A single methyltransferase YefA (RlmCD) catalyses both m^{5}U747 and m^{5}U1939 modifications in *Bacillus subtilis* 23S rRNA

Benoit Desmolaize[1], Céline Fabret[2,3], Damien Brégeon[2], Simon Rose[1], Henri Grosjean[2] and Stephen Douthwaite[1,*]

[1]Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, [2]IGM, Université Paris-Sud, CNRS-UMR 8621, 91405 Orsay, France and [3]Université Pierre et Marie Curie, Paris, F-75006

ABSTRACT

Methyltransferases that use *S*-adenosylmethionine (AdoMet) as a cofactor to catalyse 5-methyl uridine (m^{5}U) formation in tRNAs and rRNAs are widespread in Bacteria and Eukaryota, and are also found in certain Archaea. These enzymes belong to the COG2265 cluster, and the Gram-negative bacterium *Escherichia coli* possesses three paralogues. These comprise the methyltransferases TrmA that targets U54 in tRNAs, RlmC that modifies U747 in 23S rRNA and RlmD that is specific for U1939 in 23S rRNA. The tRNAs and rRNAs of the Gram-positive bacterium *Bacillus subtilis* have the same three m^{5}U modifications. However, as previously shown, the m^{5}U54 modification in *B. subtilis* tRNAs is catalysed in a fundamentally different manner by the folate-dependent enzyme TrmFO, which is unrelated to the *E. coli* TrmA. Here, we show that methylation of U747 and U1939 in *B. subtilis* rRNA is catalysed by a single enzyme, YefA that is a COG2265 member. A recombinant version of YefA functions in an *E. coli* m^{5}U-null mutant adding the same two rRNA methylations. The findings suggest that during evolution, COG2265 enzymes have undergone a series of changes in target specificity and that YefA is closer to an archetypical m^{5}U methyltransferase. To reflect its dual specificity, YefA is renamed RlmCD.

INTRODUCTION

Ribothymidine (m^{5}U) is present in tRNAs from the three domains of life (1–3) and is additionally found in the rRNAs of all Bacteria (4; http://www.ecosal.org) and certain Archaea (5). The RNAs of the best-studied bacterium, *Escherichia coli,* contain three m^{5}U's. Two of these are in 23S rRNA and one in tRNA, and these modifications are conserved to varying degrees in other bacteria. The most highly conserved m^{5}U is at position 54 in the T-loop of tRNA (6) and in *E. coli* this modification is added by the methyltransferase TrmA, which was formerly termed RumT (7). In 23S rRNA, m^{5}U at nucleotide 1939 is present in all bacterial ribosomes characterized to date (8), and its formation is catalysed by the *E. coli* enzyme RlmD (formerly RumA) (9,10). The third m^{5}U, at 23S rRNA position 747, is the least conserved of the three and is found mainly in Gram-negative Beta-, Epsilon- and Gamma Proteobacteria (3) and, in *E. coli*, this modification is added by the methyltransferase RlmC (formerly RumB) (10). All three *E. coli* m^{5}U methyltransferases use *S*-adenosyl-L-methionine (AdoMet) as the methyl group donor, and their sequence similarity indicates that they share a common evolutionary origin (3). The sequences cluster within the family of orthologous groups COG2265 (11), which contains homologues from phylogenetically distant organisms such as the tRNA m^{5}U54 methyltransferases Trm2p, found in the yeast *Saccharomyces cerevisiae* (12), and PAB0719 in the hyperthermophilic archaeon *Pyrococcus abyssi* (3).

The sites of m^{5}U modification in the RNAs of *Bacillus subtilis* are similar to those in *E. coli*—the tRNAs of *B. subtilis* contain m^{5}U54, and our earlier provisional studies on *B. subtilis* 23S rRNA indicated that nucleotides U747 and U1939 (*E. coli* numbering system) were also methylated. However, this similarity in m^{5}U location obscured the fact that there are fundamental differences in the modification mechanisms in *B. subtilis* and *E. coli*. It had already been shown that in *B. subtilis*, the tRNA

m^{5}U54 modification is added by a folate-dependent methyltransferase (TrmFO) that bears no relationship to the COG2265 family members (13), and this discovery begged the question as to how m^{5}U's might be added to the rRNA. Our searches of the *B. subtilis* genome revealed no paralogues of TrmFO, although two COG2265-type genes, *yefA* and *yfjO*, were evident. The protein products YefA and YfjO showed sufficient sequence similarity to each other, and to the *E. coli* RlmC and RlmD enzymes, to rank them as the most likely candidates for catalysis of m^{5}U modification in *B. subtilis* rRNA.

In this study, we have used Matrix Assisted Laser Desorption/Ionization (MALDI) mass spectrometry (MS) to demonstrate conclusively that m^{5}U's are present at nucleotides 747 and 1939 in *B. subtilis* 23S rRNA. The rRNAs were then analysed after inactivation of *yefA* or *yfjO* to establish whether the protein products of these genes were responsible for the m^{5}U modifications. Enzyme function was re-established by expressing plasmid-encoded recombinant protein in the *B. subtilis* null-mutant and also in an *E. coli* mutant in which all three *trmA*, *rlmC* and *rlmD* genes had been deleted. Our findings show that compared to the *E. coli* RlmC and RlmD counterparts, a functional shift has occurred during the evolution of the *B. subtilis* YefA and YfjO methyltransferases.

MATERIALS AND METHODS

Bacillus subtilis and *E. coli* strains

Bacillus subtilis 168 (*trpC2*) was used here as the reference strain with a wild-type rRNA modification pattern. The *B. subtilis* knockout strains, BG12826 (*yefA*::pMutin2) and BG12911 (*yfjO*::pMutin2), originated from the collection of the Japanese/European *B. subtilis* Functional Analysis Program, and were created by a standard single crossover-based protocol using PCR-amplified fragments of target genes cloned in the pMUTIN2 vector (14). The Δ*yefA*::Cm deletion mutant was derived from the *B. subtilis* 168 strain. The upstream 891 bp and downstream 919 bp fragments of the *yefA* gene were PCR amplified from the genomic DNA using the primers yefA-up1 (5′-GCGCTCGGAATCCCACGCGAGG) together with yefA-up2 (5′-<u>CGACCTGCAGGCATGCA AGCTCTATAATCACCTC</u>TAGTTTTAAGCGAA AG), as well as yefA-down1 (5′-<u>GCTCGAATTCACTGG CCGTCG</u>TAAATAGTCAAAAATCTATGCTGATG G) together with yefA-down2 (5′-TAGTCCCTCGCAGC TTGCCCGC). The sequences underlined in the yefA-up2 and yefA-down1 primers are homologous to the ends of the pUC19-cloned 727-bp chloramphenicol resistance cassette (15) that was amplified using the PCR primers Cm5 (5′-CGACGGCCAGTGAATTCGAGC) together with Cm3 (5′-AGCTTGCATGCCTGCAGGTCG). The three PCR products were mixed in equimolecular amounts and joined in a subsequent PCR. The resulting 2.5-kb DNA fragment was used to transform competent *B. subtilis* 168 cells prior to selection at 37°C on LB rich medium containing chloramphenicol (5 μg/ml). Chromosomal DNA was extracted from chloramphenicol-resistant clones and *yefA*

gene deletion was verified by PCR using the upstream out5 (5′-GAGTTTGACCTGATCATTGCGGC-3′) and downstream out3 (5′-GAATCTTTTCAGAATATATTG CTGG-3′) primers. One of the correct strains was dubbed 168 Δ*yefA*::Cm and was used for further study. The *B. subtilis* strain BFS2838 carrying an inactivated *trmFO* (*gid-0*) gene was kindly provided by S. Serror (European functional analysis project of *B. subtilis*; http://genome .jouy.inra.fr/cgi-bin/micado/fonct_analysis.cgi).

Escherichia coli strains used in this work are derivatives of BW 25113 [F-, Δ(*araD-araB*)567, Δ*lacZ4787*(::rrnB-3), λ$^-$, *rph-1*, Δ(*rhaD-rhaB*)568, *hsdR514*] and were obtained from the Coli Genetic Stock Center. The BW 25113 Δ*trmA*/Δ*rlmC*/Δ*rlmD* strain was obtained by sequential P1-mediated transductions of the Δ*trmA753*::*kan*, Δ*rlmD783*::*kan* and Δ*rlmC760*::*kan* alleles (16). After each transduction, the kanamycin cassette was excised by transient expression of the specific FLP recombinase (17). Deletion of the three genes was verified by PCR using primers adjacent to the inactivated sequences.

Plasmids

The 1.4-kb *yefA* ORF was amplified from genomic DNA with the PCR primers atg_Hind (5′CCCAAGCTTATGA AAATGAAACCCCCAGTAG-3′) and stop_Xba (5′-CT AGTCTAGATTATTCTTTTAATTTTATCAACACAC AG-3′) containing HindIII and XbaI sites, respectively. The PCR fragment was cloned into the SmaI-digested pUC19 vector to form pUC-*yefA* and was then digested with HindIII and XbaI before cloning the *yefA* sequence into these sites in plasmid pDG148 (18) forming pDG-*yefA*. In this vector, expression of *yefA* was under control of the *Pspac* promoter, and was induced with 0.5–1 mM IPTG. The pDG148 and pDG-*yefA* vectors were shuttled into competent 168 Δ*yefA*::Cm cells, and transformants were selected on LB plates containing chloramphenicol (5 μg/ml) and kanamycin (6 μg/ml).

The pBAD-*yefA* vector was obtained by digesting pUC-*yefA* with EcoRI and XbaI, and cloning the *yefA* fragment into the same sites in the pBAD24 vector bringing *yefA* expression under the control of the arabinose promoter, which was induced with 0.2% arabinose. Electrocompetent Δ*trmA*/Δ*rlmC*/Δ*rlmD* cells were transformed with the pBAD-*yefA* vector and selected on LB plates containing ampicillin (100 μg/ml). All recombinants were verified by sequencing.

Preparation of *B. subtilis* and *E. coli* rRNA

The *E. coli* and *B. subtilis* strains were grown at 37°C to an OD$_{600}$ of 0.6 in 200 ml LB medium containing the appropriate antibiotics. For strains complemented with recombinant copies of *yefA* and *yfjO*, IPTG was added at this stage to a final concentration of 1 mM to induce protein expression. Cells were kept at 37°C for three more hours, and were then harvested by centrifugation, washed and resuspended on ice in 3 ml of 50 mM Tris–HCl pH 7.5, 10 mM MgCl$_2$, 100 mM NH$_4$Cl before being lysed by sonication. Cell debris was removed by centrifugation and ribosomal particles were pelleted by ultracentrifugation (19) leaving bulk tRNA in the supernatant. The tRNA

and rRNA fractions were extracted with phenol and chloroform, the RNAs were recovered by ethanol precipitation and dissolved in H2O.

MALDI mass spectrometry analyses of rRNA

RNA was extracted from the ribosomal particles of the *B. subtilis* 168 wild-type, *yefA* and *yfjO* knockouts, and from the knockouts complemented with active copies of *yefA* and *yfjO*; rRNAs were additionally extracted from the *E. coli* $\Delta rlmC/\Delta rlmD/\Delta trmA$ triple knockout and from this strain expressing an active copy of *yefA* or *yfjO*. The sequence from G725 to G772 within domain II of the *B. subtilis* 23S rRNA (Figure 1A) was isolated by hybridization to a complementary 48-mer deoxy-oligonucleotide, 5′-CCCACACCTCATCCCCGCACTTT TCAACGTGCGTGGGTTCGGGCCTCC, while the sequence from C1914 to C1961 within domain IV of the *B. subtilis* 23S rRNA (Figure 1B) was hybridized to the 48-mer, 5′GTCGGAACTTACCCGACAAGGAATTTC GCTACCTTAGGACCGTTATAG. The corresponding regions in *E. coli* rRNA were isolated using similar oligonucleotides adjusted for the sequence variations shown in Figure 1. In each case, 100 pmol of total rRNA was hybridized to 500 pmol of deoxyoligonucleotide; unprotected regions of the rRNAs were digested away with mung bean nuclease and RNase A; and the rRNA sequences paired to the deoxyoligonucleotide were separated by gel electrophoresis and extracted (20,21). The rRNA sequences of around 48 nt were digested with either RNase A or RNase T1 in aqueous solution containing 3-hydroxypicolinic acid and analysed by MALDI-MS (Voyager Elite, Perspective Biosystems) recording in reflector and positive ion mode (22). Spectra were analysed using the program m/z (Proteometrics Inc). Tandem mass spectra were recorded in positive ion mode on a MicroMass MALDI Q-TOF Ultima mass spectrometer (23).

Analyses of RNAs by HPLC and ion-trap electrospray mass spectrometry

The supernatant fractions containing tRNAs and other soluble RNAs were passed through Microcon YM-100 columns (Millipore) to reduce the amount of impurities. The supernatant samples were digested to completion to form nucleosides (24), and the pellet samples containing the rRNAs were similarly treated. The resultant nucleosides were analysed using a HPLC/ES-MS ion-trap system. Briefly, the loading pump of an Agilent 1100 HPLC was run at 5 μl/min; the column was a Hypercarb 150×0.3 mm, 5 μm PGC with a pore size of 250 (Thermo Scientific), linked to an Agilent XCT Ultra 6340 ion trap. Separation was achieved on gradients of 0–90% acetonitrile in 0.1% formic acid. ES-MS spectra were recorded in negative ion mode scanning initially in the *m/z* range 220–400 for intact nucleosides and then in lower *m/z* ranges after fragmentation (24). Nucleosides from the tRNAs were additionally analysed using reversed phase HPLC as previously described (24).
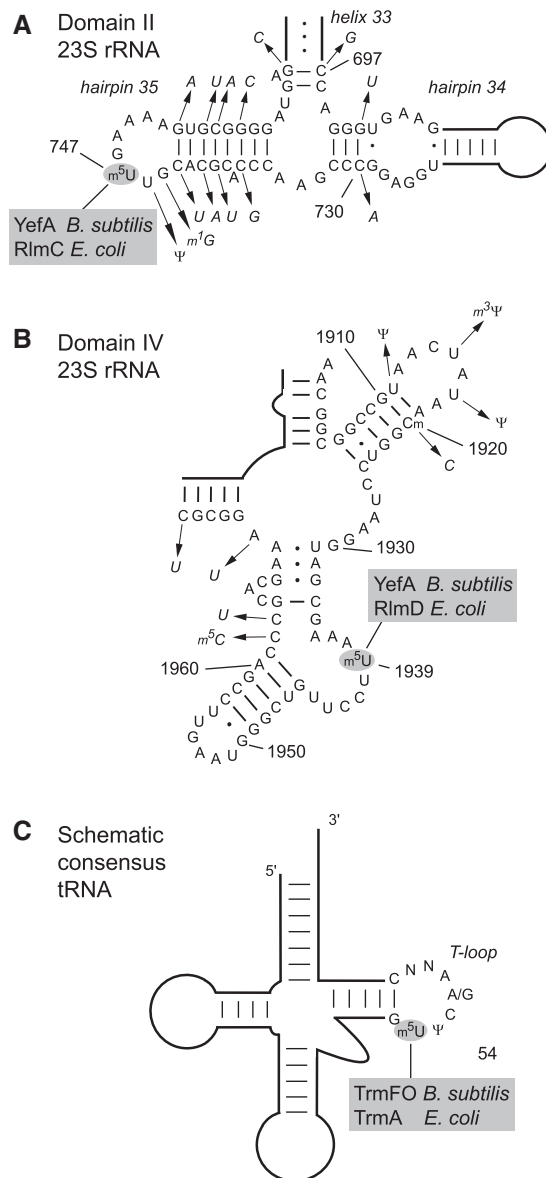

Figure 1. The three characterized sites of m⁵U modification in *B. subtilis* and *E. coli* RNAs. The 23S rRNA secondary structures are based on the *B. subtilis* sequence and nucleotide differences in *E. coli* are indicated by the arrows pointing to the italicized nucleotides. (**A**) Portion of domain II of 23S rRNA, containing hairpins 34 and 35 and part of helix 33, showing the location of m⁵U747 modified by YefA in *B. subtilis* rRNA and RlmC in *E. coli*. The m¹G745 modification is seen in *E. coli* rRNA but is absent in *B. subtilis*; the presence of pseudouridine Ψ746 could not be tested by mass spectrometry. (**B**) Region from domain IV of 23S rRNA containing m⁵U 1939 that is modified by YefA in *B. subtilis* rRNA, and by RlmD in *E. coli*. The *B. subtilis* rRNA is methylated on the ribose of nucleotide C1920 presumably by an orthologue of the enzyme TlyA (42) and this modification is missing in *E. coli*; the m⁵C1962 modification is present in the *E. coli* rRNA (47) but was not detected in *B. subtilis*; the pseudouridines Ψ1911, Ψ1915 and Ψ1917 could not be verified by mass spectrometry. (**C**) Schematic consensus of tRNA structures indicating the U54 target of the *E. coli* TrmA (48) and the structurally unrelated enzyme TrmFO that modifies the same nucleotide in *B. subtilis* using a tetrahydrofolate cofactor as the methyl donor (13,25).

RESULTS

Nucleotide modifications m⁵U747 and m⁵U1939 are present in *B. subtilis* 23S rRNA

Isolation of the *B. subtilis* 23S rRNA domain II region (Figure 1A) followed by RNase T1 digestion gives rise to a unique UUGp fragment containing nucleotide U747. In the MALDI-MS spectra of the rRNA from the

wild-type *B. subtilis* strain, this fragment produced peaks at m/z 990 and m/z 972, respectively, for the linear and cyclic phosphate forms of UUGp (Figure 2A), corresponding to the trimer plus 14 Da extra mass of one methyl group. There was no peak at m/z 976 (or at m/z 958 for the cyclic phosphate form), indicating that methylation of the sequence was stoichiometric. Tandem MS of this fragment and a longer partially digested RNase
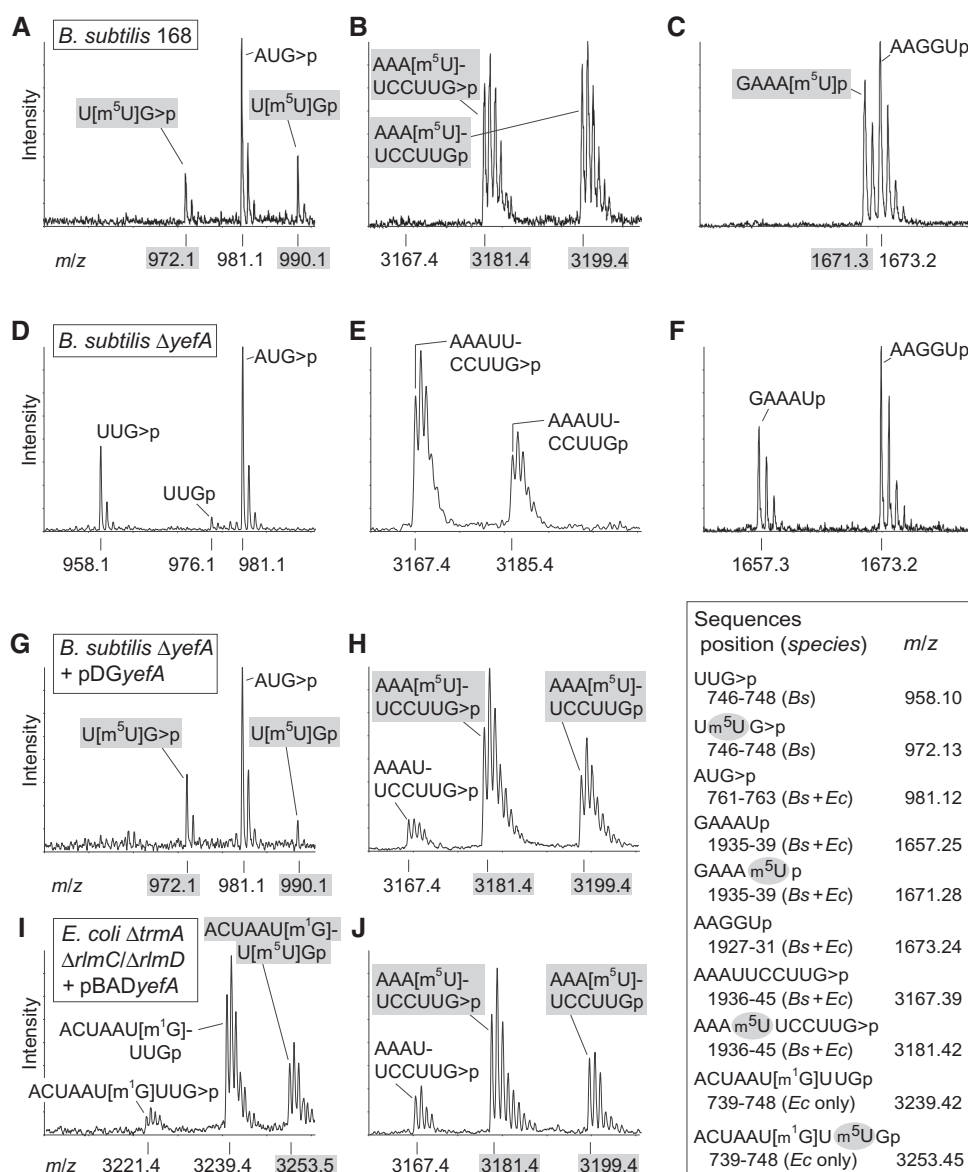
Figure 2. MALDI-MS analyses of m⁵U methylation sites in 23S rRNA. The upper row of spectra were generated from 23S rRNA of the *B. subtilis* wild-type strain 168 and show sequences containing m⁵U nucleotides (shaded in grey) at (A) U747 (in an RNase T1 oligo), (B) U1939 (RNase T1 oligo) and (C) U1939 (RNase A oligo). The corresponding spectral regions from the *B. subtilis* Δ*yefA*::Cm strain are shown in panels (D–F). The rRNA spectra from the *B. subtilis* Δ*yfjO* strain (data not shown) were identical to the wild-type rRNA. The rRNA analysed in spectra (G and H) was from *B. subtilis* Δ*yefA*::Cm strain complemented with an active copy of the *yefA* gene. (I and J) are spectra from the same rRNA regions of the *E. coli* Δ*rlmC*/Δ*rlmD*/Δ*trmA* triple mutant expressing an active copy of *yefA* (note the *E. coli* m¹G745 is resistant to RNase T1 digestion, producing a longer fragment). Before complementation with *yefA*, the rRNA from the *E. coli* Δ*rlmC*/Δ*rlmD*/Δ*trmA* mutant produced none of the peaks containing m⁵U (data not shown). The table shows relevant fragments (sequences 5′–3′) generated from digestion of *B. subtilis* (*Bs*) and/or *E. coli* (*Ec*) rRNAs with RNase A (leaving a 3′-U or -C) and RNase T1 (3′-G). RNase T1 produced a mixture of fragments with a 2′–3′-cyclic phosphate (>p) and a linear 3′-phosphate (p), the latter having an 18 Da larger mass; the theoretical monoisotopic mass/charge values (m/z) are given for the predominant peaks.

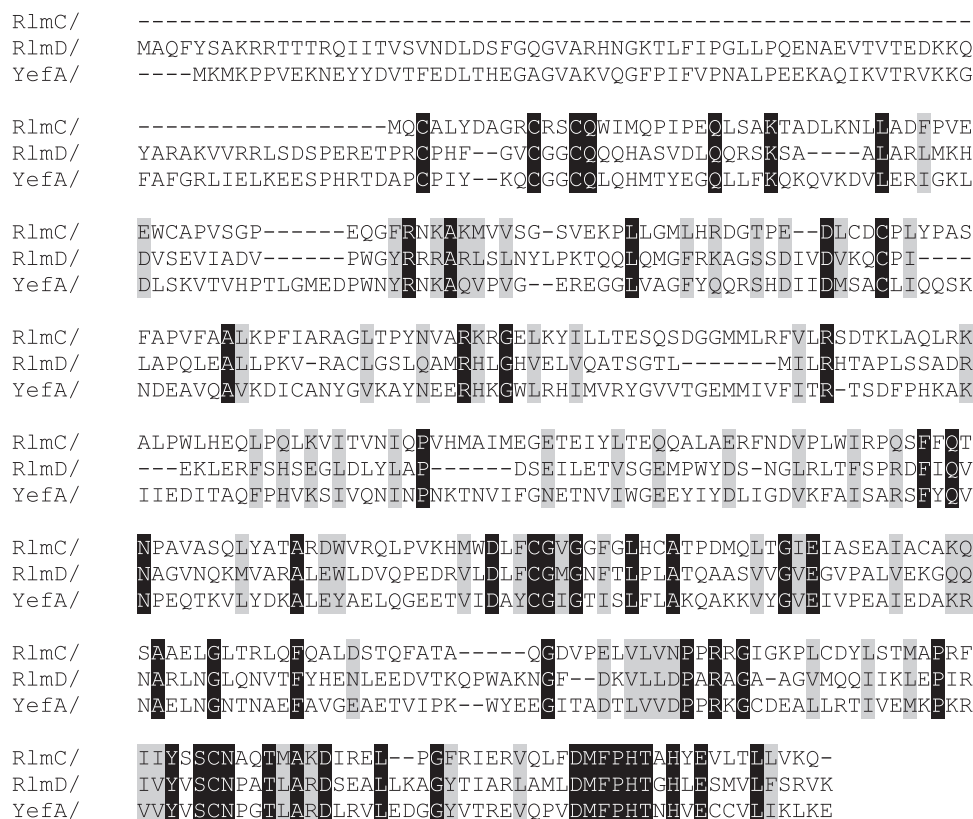T1 fragment (CACGUUGp) showed that the methyl group resided on nucleotide U747 (Supplementary Figure S1).

In domain IV of the rRNA from the wild-type *B. subtilis* strain (Figure 1B), RNase T1 digestion produced a unique decamer AAAUUCCUUGp containing nucleotide U1939. The MS peak at m/z 3199 (and at m/z 3181 with a cyclic phosphate) corresponds to the decamer plus a single methyl group (Figure 2B). The lack of any peak at m/z 3185 (and at m/z 3167 with the cyclic phosphate) indicated stoichiometric methylation of this sequence. The RNase A fragment GAAAUp (nucleotides 1935–1939) flew at m/z 1671 (Figure 2C) showing that this overlapping sequence also carried the methyl group. The exact location of the methylation site was pinpointed to nucleotide U1939 by tandem MS analysis of this sequence using the slightly longer fragment GAAAU Up formed by partial RNase A digestion (Supplementary Figure S2).

These regions of the *B. subtilis* rRNA were analysed by primer extension, which did not reveal any reverse transcriptase stops or pauses at nucleotides U747 and U1939 (data not shown), and ruled out that the methylations could be at the N3 of the bases or the 2′-OH of the riboses (21). These findings were fully consistent with the location of methyl groups added by YefA being at the C5-positions of nucleotides U747 and U1939.

Candidate methyltransferases for the m⁵U747 and m⁵U1939 modifications

A BLAST search of the *B. subtilis* genome using the sequence of the tRNA-modifying enzyme TrmFO as the query revealed no paralogue that might function in rRNA modification (25). Similar searches using the *E. coli* 23S rRNA m⁵U methyltransferases RlmC and RlmD, revealed two *B. subtilis* orthologues, YefA and YfjO. These two putative *B. subtilis* methyltransferases are 43% identical (68% similar) in their amino acid sequences (Supplementary Figure S3) and are clearly paralogues. Comparison with the *E. coli* methyltransferases revealed that YefA displays 22% and 30% identity (40% and 49% similarity) to RlmC and RlmD, respectively (Figure 3); YjfO is slightly less close to the *E. coli* sequences and shows corresponding identities of 19% and 24% (42% and 47% similarity). Phylogenetically, all the four proteins YefA, YfjO, RlmC and RlmD fit into the same cluster of orthologous groups COG2265 (11). The *yefA* and *yfjO* genes were therefore candidates (and, in fact, the only likely candidates) for encoding enzymes

```
RlmC/   ------------------------------------------------------------
RlmD/   MAQFYSAKRRTTTRQIITVSVNDLDSFGQGVARHNGKTLFIPGLLPQENAEVTVTEDKKQ
YefA/   ----MKMKPPVEKNEYYDVTFEDLTHEGAGVAKVQGFPIFVPNALPEEKAQIKVTRVKKG

RlmC/   ------------------MQCALYDAGRCRSCQWIMQPIPEQLSAKTADLKNLLADFPVE
RlmD/   YARAKVVRRLSDSPERETPRCPHF--GVCGGCQQQHASVDLQQRSKSA----ALARLMKH
YefA/   FAFGRLIELKEESPHRTDAPCPIY--KQCGGCQLQHMTYEGQLLFKQKQVKDVLERIGKL

RlmC/   EWCAPVSGP------EQGFRNKAKMVVSG-SVEKPLLGMLHRDGTPE--DLCDCPLYPAS
RlmD/   DVSEVIADV------PWGYRRRARLSLNYLPKTQQLQMGFRKAGSSDIVDVKQCPI----
YefA/   DLSKVTVHPTLGMEDPWNYRNKAQVPVG--EREGGLVAGFYQQRSHDIIDMSACLIQQSK

RlmC/   FAPVFAALKPFIARAGLTPYNVARKRGELKYILLTESQSDGGMMLRFVLRSDTKLAQLRK
RlmD/   LAPQLEALLPKV-RACLGSLQAMRHLGHVELVQATSGTL-------MILRHTAPLSSADR
YefA/   NDEAVQAVKDICANYGVKAYNEERHKGWLRHIMVRYGVVTGEMMIVFITR-TSDFPHKAK

RlmC/   ALPWLHEQLPQLKVITVNIQPVHMAIMEGETEIYLTEQQALAERFNDVPLWIRPQSFFQT
RlmD/   ---EKLERFSHSEGLDLYLAP------DSEILETVSGEMPWYDS-NGLRLTFSPRDFIQV
YefA/   IIEDITAQFPHVKSIVQNINPNKTNVIFGNETNVIWGEEYIYDLIGDVKFAISARSEYQV

RlmC/   NPAVASQLYATARDWVRQLPVKHMWDLFCGVGGFGLHCATPDMQLTGIEIASEAIACAKQ
RlmD/   NAGVNQKMVARALEWLDVQPEDRVLDLFCGMGNFTLPLATQAASVVCGVEGVPALVEKGQQ
YefA/   NPEQTKVLYDKALEYAELQGEETVIDAYCGIGTISLFLAKQAKKVYGVEIVPEAIEDAKR

RlmC/   SAAELGLTRLQFQALDSTQFATA-----QGDVPELVLVNPPRRGIGKPLCDYLSTMAPRF
RlmD/   NARLNGLQNVTFYHENLEEDVTKQPWAKNGF--DKVLLDPARAGA-AGVMQQIIKLEPIR
YefA/   NAELNGNTNAEFAVGEAETVIPK--WYEEGITADTLVVDPPRKGCDEALLRTIVEMKEKR

RlmC/   IIYSSCNAQTMAKDIREL--PGFRIERVQLFDMFPHTAHYEVLTLLVKQ-
RlmD/   IVYVSCNPATLARDSEALLKAGYTIARLAMLDMFPHTGHLESMVLFSRVK
YefA/   VVYVSCNPGTLARDLRVLEDGGYVTREVQPVDMFPHTNHVECCVLIKLKE
```

Figure 3. Amino acid sequence alignment of m⁵U methyltransferases. The *E. coli* RlmC and RlmD enzymes, which are respectively specific for 23S rRNA nucleotides U747 and U1939, are aligned with the *B. subtilis* YefA sequence. Identical amino acid residues are highlighted in black, and grey indicates highly similar residues. Full-length sequences are shown; RlmC lacks the N-terminal extension seen in RlmD and YefA. The common core sequences were used for calculating percentages of amino acid identity and similarity.

that could catalyse the m^{5}U modifications in *B. subtilis* 23S rRNA.

Identifying the enzyme(s) responsible for the *B. subtilis* m^{5}U747 and m^{5}U1939 modifications

Strains BG12826 (*yefA*::pMutin2) and BG12911 (*yfjO*:: pMutin2) were procured and then tested by PCR to established whether gene inactivation had been successful. The BG12911 (*yfjO*::pMutin2) strain produced PCR fragment patterns that were fully consistent with targeted inactivation of the *yfjO* gene; however, the (*yefA*::pMutin2) strain showed anomalies indicating that genes neighbouring *yefA* on the chromosome might also have been disturbed. We therefore made another *yefA* knockout ($\Delta yefA$::Cm) in-house using a different procedure; subsequent PCR analysis (data not shown) confirmed that *yefA* had been inactivated without disruption of any other gene.

Analysis of the rRNA from the $\Delta yefA$::Cm strain by MALDI-MS showed that methylation had been lost at both U747 (Figure 2D) and U1939 (Figure 2E and F). Parenthetically, we note that the initial *yefA* knockout, BG12826 (*yefA*::pMutin2), had also lost the methylations at both of these nucleotides (data not shown); however, we continued all recombinant work with our own $\Delta yefA$::Cm strain. Transforming this latter strain with an active copy of *yefA* on a plasmid restored to a significant degree the methylation at both U747 (Figure 2G) and U1939 (Figure 2H).

The dual methylation function of YefA was additionally tested in a strain of *E. coli* in which all three m^{5}U methyltransferase genes had been deleted. In this heterologous expression system, methylation was partially restored at both U747 (Figure 2I) and U1939 positions in the *E. coli* 23S rRNA (Figure 2J) with the latter site being more efficiently methylated. Methylated uridines from *E. coli* and *B. subtilis* tRNAs were analysed using an HPLC/ES-MS ion-trap system, which verified that the methyl group was on the C5-position of the uracil base (Supplementary Figure S4). Further HPLC analysis showed that YefA is not involved in m^{5}U modification of tRNAs (Supplementary Figure S5).

Parallel sets of experiments were carried out in attempts to identify the function of the YfjO enzyme. In the BG12911 strain, in which this enzyme is missing, nucleotides m^{5}U747 and m^{5}U1939 remained fully methylated. Furthermore, HPLC analyses of tRNAs from the *B. subtilis* strains showed that, similar to YefA, YfjO plays no role in tRNA m^{5}U modification (Supplementary Figure S5). At present therefore, the methylation target for the YfjO enzyme remains unknown.

DISCUSSION

The rRNAs of all living organisms are post-transcriptionally modified to improve their performance in protein synthesis (2,26), although there are marked differences in the mechanisms by which rRNA modification is achieved. In Archaea and Eukaryota, most rRNA modifications are uridine isomerizations and 2'-*O*-ribose methylations (8). These modifications are generally added by enzyme complexes that are guided to their nucleotide targets by complementary sequences in the sRNAs of Archaea (27,28) and the snoRNAs of Eukaryota (29–31), although some non-guided modifications have also been noted (32,33). In the rRNAs of Bacteria, uridine isomerizations and 2'-*O*-methylations also occur, although base methylation is the most frequent type of modification (4,34,35). None of the bacterial rRNA modifications requires the help of guide RNAs, and the general view has been that each modification is added by a single-specific enzyme that is capable of independently finding its target nucleotide.

On the whole, this view has been substantiated by studies of the numerous rRNA modifying enzymes in *E. coli* [reviewed in (34)] with only a few amendments being necessary. For instance, a handful of enzymes including the pseudouridine synthase RluD (20,36), the highly conserved m^{6}A dimethyltransferase RsmA/KsgA (37–39) and the m^{5}C methyltransferase RsmF of *Thermus thermophilus* (40) modify multiple sites that are immediately adjacent (RsmA) or within several nucleotides of each other on the bacterial rRNAs (RluD and RsmF). It could be argued that these enzymes engage in a single binding event with the rRNA during which a second or third nucleotide is accommodated into the active site without the need for the enzyme to dissociate from the substrate. Rare exceptions include the 2'-*O*-methyltransferase TlyA from *Mycobacterium tuberculosis*, which in separate reactions modifies nucleotides on both subunits of the ribosome (41), RluC which catalyses pseudouridine formation at distant locations in *E. coli* 23S rRNA (42), and RluA which forms pseudouridines in both the tRNAs and 23S rRNA of *E. coli* (43).

The data presented here show that *B. subtilis* YefA is able to recognize, bind and modify two targets that are at separate locations. From the MS analyses, it can be seen that *B. subtilis* 23S rRNA possesses the modifications m^{5}U 747 (Figure 2A) and m^{5}U1939 (Figure 2B and C) and that methylation is stoichiometric at each of the two nucleotides. Both methylations are lost upon inactivation of *yefA* and both are then restored upon complementation with an active copy of this gene. However, as it could formally have been argued that recognition by YefA of the two sites was facilitated by an auxiliary *B. subtilis* factor or was perhaps due to some peculiarity in the *B. subtilis* rRNA, a strain of *E. coli* was constructed in which all three of its m^{5}U methyltransferase genes (*rlmC, rlmD* and *trmA*) were deleted. Upon heterologous expression of *yefA* in this *E. coli* strain, methylation at U747 and U1939 was regained showing that the specific recognition of dual targets by YefA is a property inherent within the structure of this enzyme. Analyses of the tRNA component of these cells showed that YefA is not involved in tRNA methylation. Thus, in *B. subtilis,* YefA is specific for the 23S rRNA modifications m^{5}U747 and m^{5}U1939, while modification at m^{5}U54 tRNAs is added exclusively by the folate-dependent methyltransferase TrmFO (13,25).

In wild-type *E. coli*, the C5 positions of nucleotides U747 and U1939 are stoichiometrically methylated by the independent action of the enzymes RlmC and RlmD (9,10). The requirement for two autonomous m^{5}U rRNA

methyltransferases in *E. coli,* while a single enzyme can do the same job in *B. subtilis*, raises the question as to how the functions of YefA, RlmC and RlmD evolved. One likely scenario, consistent with phylogenetic analyses (3,5), would be that the contemporary *B. subtilis* YefA enzyme arose from a more promiscuous multi-site, AdoMet-dependent m⁵U rRNA methyltransferase that was an earlier (and now perhaps extinct) member of the COG2265 cluster of orthologous genes. Similarly, in *E. coli* and related species, single-site specificity evolved following duplication of an ancestral COG2265 gene giving rise to the *rlmC* and *rlmD* paralogues. The limited amino acid identity between the three enzyme sequences (Figure 3) favours the idea that separation of RlmC, RlmD and YefA was indeed an ancient event. In the context of functional shifts, we note that while YefA stoichiometrically methylates both of its *B. subtilis* U747 and U1939 targets (Figure 2A and B) and the recombinant YefA methylates with high efficiency at U1939 in *B. subtilis* (Figure 2H) and *E. coli* (Figure 2J), the enzyme was clearly less effective at recognizing U747 in *E. coli* where an appreciable residue of unmethylated nucleotide remained (Figure 2I). A precedence for target shifts in m⁵U RNA methyltransferases has already been observed in the archaeon *Pyrococcus abyssi* where two functional RlmD-like paralogues are present, and neither of these modified U1939 (3,5).

Although m⁵U RNA methyltransferases have changed their modification targets during evolution, the requirement for m⁵U modification at specific locations has remained remarkably constant. The majority of organisms appear to possess a homologue of either TrmA or TrmFO for m⁵U54 modification in tRNA (25). Likewise, the majority of bacteria encoded at least one RlmD gene, which fits with the observation that all bacteria studied to date possess the 23S rRNA m⁵U1939 methylation and many also have the m⁵U747 modification. The conservation of modified nucleotides at identical positions in rRNA would suggest that they play essential roles. Nevertheless, the *E. coli* knockout strain grew comfortably (in rich medium) despite the loss of all three of its m⁵U's and, even though the strain would be expected to be at a disadvantage if subjected to stresses including antibiotics (44), its resilience was noteworthy. Similarly, the deletion of *yefA* (or *yfjO*) in *B.subtilis* had no major effect on the viability of the strains under standard laboratory growth conditions.

A functional role for the paralogue YfjO remains to be determined. The YfjO enzyme does not play a role in m⁵U methylation of *B. subtilis* tRNAs (Supplementary Figure S4). The HPLC and ion-trap-MS approaches (Supplementary Figures S4 and S5) are less well suited to the analysis of m⁵U nucleotides in the larger rRNAs, where this modification makes up a much smaller a proportion of the total uridine complement and, although we did show by MALDI-MS that YfjO does not modify within the 747 or 1939 regions of 23S rRNA, we have not ruled out that this enzyme might have an rRNA modifying function. In fact, four *rlmD* orthologues and one *trmFO* orthologue are evident in the genome of *Acholeplasma laidlawii* (45), which belongs to the Mollicutes and is distantly related to *B. subtilis*, and this

fits well with an earlier analysis that indicated up to six m⁵U nucleotides at unidentified locations in the *A. laidlawii* 23S rRNA (46). The present evidence is consistent with certain members of the Gram-positive Firmicutes (which include *Bacillus* and *Acholeplasma*) having, in addition to U747 and U1939, m⁵U modification(s) in at least one other region of the 23S rRNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Björk,G.R. and Hagervall,T.G. (2005) Transfer RNA modification. In Böck,A., Curtis,R., Kaper,J.B., Neidhardt,T., Nyström,T. and Squires,C. (eds), *Escherichia coli and Salmonella*. ASM Press, Washington DC, p. 4.6.2.
2. Grosjean,H. (2009) *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution*. Landes Biosciences, Austin Texas.
3. Urbonavičius,J., Auxilien,S., Walbott,H., Trachana,K., Golinelli-Pimpaneau,B., Brochier-Armanet,C. and Grosjean,H. (2008) Acquisition of a bacterial RumA-type tRNA(uracil-54, C5)-methyltransferase by Archaea through an ancient horizontal gene transfer. *Mol. Microbiol.*, **67**, 323–335.
4. Ofengand,J. and Del Campo,M. (2005) Modified nucleotides of *E. coli* ribosomal RNA. In Böck,A., Curtis,R., Kaper,J.B., Neidhardt,T., Nyström,T. and Squires,C. (eds), *Escherichia coli and Salmonella*. ASM Press, Washington DC, p. 4.6.1.
5. Auxilien,S., Rasmussen,A., Rose,S., Brochier-Armanet,C., Husson,C., Fourmy,D., Grosjean,H. and Douthwaite,S. (2011) Specificity shifts in the rRNA and tRNA nucleotide targets of archaeal and bacterial m⁵U methyltransferases. *RNA*, **17**, 45–53.
6. Jühling,F., Mörl,M., Hartmann,R.K., Sprinzl,M., Stadler,P.F. and Pütz,J. (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.*, **37**, D159–D162.
7. Björk,G.R. (1975) Transductional mapping of gene *trmA* responsible for the production of 5methyluridine in

transfer ribonucleic acid of *Escherichia coli*. *J. Bacteriol.*, **124**, 92–98.

8. Piekna-Przybylska,D., Decatur,W.A. and Fournier,M.J. (2008) The 3D rRNA modification maps database: with interactive tools for ribosome analysis. *Nucleic Acids Res.*, **36**, D178–D183.

9. Agarwalla,S., Kealey,J.T., Santi,D.V. and Stroud,R.M. (2002) Characterization of the 23S ribosomal RNA m^{5}U1939 methyltransferase from *Escherichia coli*. *J. Biol. Chem.*, **277**, 8835–8840.

10. Madsen,C.T., Mengel-Jorgensen,J., Kirpekar,F. and Douthwaite,S. (2003) Identifying the methyltransferases for m^{5}U747 and m^{5}U1939 in 23S rRNA using MALDI mass spectrometry. *Nucleic Acids Res.*, **31**, 4738–4746.

11. Tatusov,R.L., Fedorova,N.D., Jackson,J.D., Jacobs,A.R., Kiryutin,B., Koonin,E.V., Krylov,D.M., Mazumder,R., Mekhedov,S.L., Nikolskaya,A.N. *et al.* (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, **4**, 41.

12. Nordlund,M.E., Johansson,J.O., von Pawel-Rammingen,U. and Byström,A.S. (2000) Identification of the TRM2 gene encoding the tRNA(m^{5}U54)methyltransferase of *Saccharomyces cerevisiae*. *RNA*, **6**, 844–860.

13. Urbonavičius,J., Skouloubris,S., Myllykallio,H. and Grosjean,H. (2005) Identification of a novel gene encoding a flavin-dependent tRNA: m^{5}U methyltransferase in bacteria-evolutionary implications. *Nucleic Acids Res.*, **33**, 3955–3964.

14. Vagner,V., Dervyn,E. and Ehrlich,S.D. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology*, **144**, 3097–3104.

15. Trieu-Cuot,P., de Cespedes,G. and Horaud,T. (1992) Nucleotide sequence of the chloramphenicol resistance determinant of the streptococcal plasmid pIP501. *Plasmid*, **28**, 272–276.

16. 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.

17. 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.

18. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Press, Cold Spring Harbor, NY.

19. Douthwaite,S., Powers,T., Lee,J.Y. and Noller,H.F. (1989) Defining the structural requirements for a helix in 23S ribosomal RNA that confers erythromycin resistance. *J. Mol. Biol.*, **209**, 655–665.

20. Andersen,T.E., Porse,B.T. and Kirpekar,F. (2004) A novel partial modification at C2501 in *Escherichia coli* 23S ribosomal RNA. *RNA*, **10**, 907–913.

21. Douthwaite,S. and Kirpekar,F. (2007) Identifying modifications in RNA by MALDI mass spectrometry. *Methods Enzymol.*, **425**, 1–20.

22. Kirpekar,F., Douthwaite,S. and Roepstorff,P. (2000) Mapping posttranscriptional modifications in 5S ribosomal RNA by MALDI mass spectrometry. *RNA*, **6**, 296–306.

23. Kirpekar,F. and Krogh,T.N. (2001) RNA fragmentation studied in a matrix-assisted laser desorption/ionisation tandem quadropole/orthogonal time-of-flight mass spectrometer. *Rapid Commun. Mass Sp.*, **15**, 8–14.

24. Giessing,A.M., Jensen,S.S., Rasmussen,A., Hansen,L.H., Gondela,A., Long,K., Vester,B. and Kirpekar,F. (2009) Identification of 8-methyladenosine as the modification catalyzed by the radical SAM methyltransferase Cfr that confers antibiotic resistance in bacteria. *RNA*, **15**, 327–336.

25. Urbonavičius,J., Brochier-Armanet,C., Skouloubris,S., Myllykallio,H. and Grosjean,H. (2007) In vitro detection of the enzyme activity of folate-dependent tRNA(U54, C5)methyltransferase. *Methods in Enzymology*, **425**, 103–119.

26. Decatur,W.A. and Fournier,M.J. (2002) rRNA modifications and ribosome function. *Trends Biochem. Sci.*, **27**, 344–351.

27. Dennis,P.P., Omer,A. and Lowe,T. (2001) A guided tour: small RNA function in Archaea. *Mol Microbiol.*, **40**, 509–519.

28. Tran,E., Brown,J. and Maxwell,E.S. (2004) Evolutionary origins of the RNA-guided nucleotide-modification complexes: from the primitive translation apparatus? *Trends Biochem. Sci.*, **29**, 343–350.

29. Bachellerie,J.P. and Cavaille,J. (1997) Guiding ribose methylation of rRNA. *Trends Biochem. Sci.*, **22**, 257–261.

30. Kiss,T. (2001) Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs. *EMBO J.*, **20**, 3617–3622.

31. Lafontaine,D.L. and Tollervey,D. (1998) Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. *Trends Biochem. Sci.*, **23**, 383–388.

32. Bonnerot,C., Pintard,L. and Lutfalla,G. (2003) Functional redundancy of Spb1p and a snR52dependent mechanism for the 2′-O-ribose methylation of a conserved rRNA position in yeast. *Mol. Cell.*, **12**, 1309–1315.

33. Decatur,W.A. and Schnare,M.N. (2008) Different mechanisms for pseudouridine formation in yeast 5S and 5.8S rRNAs. *Mol. Cell Biol.*, **28**, 3089–3100.

34. 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.

35. Cantara,W.A., Crain,P.F., Rozenski,J., McCloskey,J.A., Harris,K.A., Zhang,X., Vendeix,F.A., Fabris,D. and Agris,P.F. (2011) The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res.*, **39**, D195–D201.

36. Ero,R., Leppik,M., Liiv,A. and Remme,J. (2010) Specificity and kinetics of 23S rRNA modification enzymes RlmH and RluD. *RNA*, **16**, 2075–2084.

37. Helser,T.L., Davies,J.E. and Dahlberg,J.E. (1972) Mechanism of kasugamycin resistance in *Escherichia coli*. *Nat. New Biol.*, **235**, 6–9.

38. Demirci,H., Murphy,F.t., Belardinelli,R., Kelley,A.C., Ramakrishnan,V., Gregory,S.T., Dahlberg,A.E. and Jogl,G. (2010) Modification of 16S ribosomal RNA by the KsgA methyltransferase restructures the 30S subunit to optimize ribosome function. *RNA*, **16**, 2319–2324.

39. Desai,P.M., Culver,G.M. and Rife,J.P. (2011) Site-directed mutants of 16S rRNA reveal important RNA domains for KsgA function and 30S subunit assembly. *Biochemistry*, **50**, 854–863.

40. Demirci,H., Larsen,L.H., Hansen,T., Rasmussen,A., Cadambi,A., Gregory,S.T., Kirpekar,F. and Jogl,G. (2010) Multi-site-specific 16S rRNA methyltransferase RsmF from *Thermus thermophilus*. *RNA*, **16**, 1584–1596.

41. Johansen,S.K., Maus,C.E., Plikaytis,B.B. and Douthwaite,S. (2006) Capreomycin binds across the ribosomal subunit interface using *tlyA*-encoded 2′-O-methylations in 16S and 23S rRNAs. *Mol. Cell.*, **23**, 173–182.

42. Conrad,J., Sun,D., Englund,N. and Ofengand,J. (1998) The *rluC* gene of *Escherichia coli* codes for a pseudouridine synthase that is solely responsible for synthesis of pseudouridine at positions 955, 2504, and 2580 in 23S ribosomal RNA. *J. Biol, Chem.*, **273**, 18562–18566.

43. Wrzesinski,J., Nurse,K., Bakin,A., Lane,B.G. and Ofengand,J. (1995) A dual-specificity pseudouridine synthase: an *Escherichia coli* synthase purified and cloned on the basis of its specificity for Ψ746 in 23S RNA is also specific for Ψ32 in tRNA^{Phe}. *RNA*, **1**, 437–448.

44. Persaud,C., Lu,Y., Vila-Sanjurjo,A., Campbell,J.L., Finley,J. and O'Connor,M. (2010) Mutagenesis of the modified bases, m^{5}U1939 and Ψ2504, in *Escherichia coli* 23S rRNA. *Biochem. Biophys. Res. Commun.*, **392**, 223–227.

45. Kovaleva,G.Y., Kovaleva,G.Y., Kazanov,M.D., Malko,D.B., Vitreschak,A.G., Sernova,N.V., Gelfand,M.S., Lazarev,V.N., Levitskii,S.A., Basovskii,Y.I. *et al.* (2007) *Acholeplasma laidlawii* complete genome. *EMBL accession number CP000896*.

46. Hsuchen,C.C. and Dubin,D.T. (1980) Methylation patterns of mycoplasma transfer and ribosomal ribonucleic acid. *J. Bacteriol.*, **144**, 991–998.

47. Purta,E., O'Connor,M., Bujnicki,J. and Douthwaite,S. (2008) YccW is the m^{5}C methyltransferase specific for 23S rRNA nucleotide 1962. *J. Mol. Biol.*, **383**, 641–651.

48. Ny,T., Lindstrom,H.R., Hagervall,T.G. and Björk,G.R. (1988) Purification of transfer RNA m^{5}U54-methyltransferase from *Escherichia coli*. Association with RNA. *Eur. J. Biochem.*, **177**, 467–475.