negative ion of the RNase T_1-digested fragment containing mcmo^{5}Um34 (*m/z* 850.10) was used as the precursor ion for CID. The product ions were assigned as described previously (66). Sequences of parent ion and assigned product ions are described on the upper left side of this panel. (**D**) Nucleoside analysis of *E. coli* tRNA^{Ser1}. Top panel: UV chromatogram at 254 nm of the four major nucleosides (C, U, G and A). Second panel: mass chromatograms of the protonated nucleosides (MH^+) on the base peak chromatogram (BPC). Third to bottom panels: XICs for cmo^{5}U (*m/z* 319.08), mcmo^{5}U (*m/z* 333.09) and mcmo^{5}Um (*m/z* 347.11), respectively. The peak marked with an asterisk represents unspecific peak. (**E**) CID spectrum of mcmo^{5}Um nucleoside. The protonated mcmo^{5}Um (MH^+, *m/z* 347.11) was used as the precursor ion for CID. The N-glycoside bond cleaved to generate the base-related ion (BH_2^+) and other product ions are assigned on the chemical structure.

ions (Figure 2A and Supplementary Figure S1) and further probed by collision-induced dissociation (CID) to determine the position of each modification (Supplementary Figure S2). Frequencies of the wobble modifications in these six tRNA species were calculated from the peak areas of mass chromatograms (Figure 2B). The wobble modifications of $tRNA^{Ala1}$, $tRNA^{Pro3}$, $tRNA^{Ser1}$ and $tRNA^{Thr4}$ consisted of ≈80% mcmo^{5}U and ≈20% cmo^{5}U. By contrast, over 90% of $tRNA^{Leu3}$ and $tRNA^{Val1}$ molecules contained cmo^{5}U, whereas the remaining 10% contained mcmo^{5}U or unmodified U. In $tRNA^{Ala1}$, we observed partial modification (≈11%) of 2′-*O*-methyluridine (Um) at position 32 (Figure 2A), which was confirmed by CID analysis (Supplementary Figure S2A).

Discovery of mcmo^{5}Um as a novel modification

At the wobble position of $tRNA^{Ser1}$, we found a novel minor modification with a molecular mass 14 Da larger than that of mcmo^{5}U (Figure 2A). CID analysis of this fragment (Figure 2C) revealed a specific product ion (a-B4) lacking the 5-methoxycarbonylmethoxyuracil-base (mcmo^{5}U-base), strongly suggesting that the 2′ hydroxyl group of the ribose was methylated. To confirm this result, we analyzed total nucleosides of $tRNA^{Ser1}$ and detected a proton-adduct (MH^+) of this minor nucleoside (*m/z* 347) (Figure 2D) which we then probed by CID analysis (Figure 2E). We clearly detected a base-related fragment (BH_2^+) with *m/z* 201, which also appears in mcmo^{5}U, confirming that the ribose portion is methylated. We also verified the absence of this modification in a knockout strain of *trmL*, which encodes a 2′-*O*-methyltransferase that targets position 34 (see Figure 4C). Taken together, we conclude that the novel modification found in $tRNA^{Ser1}$ is 5-methoxycarbonylmethoxy-2′-*O*-methyluridine (mcmo^{5}Um) (Figure 1A,B). About 4.2% of $tRNA^{Ser1}$ molecules contain this modification (Figure 2A,B). Given that 2′-*O*-methylation stabilizes C3′-*endo* ribose pucker conformation (48), mcmo^{5}Um is likely to allow $tRNA^{Ser1}$ to recognize UCG codon more efficiently than mcmo^{5}U.

Frequency of mcmo^{5}U in cellular tRNAs estimated by shotgun approach

Although mcmo^{5}U was detected in the four tRNAs with a frequency of ≈80% (Figure 2B), it is possible that the methyl ester of mcmo^{5}U was partially hydrolyzed during isolation of individual tRNAs by RCC. To exclude the effects of the inevitable hydrolysis of mcmo^{5}U during tRNA isolation, we analyzed the wobble modifications in total tRNAs using a shotgun approach. In order to profile the huge number of RNA fragments derived from all species of tRNAs, we prepared crude *E. coli* tRNA fractions from different growth phases, digested them with RNase T_1 and subjected the digested material to LC/MS (Figure 3A). Among the four tRNAs bearing mcmo^{5}U, we clearly detected an anticodon-containing fragment of $tRNA^{Pro3}$ and $tRNA^{Thr4}$ containing mcmo^{5}U or cmo^{5}U, each of which has a unique molecular mass and does not overlap with other fragments. Judging from the peak area of mass chromatograms for

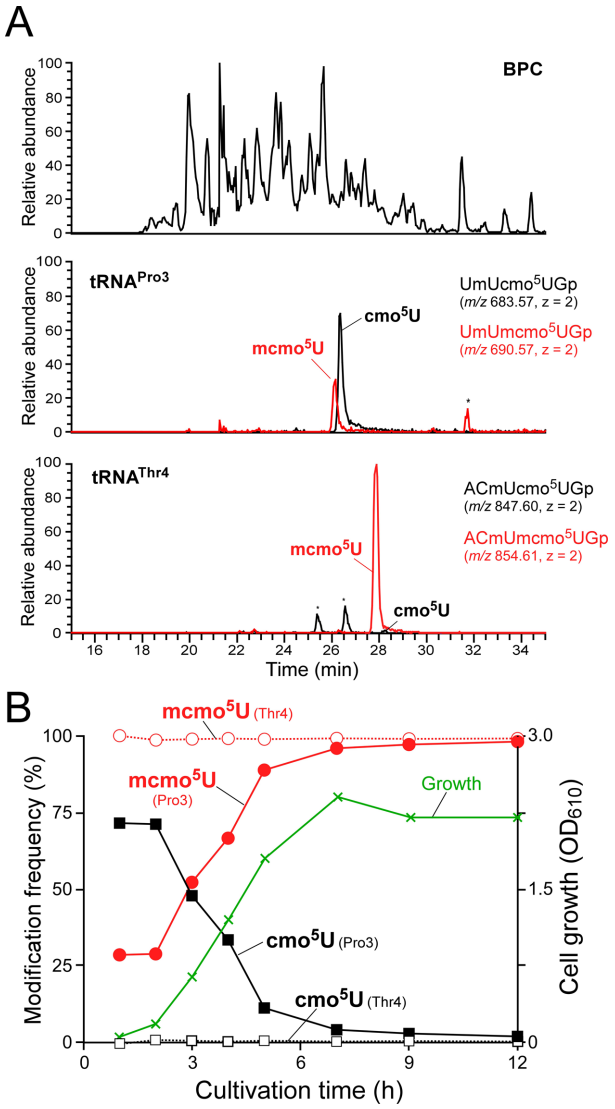

Figure 3. Shotgun analysis of tRNA fragments and growth phase-dependent alteration of mcmo^{5}U. (**A**) Shotgun analysis of total tRNA digested by RNase T_1. Top panel: base peak chromatogram (BPC) of RNase T_1-digested total tRNA isolated from early log phase (2 h) *E. coli*. Second panel: XICs for doubly-charged negative ions of the cmo^{5}U34-containing fragment (UmUcmo^{5}UGp, *m/z* 683.57) and the mcmo^{5}U34-containing fragment (UmUmcmo^{5}UGp, *m/z* 690.57) from $tRNA^{Pro3}$. Bottom panel: XICs for doubly-charged negative ions of the cmo^{5}U34-containing fragment (ACmUcmo^{5}UGp, *m/z* 847.60) and the mcmo^{5}U34-containing fragment (ACmUmcmo^{5}UGp, *m/z* 854.61) from $tRNA^{Thr4}$. (**B**) mcmo^{5}U level is altered by growth phase. $tRNA^{Pro3}$ (filled circle and rectangle) or $tRNA^{Thr4}$ (open circle and rectangle) were calculated from the peak area ratio of XICs for the fragments containing mcmo^{5}U34 and cmo^{5}U34 at different cultivation times. Cell growth (green line) was monitored by absorbance at 610 nm.

UmUmcmo^{5}UGp (*m/z* 690.575) and UmUcmo^{5}UGp (*m/z* 683.567) in $tRNA^{Pro3}$ isolated from cells in stationary phase, mcmo^{5}U was present in $tRNA^{Pro3}$ with a frequency of 98% (Figure 3B). Similarly, $tRNA^{Thr4}$ isolated from cells in stationary phase contained nearly 100% mcmo^{5}U (Figure 3B). These results suggest that about 20% of the cmo^{5}U detected in $tRNA^{Pro3}$ and $tRNA^{Thr4}$ (Figure 2B) was generated by artificial hydrolysis of mcmo^{5}U during tRNA iso-

lation. Given that mcmo^{5}U was present in the other two isolated tRNAs (Ala1 and Ser1) at a frequency of ≈80%, nearly 100% mcmo^{5}U should be present in all four tRNAs in stationary-phase *E. coli*.

Growth phase-dependent alteration of mcmo^{5}U in tRNA^{Pro3}

To investigate the effect of growth phase on mcmo^{5}U, we prepared total tRNA from *E. coli* cells harvested at various growth phases. The total tRNA was digested with RNase T_1, and analyzed by LC/MS to monitor the RNA fragment containing mcmo^{5}U or cmo^{5}U derived from tRNA^{Pro3} (Figure 3A). Based on the ratio of the two fragment peaks, we calculated the frequency of mcmo^{5}U over the course of *E. coli* cultivation (Figure 3B and Supplementary Figure S3A). In early log phase (1–2 h after inoculation), mcmo^{5}U was present in tRNA^{Pro3} with a frequency of ≈30%, whereas cmo^{5}U was present in the remaining molecules. The mcmo^{5}U frequency gradually increased as growth proceeded, reaching nearly 100% in late log and stationary phase. By contrast, mcmo^{5}U frequency in tRNA^{Thr4} was consistently high (>99%) in all growth phases (Figure 3B and Supplementary Figure S3B).

The shotgun analysis failed to detect specific fragments bearing mcmo^{5}U or cmo^{5}U derived from tRNA^{Ala1} and tRNA^{Ser1}, because the molecular masses of these fragments overlapped with those of other fragments. To investigate the modification status of these tRNAs in log phase, we isolated four tRNAs containing mcmo^{5}U from *E. coli* cells harvested in mid-log phase ($OD_{600} = 0.4$). Mass spectrometric analysis revealed that both tRNA^{Ala1} and tRNA^{Ser1} contained mcmo^{5}U with frequencies of 82% and 84%, respectively (Supplementary Figure S4). Because 20% of mcmo^{5}U is hydrolyzed and converted to cmo^{5}U during isolation of tRNAs by RCC (Figure 2B), we concluded that both tRNA^{Ala1} and tRNA^{Ser1} were fully modified with mcmo^{5}U in mid-log phase. Meanwhile, tRNA^{Pro3} and tRNA^{Thr4} contained mcmo^{5}U at frequencies of 34% and 80%, respectively (Supplementary Figure S4). These results are consistent with those observed in the shotgun analysis (Figure 3B). Based on these findings, we conclude that mcmo^{5}U content in tRNA^{Pro3} is dependent on growth phase, whereas the other three tRNAs are fully modified with mcmo^{5}U during all phases of cell growth.

Identification of a gene responsible for terminal methylation of mcmo^{5}U

To identify the gene encoding the AdoMet-dependent methyltransferase that methylates cmo^{5}U to form mcmo^{5}U, we conducted a genome-wide screen to identify genes responsible for RNA modifications. The method we employed, ribonucleome analysis, uses a reverse genetic approach combined with RNA-MS (44). This approach has been used to successfully identify many genes responsible for tRNA/rRNA modifications among uncharacterized genes in *E. coli* (40,46,49–52) and *Saccharomyces cerevisiae* (53–56). Screening of *E. coli* knockout strains revealed that mcmo^{5}U (m/z 333) was completely absent in the Δ*smtA* strain; instead, the level of cmo^{5}U (m/z 319) was higher than that in the wild-type (Figure 4A). When *smtA* was introduced on a plasmid (psmtA) into the Δ*smtA* strain, formation of mcmo^{5}U (m/z 333) was restored (Figure 4A). These data indicate that SmtA is a methyltransferase that produces mcmo^{5}U from cmo^{5}U.

Next, we isolated tRNA^{Ala1} and tRNA^{Ser1} from the Δ*smtA* strain, and analyzed their wobble modifications by LC/MS (Figure 4B,C). As expected, mcmo^{5}U-containing fragments were completely converted to cmo^{5}U-containing fragments in both strains.

As mentioned above, we discovered mcmo^{5}Um (Figure 1A) as a minor modification in tRNA^{Ser1}. We confirmed the absence of mcmo^{5}Um in both Δ*smtA* and Δ*trmL* (Figure 4C). *trmL* encodes a 2′-*O*-methyltransferase responsible for 2′-*O*-methylation of cmnm^{5}Um34 of tRNA^{Leu4} and Cm34 of tRNA^{Leu5} (57). Therefore, we conclude that mcmo^{5}Um in tRNA^{Ser1} is generated by 2′-*O*-methylation of mcmo^{5}U by TrmL.

In vitro reconstitution of mcmo^{5}U mediated by CmoM

To determine whether SmtA actually has methyltransferase activity, we conducted *in vitro* reconstitution of mcmo^{5}U formation by recombinant SmtA. In this experiment, tRNA^{Ser1} bearing cmo^{5}U was isolated from the Δ*smtA* strain and used as a substrate. We clearly detected mcmo^{5}U in the tRNA^{Ser1} only in the presence of both recombinant SmtA and AdoMet (Figure 5A). CID analysis of the anticodon-containing fragment confirmed the methylation occurred at the wobble position (Figure 5B). This result demonstrated that SmtA is an AdoMet-dependent methyltransferase that transfers a methyl group to cmo^{5}U34 of tRNAs to form mcmo^{5}U34. Based on the enzymatic activity, we renamed this gene CmoM (cmo^{5}U methyltransferase).

Characterization of CmoM

CmoM belongs to the Class I AdoMet-dependent methyltransferase (MTase) family, whose members contain a Rossmann-fold as a characteristic structural motif (Figure 6A) (58–60). The high-resolution crystal structure of *E. coli* CmoM (SmtA) (PDB ID:4HTF) revealed that CmoM forms a homodimer, and that each subunit contains one molecule each of AdoMet, sulfate, acetate and 2-mercaptoethanol as ligands (Figure 6C and Supplementary Figure S5). Based on this structure, we designed eight mutant *cmoM* constructs bearing single amino-acid alterations (Figure 6A). The mutated residues, which are conserved in *cmoM* homologs, are located at the catalytic site where AdoMet is bound (Figure 6C). The Δ*cmoM* strain was transformed with each of the mutant constructs, and total RNA extracted from each construct was digested with RNase T_1 and subjected to LC/MS to detect the RNA fragment of tRNA^{Pro3} bearing mcmo^{5}U or cmo^{5}U (Figure 6B). The positive and negative controls behaved as expected: mcmo^{5}U was fully restored by wild-type *cmoM*, whereas no mcmo^{5}U was formed in cells transfected with an empty vector. Little or no mcmo^{5}U was observed in the Δ*cmoM* strain introduced by the mutant constructs R26A, D73A, W124A, Y150A, R209A, D213A and R246A, whereas mcmo^{5}U formation was partially restored by the Y247A mutant. The

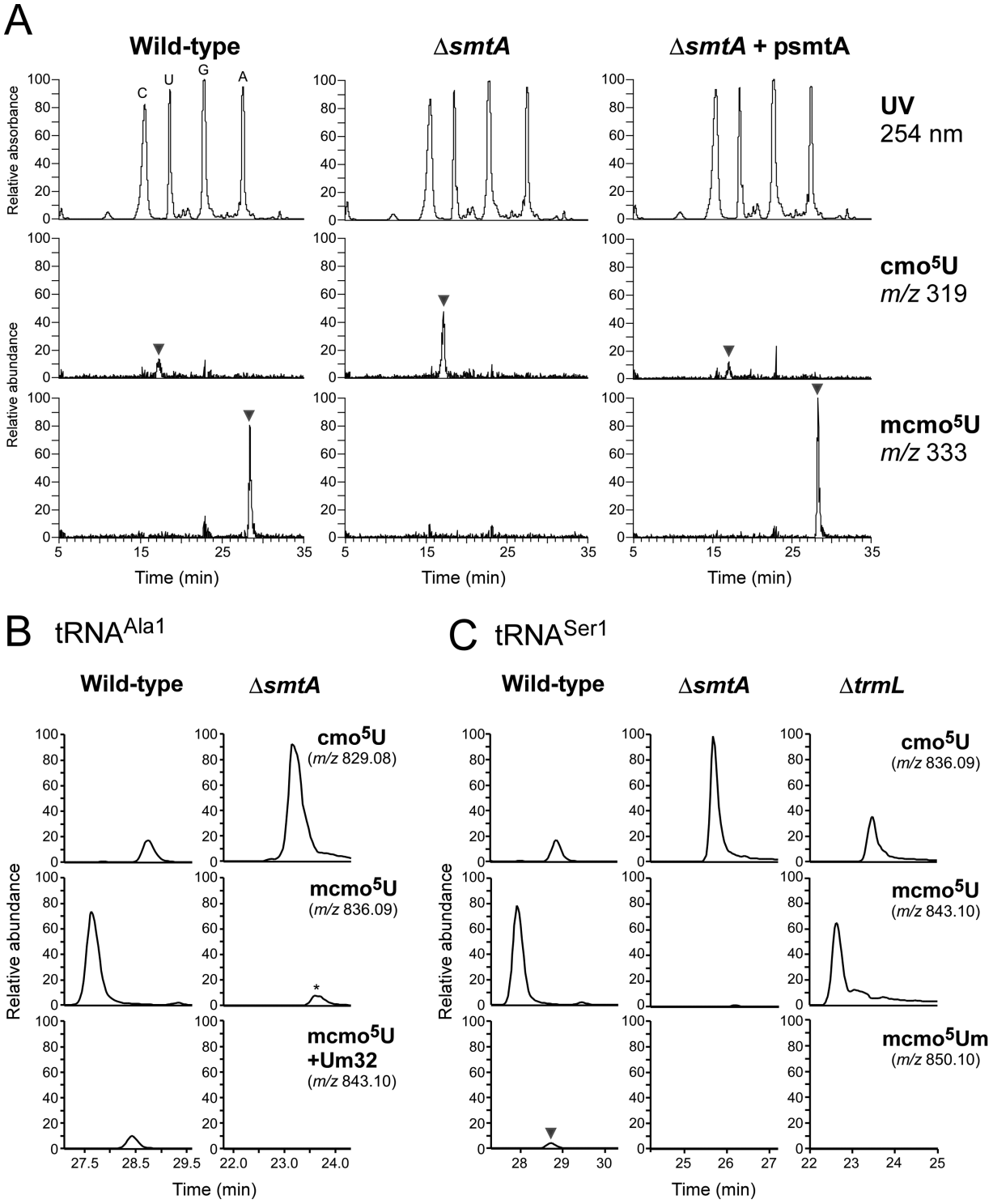

Figure 4. Reverse genetic approach identified a gene responsible for terminal methylation of mcmo^{5}U. (**A**) Nucleoside analyzes by LC/MS using reverse phase chromatography of total RNA from wild-type (left panels), Δ*smtA* (middle panels) and Δ*smtA* rescued with psmtA (right panels). Top panels: UV trace at 254 nm. Second and bottom panels: XICs for cmo^{5}U (*m/z* 319) and mcmo^{5}U (*m/z* 333), respectively. Intensity of each peak was normalized to that of cyclic t^{6}A (*m/z* 395). (**B**) Mass chromatograms of RNase T_1-digested fragments containing cmo^{5}U and its derivatives from tRNA^{Ala1} isolated from wild-type (left panels) and Δ*smtA* (right panels) strains. Top, middle and bottom panels: XICs for doubly-charged negative ions of the cmo^{5}U34-containing fragments (CUUcmo^{5}UGp, *m/z* 829.08), the mcmo^{5}U34-containing fragments (CUUmcmo^{5}UGp, *m/z* 836.09) and the Um32/mcmo^{5}U34-containing fragment (CUmUmcmo^{5}UGp, *m/z* 843.10), respectively. The peak marked with an asterisk represent the Um32/cmo^{5}U-containing fragment (CUmUcmo^{5}UGp, *m/z* 836.09). (**C**) Mass chromatograms of RNase T_1-digested fragments containing cmo^{5}U and its derivatives from tRNA^{Ser1} isolated from wild-type (left panels), Δ*smtA* (middle panels) and Δ*trmL* (right panels) strains. Top, middle and bottom panels: XICs for doubly-charged negative ions of the cmo^{5}U34-containing fragments (UCmUcmo^{5}UGp, *m/z* 836.09), the mcmo^{5}U34-containing fragments (UCmUmcmo^{5}UGp, *m/z* 843.10) and the mcmo^{5}Um34-containing fragment (UCmUmcmo^{5}UmGp, *m/z* 850.10), respectively.

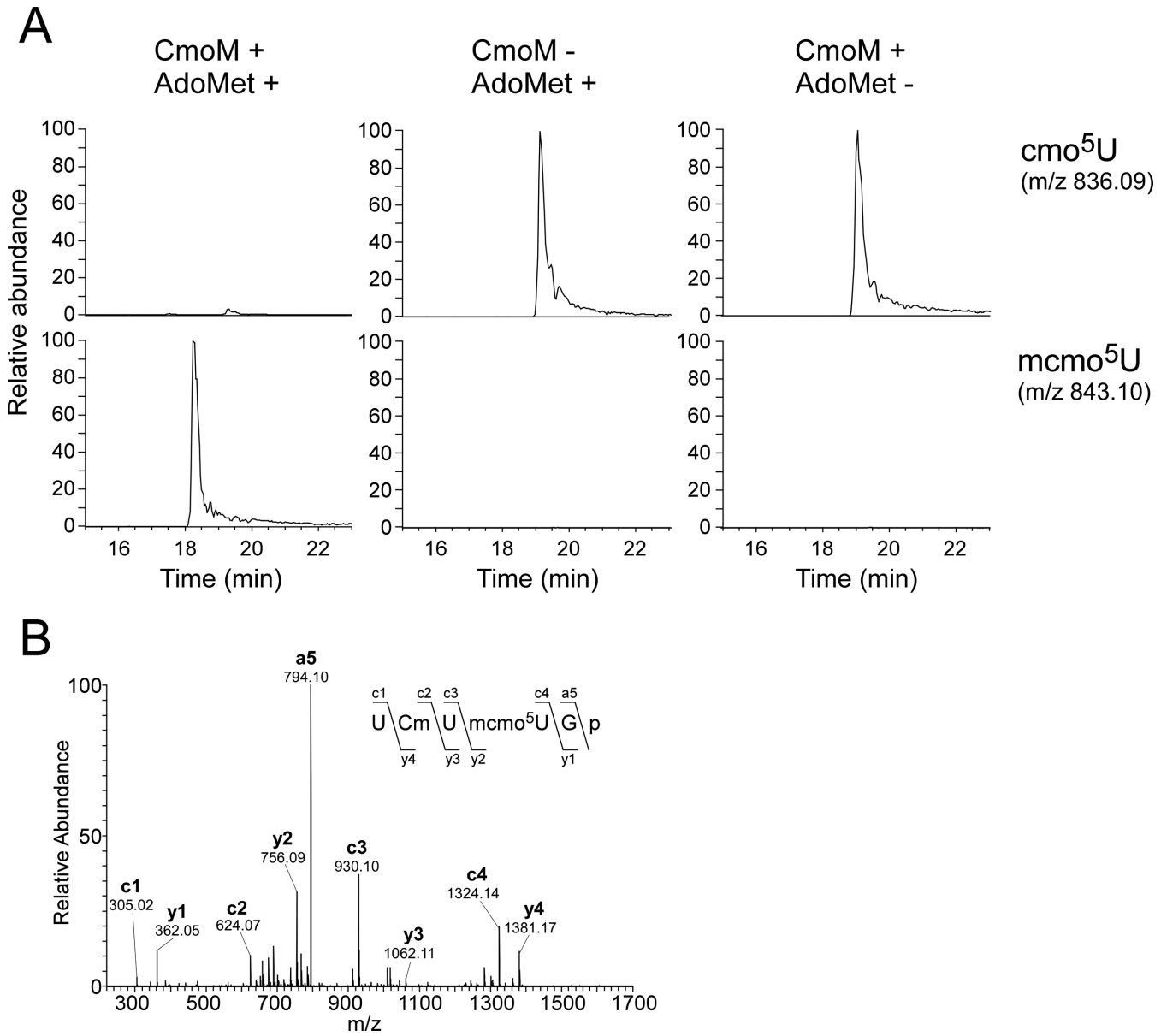

Figure 5. *In vitro* reconstitution of cmo^{5}U methylation by recombinant CmoM. (**A**) *E. coli* tRNA^{Ser1} bearing cmo^{5}U isolated from the Δ*cmoM* strain was incubated in the presence or absence of recombinant CmoM with or without AdoMet. Top and bottom panels: XICs for doubly-charged negative ions of the cmo^{5}U34-containing fragments (UCmUcmo^{5}UGp, *m/z* 836.09) and the mcmo^{5}U34-containing fragments (UCmUmcmo^{5}UGp, *m/z* 843.10), respectively. (**B**) A CID spectrum of RNase T_1-digested fragment of *E. coli* tRNA^{Ser1} incubated in the presence of recombinant CmoM with AdoMet. The doubly-charged negative ion of the mcmo^{5}U34-containing fragment (UCmUmcmo^{5}UGp, *m/z* 843.10) was used as the precursor ion for CID. The product ions were assigned according to the literature (66). Sequences of parent ion and assigned product ions are described upper left side in this panel.

data indicate that the highly conserved residues in the catalytic center are essential for normal methyltransferase activity.

Terminal methylation of mcmo^{5}U contributes to the decoding process

To investigate the functional role of the terminal methylation of mcmo^{5}U, we constructed dual-luciferase reporters based on the RF2 recoding system (32,41). *Renilla* luciferase (Rluc) was fused to firefly luciferase (Fluc) by a linker sequence bearing the +1 frameshift signal of the RF2 recoding site (Figure 7A). The original UGA codon at the frameshift site was replaced with GCG or UCG to examine the decoding abilities of tRNA^{Ala1} and tRNA^{Ser1}, respectively. The GCG codon is exclusively deciphered by tRNA^{Ala1} (mcmo^{5}UGC), whereas the UCG codon is redundantly recognized by tRNA^{Ser1}(mcmo^{5}UGA) and tRNA^{Ser2}(CGA) (Supplementary Figure S6) (14). We also prepared a control reporter construct lacking the +1 frameshift site (zero frame) (Figure 7A). Each of these reporters was introduced into wild-type (WT), Δ*cmoM* and Δ*cmoB* strains. The decoding ability of the test codon at the frameshift site is reflected by the +1 frameshift activity. Because the +1 frameshift activity in this system is pro-

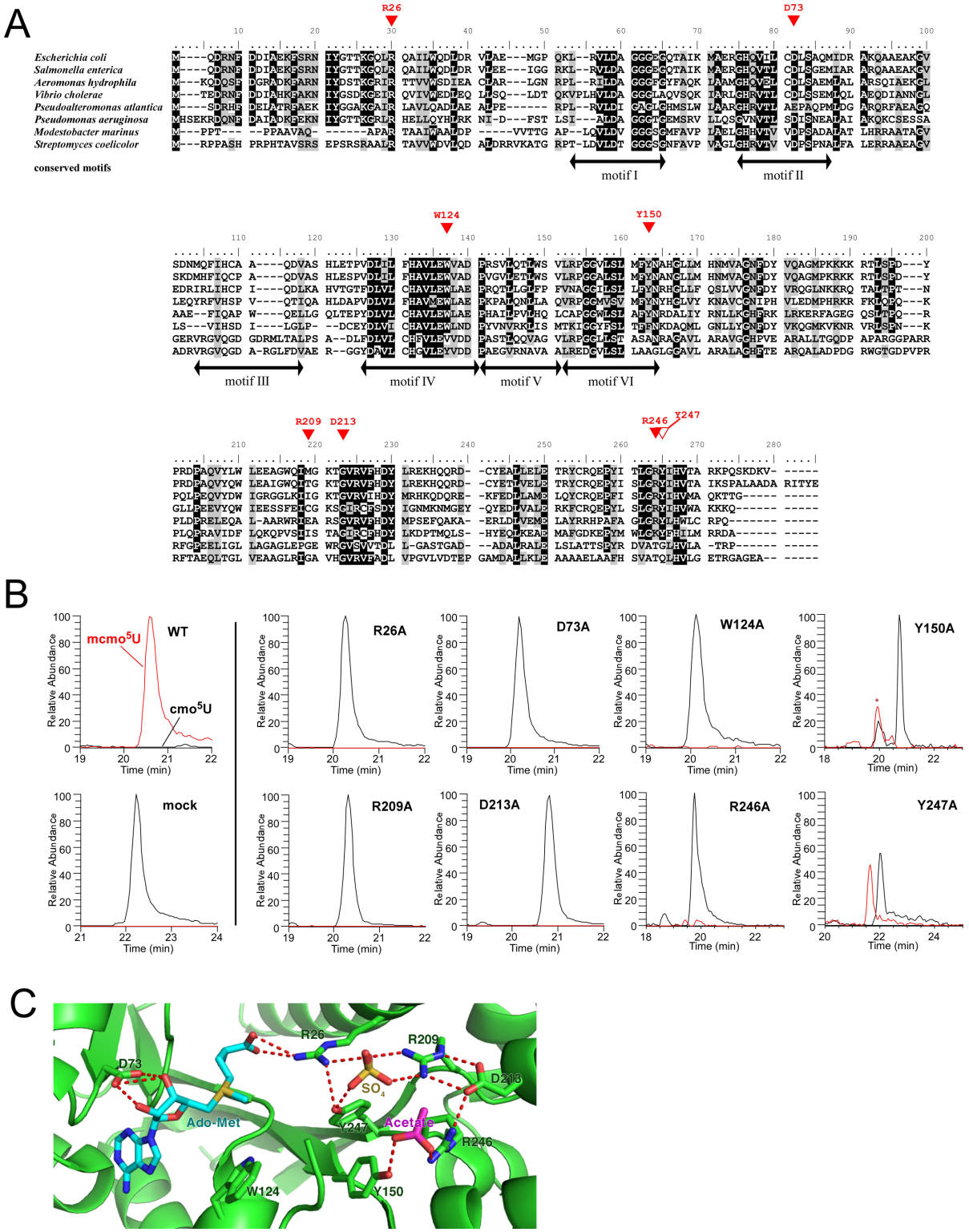

Figure 6. Characterization of CmoM. (**A**) Sequence alignment of CmoM homologs from six γ-proteobacteria, *Escherichia coli* (NP_415441.1), *Salmonella enterica* (NP_455477.1), *Aeromonas hydrophila* (YP_856903.1), *Vibrio cholerae* (NP_231353.1), *Pseudoalteromonas atlantica* (ABG40695.1), *Pseudomonas aeruginosa* (NP_253478.1), and two Actinobacteria, *Modestobacter marinus* (WP_014741474.1), *Streptomyces coelicolor* (WP_011028138.1). Identical or similar residues are shaded in black or gray, respectively. Red triangles indicate residues that are essential (filled) or non-essential (open) for generic complementation. Motifs I to VI are conserved in Class I AdoMet-dependent methyltransferases. (**B**) XICs for doubly-charged negative ions of the cmo^{5}U34-containing fragment (black lines, UmUcmo^{5}UGp, *m/z* 683.57) and the mcmo^{5}U34-containing fragment (red lines, UmUmcmo^{5}UGp, *m/z* 690.57) from tRNA^{Pro3} in the Δ*cmoM* strain rescued by plasmid-encoded wild-type *cmoM* or its mutant derivatives. The peak marked with an asterisk represents unspecific peak. (**C**) Close-up view of the AdoMet-binding site in the crystal structure of *E. coli* CmoM (PDB ID: 4HTF) containing ligands, AdoMet, acetate and sulfate. Predicted hydrogen bonds between ligands and CmoM are indicated by red dotted lines.

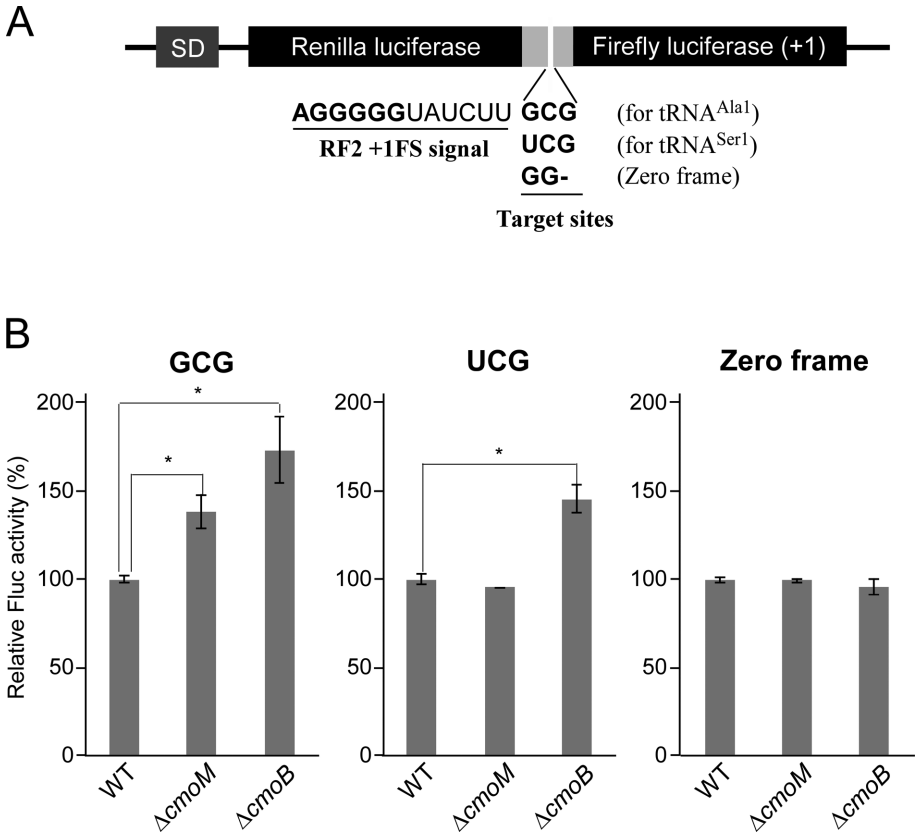

Figure 7. Terminal methylation of mcmo^{5}U contributes to GCG decoding. (**A**) Schematic depiction of the dual-luciferase reporter constructs based on the RF2 recoding system. SD, Shine-Dalgarno sequence. *Renilla* and firefly luciferases were fused with a linker containing the +1 frameshift signal of the RF2 recoding site. The frameshift target site was replaced with a GCG codon for tRNA^{Ala1}, a UCG codon for tRNA^{Ser1}or GG for zero frame (used as a control). (**B**) Relative pausing activity at the frameshift site with GCG (left), UCG (middle) or zero frame (right) was calculated based on relative Fluc activity normalized to Rluc activity in wild-type, Δ*cmoM* and Δ*cmoB* strains. Data are presented as means ± SD (n = 4). *, $P < 0.01$ versus control (Student's t-test).

moted by the 'hungry' A-site, the ability of the cognate tRNA to decode the test codons can be estimated indirectly. The +1 frameshift activity was calculated from the Fluc signal against the Rluc signal (F/R value). No difference in the frameshift activity of the zero frame construct, used as a negative control, was observed in the three strains (Figure 7B). In the Δ*cmoB* strain, in which all mcmo^{5}U/cmo^{5}U should be converted to ho^{5}U, we observed clear stimulation of +1 frameshift activity at both GCG and UCG codons (Figure 7B), indicating that the carboxymethyl group of cmo^{5}U contributes to decoding of G-ending codons. This result is consistent with the previous reports (32,61). On the other hand, in the Δ*cmoM* strain, we detected a slight but significant stimulation of frameshift activity at the GCG codon (Figure 7B), but not at UCG. This observation indicates that the terminal methyl group of mcmo^{5}U is partially involved in GCG decoding by tRNA^{Ala1}. The absence of any reduction in UCG decoding in Δ*cmoM* strain can be explained by the fact that this codon is redundantly deciphered by tRNA^{Ser1} and tRNA^{Ser2}; tRNA^{Ser2} might have compensated for the reduced decoding ability of hypomodified tRNA^{Ser1} in the Δ*cmoM* strain.

DISCUSSION

It is difficult to estimate the exact frequency of an RNA modification with unstable chemical structure, such as an ester group, in individual tRNAs, because such modifications are easily hydrolyzed during tRNA isolation (42). Although the presence of mcmo^{5}U in cellular tRNAs was reported previously, the exact frequency of this modification in individual tRNAs remained unknown (14,22). Using the shotgun approach, we showed here that mcmo^{5}U is present in nearly 100% of tRNA^{Pro3} and tRNA^{Thr4} molecules isolated from stationary-phase *E. coli* (Figure 3B). This method can be applied to analysis of other RNA modifications with unstable chemical structures from various sources, including 5-methoxycarbonylmethyluridine (mcm^{5}U), wybutosine (yW), cyclic N^6-threonylcarbamoyladenosine (ct^{6}A), glutamylqueuosine (GluQ) and their derivatives. In addition, we used RCC to estimate the fraction of mcmo^{5}U hydrolyzed during tRNA isolation. Judging from the mcmo^{5}U frequency (≈80%) in the isolated tRNAs, we concluded that ≈20% of mcmo^{5}U was converted to cmo^{5}U during tRNA isolation (Figure 2B), indicating that all four tRNAs (tRNA^{Ala1}, tRNA^{Ser1}, tRNA^{Pro3} and tRNA^{Thr4}) are fully modified with mcmo^{5}U in stationary-phase *E.*

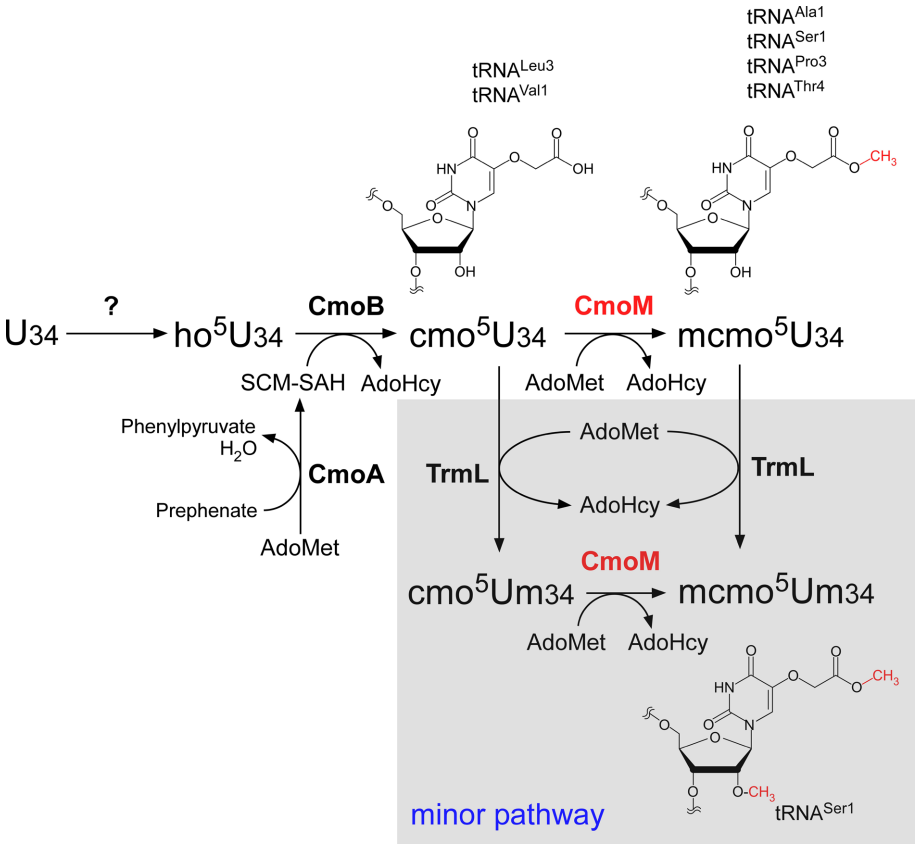

Figure 8. Biosynthesis of mcmo^{5}U. In *E. coli*, U34 of six tRNAs responsible for decoding NCN codons is modified to ho^{5}U34 by an unknown pathway, and subsequently modified to cmo^{5}U in a reaction catalyzed by CmoM. CmoB uses SCM-SAH as a substrate and transfers its carboxymethyl group to ho^{5}U34 on tRNAs. SCM-SAH is synthesized from prephenate and AdoMet catalyzed by CmoA. Four tRNAs (Ala1, Ser1, Pro3 and Thr4) are further modified to mcmo^{5}U by CmoM, using AdoMet as a substrate. mcmo^{5}U frequency is altered by growth phase only in tRNA^{Pro3}. In a minor pathway, mcmo^{5}U34 in tRNA^{Ser1} is further methylated by TrmL to yield mcmo^{5}Um34. Alternatively, cmo^{5}U34 could be first converted to cmo^{5}Um34, then to mcmo^{5}Um34.

coli. The four tRNAs containing mcmo^{5}U all have G35 at the second letter of the anticodon, and therefore specify NCN codons, implying that CmoM preferentially recognizes G35. However, because mcmo^{5}U was present (albeit at a low frequency, <10%) in tRNA^{Leu3} and tRNA^{Val1}, which do not have G35 (Figure 2B), this residue is not an essential determinant for CmoM.

By applying the shotgun approach to total tRNA, we observed growth phase-dependent alteration of mcmo^{5}U in tRNA^{Pro3} (Figure 3B). In all phases of cell growth, cmo^{5}U was fully incorporated into this tRNA. In early log phase, cmo^{5}U of tRNA^{Pro3} was partially modified by CmoM to yield mcmo^{5}U with a frequency of 30%. As growth proceeded, the level of mcmo^{5}U gradually increased, and the level of cmo^{5}U concomitantly decreased. At late log and stationary phases, tRNA^{Pro3} was fully modified with mcmo^{5}U. According to the GEO profile database (ID: 35525121 and 35543521) (62), the steady-state level of *cmoM* mRNA is temporarily elevated in early log phase, and is expressed at a constant in late log and stationary phases. Therefore, we speculate that hypomodification of mcmo^{5}U in tRNA^{Pro3} might be due to slow methylation by CmoM that fails to catch up with fast production of tRNA in early log phase; as growth rate decreases, mcmo^{5}U accumulates gradually. By contrary, tRNA^{Thr4} was fully modified with mcmo^{5}U in all growth phases (Figure 3B). In addition, a high level of mcmo^{5}U was also found in tRNA^{Ala1}, tRNA^{Ser1} as well as tRNA^{Thr4} isolated from log phase *E. coli*. These results imply that cmo^{5}U34 is a better substrate for CmoM in these three tRNAs than in tRNA^{Pro3}. The growth phase dependency of mcmo^{5}U in tRNA^{Pro3} might be involved in regulatory decoding of CCN codons in growth phase-specific gene expression. Further studies will be necessary to examine this speculation.

Because Δ*cmoM* exhibited no obvious growth phenotype (data not shown)(63), the functional role of the terminal methyl group of mcmo^{5}U may be limited. To characterize this modification, we employed a reporter assay based on the RF2 recoding system to estimate the ability to decode GCG and UCG codons in the presence or absence of *cmoM* (Figure 7B). The +1 frameshift activity at a GCG codon increased specifically in the absence of *cmoM*, indicating that mcmo^{5}U facilitates decoding of GCG by tRNA^{Ala1}. However, we observed no change in UCG decoding in the absence of *cmoM*, because the UCG codon is redundantly recognized by the other isoacceptor. Similarly, it is difficult to assess the decoding ability of other NCN codons in the absence of *cmoM*, because NCA codons are recognized exclu-

sively by Watson–Crick base pairs, and other NCN codons are redundantly deciphered by two isoacceptors (14). However, in light of our findings, it is reasonable to assume that the terminal methyl group of mcmo^{5}U contributes to NCN decoding in general.

According to the crystal structure of the 30S ribosomal subunit in complex with ASL of tRNA^{Val} bearing the cmo^{5}UAC anticodon and its cognate codons (33), the carboxylate of cmo^{5}U forms a hydrogen bond with the N6 amino group of A35 (the second letter of the anticodon). This interaction is one component of the intramolecular hydrogen bonding network that pre-structures the anticodon loop, so that cmo^{5}U can pair with all four bases at the third letters of codons. We showed here that mcmo^{5}U is primarily present in tRNAs with anticodons containing G35. When cmo^{5}U is present in these tRNAs at the ribosome A-site, the carboxylate of cmo^{5}U cannot form a hydrogen bond with G35. To make the matter worse, cmo^{5}U might be destabilized due to electrostatic repulsion between the carboxylate and the O6 carbonyl oxygen of G35, both of which are negatively charged. The terminal methyl group of mcmo^{5}U neutralizes the negative charge of cmo^{5}U carboxylate, suggesting that mcmo^{5}U is involved in stabilizing the wobble base in the anticodons containing G35.

The crystal structure of CmoM also reveals AdoMet and other ligands bound to the positively-charged surface, which might be involved in tRNA recognition (Supplementary Figure S5). Seven residues essential for mcmo^{5}U formation, R26, D73, W124, Y150, R209, D213 and R246, which we identified in this study, reside near the ligand-binding site on the positively charged surface of CmoM (Figure 6C). R26 and D73 play a critical role in positioning AdoMet by forming hydrogen bonds. D73 is a conserved carboxylate in motif II of AdoMet MTase (Figure 6A) (58,59). W124 may participate in AdoMet binding via a stacking interaction with the adenine base of AdoMet. Given that W124 is in close proximity to the methyl group of AdoMet, it might be involved in the interaction with the cmo^{5}U base of tRNA and thereby facilitate the cmo^{5}U methylation. R26 forms a network of hydrogen bonds with two other essential residues, R209 and Y247, along with a sulfate. R209 extends the network to D213, which interacts with R246. R246 and Y150 to form the binding site for an acetate. The functional roles of the sulfate and the acetate bound to the catalytic site remain unknown, but these ligands may act as mimics for the phosphate group of tRNA bound to CmoM.

Phylogenetic analysis revealed that *cmoM* is present in γ-proteobacteria, actinobacteria and a few species in other bacterial clades (Figure 6A). Because actinobacteria doesn't have homologs of *cmoB*, cmo^{5}U is not predicted to be present in this organism, indicating that actinobacterial counterpart is not a functional homolog of *cmoM*. Consistent with this, two essential residues, W124 and Y150, in *E. coli* CmoM are not conserved in there organisms (Figure 6A). Thus, CmoM and mcmo^{5}U are mainly distributed in γ-proteobacteria. Presence of *cmoM* homologs in *Spirochaeta cellobiosiphila* (Spirochaeta), *Zetaproteobacteria bacterium* (ζ-proteobacteria) and *Paenibacillus sophorae* (Firmicutes) indicates horizontal gene transfer of *cmoM* from γ-proteobacteria to these species.

In the biogenesis of cmo^{5}U and mcmo^{5}U (Figure 8), six tRNA species (tRNA^{Ala1}, tRNA^{Leu3}, tRNA^{Pro3}, tRNA^{Ser1}, tRNA^{Thr4} and tRNA^{Val1}) first undergo hydroxylation at C5 of the uracil base in U34 to yield ho^{5}U34; the enzyme and substrate involved in this process are unknown. Next, ho^{5}U34 is further modified by CmoB using SCM-SAH as a substrate to yield cmo^{5}U34. SCM-SAH is generated from AdoMet and prephenate in a reaction catalyzed by CmoA. For four tRNA species (tRNA^{Ala1}, tRNA^{Pro3}, tRNA^{Ser1} and tRNA^{Thr4}), cmo^{5}U34 is methylated by CmoM in the presence of AdoMet to yield mcmo^{5}U34. Only in tRNA^{Ser1}, small portion of mcmo^{5}U34 is further methylated by TrmL to yield mcmo^{5}Um34. Alternatively, cmo^{5}U34 could be first converted to cmo^{5}Um34, then to mcmo^{5}Um34. Growth phase-dependent alteration of mcmo^{5}U34 takes place in tRNA^{Pro3}, implying a possible mechanism of translational control mediated by the regulatory decoding efficiency of CCN codons.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to the members of the Suzuki laboratory, in particular Yuriko Sakaguchi, for technical support and many insightful discussions.

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 Y.S. and S.K.]. Funding for open access charge: Grants-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports, and Culture of Japan [to T.S.].
Conflict of interest statement. None declared.

REFERENCE

1. Machnicka,M.A., Milanowska,K., Osman Oglou,O., Purta,E., Kurkowska,M., Olchowik,A., Januszewski,W., Kalinowski,S., Dunin-Horkawicz,S., Rother,K.M. *et al.* (2013) MODOMICS: a database of RNA modification pathways–2013 update. *Nucleic Acids Res.*, **41**, D262–D267.
2. Yokoyama,S. and Nishimura,S. (1995) In: Soll,DRUL (ed). *tRNA: Structure, Biosynthesis, and Function*. American Society for Microbiology, Washington, D.C., pp. 207–224.
3. Bjork,G. (1995) In: Soll,DRUL (ed). *tRNA: Structure, Biosynthesis, and Function*. American Society for Microbiology, Washington, D.C., pp. 165–205.
4. Suzuki,T. (2005) In: Grosjean,H (ed). *Fine-Tuning of RNA Functions by Modification and Editing*. GmbH & Co. KG, Springer-Verlag Berlin and Heidelberg, Vol. **12**, pp. 23–69.
5. Crick,F.H. (1966) Codon–anticodon pairing: the wobble hypothesis. *J. Mol. Biol.*, **19**, 548–555.
6. Barrell,B.G., Anderson,S., Bankier,A.T., de Bruijn,M.H., Chen,E., Coulson,A.R., Drouin,J., Eperon,I.C., Nierlich,D.P., Roe,B.A. *et al.* (1980) Different pattern of codon recognition by mammalian mitochondrial tRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 3164–3166.
7. Inagaki,Y., Kojima,A., Bessho,Y., Hori,H., Ohama,T. and Osawa,S. (1995) Translation of synonymous codons in family boxes by Mycoplasma capricolum tRNAs with unmodified uridine or adenosine at the first anticodon position. *J. Mol. Biol.*, **251**, 486–492.

8. Suzuki,T., Nagao,A. and Suzuki,T. (2011) Human Mitochondrial tRNAs: Biogenesis, Function, Structural Aspects, and Diseases. *Annu. Rev. Genet.*, **45**, 299–329.
9. Suzuki,T. and Suzuki,T. (2014) A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. *Nucleic Acids Res.*, **42**, 7346–7357.
10. Yokoyama,S., Watanabe,T., Murao,K., Ishikura,H., Yamaizumi,Z., Nishimura,S. and Miyazawa,T. (1985) Molecular mechanism of codon recognition by tRNA species with modified uridine in the first position of the anticodon. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 4905–4909.
11. Agris,P.F., Soll,D. and Seno,T. (1973) Biological function of 2-thiouridine in Escherichia coli glutamic acid transfer ribonucleic acid. *Biochemistry*, **12**, 4331–4337.
12. Kirino,Y., Yasukawa,T., Ohta,S., Akira,S., Ishihara,K., Watanabe,K. and Suzuki,T. (2004) Codon-specific translational defect caused by a wobble modification deficiency in mutant tRNA from a human mitochondrial disease. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 15070–15075.
13. Kurata,S., Weixlbaumer,A., Ohtsuki,T., Shimazaki,T., Wada,T., Kirino,Y., Takai,K., Watanabe,K., Ramakrishnan,V. and Suzuki,T. (2008) Modified uridines with C5-methylene substituents at the first position of the tRNA anticodon stabilize U.G wobble pairing during decoding. *J. Biol. Chem.*, **283**, 18801–18811.
14. Bjork,G.R. and Hagervall,T.G. (2014) Transfer RNA Modification: Presence, Synthesis and Function. *EcoSal Plus*, **1**, doi:10.1128/ecosalplus.ESP-0007-2013.
15. Yaniv,M. and Barrell,B.G. (1969) Nucleotide sequence of E. coli B tRNA1-Val. *Nature*, **222**, 278–279.
16. Harada,F., Kimura,F. and Nishimura,S. (1969) Nucleotide sequence of valine tRNA 1 from Escherichia coli B. *Biochim. Biophys. Aacta*, **195**, 590–592.
17. Murao,K., Saneyoshi,M., Harada,F. and Nishimura,S. (1970) Uridin-5-oxy acetic acid: a new minor constituent from E. coli valine transfer RNA I. *Biochem. Biophys. Res. Commun.*, **38**, 657–662.
18. Ishikura,H., Yamada,Y. and Nishimura,S. (1971) Structure of serine tRNA from Escherichia coli. I. Purification of serine tRNA's with different codon responses. *Biochim. Biophys. Acta*, **228**, 471–481.
19. Williams,R.J., Nagel,W., Roe,B. and Dudock,B. (1974) Primary structure of E. coli alanine transfer RNA: relation to the yeast phenylalanyl tRNA synthetase recognition site. *Biochem. Biophys. Res. Commun.*, **60**, 1215–1221.
20. Sørensen,M.A., Elf,J., Bouakaz,E., Tenson,T., Sanyal,S., Bjork,G.R. and Ehrenberg,M. (2005) Over expression of a tRNALeu isoacceptor changes charging pattern of leucine tRNAs and reveals new codon reading. *J. Mol. Biol.*, **354**, 16–24.
21. Kuchino,Y., Yabusaki,Y., Mori,F. and Nishimura,S. (1984) Nucleotide sequences of three proline tRNAs from Salmonella typhimurium. *Nucleic Acids Res.*, **12**, 1559–1562.
22. Pope,W.T., Brown,A. and Reeves,R.H. (1978) The identification of the tRNA substrates for the supK tRNA methylase. *Nucleic Acids Res.*, **5**, 1041–1058.
23. Pope,W.T. and Reeves,R.H. (1978) Purification and characterization of a tRNA methylase from Salmonella typhimurium. *J. Bacteriol.*, **136**, 191–200.
24. Lesiewicz,J. and Dudock,B. (1977) In vitro methylation of E. coli alanine trna with homologous E. coli methylase. *Fed. Proc.*, **36**, 705.
25. Murao,K., Hasegawa,T. and Ishikura,H. (1976) 5-methoxyuridine: a new minor constituent located in the first position of the anticodon of tRNAAla, tRNAThr, and tRNAVal from Bacillus subtilis. *Nucleic Acids Res.*, **3**, 2851–2860.
26. Takai,K., Takaku,H. and Yokoyama,S. (1996) Codon-reading specificity of an unmodified form of Escherichia coli tRNA1Ser in cell-free protein synthesis. *Nucleic Acids Res.*, **24**, 2894–2899.
27. Phelps,S.S., Malkiewicz,A., Agris,P.F. and Joseph,S. (2004) Modified Nucleotides in tRNALys and tRNAVal are Important for Translocation. *J. Mol. Biol.*, **338**, 439–444.
28. Samuelsson,T., Elias,P., Lustig,F., Axberg,T., Fölsch,G., Akesson,B. and Lagerkvist,U. (1980) Aberrations of the classic codon reading scheme during protein synthesis in vitro. *J. Biol. Chem.*, **255**, 4583–4588.
29. Takemoto,T., Takeishi,K., Nishimura,S. and Ukita,T. (1973) Transfer of valine into rabbit haemoglobin from various isoaccepting species of valyl-tRNA differing in codon recognition. *Eur. J. Biochem.*, **38**, 489–496.
30. Gabriel,K., Schneider,J. and McClain,W.H. (1996) Functional evidence for indirect recognition of G.U in tRNA(Ala) by alanyl-tRNA synthetase. *Science*, **271**, 195–197.
31. Nasvall,S., Chen,P. and Bjork,G.R. (2004) The modified wobble nucleoside uridine-5-oxyacetic acid in tRNA(cmo5UGG)(Pro) promotes reading of all four proline codons in vivo. *RNA*, **10**, 1662–1673.
32. Nasvall,S.J., Chen,P. and Bjork,G.R. (2007) The wobble hypothesis revisited: Uridine-5-oxyacetic acid is critical for reading of G-ending codons. *RNA*, **13**, 2151–2164.
33. Weixlbaumer,A., Murphy,F.V., Dziergowska,A., Malkiewicz,A., Vendeix,F.A.P., Agris,P.F. and Ramakrishnan,V. (2007) Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nat. Struct. Mol. Biol.*, **14**, 498–502.
34. Bjork,G.R. (1980) A novel link between the biosynthesis of aromatic amino acids and transfer RNA modification in Escherichia coli. *J. Mol. Biol.*, **140**, 391–410.
35. Hagervall,T.G., Jönsson,Y.H., Edmonds,C.G., McCloskey,J.A. and Bjork,G.R. (1990) Chorismic acid, a key metabolite in modification of tRNA. *J. Bacteriol.*, **172**, 252–259.
36. Byrne,R.T., Whelan,F., Aller,P. and Bird,L.E. (2013) S-Adenosyl-S-carboxymethyl-L-homocysteine: a novel cofactor found in the putative tRNA-modifying enzyme CmoA. *Acta Crystallogr. D Biol. Crystallogr.*, **D69**, 1090–1098.
37. Kim,J., Xiao,H., Bonanno,J.B., Kalyanaraman,C., Brown,S., Tang,X., Al-Obaidi,N.F., Patskovsky,Y., Babbitt,P.C., Jacobson,M.P. *et al.* (2013) Structure-guided discovery of the metabolite carboxy-SAM that modulates tRNA function. *Nature*, **498**, 123–126.
38. Kawakami,K., Jönsson,Y.H., Bjork,G.R., Ikeda,H. and Nakamura,Y. (1988) Chromosomal location and structure of the operon encoding peptide-chain-release factor 2 of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 5620–5624.
39. 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**, 20060008.
40. 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.
41. Curran,J.F. (1993) Analysis of effects of tRNA:message stability on frameshift frequency at the Escherichia coli RF2 programmed frameshift site. *Nucleic Acids Res.*, **21**, 1837–1843.
42. Miyauchi,K., Kimura,S. and Suzuki,T. (2013) A cyclic form of N6-threonylcarbamoyladenosine as a widely distributed tRNA hypermodification. *Nat. Chem. Biol.*, **9**, 105–111.
43. Miyauchi,K., Ohara,T. and Suzuki,T. (2007) Automated parallel isolation of multiple species of non-coding RNAs by the reciprocal circulating chromatography method. *Nucleic Acids Res.*, **35**, e24.
44. 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.
45. Sakaguchi,Y., Miyauchi,K., Kang,B.I. and Suzuki,T. (2015) Nucleoside Analysis by Hydrophilic Interaction Liquid Chromatography Coupled with Mass Spectrometry. *Methods Enzymol.*, **560**, 19–28.
46. Kimura,S., Miyauchi,K., Ikeuchi,Y., Thiaville,P.C., Crecy-Lagard,V. and Suzuki,T. (2014) Discovery of the beta-barrel-type RNA methyltransferase responsible for N6-methylation of N6-threonylcarbamoyladenosine in tRNAs. *Nucleic Acids Res.*, **42**, 9350–9365.
47. Hirabayashi,N., Sato,N.S. and Suzuki,T. (2006) Conserved Loop Sequence of Helix 69 in Escherichia coli 23 S rRNA Is Involved in A-site tRNA Binding and Translational Fidelity. *J. Biol. Chem.*, **281**, 17203–17211.
48. Kawai,G., Yamamoto,Y., Kamimura,T., Masegi,T., Sekine,M., Hata,T., Iimori,T., Watanabe,T., Miyazawa,T. and Yokoyama,S. (1992) Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2′-hydroxyl group. *Biochemistry*, **31**, 1040–1046.
49. 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.
50. 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.
51. 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.
52. Kimura,S., Ikeuchi,Y., Kitahara,K., Sakaguchi,Y. and Suzuki,T. (2012) Base methylations in the double-stranded RNA by a fused methyltransferase bearing unwinding activity. *Nucleic Acids Res.*, **40**, 4071–4085.
53. 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.
54. 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.
55. 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.
56. Noma,A., Yi,S., Katoh,T., Takai,Y. and Suzuki,T. (2011) Actin-binding protein ABP140 is a methyltransferase for 3-methylcytidine at position 32 of tRNAs in Saccharomyces cerevisiae. *RNA*, **17**, 1111–1119.
57. Benítez-Páez,A., Villarroya,M., Douthwaite,S., Gabaldón,T. and Armengod,M.-E. (2010) YibK is the 2′-O-methyltransferase TrmL that modifies the wobble nucleotide in Escherichia coli tRNA(Leu) isoacceptors. *RNA*, **16**, 2131–2143.
58. Martin,J.L. and McMillan,F.M. (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol.*, **12**, 783–793.
59. Schubert,H.L., Blumenthal,R.M. and Cheng,X. (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.*, **28**, 329–335.
60. Kozbial,P.Z. and Mushegian,A.R. (2005) Natural history of S-adenosylmethionine-binding proteins. *BMC Struct. Biol.*, **5**, 19.
61. Takai,K., Okumura,S., Hosono,K., Yokoyama,S. and Takaku,H. (1999) A single uridine modification at the wobble position of an artificial tRNA enhances wobbling in an Escherichia coli cell-free translation system. *FEBS Lett.*, **447**, 1–4.
62. Barrett,T., Wilhite,S.E., Ledoux,P., Evangelista,C., Kim,I.F., Tomashevsky,M., Marshall,K.A., Phillippy,K.H., Sherman,P.M., Holko,M. *et al.* (2012) NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.*, **41**, D991–D995.
63. Yamanaka,K., Ogura,T., Niki,H. and Hiraga,S. (1995) Characterization of the smtA gene encoding an S-adenosylmethionine-dependent methyltransferase of Escherichia coli. *FEMS Microbiol. Lett.*, **133**, 59–63.
64. Sprinzl,M. and Vassilenko,K.S. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **33**, D139–D140.
65. Lund,E. and Dahlberg,J.E. (1977) Spacer transfer RNAs in ribosomal RNA transcripts of E. coli: processing of 30S ribosomal RNA in vitro. *Cell*, **11**, 247–262.
66. McLuckey,S.A., Van Berkel,G.J. and Glish,G.L. (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.*, **3**, 60–70.