

Formation of m²G6 in *Methanocaldococcus jannaschii* tRNA catalyzed by the novel methyltransferase Trm14

Sheena Menezes¹, Kirk W. Gaston², Kady L. Krivos², Ethel E. Apolinario³, Norbert O. Reich^{1,4}, Kevin R. Sowers³, Patrick A. Limbach² and John J. Perona^{1,4,*}

¹Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106-9510,

²Rieveschl Laboratories for Mass Spectrometry, Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221, ³Department of Marine Biotechnology, Institute of Marine and Environmental Technology, University of Maryland Baltimore County, Baltimore, MD 21202 and ⁴Interdepartmental Program in Biomolecular Science and Engineering, University of California, Santa Barbara, CA 93106-9510, USA

Received April 20, 2011; Revised May 22, 2011; Accepted May 23, 2011

ABSTRACT

The modified nucleosides N²-methylguanosine and N²-dimethylguanosine in transfer RNA occur at five positions in the D and anticodon arms, and at positions G6 and G7 in the acceptor stem. Trm1 and Trm11 enzymes are known to be responsible for several of the D/anticodon arm modifications, but methylases catalyzing post-transcriptional m²G synthesis in the acceptor stem are uncharacterized. Here, we report that the MJ0438 gene from *Methanocaldococcus jannaschii* encodes a novel S-adenosylmethionine-dependent methyltransferase, now identified as Trm14, which generates m²G at position 6 in tRNA^{Cys}. The 381 amino acid Trm14 protein possesses a canonical RNA recognition THUMP domain at the amino terminus, followed by a γ -class Rossmann fold amino-methyltransferase catalytic domain featuring the signature NPPY active site motif. Trm14 is associated with cluster of orthologous groups (COG) 0116, and most closely resembles the m²G10 tRNA methylase Trm11. Phylogenetic analysis reveals a canonical archaeal/bacterial evolutionary separation with 20–30% sequence identities between the two branches, but it is likely that the detailed functions of COG 0116 enzymes differ between the archaeal and bacterial domains. In the archaeal branch, the protein is found exclusively in thermophiles. More distantly related Trm14 homologs were also identified in eukaryotes known to possess the m²G6 tRNA modification.

INTRODUCTION

A large fraction of the over 100 modified nucleosides identified in RNA molecules are found in tRNA, substantially expanding the chemical diversity of the family (1,2). Methylation of the four canonical nucleotide bases at a wide variety of positions is particularly prominent among the known modifications (3). Among other functions, methylated tRNA bases are known to help stabilize tRNA structure (4), ensure decoding specificity during translation on the ribosome (5) and provide determinants for efficiency and selectivity in aminoacylation (6–9). All methyl groups added to tRNA bases are introduced by specific tRNA methyltransferase enzymes. In the great majority of cases, the methyl group is derived from S-adenosylmethionine (SAM), although examples are known in which folate and flavin adenine dinucleotide (FAD) are used instead (10). The SAM-dependent tRNA methyltransferases are divisible into five structural classes, each possessing a distinct α/β topology in the active site domain (11). Among these, the majority of the known enzymes possess the Class I structure featuring the ancient Rossmann fold.

Although the N²-methylguanosine modification has been found at positions 6, 7, 9, 10, 18, 26 and 27 of tRNA (1,12), the only modifying enzymes that have been characterized are those associated with methylation at positions G10, G26 and G27 (13–16). These three nucleotides are located in the tRNA vertical arm at the junction of the D and anticodon stems, where the structure deviates from canonical A-form helical geometry owing to a non-contiguous sugar–phosphate backbone and non-Watson–Crick hydrogen-bonding arrangement at the bridging 26–44 pair. The enzyme catalyzing dimethylation at G26, Trm1, produces the monomethylated

*To whom correspondence should be addressed. Tel: +805 893 7389; Fax: +805 893 4120; Email: perona@chem.ucsb.edu

species as a reaction intermediate, and possesses the Class I structure together with a unique 140 amino acid α/β C-terminal domain of unknown function (13,14). In at least one case, the *Aquifex aeolicus* Trm1, the enzyme exhibits multisite recognition properties: it is able to generate dimethylguanosine at both G26 and G27 (13). The archaeal enzyme catalyzing m^2G10 formation, Trm11, possesses an N-terminal THUMP domain (found in thioridine synthases, RNA methyltransferases and pseudouridine synthases) linked to a canonical Class I catalytic domain (15). Like m^2G26 synthesis by Trm1, Trm11-catalyzed formation of m^2G10 also proceeds through a monomethylated intermediate. In yeast, m^2G10 formation requires a second protein, designated Trm112, and possibly additional factors as well (16). Interestingly, the sequence and structural features in the tRNA which are required for monomethylation of G10 and dimethylation of G26 are the same, suggesting independent evolution of Trm1 and Trm11 toward their distinct positional specificities.

We sought to further our understanding of m^2G incorporation into tRNA by examining archaeal genomes for additional candidate enzymes. We focused our search primarily on methanogens, since biochemical studies have identified the presence of m^2G in tRNA isolated from the hyperthermophiles *Methanocaldococcus jannaschii*, *Methanococcus thermolithotrophicus*, *Methanococcus igneus* and *Methanopyrus kandleri* (17,18), the mesophiles *Methanococcus vanielli* and *Methanococcus maripaludis* (17), and the psychrotolerant *Methanococcus burtonii* (19). Among other Archaea, m^2G has been detected in tRNA isolated from *Pyrococcus furiosus*, *Thermoproteus neutrophilus*, *Haloferax volcanii* (previously named *Halobacterium volcanii*) and *Pyrolobus fumarii* (20–23). Study of archaeal tRNA modification has provided insight into the roles of the modified nucleosides in stabilizing tRNA structure at the elevated temperatures where hyperthermophiles flourish. In general, it appears that archaeal tRNA modifications are more similar to their counterparts in eukarya than in bacteria, yet are often simpler in structure than those present in either of the other domains (17,19,24,25).

Identification of novel tRNA methylases in methanogens is also of interest because, together with the related non-methanogenic euryarchaeote *Archaeoglobus fulgidus*, these organisms are unique in retaining an ancient two-step pathway for formation of Cys-tRNA^{Cys} (26). Phosphoseryl (Sep)-tRNA^{Cys} is first generated by phosphoseryl-tRNA synthetase (SepRS), and this intermediate is then processed by the sulfur-donating PLP-dependent enzyme SepCysS to generate Cys-tRNA^{Cys}. Interestingly, a subset of methanogens retain the canonical CysRS as well, thereby possessing redundant pathways for Cys-tRNA^{Cys} formation. A direct link between Cys-tRNA^{Cys} formation and tRNA methylation has been demonstrated: SepRS enzymes from both *M. jannaschii* and *Methanosarcina mazei*, and the CysRS enzyme from *M. mazei*, each require SAM-dependent formation of m^1G37 by Trm5 for efficient aminoacylation of tRNA^{Cys} (7–9,27). The strong phylogenetic covariation of SepRS with methanogenesis (28), the requirement for

SAM-dependent tRNA methylation by both SepRS and CysRS, the 2-fold elevated cysteine levels in methanogen proteins as compared with other Archaea (29), the tRNA-linked cysteine biosynthesis pathway in some methanogens (26), and the dependence of SepCysS function on an unknown sulfur donor *in vivo* (8), together strongly suggest the existence of uncharacterized linkages among sulfur metabolism, protein synthesis and methanogenesis.

Using a bioinformatics-based approach, we identified a cluster of orthologous genes (COG) 0116 in genomes of hyperthermophilic methanogens, and in some other Archaea, Bacteria and Eukarya (30). Although COG 0116 is annotated as an N⁶-adenine-specific DNA methylase (31), the presence of a THUMP domain common to RNA methylases, thioridine synthases and pseudouridine synthases instead strongly suggests activity toward RNA substrates (32). Cloning and characterization of the MJ0438 gene product from *M. jannaschii*, a representative member of the archaeal clade, demonstrates that this protein catalyzes formation of m^2G6 in *M. jannaschii* tRNA^{Cys}. This is the first identification of the enzyme responsible for m^2G6 formation in tRNA.

MATERIALS AND METHODS

Expression and purification of recombinant Trm14

Methanocaldococcus jannaschii genomic DNA was purchased from the American Type Culture Collection, and used as template for polymerase chain reaction (PCR) amplification of the open reading frame corresponding to the MJ 0438 gene, which we designate *trm14*. The primers used were: 5'-AAACTGCATATGGATTACTATGTTCACCTATCC (forward primer) and 5'-CCGCTCGAGAGTTAAATAAACACC (reverse primer). The reaction was carried out using 100 ng of genomic DNA, 20 μ M of each primer (forward and reverse), 250 μ M dNTPs, 10 \times Pfu turbo reaction buffer and 2.5 U of Pfu turbo DNA polymerase, at 51°C and 30 s per cycle for a total of 32 cycles. The PCR product containing the *trm14* gene was digested with NdeI and XhoI, and inserted into the pet22b⁺ vector (Novagen) for expression of C-terminal His-tagged protein in *Escherichia coli* Rosetta2(DE3) pLysS cells. Cells were grown at 37° in Luria Broth (LB) medium supplemented with 100 μ g/ml ampicillin and 34 μ g/ml chloramphenicol. Cells were induced when the optical density at 600 nm reached 0.7, by addition of isopropyl- β -D-thiogalactoside (IPTG) to a final concentration of 1 mM, and were then grown for an additional 6 h prior to harvesting.

Cells expressing Trm14 were resuspended in a buffer containing 50 mM NaH₂PO₄ (pH 7.8), 0.8 M NaCl, 15 mM β -mercaptoethanol and 10 mM imidazole, and were disrupted by sonication. The lysate was applied to a Ni-NTA agarose (Qiagen) column. Two wash steps were performed at increasing imidazole concentrations of 45 mM and 60 mM, prior to elution of Trm14 in buffer containing 250 mM imidazole. The enzyme was then dialyzed into a solution containing 50 mM NaH₂PO₄ (pH 7.8), 150 mM NaCl, 15 mM β -mercaptoethanol and 50% glycerol, and stored at –20°C. Trm14 was recovered

at ~98% purity as judged on Coomassie-stained SDS–polyacrylamide gels, at a yield of roughly 2 mg of purified protein per liter of culture.

Preparation of tRNA substrates

Wild-type and mutant forms of *M. jannaschii* tRNA^{Cys}, *M. jannaschii* tRNA^{Asp} and *M. jannaschii* tRNA^{Pro(TGG)} were transcribed from synthetic duplex DNA templates. The templates were each first synthesized from two overlapping synthetic deoxyribonucleotides (purchased from Fisher Operon). For the wild-type tRNA^{Cys} gene, the oligonucleotides used were: 5'-AAT TCC TGC AGT AAT ACG ACT CAC TAT AGC CGG GGT AGT CTA GGG GCT AGG CAG CGG ACT G (forward primer), and 5'-TGG AGC CGG GGG TGG GAT TTG AAC CCA CGT AAG GCG GAT CTG CAG TCC GCT GCC TAGC (reverse primer), where the underlined portions represent the overlap region. Primers used for the other tRNAs are provided in Supplementary Data. Overlapping DNAs were extended using Klenow fragment of *E. coli* DNA polymerase I, as described (33), and were recovered by ethanol precipitation. *In vitro* transcription reactions were then performed as described (33). Typically, 0.1 mg of DNA template was used in the transcription reaction to generate 1 mg of tRNA. tRNA^{Cys} was purified by gel extraction and stored as an ethanol precipitate or in purified water.

In vitro methylation reactions

Except where otherwise specified, methylation reactions included the purified Trm14 protein at a final concentration of 10 μ M. Reactions were performed at 52°C in 0.2 M Tris–HCl (pH 8.0), 0.8 mM DTT, 1.2 mM MgCl₂, 20 mM KCl, 4.8 μ g/ml bovine serum albumen (BSA), using 1 μ M *M. jannaschii* tRNA^{Cys} transcript as the substrate and 100 μ M SAM [a mixture containing 7.55 μ M of 68 Ci/mmol [³H]SAM (Perkin Elmer) and 92.45 μ M unlabeled SAM (Sigma)]. Prior to methylation, the tRNA was heated at 85°C for 3 min and allowed to renature by addition of 5 mM MgCl₂ followed by slow cooling to room temperature. At various time points, 5 μ l aliquots were removed, precipitated with 5% (w/v) trichloroacetic acid (TCA) on filter pads, washed with ethanol and ether and air-dried before scintillation counting. These methylation conditions were also used to assay methylation of tRNA^{Cys} mutants and WT tRNA^{Pro} and tRNA^{Asp} substrates.

Modified nucleotide analysis

Radiolabeled tRNA^{Cys} was synthesized *in vitro* as described above, except that the transcription reactions also included 125 μ Ci [α -³²P] GTP (3000 Ci/mmol). Trm14 reactions were then performed, and the labeled, methylated tRNA recovered by phenol/chloroform and ethanol precipitation. Thirty μ g of methylated tRNA was resuspended in 5 μ l water and digested with 3 μ l nuclease P1 (1 mg/ml) at 37°C for 1 h to generate 5'-phosphorylated nucleotides. A 2 μ l aliquot was then applied to a thin layer cellulose (TLC) plate (Sigma, 801 063). Chromatography in the first dimension was carried out in isobutyric acid/ammonia/water (66/1/33 [v/v/v])

(solvent A), while the second dimension was developed with isopropanol/concentrated HCl/water (68/18/14 [v/v/v]) (solvent C). TLC plates were analyzed by phosphorimaging, and the modified and unmodified nucleotides were identified using reference maps (34).

For liquid chromatography–mass spectrometry (LC–MS) analysis of modified nucleosides, nuclease P1 (Sigma), snake venom phosphodiesterase I (Worthington Biochemical Corporation) and Antarctic phosphatase (New England Biolabs) were used to digest Trm14-reacted transcripts and control unmethylated transcripts to nucleosides. The nucleosides were separated using a Hitachi D-7000 HPLC with a Hitachi L-7400 UV detector at 0.3 ml/min at room temperature on an LC-18 S 2.1 \times 250 mm column from Supelco using a gradient of 5 mM ammonium acetate pH 5.3 and acetonitrile:water (40:60 v/v) as described (35). The eluent from the column was split with two-thirds flowing into the UV detector and one-third flowing into a Thermo LTQ-XL (Waltham, MA, USA) equipped with an electrospray ionization (ESI) source. Mass spectra were recorded in the positive ion mode at a capillary temperature of 275°C, spray voltage of 3.7–4 kV and sheath gas, auxiliary gas, and sweep gas set to 45, 25 and 10 arbitrary units, respectively.

LC–MS/MS analysis of RNase digests

Transcripts were incubated with 0.01 U of RNase A per microgram of RNA in 20 mM ammonium acetate for 2 h at 37°C. Digestion products from 1 μ g of RNA were separated using a Thermo Surveyor HPLC system with an Xterra C18 1.0 \times 150 mm column (Waters) at room temperature with a flow rate of 40 μ l/min. Before each run the column was equilibrated for 15 min at 95% Buffer A [200 mM 1,1,1,3,3,3-hexafluoroisopropanol (HFIP), 8.15 mM triethylamine (TEA)] and 5% Buffer B (200 mM HFIP, 8.15 mM TEA:methanol 50:50 v:v). Both buffers were adjusted to pH 7.0. The gradient used was 5% B for 5 min, 30% B at 7 min and 95% B at 50 min and held at 95% B for 5 min. The eluent was directed into a Thermo LTQ-XL. Operating parameters were capillary temperature of 275°C, spray voltage of 4.5 kV, sheath gas at 25 arbitrary units, auxiliary gas at 14 arbitrary units and sweep gas to 10 arbitrary units. Collision-induced dissociation (CID) tandem mass spectrometry (normalized collision energy 35%) was used in data-dependent mode to obtain sequence information from the RNase A digestion products. The data-dependent scan was recorded based on the most abundant ion, and each ion selected for CID was analyzed for 30 scans or 30 s before it was added to a dynamic exclusion list for 30 s.

Cell strain and growth

Methanocaldococcus jannaschii strain JAL-1 was obtained from the Oregon Collection of Methanogens (OCM 168 = DSM 2661) and maintained in the laboratory as frozen stocks in DSM 282 medium with 25% glycerol. Cultures were routinely grown in liquid medium (DSM 282) prepared anaerobically by a modification of the Hungate technique (36). Medium was dispensed in 25 ml aliquots into 160-ml serum bottles, and sealed under an

$\text{N}_2\text{-CO}_2$ (4:1) atmosphere with butyl rubber stoppers secured with aluminum crimp seals. $\text{Na}_2\text{S}\cdot 9\text{H}_2\text{O}$ was added to a final concentration of 0.025% immediately prior to inoculation. After inoculation the gas phase was replaced with $\text{H}_2\text{-CO}_2$ (80:20) pressurized to 101 kPa. Cultures were incubated on a rotary shaker at 250 rpm at 83°C. Growth was routinely monitored by measuring methane production and gas phase was repressurized with $\text{H}_2\text{-CO}_2$ (80:20) whenever a vacuum was detected from depletion of substrate. For preparation of cell material, cultures were scaled up by inoculating six 25 ml cultures into 16 l of selenium enriched medium in a 20 l Bioflo IV bioreactor (New Brunswick Scientific, Edison NJ) (37). The culture was subsequently transferred anaerobically to 210 l of medium in a 250 l bioreactor (New Brunswick Scientific). Both bioreactors were maintained at 83°C and pH 6.0 with a $\text{H}_2\text{-CO}_2$ (80:20) flow-rate of 0.2 vvm at 1.7×10^5 Pa. Cells from a 250 l fermentor were harvested in late exponential growth by centrifugation in a 2 l capacity continuous flow centrifuge (CEPA Carl Padberg Zentrifugenbau GmbH, Germany) and stored in liquid nitrogen.

Methylation activity in cell extracts

Methanocaldococcus jannaschii cells (2.5 g) were resuspended in 3 ml of a solution containing 20 mM Tris-HCl (pH 7.4) and 50 mM NaCl, and were frozen overnight at -20°C. The cells were thawed, resuspended and sonicated at 50% duty cycle for 2 min. The lysate was then ultracentrifuged at 80 000 rcf using a 70.1 Ti rotor (Beckman Coulter). The supernatant was recovered and concentrated to 2 ml using spin columns with a 30 000 Da cut-off to partially remove endogenous tRNA, thereby reducing background methylation activity. The extract was then dialyzed into a buffer containing 10 mM Tris-HCl (pH 7.3), 20 mM β -mercaptoethanol, 1 mM EDTA, 4 mM MgCl_2 and 15% glycerol for 8 h at 4°C, and concentrated to 100 mg/ml protein, as estimated from absorbance at 280 nm.

Methylation reactions were conducted in 50 μ l volumes using conditions detailed above, 1 μ M WT or G6A/C67U tRNA^{Cys} transcript as substrate and 200 μ g (2 μ l) of protein from *M. jannaschii* extract as the source of Trm14 activity. *Escherichia coli* S100 cell extracts were prepared by standard approaches, and methylation reactions were conducted using extract equivalent to 500 μ g protein, 40 U of anti-RNASin, 3 μ M WT or G6A/C67U *M. jannaschii* tRNA^{Cys} and 50 μ M SAM in 50 μ l reactions carried out at 37°C (38).

Bioinformatics

Initial comparative analysis of a limited subset of methanogen genomes was performed using STRING (v 6.2) (39). Sequence homology searches to the omniome or selected subsets of the omniome were performed with BLAST (40). For phylogenetic analysis, the multiple sequence alignment for a subset of proteins within COG 0116 was carried out using ClustalW (41). The phylogenetic tree based on these sequences was constructed with MEGA5 (42), using the neighbor-joining method with the

BioNJ(JTT) model. The bootstrap test (1000 replicates) was performed to validate the stability of the nodes.

RESULTS AND DISCUSSION

Identification, recombinant expression and purification of Trm14

A bioinformatics-based approach to identify new putative methylase activities in methanogens was taken. To link the search to unique aspects of protein synthesis in these organisms, we elected to probe genomes for the presence of RNA methylases that covary with the presence or absence of a canonical CysRS enzyme. The initial search in six methanogen genomes was carried out using STRING, a bioinformatics resource developed to identify physical and functional interactions among proteins (39). The genomes probed were: *M. jannaschii*, *M. kandleri*, *M. thermoautotrophicus*, *M. mazei*, *M. acetivorans* and *M. maripaludis*. SQL (structured query language) queries within STRING were used to interrogate these six genomes to determine which COGs were present in all three of the organisms *M. jannaschii*, *M. kandleri* and *M. thermoautotrophicus*, and absent in each of the organisms *M. mazei*, *M. acetivorans* and *M. maripaludis*. The latter group possesses both SepRS and CysRS, while the former group possesses only SepRS.

Eight COGs were identified by this approach, providing a basis for abundant further analysis to identify aspects of methanogen metabolism related to how cysteine is incorporated into proteins. For this study, we focused on COG 1818, annotated as a predicted RNA binding protein containing the THUMP domain. Seven proteins in *M. jannaschii*, *M. kandleri* and *M. thermoautotrophicus* are found in COG 1818. All seven proteins were then used as queries in BLAST searches of all methanogen genomes. The search in which a COG 1818 protein from *M. kandleri* (2320.Q8TVD0) was used as query revealed a predicted N⁶-adenine specific RNA methylase containing a THUMP domain (MK0969) also from *M. kandleri*. Subsequent BLAST searches of all methanogen genomes using MK0969 as query sequence then revealed that a highly similar protein is present in a subset of these organisms, all of which lack CysRS (Supplementary Data). Some methanogens, such as *M. burtonii* and *M. thermoautotrophicus*, lack both CysRS and the MK0969 homolog. An MK0969 homolog is never present when both SepRS and CysRS are found in a methanogen genome. These characteristics suggested that the function of the protein might be related to the loss of CysRS from a methanogen genome. Phylogenetic analyses suggest that the last universal common ancestor (LUCA) possessed both CysRS and SepRS; it appears that SepRS has since been lost everywhere except in most methanogens and the related *A. fulgidus* (43).

In this manner, we identified eight methanogens possessing a predicted RNA methylase homologous to MK0969, all of which are thermophiles or hyperthermophiles (Supplementary Data). These proteins are members of COG 0116. We elected to study the MJ0438 homolog from *M. jannaschii*, because this is the best-studied

methanogen in the group and can be cultured in the laboratory (36,37). MJ0438 possesses C-terminal methylase catalytic motifs identifying it as a member of the Rossmann fold methylase family, and an N-terminal THUMP domain. Based on these characteristics, MJ0438 was previously identified as a representative of a large monophyletic family of proteins involved in the N-methylation of RNA bases (44). It is renamed here as Trm14.

We cloned the *trm14* gene by PCR amplification of *M. jannaschii* genomic DNA, inserted it into the *E. coli* expression vector pet22b⁺ for expression of His-tagged protein in Rosetta2(DE3) pLysS cells, and purified the protein using metal ion affinity chromatography. The purified 381 amino acid protein showed an apparent molecular mass of 43 kDa on sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) which is consistent with the predicted mass from the gene sequence (Figure 1A).

tRNA methyltransferase activity of Trm14

We tested the activity of recombinant, purified Trm14 on a variety of DNA and RNA substrates, using an assay in which the methyl group from [³H]SAM is incorporated into nucleic acid, and the acid-precipitable labeled product then isolated on filter paper and quantitated by liquid scintillation counting. As expected given the presence of a THUMP domain in the protein, no

methyltransferase activity was detected toward a 56 nt single-stranded oligodeoxynucleotide (Figure 1B). Plasmid DNA did not function as a Trm14 substrate, nor did a 37-mer chemically synthesized RNA corresponding to a bacterial tRNA^{Gln} 3'-half molecule (data not shown). To examine whether Trm14 possesses activity toward rRNA, total native RNA was isolated from *M. mazei*, which lacks a Trm14 homolog, and the rRNA was separated from tRNA by gel filtration chromatography. Large and small-subunit rRNAs in *M. mazei* and *M. jannaschii* are 77% and 87% identical, respectively. No activity was detected toward *M. mazei* rRNA even at very high concentrations of both Trm14 and substrate (10 μM Trm14; 1 μM rRNA) (Figure 1B).

In contrast, transfer of ³H-methyl groups from SAM was readily detected using a purified and renatured *M. jannaschii* tRNA^{Cys} transcript as substrate (Figures 1B, 1C and 2). Methylation reaction conditions were optimized with respect to the identity and concentration of the buffer, pH of the reaction and temperature: low methylation activity was observed at 37°C, while the activity was considerably higher at 52°C, the highest temperature compatible with maintenance of tertiary structure in the tRNA transcript. Under conditions of enzyme molar excess, we observed that the activity toward *M. jannaschii* tRNA^{Cys} approaches the incorporation of one methyl group per tRNA molecule (Figures 1B and 2). An *M. jannaschii* tRNA^{Asp} transcript also could be methylated, albeit with reduced efficiency as compared with tRNA^{Cys} (Figure 2C). In contrast, methylation of tRNA^{Pro(TGG)} was detectable only slightly above background levels. Together, the data show that Trm14 is able to specifically catalyze methyl group incorporation into tRNA *in vitro*, with variation in the efficiency of methylation depending on the identity of the isoacceptor species.

To more rigorously assess the function of Trm14 as a tRNA methylase, we carried out methylation reactions using *M. jannaschii* tRNA^{Cys} at a variety of protein concentrations and with saturating concentrations (50 μM) of SAM (Figure 2A and B). Only very weak activity could be detected under multiple-turnover conditions corresponding to 100 nM enzyme and 1 μM tRNA. However, robust methylation was observed when the enzyme is present in molar excess over its tRNA substrate. The concentration of Trm14 was varied while maintaining tRNA^{Cys} at 1 μM, and reaction time courses were monitored in each case (Figure 2A). Replot of the variation in the extracted *k*_{obs} with Trm14 concentration shows saturation, with a maximum value for the rate constant of about 0.05 min⁻¹ (Figure 2B).

The robust concentration-dependent behavior of recombinant Trm14 in catalyzing methylation of tRNA^{Cys} provides substantial confidence that this activity is physiologically relevant. The data also suggest that Trm14 may function as a single-turnover catalyst *in vivo*. This is consistent with the notion that tRNA structure in hyperthermophiles may be stabilized by protein binding: the relatively high intracellular Trm14 concentrations that may be needed for tRNA methylation may also serve to protect the molecule from thermal denaturation (25). However, an alternative rationale for the very weak

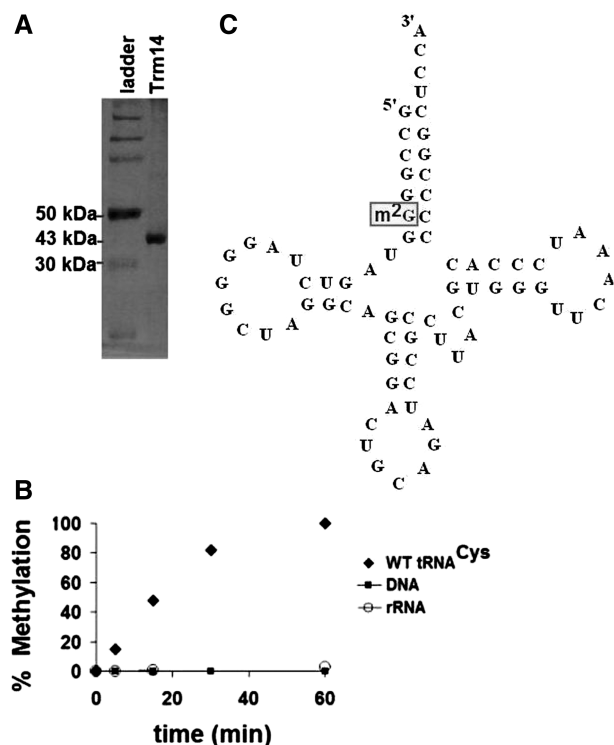


Figure 1. (A) SDS gel showing purified, recombinant Trm14 produced from *E. coli*. (B) Time courses for methylation of nucleic acid substrates by Trm14. The DNA substrate consisted of a 56-mer oligodeoxynucleotide. rRNA was purified from *M. mazei* cells. (C) Cloverleaf secondary structure representation of *M. jannaschii* tRNA^{Cys}.

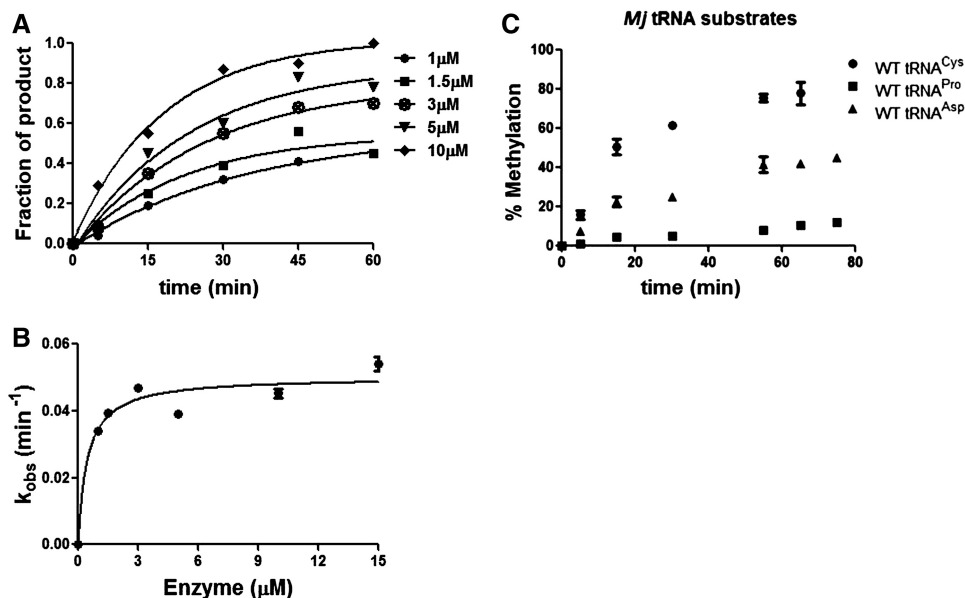


Figure 2. (A) Kinetic analysis of *M. jannaschii* Trm14 toward an unmodified tRNA^{Cys} transcript. Single-turnover reaction time courses carried out at 50 μM SAM, 1 μM tRNA^{Cys} transcript and varying concentrations of Trm14. For each time course k_{obs} is derived by fitting the data to a single exponential function. (B) Replot of k_{obs} from part A, as a function of enzyme concentration. (C) Comparative methylation of tRNA transcripts by Trm14.

activity of Trm14 under steady-state conditions is that another factor or condition present *in vivo* might be missing in these *in vitro* experiments. For example, the archaease protein from the hyperthermophile *Pyrococcus abyssi* increases the substrate specificity of a tRNA m⁵C methyltransferase (45); a homolog of this protein exists in the *M. jannaschii* genome. Comparable time courses for SAM-dependent methyl group incorporation were reported for m²G26 and m²G10 formation by *Pyrococcus horikoshii* Trm1 and *P. abyssi* Trm11, respectively (14,46).

Trm14 is a novel m²G6 tRNA methylase

We employed a combination of LC-MS/MS analysis and two-dimensional thin-layer chromatography (2D-TLC) to identify the product of methyl transfer by Trm14. Methylated and unmethylated tRNA^{Cys} transcripts were first digested to nucleosides with RNase P1, snake venom phosphodiesterase and Antarctic phosphatase, and the resulting mixture was separated and analyzed by LC-MS. Splitting the flow from the high pressure liquid chromatograph (HPLC) into a UV monitor and mass spectrometer allowed monitoring of elution time together with the mass spectra of the peaks (35). Nine UV-absorbing peaks were identified from the methylated tRNA (Figure 3A): the four standard nucleosides of RNA and DNA, together with one modified nucleoside. The modified nucleoside eluted at 26.07 min with a molecular mass of 297 Da ($[M+H]^+$ m/z 298). No other species at this m/z were detected in the chromatogram (Figure 3B). The selected ion chromatogram (SIC) of the modified base ion, m/z 166 (BH_2^+) was also recorded (Figure 3C). The mass difference of 132 u between the molecular ion and base ion is consistent with location of the modification on the base and not the ribose of the nucleoside. These m/z

values and the elution time best match a methylated guanosine with an expected m/z of 298 for the nucleoside and 166 for the base. Either m¹G or m²G match the elution time within the experimental precision of the measurements.

To distinguish whether the modified base is m¹G or m²G, we employed 2D-TLC analysis. The *M. jannaschii* tRNA^{Cys} transcript was synthesized by T7 RNA polymerase in reaction mixtures supplemented with [α -³²P] GTP for internal labeling. The transcript was purified, refolded *in vitro* and methylated by Trm14. Next, the methylated transcript was digested with P1 nuclease to generate 5'-monophosphorylated nucleosides, which were then separated using two-dimensional thin-layer chromatography (2D-TLC) (Figure 4). Comparison of the migration patterns of labeled 5'-monophosphorylated nucleosides in unmethylated and methylated samples revealed the presence of a single modified base, and comparison with established reference maps in this solvent system demonstrates that the modified base is m²G (34). All of the methylated guanosine nucleosides found in unfractionated *M. jannaschii* tRNA hydrosylates, and that can be generated by the activity of a single methyltransferase, are present on the reference maps (17). The plateau methylation level equivalent to transfer of one methyl group per tRNA^{Cys} substrate molecule (Figure 2) is consistent with identification of a monomethylated product.

To identify the position of m²G formation in the tRNA, transcripts were digested with RNase A and subsequently analyzed by LC-MS/MS (47–52). Upon elution from the C18 column, the digestion products were visualized via the total ion chromatogram (TIC) (Figure 5A). Comparing the TIC of the unmodified and modified RNAs (Figure 5A and 5B) revealed an additional digestion product in the modified sample eluting between

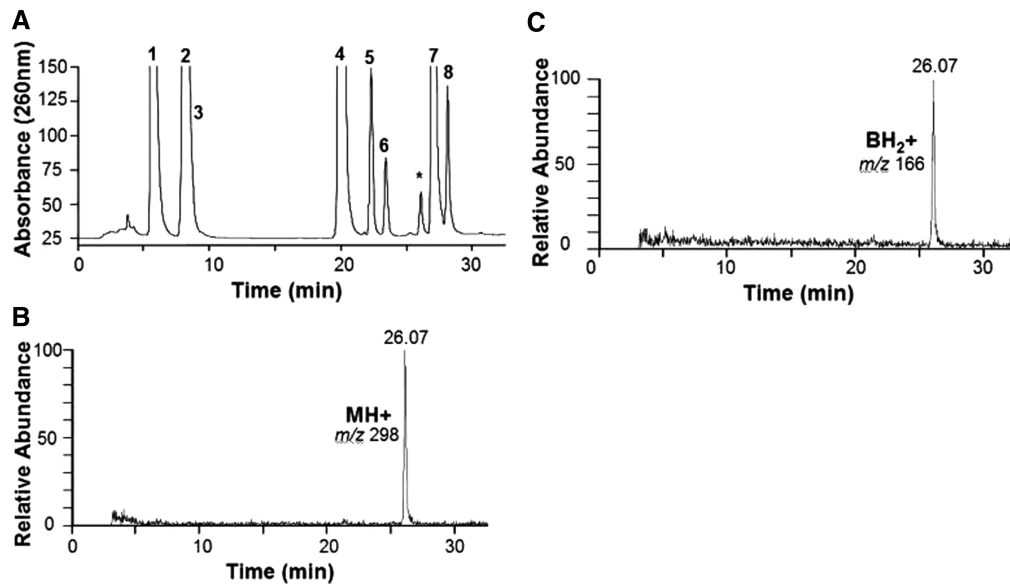


Figure 3. LC-MS analysis of the nucleosides contained in *M. jannaschii* tRNA^{Cys} after reaction with Trm14. (A) UV chromatogram of the nucleosides contained in the Trm14-reacted tRNA. The numbers found in this chromatogram designate the nucleosides: 1-C, 2-U, 3-dC, 4-G, 5-dT, 6-dG, 7-A, 8-dA and asterisks (a modified nucleoside later identified as m²G). (B) SIC of the nucleoside ion (MH⁺ *m/z* 298) found to elute as the same time as the modified nucleoside asterisks at 26.07 min. (C) SIC of the base ion (BH₂⁺ *m/z* 166) that is found at the same elution time as the molecular ion found in B.

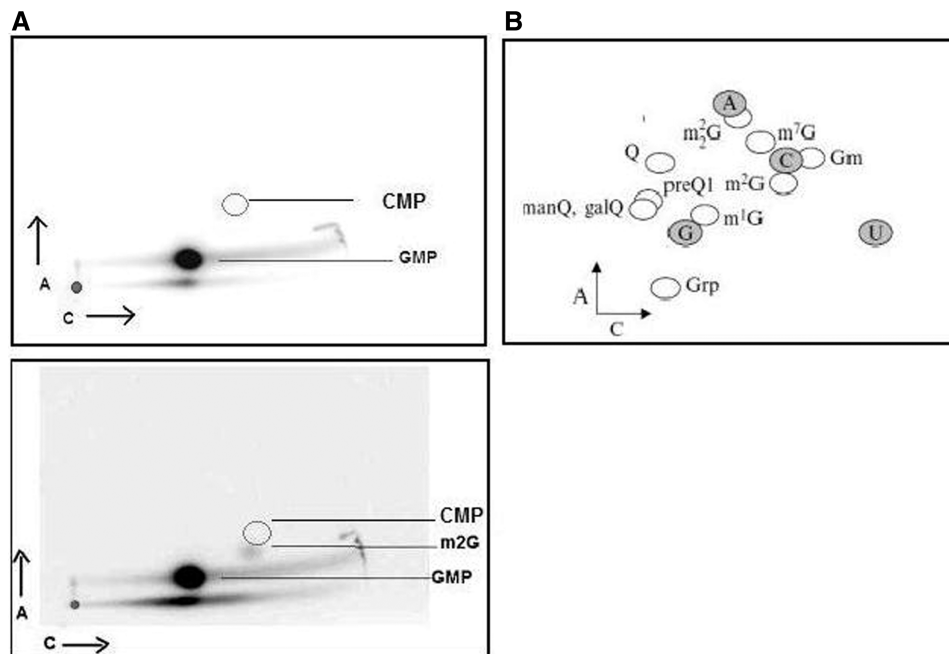


Figure 4. 2D TLC separation of a P1 nuclease hydrolysate of tRNA^{Cys} methylated with Trm14. Standard fingerprints are provided for comparison. (A) TLC separation of the P1 hydrolysate from unmethylated (top) and Trm14-methylated samples (bottom). (B) Reference map of modified GMPs in the same solvent system used in panel (A). These data indicate that m²G and not m¹G is the modified base produced in the Trm14 reaction.

22 and 23 min that is not present in the unmethylated control. The major peak eluting around 22 min in both sample was analyzed (Figure 5C) and found to contain two digestion products that co-elute with *m/z* 851.25 and 654.75 and correspond to the doubly charged ions for GGGGUp and AAAUp, respectively (Figure 5D).

The additional digestion product eluting between 22 and 23 min that appears when the RNA is treated with Trm14 was analyzed (Figure 5E) and was found to contain a doubly charged ion at *m/z* 858.17 (Figure 5F). The difference in *m/z* values between the closely eluting *m/z* 851 peak (GGGGUp) and this *m/z* 858 peak is

consistent with the addition of a methyl group to the GGG GUp digestion product. To determine the location of this methyl group, data-dependent CID for tandem mass spectrometry (LC-MS/MS) was performed. Fragmentation of the m/z 851 digestion product was consistent with the sequence GGGGUp with the identification of many of the expected fragmentation products (y -, w - and c -ions) (Figure 6A). Fragmentation of the m/z 858 digestion product shows a mass shift in the c_3 , c_4 , y_3 , w_3 and w_4 fragments that is consistent with the addition of 14 u on the guanosine in the third position in the RNase A digestion product (Figure 6B). Combined with the previous nucleoside analyses, these results are consistent with the sequence GGm²GGUp for this RNase digestion product. The digestion product with the sequence of GGm²GGUp can only be located in one part of the sequence in this tRNA transcript, position 6 of the tRNA (Figure 6C).

To further confirm that Trm14 catalyzes m²G formation at position 6 of tRNA^{Cys}, we synthesized a variant *M. jannaschii* tRNA^{Cys} transcript incorporating A6-U67 in place of the wild-type G6-C67 base pair. Methylation time courses with 1 μ M mutant tRNA, saturating levels of SAM and molar excess of Trm14, conditions identical to those used for the wild-type transcript, showed extremely low activity (Figure 7). In contrast, the A6-U67 tRNA

variant was an excellent substrate for *M. jannaschii* Trm5, which requires a properly folded tertiary core region for efficient formation of m¹G37 (data not shown) (38). Thus, as expected, replacement of the G6-C67 acceptor stem base pair with the isosteric A6-U67 does not compromise the integrity of the tRNA structure. These experiments support the identification of m²G6 as the site of Trm14 methylation. Further, Trm14-catalyzed methylation of tRNA^{Asp} is consistent with the presence of G6 in this species, while the extremely weak activity toward tRNA^{Pro} likely arises from the presence of U6 in this isoacceptor (Figure 2C).

A tRNA^{Cys} m²G6-forming activity in native *M. jannaschii* cell-free extract

To further validate the assignment of MJ0438 as a tRNA methylase, we examined whether cell-free lysates of *M. jannaschii* possess an equivalent activity. We cultivated *M. jannaschii* strain JAL-1 under anaerobic conditions in a 250l bioreactor, recovered cells by centrifugation, and prepared S100-fractionated lysates by sonication. Background incorporation levels are very low using this extract as the source of enzyme, ³H-labeled SAM as methyl-group donor, and no added tRNA (Figure 8A). Inclusion of 1 μ M *M. jannaschii* tRNA^{Cys} transcript in the reactions then revealed that the extract possesses

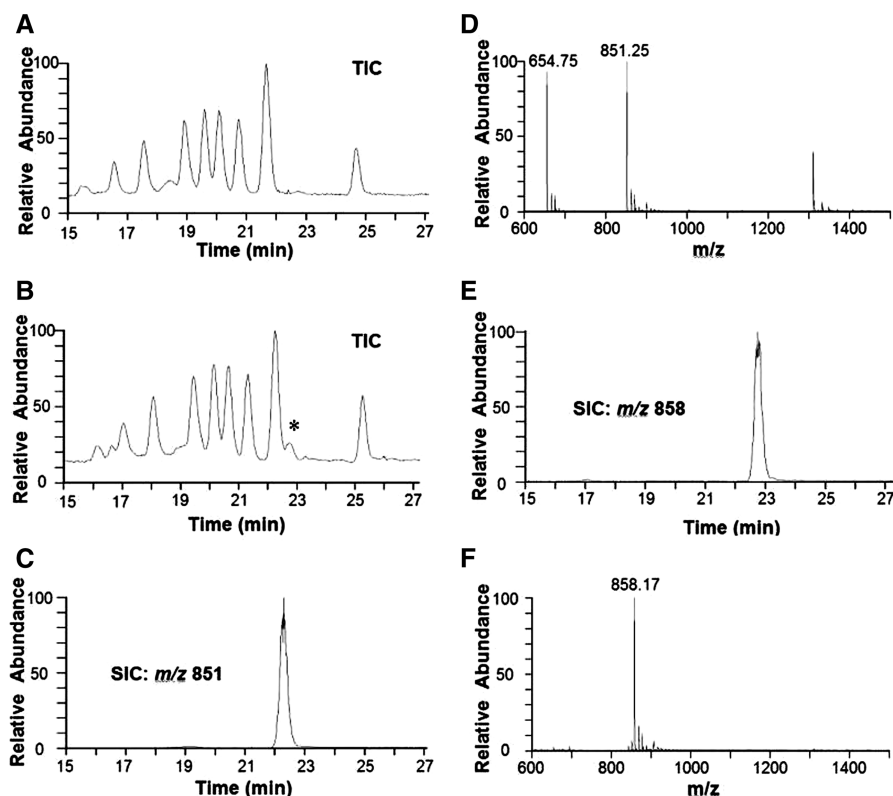


Figure 5. LC-MS analysis of RNaseA digestion products of Trm14-reacted and control unreacted tRNA^{Cys} transcripts. (A) TIC of RNase A-digested RNA transcripts. (B) TIC of RNase A-digested RNA transcripts treated with Trm14. There is clearly an additional digestion product (asterisks). (C) SIC of the unmodified digestion product with the sequence GGGGUp at m/z 851 (doubly charged). This product was found in the Trm14 treated and untreated samples. (D) Mass spectra of the digestion products at the elution time of the m/z 851 product. There is a digestion product with the sequence AAAUp that co-elutes with m/z 654.75 (doubly charged). (E) SIC of the modified digestion product with the sequence GGm²GGUp at m/z 858 (doubly charged). (F) Mass spectra of the digestion products at the elution time of the m/z 858 product.

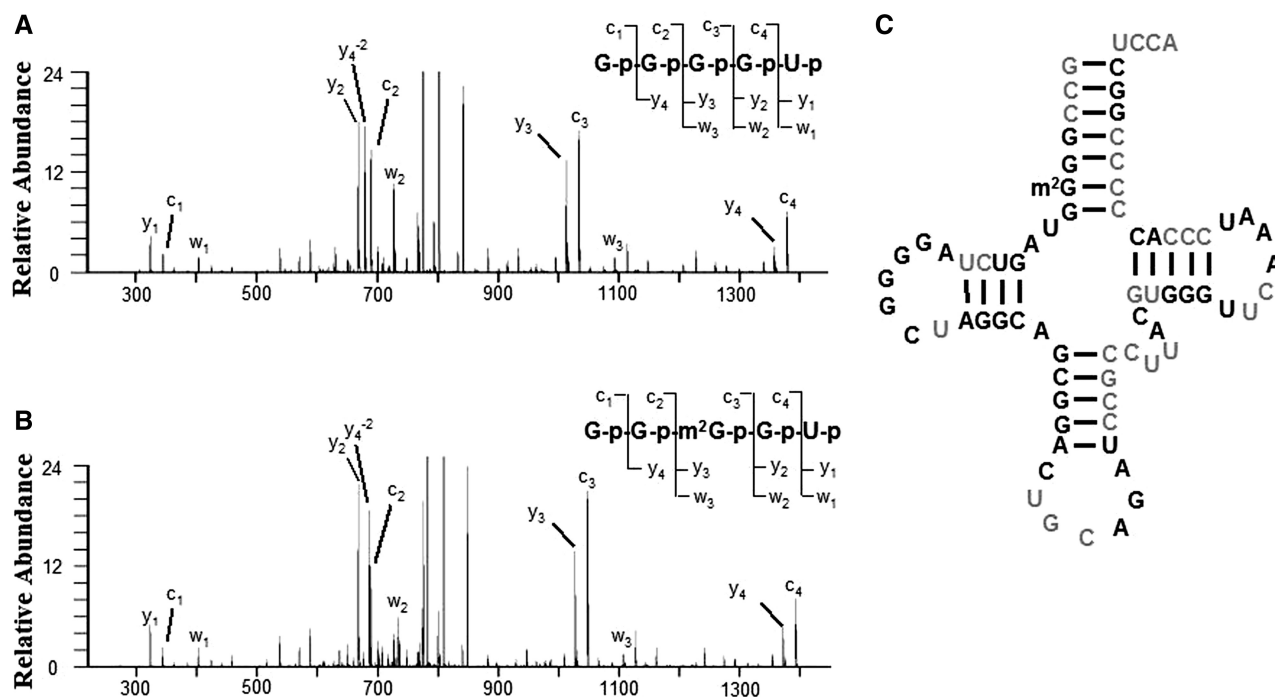


Figure 6. LC-MS/MS analysis of RNase A digestion products of Trm14-reacted RNA transcripts. (A) CID spectrum of the unmodified digestion product at m/z 851. Expected c- and y-type fragment ions are seen for all possible fragments and w-type ions are seen for three of the four possible fragments. This confirms the sequence of GGGGUp for the m/z 851 product. (B) CID spectrum of the modified digestion product at m/z 858. The same fragment ions seen for the unmodified product are seen here, however, the ions containing the third G in the sequence are shifted by 1 U, indicating a methyl group has been added to this base. This data confirms that the sequence of this digestion product is GGm²GGUp. (C) The secondary structure of the tRNA transcript used in these studies with the location of the methylated base indicated. Boldface nucleotides indicate those that were confirmed by LC-MS/MS. In gray are bases that were not found in fragments from the RNaseA analysis due to their small size.

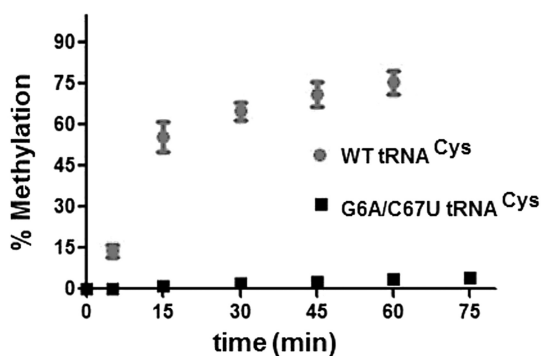


Figure 7. Methylation time courses by Trm14 under single turnover conditions for WT tRNA^{Cys} and the G6A/C67U base pair substitution mutant. Error bars indicate the standard deviation derived from two replicate time courses.

robust tRNA methylase activity. Importantly, plateau methylation levels were decreased nearly 3-fold when the A6-U67 tRNA^{Cys} variant was substituted for WT tRNA^{Cys} as the substrate. Thus, a substantial portion of the activity in native cell lysates is sensitive to the identity of the base pair at position 6–67 in a manner consistent with the *in vitro* activity of Trm14. The remaining SAM-dependent activity toward the A6-U67 transcript could be associated with formation of m¹A, m⁵C, m⁵G, m⁶A, m²G or m²G, as all of these modified bases were

previously identified in *M. jannaschii* tRNA (17). These findings suggest that Trm14 is expressed *in vivo* and is responsible for a significant portion of the observed activity.

As a further control, we tested an *E. coli* S100 extract for its capacity to methylate an exogenously added *M. jannaschii* tRNA^{Cys} transcript. While this preparation is indeed capable of heterologous methylation at levels ~2-fold lower than observed for *E. coli* tRNA^{Cys}, the level of methyl group incorporation is insensitive to the identity of the 6–67 bp (Figure 8B). This is consistent with the known absence of methylation at the 6–67 bp in any *E. coli* tRNA (1), and with our inability to detect the presence of a Trm14 homolog in the genome.

Phylogenetic distribution of Trm14-like proteins

All proteins identified as members of COG 0116 (31) encompass two domains: a THUMP (thiouridine synthases, RNA methyltransferases and pseudouridine synthases) domain (32), and a C-terminal SAM-dependent methyltransferase Rossmann fold. A conserved polypeptide 40 amino acids in length is present at the N-terminus, but this portion of the enzyme possesses no detectable sequence similarity with other proteins. Alignment of a subset of COG 0116 proteins, together with other candidates identified in BLAST searches of the omniome using *M. jannaschii* Trm14 as the query sequence, reveals similarities across the entire structure (Figure 9).

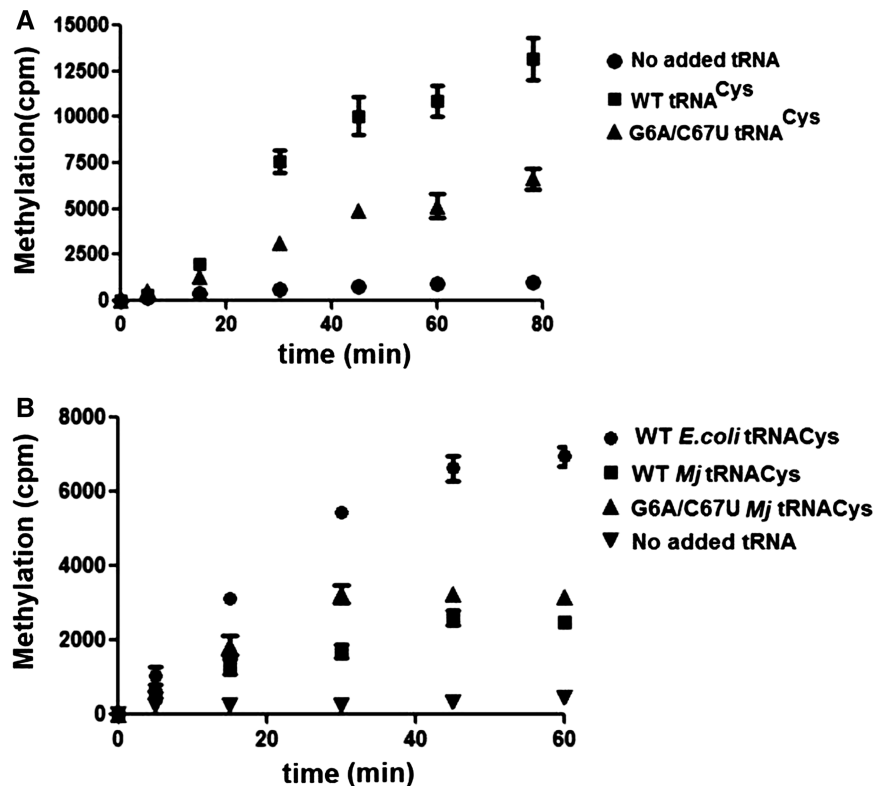


Figure 8. (A) Time courses for methylation of WT or G6A/C67U *M. jannaschii* tRNA^{Cys} transcripts using native cell-free *M. jannaschii* extract as the source of enzyme. (B) Time courses for methylation of WT or G6A/C67U *M. jannaschii* tRNA^{Cys} transcripts, and for methylation of an *E. coli* tRNA^{Cys} transcript, using native cell-free *E. coli* extract as the source of enzyme. In each panel, error bars indicate the standard deviation derived from two replicate time courses.

A phylogenetic tree of the COG 0116 family based on this alignment clearly distinguishes three main branches corresponding to the bacterial, archaeal and eukaryotic domains (Figure 10). The eukaryotic branch is more distantly related to the bacterial or archaeal branches than the latter two branches are to each other.

The separation between the bacterial and archaeal branches, and the closer connection of the eukaryotic branch to the archaeal, suggests a canonical vertical gene flow pattern. However, there are several exceptions. First, in the hyperthermophilic bacterium *Thermus thermophilus*, the protein that is most similar to *M. jannaschii* Trm14 is more closely related to proteins in the archaeal branch than to those of other bacteria (29% identity to the *M. jannaschii* ortholog; Figures 9 and 10). Second, the tightly grouped eukaryotic proteins depicted in Figure 10 are confined to chordates. However, BLAST searches with *M. jannaschii* Trm14 as query also detect other eukaryotic sequences, including from the unicellular species *Paulinella chromatophora* and *Thalassiosira pseudonana*, the nematode *Caenorhabditis remanei*, and the plants *Ricinus communis* and *Populus trichocarpa*. The sequences of these proteins are not closely related to those of the chordates, and are instead dispersed in the bacterial domain. If all these proteins are indeed homologs with a common evolutionary origin, then significant horizontal gene transfer may have occurred. Further analysis

will be required to examine this possibility in greater depth.

The Trm14 THUMP domain has previously been found in enzymes modifying tRNA at diverse sites. At position 8, both an archaeal cytosine deaminase activity and ThiI-catalyzed 4-thiouridine formation are associated with this domain (53,54). The tRNA positional specificity of THUMP domains also includes position 12 (N⁴-acetylcytidine formation by Tan1) (55), position 55 (pseudouridine formation by human Pus10) (56) and position 10 (m²G10 formation by Trm11) (15). Interestingly, m²G26 formation is catalyzed by a Rossmann-fold class methyltransferase, Trm1, which does not possess a THUMP domain (13,14,46). In general, THUMP domain enzymes have previously been found associated only with modification in the globular tRNA core, but target diverse positions within this region. The association of THUMP with catalysis of m²G6 formation by Trm14, then, expands its known positional specificity. A separately expressed THUMP domain possessed only weak tRNA binding activity (57); in general, how THUMP functions to target a wide variety of modification activities throughout the globular core region remains unknown.

Within the Trm14 catalytic domain, the alignments reveal conservation of the SAM binding site, which includes residues in motifs I, II and III, as well as the conserved catalytic sequence NPPY in motif IV. The common motifs

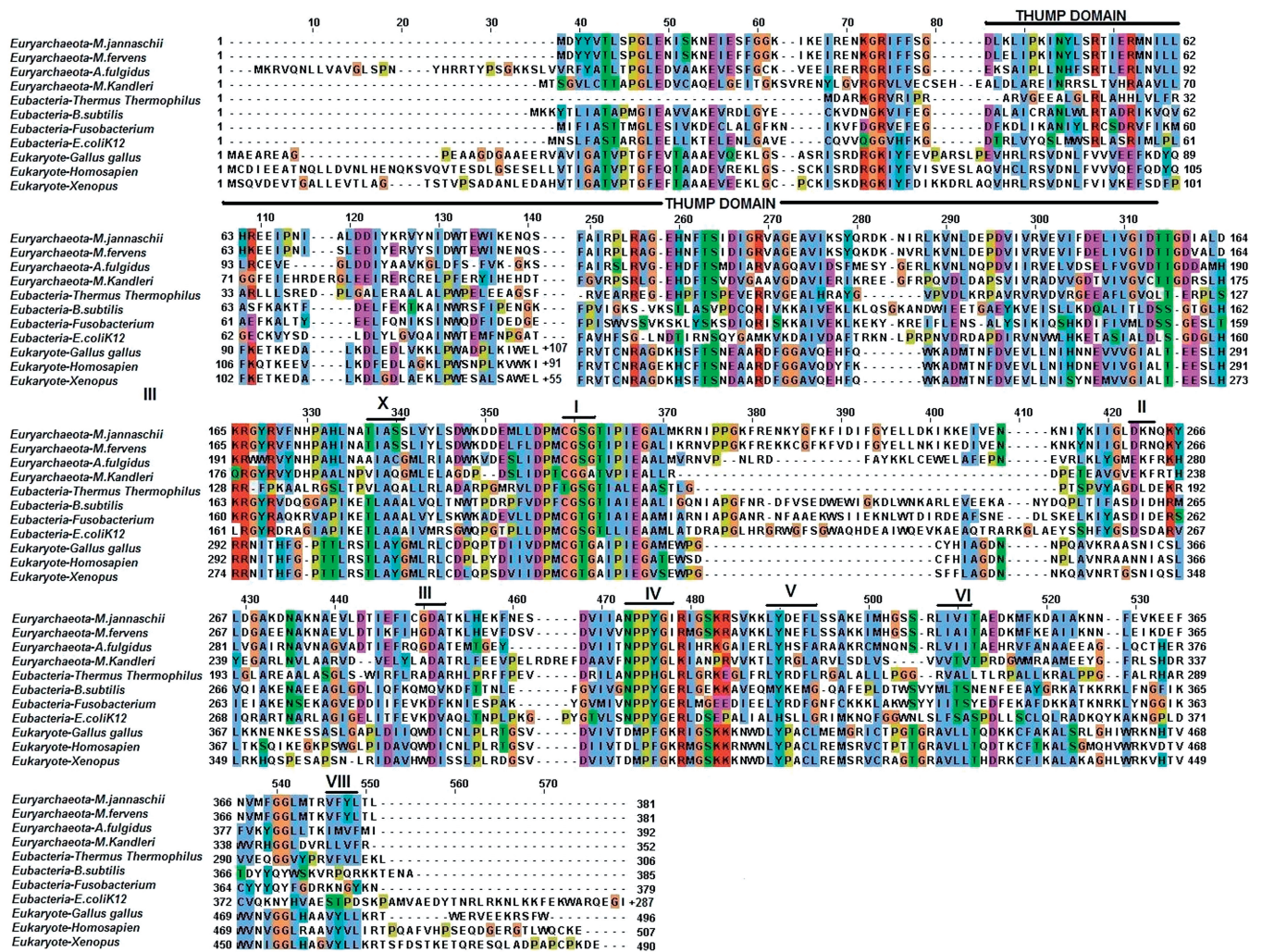


Figure 9. Multiple sequence alignment of Trm14-like proteins. The positions of the THUMP domain and the motifs in the class I methyltransferase catalytic domain (X, I, II, III, IV, V, VI, VIII) are indicated. Accession numbers for sequences are provided in the Supplementary Data.

of the amino methyltransferase family are present in the sequential order X, I, II, III, IV, V, VI, VII and VIII, indicating that the COG 0116 proteins belong to the γ class of these enzymes (58). More detailed examination of the alignment shows that the proteins in the archaeal and bacterial branches possess some consistent distinctions in the conserved motif sequences. In motif II, while proteins from both branches possess a conserved negatively charged residue (usually Asp) that interacts with the SAM ribose group, most members of the bacterial branch contain Asp-X-Asp/Glu while the archaeal enzymes lack the second negatively charged residue. In motif IV, almost all Trm14 proteins contain the signature NPPY motif of the group γ aminomethyltransferases, which is followed by Gly. The exceptions include *T. thermophilus* and the *Pyrococci* species with ‘NPPH’ and ‘NLPY’ as the catalytic motifs, respectively. Finally, motif V contains the consensus Leu-Tyr-X-X-Phe/Leu/Ile/Met. This motif can be detected in most members of the archaeal branch, but the slightly different signature Leu-Tyr-X-X-(Phe/Leu/Ile/Met)-Gly is present in the bacterial

branch, with the Gly particularly well conserved. The Phe in motif V, which directly contacts the adenine moiety of SAM, is sometimes present when a spatially conserved Phe is missing at the position immediately N-terminal to the well-conserved Cys–Gly–Ser–Gly motif I sequence. Such replacement, when it occurs in the Trm14 family, is limited to the bacterial enzymes. The other common catalytic motifs III, VI, VII, VIII and X are not clearly distinguished between the two branches.

Among known tRNA methylases, the closest match of MJ0438 (18–26% sequence identity) is to the m_2^2 G10-forming Trm11 tRNA methylase, for which the ortholog exists in *M. jannaschii* (MJ0710) (40). MJ0710 and MJ0438 are similarly organized with N-terminal THUMP and C-terminal catalytic motifs, but differ in the motif IV active-site sequence motif (DPPY in Trm11; NPPY in MJ0438). The phylogenetic distribution of Trm14 differs from that of Trm11, since the latter is found in Archaea and Eukarya only, and is absent in Bacteria. No crystal structure of Trm11 is yet available.

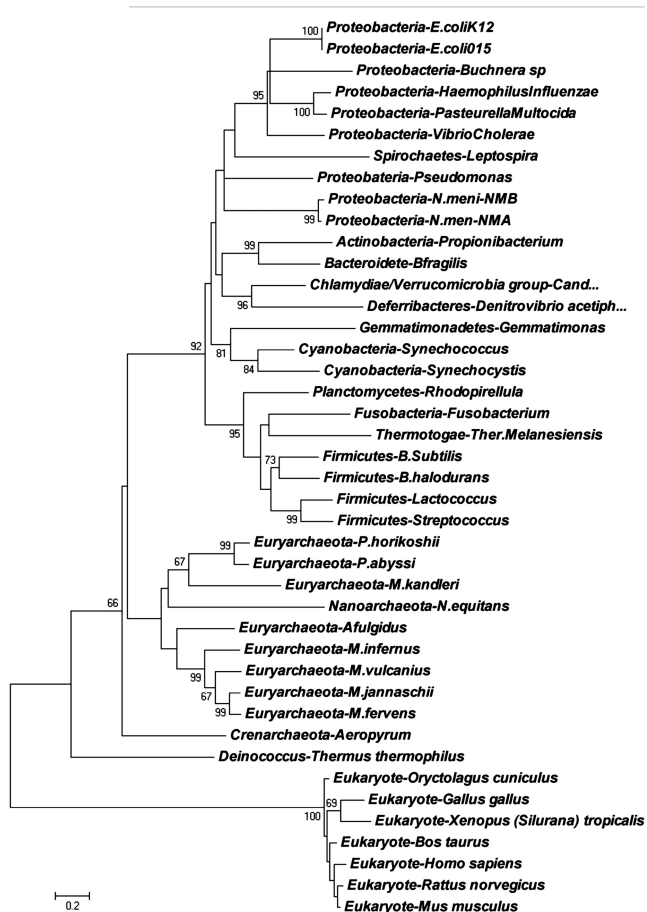


Figure 10. Unrooted phylogenetic tree of Trm14-related proteins. The tree is depicted in traditional rectangular format, and the evolutionary history was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Bootstrapping values above 65% are displayed. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method and are in units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the data set (complete deletion option). There were a total of 206 positions in the final data set. Accession numbers for sequences are provided in the Supplementary Data.

Functional roles of Trm14-like proteins

M. jannaschii Trm14 is tightly grouped with likely orthologs in the *Methanocaldococcus* family (65–87% sequence identities), and with other Archaea such as *A. fulgidus*, *Aeropyrum pernix*, *M. kandleri*, *P. abyssi*, *P. horikoshii* and *Nanoarchaeum equitans* (30–49% identities). Further, all organisms in the archaeal clade that possess Trm14 are thermophiles. Based on the high degree of sequence similarity, we expect that the Trm14 function in catalyzing tRNA m²G6 formation is very likely conserved among these archaeal proteins, and may play a role in stabilizing the structure of the RNA. The position of the modification appears unusual in this

respect, since previously well-characterized examples of modification-derived thermal protection in tRNA are for nucleotides in the T-loop/D-loop interaction region at the outside globular corner of the L-shape (25). Possibly, m²G6 may function synergistically with these core-region modifications to stabilize the molecule against thermal stress.

Another possibility is that Trm14-catalyzed m²G6 formation in tRNA^{Cys} may play a role in modulating the aminoacylation efficiency of SepRS. It is already known that the presence of m¹G37 in the tRNA^{Cys} anticodon loop enhances aminoacylation by *M. jannaschii* SepRS, but k_{cat}/K_m for phosphoserine remains lower than for many tRNA synthetases (9). It is of particular interest that the SAM cofactor required for methylation is synthesized *in vivo* from methionine and hence is metabolically linked to the flux of sulfur, including its assimilation from the environment. Methanogen proteins possess on average twice as much cysteine as those of other Archaea (29), regardless of whether Cys-tRNA^{Cys} formation is accomplished by redundant SepRS and CysRS pathways, or by SepRS alone. We speculate that control of SepRS (or CysRS) activity by SAM-dependent methylation of tRNA^{Cys} might provide the methanogen cell with a regulatory mechanism by which protein synthesis rates can respond to environmental sulfur levels or to concentrations of sulfur metabolites.

Sequence identities among Trm14-like proteins within the bacterial branch vary from 30% to 81%, but the level of identity between bacterial and archaeal proteins is in the 27–32% range. COG0116 is broadly distributed in at least 13 of the 24 different NCBI-classified phyla: the Actinobacteria, Bacteroidetes/Chlorobi group, Cyanobacteria (blue-green algae), Proteobacteria (purple bacteria and relatives), Spirochaetes, Thermotogae, Firmicutes (Gram-positive bacteria), Fusobacteria, Planctomycetes, Chlamydiae/Verrucomicrobia group, Deferribacteres, Deinococcus-Thermus and Gemmatimonadetes. As noted above, the bacterial and archaeal Trm14-like proteins exhibit conserved divergence in some key catalytic motifs, although this is not predictive with respect to whether the RNA methylation target of the bacterial enzymes is conserved as m²G6 in tRNA.

However, evidence that function may diverge in the bacterial domain is available from tRNA sequence information. While sequences determined at the RNA level are available for a variety of bacteria (1), the only example known in which m²G6 is present in tRNA is *T. thermophilus*, which possesses a Trm14-like protein that groups with the archaeal clade (Figure 10). This concordance suggests that the *T. thermophilus* protein is an ortholog of the thermophilic archaeal enzymes. However, for bacteria known to lack m²G6 in tRNA, the Trm14-like protein that we have identified must possess a different function. An interesting example is provided by *E. coli* and some related bacteria. In these organisms, the COG 0116 protein is much larger than *M. jannaschii* Trm14 and its related orthologs (Figure 9). The *E. coli* protein, the product of the *ycbY* gene, possesses an N-terminal THUMP domain followed by two consecutive catalytic

domains belonging to COG 0116 and, in the C-terminal portion of the protein, to COG 1092 (59). Knockout of *ycbY* (now renamed *rlmL*) showed that the enzyme catalyzes m²G2445 formation in 23S rRNA *in vivo*. The COG 1092 family contains several representatives in *M. jannaschii* (MJ1653 and MJ1649, each unlinked to MJ0438), and there are also several examples in which it is known that this protein methylates rRNA (60). Hence, the COG 0116 portion of *E. coli rlmL* may possess an entirely different function or may assist the rRNA methylation in an unknown way. No other members of the COG 0116 family from the archaeal, bacterial or eukaryal domains have been characterized.

The m²G6 modification has also been found in the tRNA of eukaryotes (1), including chordates such as *Homo sapiens* and *Xenopus laevis* that possess Trm14-like proteins more distantly related to the archaeal and bacterial enzymes. The chordate enzymes possess an additional 20–40 amino acid extension at their N-termini, and also conserve several other structural features in the catalytic domain not present in the bacterial and archaeal proteins (Figure 9). Further, these proteins either possess variants of the NPPY motif with substitutions at one or two positions, or do not conserve any of the four amino acids. The precise role of the active-site motif D/N/S-PP-Y/W/F/H is unknown, however, and the lack of conservation of any titratable side-chain group suggests that it does not directly participate in acid-base catalysis (3,61,62). Thus, the precise sequence of this motif has no clear predictive value, certainly not with respect to RNA substrate specificity. While the presence of m²G6 in chordate tRNA is provocative, direct experimental characterization of the Trm14-like protein in one or more of these organisms is required to assess whether it is indeed responsible for the activity. Previous studies monitoring formation of m²G6 in *Xenopus* tRNA *in vivo* showed that the activity is insensitive to alterations in the tRNA tertiary structure (63). In general, the identification of archaeal and candidate eukaryal m²G6-forming enzymes opens the possibility for more detailed structure–function studies to address questions of catalytic mechanism and tRNA sequence requirements for function. The possible roles of m²G6 in serving as a modulator of elongation factor affinity for aminoacyl-tRNA or to regulate function on the ribosome are also entirely unknown at this time.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Javin Oza for providing the *M. mazei* ribosomal RNA, David Ardell for insights regarding phylogenetics and Annia Rodriguez-Hernandez and Hari Bhaskaran for helpful discussions.

FUNDING

National Institutes of Health (GM58843 to P.A.L., GM53763 to J.J.P.); the National Science Foundation (CHE0910751 to P.A.L.); Department of Energy, Energy Biosciences Program (DE-FG02-93ER20106 to K.R.S.). Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

- Juhling,S., Morl,M., Hartmann,R.K., Sprinzl,M., Stadler,P.F. and Putz,J. (2009) tRNAdb2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.*, **37**, D159–D162.
- Phizicky,E.M. and Alfonzo,J.D. (2010) Do all modifications benefit all tRNAs? *FEBS Lett.*, **584**, 265–271.
- Hou,Y.-M. and Perona,J.J. (2010) Stereochemical mechanisms of tRNA methyltransferases. *FEBS Lett.*, **584**, 278–286.
- Sakurai,M., Ohtsuki,T. and Watanabe,K. (2005) Modification at position 9 with 1-methyladenosine is crucial for structure and function of nematode mitochondrial tRNAs lacking the entire T-arm. *Nucleic Acids Res.*, **33**, 1653–1661.
- Lee,C., Kramer,G., Graham,D.E. and Appling,D.R. (2007) Yeast mitochondrial initiator tRNA is methylated at guanosine 37 by the Trm5-encoded tRNA (guanine-N1-)-methyltransferase. *J. Biol. Chem.*, **282**, 27744–27753.
- Putz,J., Florentz,C., Benseler,F. and Giege,R. (1994) A single methyl group prevents the mischarging of a tRNA. *Nature Struct. Biol.*, **1**, 580–582.
- Hauenstein,S.I., Hou,Y.-M. and Perona,J.J. (2008) The homotetrameric phosphoseryl-tRNA synthetase from *Methanosarcina mazei* exhibits half-of-the-sites activity. *J. Biol. Chem.*, **283**, 21997–22006.
- Hauenstein,S. and Perona,J.J. (2008) Redundant synthesis of cysteinyl-tRNA^{Cys} in *Methanosarcina mazei*. *J. Biol. Chem.*, **283**, 22007–22017.
- Zhang,C.M., Liu,C., Slater,S. and Hou,Y.-M. (2008) Aminoacylation of tRNA with phosphoserine for synthesis of cysteinyl-tRNA^{Cys}. *Nature Struct. Mol. Biol.*, **15**, 507–514.
- Urbonavicius,J., Skouloubri,S., Myllykallio,H. and Grosjean,H. (2005) Identification of a novel gene encoding a flavin-dependent tRNA:m5U methyltransferase in bacteria—evolutionary implications. *Nucleic Acids Res.*, **33**, 3955–3964.
- Schubert,H.L., Blumenthal,R.M. and Cheng,X. (2003) Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.*, **28**, 329–335.
- Auffinger,P. and Westhof,E. (1998) Location and distribution of modified nucleotides in tRNA. In: Grosjean,H. and Benne,R. (eds), *Modification and Editing of RNA*. American Society for Microbiology, Washington, DC, pp. 569–576.
- Awai,T., Kimura,S., Tomikawa,C., Ochi,A., Ihsanawati, Bessho,Y., Yokoyama,S., Ohno,S., Nishikawa,K., Yokogawa,T. et al. (2009) *Aquifex aeolicus* tRNA (N², N²-guanine)-dimethyltransferase (Trm1) catalyzes transfer of methyl groups not only to guanine 26 but also to guanine 27 in tRNA. *J. Biol. Chem.*, **284**, 20467–20478.
- Ihsanawati, Nishimoto,M., Higashijima,K., Shirouzu,M., Grosjean,H., Bessho,Y. and Yokoyama,S. (2008) Crystal structure of tRNA N²,N²-guanosine dimethyltransferase Trm1 from *Pyrococcus horikoshii*. *J. Mol. Biol.*, **383**, 871–884.
- Armengaud,J., Urbonavicius,J., Fernandez,B., Chaussinand,G., Bujnicki,J.M. and Grosjean,H. (2004) N²-methylation of guanosine at position 10 in tRNA is catalyzed by a THUMP domain-containing, S-adenosylmethionine-dependent methyltransferase, conserved in archaea and eukaryotes. *J. Biol. Chem.*, **279**, 37142–37152.
- Purushothaman,S.K., Bujnicki,J.M., Grosjean,H. and Lapeyre,B. (2005) Trm11p and Trm112p are both required for the formation of 2-methylguanosine at position 10 in yeast tRNA. *Mol. Cell. Biol.*, **25**, 4539–4570.

17. McCloskey, J.A., Graham, D.E., Zhou, S., Crain, P.F., Ibba, M., Konisky, J., Söll, D. and Olsen, G.J. (2001) Post-transcriptional modification in archaeal tRNAs: identities and phylogenetic relations of nucleotides from mesophilic and hyperthermophilic *Methanococcales*. *Nucleic Acids Res.*, **29**, 4699–4706.
18. Sauerwald, A., Sitaramaiah, D., McCloskey, J.A., Söll, D. and Crain, P.F. (2005) N⁶-acetyladenosine: a new modified nucleoside from *Methanopyrus kandleri* tRNA. *FEBS Lett.*, **579**, 2807–2810.
19. Noon, K.R., Guymon, R., Crain, P.F., McCloskey, J.A., Thomm, M., Lim, J. and Cavicchioli, R. (2003) Influence of temperature on tRNA modification in archaea: *Methanococcoides burtonii* (optimum growth temperature [T_{opt}], 23°C) and *Stetteria hydrogenophila* (T_{opt}, 95°C). *J. Bacteriol.*, **185**, 5483–5490.
20. Edmonds, C.G., Crain, P.F., Gupta, R., Hashizume, T., Hocart, C.H., Kowalak, J.A., Pomerantz, S.C., Stetter, K.O. and McCloskey, J.A. (1991) Post-transcriptional modification of tRNA in thermophilic archaea (archaeobacteria). *J. Bacteriol.*, **173**, 3138–3148.
21. Kowalak, J.A., Dalluge, J.J., McCloskey, J.A. and Stetter, K.O. (1994) The role of post-transcriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry*, **33**, 7869–7876.
22. McCloskey, K.A., Liu, X.-H., Crain, P.F., Bruenger, E., Guymon, R., Hashizume, T. and Stetter, K.O. (2000) Posttranscriptional modification of transfer RNA in the submarine hyperthermophile *Pyrolobus fumarii*. *Nucl. Acids Symp. Series*, **44**, 267–268.
23. Gupta, R. (1984) *Halobacterium volcanii* tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. *J. Biol. Chem.*, **259**, 9461–9471.
24. Zhou, S., Sitaramaiah, D., Noon, K.R., Guymon, R., Hashizume, T. and McCloskey, J.A. (2004) Structures of two new ‘minimalist’ modified nucleosides from archaeal tRNA. *Bioorg. Chem.*, **32**, 82–91.
25. Grosjean, H., and Oshima, T. (2007) How nucleic acids cope with high temperature. In: *Physiology and Biochemistry of Extremophiles*. ASM Press, Washington, DC, pp. 39–58.
26. Sauerwald, A., Zhu, W., Major, T.A., Roy, H., Palioura, S., Jahn, D., Whitman, W.B., Yates, J.R., Ibba, M. and Söll, D. (2005) RNA-dependent cysteine biosynthesis in archaea. *Science*, **307**, 1969–1972.
27. Christian, T., Evilia, C., Williams, S. and Hou, Y.-M. (2004) Distinct origins of tRNA(m¹G37)methyltransferase. *J. Mol. Biol.*, **339**, 707–719.
28. O’Donoghue, P., Sethi, A., Woese, C.R. and Luthey-Schulten, Z.A. (2005) The evolutionary history of Cys-tRNA^{Cys} formation. *Proc. Natl Acad. Sci. USA*, **102**, 19003–19008.
29. Klipcan, L., Frenkel-Morgenstern, M. and Sefror, M.G. (2008) Presence of tRNA-dependent pathways correlates with high cysteine content in methanogenic Archaea. *Trends Genet.*, **24**, 59–63.
30. Makarova, K.S., Sorokin, A.V., Novichkov, P.S., Wolf, Y.I. and Koonin, E.V. (2007) Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol. Direct.*, **2**, 33.
31. Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, E.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N. *et al.* (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*, **4**, 41.
32. Aravind, L. and Koonin, E.V. (2001) THUMP—a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases, and RNA methylases. *Trends Biochem. Sci.*, **26**, 215–217.
33. Sherlin, L.D., Bullock, T.L., Nissan, T.A., Perona, J.J., Lariiviere, F.J., Uhlenbeck, O.C. and Scaringe, S.A. (2001) Chemical and enzymatic synthesis of tRNAs for high-throughput crystallization. *RNA*, **7**, 1671–1678.
34. Grosjean, H., Droogmans, L., Roovers, M. and Keith, G. (2007) Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. *Meth. Enzymol.*, **425**, 57–101.
35. Pomerantz, S.C. and McCloskey, J.A. (1990) Analysis of RNA hydrolyzates by liquid chromatography-mass spectrometry. *Methods Enzymol.*, **193**, 796–824.
36. Sowers, K.R. and Watts, J.E.M. (2006) Methods for the study of strictly anaerobic microorganisms. In: Oren, A. and Rainey, F.A. (eds), *Methods in Microbiology—Extremophiles*, Vol. 35. Elsevier/Academic Press, Oxford, pp. 757–782.
37. Mukhopadhyay, B., Johnson, E.F. and Wolfe, R.S. (1999) Reactor-scale cultivation of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* to high cell densities. *Appl. Environ. Microbiol.*, **65**, 5059–5065.
38. Christian, T. and Hou, Y.-M. (2007) Distinct determinants of tRNA recognition by the TrmD and Trm5 methyltransferases. *J. Mol. Biol.*, **373**, 623–632.
39. Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M. *et al.* (2009) STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.*, **37**, D412–D416.
40. Altschul, A., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
41. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, **31**, 3497–3500.
42. Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molec. Biol. Evolution*, **24**, 1596–1599.
43. O’Donoghue, P., Sethi, A., Woese, C.R. and Luthey-Schulten, Z. (2005) The evolutionary history of Cys-tRNA^{Cys} formation. *Proc. Natl Acad. Sci. USA*, **102**, 19003–19008.
44. Anantharaman, V., Koonin, E.V. and Aravind, L. (2002) Comparative genomics and evolution of proteins involved in RNA metabolism. *Nucleic Acids Res.*, **30**, 1427–1464.
45. Auxilien, S., Khadali, F., Rasmussen, A., Douthwaite, S. and Grosjean, H. (2007) Archease from *Pyrococcus abyssi* improves substrate specificity and solubility of a tRNA m²C methyltransferase. *J. Biol. Chem.*, **285**, 18505–18515.
46. Urbonavicius, J., Armengaud, J. and Grosjean, H. (2006) Identity elements required for enzymatic formation of N²,N²-dimethylguanosine from N²-monomethylated derivative and its possible role in avoiding alternative conformations in archaeal tRNA. *J. Mol. Biol.*, **357**, 387–399.
47. Gregson, J.M. and McCloskey, J.A. (1997) Collision-induced dissociation of protonated guanine. *Int. J. Mass Spec. Ion Proc.*, **165–166**, 475–485.
48. Kowalak, J.A., Pomerantz, S.C., Crain, P.F. and McCloskey, J.A. (1993) A novel method for the determination of post-transcriptional modification in RNA by mass spectrometry. *Nucleic Acids Res.*, **21**, 4577–4585.
49. McCloskey, J.A., Graham, D.E., Zhou, S., Crain, P.F., Ibba, M., Konisky, J., Soll, D. and Olsen, G.J. (2001) Post-transcriptional modification in archaeal tRNAs: Identities and phylogenetic relations of nucleotides from mesophilic and hyperthermophilic methanococcales. *Nucleic Acids Res.*, **29**, 4699–4706.
50. Meng, Z. and Limbach, P.A. (2004) RNase mapping of intact nucleic acids by electrospray ionization fourier transform ion cyclotron resonance mass spectrometry (ESI-FTICRMS) and 18O labeling. *Int. J. Mass Spec.*, **234**, 37–44.
51. Mandal, D., Köhrer, C., Su, D., Russell, S.P., Krivos, K., Castleberry, C.M., Blum, P., Limbach, P.A., Söll, D. and RajBhandary, U.L. (2010) Agmatidine, a modified cytidine in the anticodon of archaeal tRNA(Ile), base pairs with adenosine but not with guanosine. *Proc. Natl Acad. Sci. USA*, **107**, 2872–2877.
52. Anton, B.P., Russell, S.P., Vertrees, J., Kasif, S., Raleigh, E.A., Limbach, P.A. and Roberts, R.J. (2010) Functional characterization of the YmcB and Yqev tRNA methyltransferases of *Bacillus subtilis*. *Nucleic Acids Res.*, **38**, 6195–6205.
53. Randau, L., Stanley, B.J., Kohlway, A., Mechta, S., Xiong, Y. and Söll, D. (2009) A cytidine deaminase edits C to U in transfer RNAs in Archaea. *Science*, **324**, 657–659.
54. Waterman, D.G., Ortiz-Lombardia, M., Fogg, M.J., Koonin, E.V. and Antson, A.A. (2006) Crystal structure of *Bacillus anthracis* ThiI, a tRNA-modifying enzyme containing the predicted RNA-binding THUMP domain. *J. Mol. Biol.*, **356**, 97–110.

55. Johansson, M.J. and Byström, A.S. (2004) The *Saccharomyces cerevisiae* TAN1 gene is required for N⁴-acetylcytidine formation in tRNA. *RNA*, **10**, 712–719.
56. McCleverty, C.J., Hornsby, M., Spraggon, G. and Kreusch, A. (2007) Crystal structure of human Pus10, a novel pseudouridine synthase. *J. Mol. Biol.*, **373**, 1243–1254.
57. Gabant, G., Auxilien, S., Tuszyńska, I., Locard, M., Gajda, M.J., Chaussinand, G., Fernandez, B., Dedieu, A., Grosjean, H., Golinelli-Pimpaneau, B. *et al.* (2006) THUMP from archaeal m²G10 methyltransferase, a genuine autonomously folding domain. *Nucleic Acids Res.*, **34**, 2483–2494.
58. Malone, T., Blumenthal, R.M. and Cheng, X. (1995) Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. *J. Mol. Biol.*, **253**, 618–632.
59. Lesnyak, D.V., Sergiev, P.V., Bogdanov, A.A. and Dontsova, O.A. (2006) Identification of *Escherichia coli* m²G methyltransferases: I. the *yebY* gene encodes a methyltransferase specific for G2445 of the 23 S rRNA. *J. Mol. Biol.*, **364**, 20–25.
60. Sunita, S., Tkaczuk, K.L., Purta, E., Kasprzak, J.M., Douthwaite, S., Bujnicki, J.M. and Sivaraman, J. (2008) Crystal structure of the *Escherichia coli* 23S rRNA:m⁵C methyltransferase RlmI (YccW) reveals evolutionary links between RNA modification enzymes. *J. Mol. Biol.*, **383**, 652–666.
61. Goedecke, K., Pignot, M., Goody, R.S., Scheidig, A.J. and Weinhold, E. (2001) Structure of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nat. Struct. Biol.*, **8**, 121–125.
62. Christian, T., Lahoud, G., Liu, C., Hoffmann, K., Perona, J.J. and Hou, Y.M. (2010) Mechanism of N-methylation by the tRNA m¹G37 methyltransferase Trm5. *RNA*, **16**, 2484–2492.
63. Grosjean, H., Edqvist, J., Stråby, K.B. and Giegé, R. (1996) Enzymatic formation of modified nucleosides in tRNA: dependence on tRNA architecture. *J. Mol. Biol.*, **255**, 67–85.