Peer-reviewed version available at *Biomolecules* **2020**, *10*, 587; doi:10.3390/biom10040587

Type of the Paper (Article)

Reductive Evolution and Diversification of C5-Uracil Methylation in the Nucleic Acids of Mollicutes

Pascal Sirand-Pugnet*[1], Damien Brégeon[2], Laure Béven[1], Catherine Goyenvalle[2], Alain Blanchard[1], Simon Rose[3], Henri Grosjean[4], Stephen Douthwaite[3], Djemel Hamdane*[5] and Valérie de Crécy-Lagard*[6,7]

[1] Univ. Bordeaux, INRAE, UMR BFP, 33882, Villenave d'Ornon, France
[2] Sorbonne Université, IBPS, Biology of Aging and Adaptation, 7 quai Saint Bernard, F-75252 Paris cedex 05, France
[3] Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark.
[4] Institute for Integrative Biology of the Cell (I2BC), Commissariat à l'énergie atomique et aux énergies alternatives, CNRS, Université Paris-Sud, Université Paris-Saclay, 91198 Gif-sur-Yvette cedex, France.
[5] Laboratoire de Chimie des Processus Biologiques, CNRS-UMR 8229, Collège De France, Sorbonne Université, 11 place Marcelin Berthelot, 75231 Paris Cedex 05 ,France.
[6] Department of Microbiology and Cell Science, University of Florida, Gainesville, FL, 32611, USA.
[7] University of Florida, Genetics Institute, Gainesville, Florida, 32610, USA.

* Correspondence: pascal.sirand-pugnet@inrae.fr; +33 5 57 12 23 59 (P.S.P.); djemel.hamdane@college-de-france.fr; +33 1 44 27 12 54 (D.H.); vcrecy@ufl.edu; +1 352-392-9416 (V.C.L.)

Abstract: The C5-methylation of uracil to form 5-methyluracil (m^5U) is a ubiquitous base modification of nucleic acids. Four enzyme families have converged to catalyze this methylation using different chemical solutions. Here, we investigate the evolution of 5-methyluracil synthase families in *Mollicutes*, a class of bacteria that has undergone extensive genome erosion. Many mollicutes have lost some of the m^5U methyltransferases present in their common ancestor. Cases of duplication and subsequent shift of function are also described. For example, most members of the Spiroplasma subgroup, use the ancestral tetrahydrofolate-dependent TrmFO enzyme, to catalyze the formation of m^5U54 in tRNA, while a TrmFO paralog (termed RlmFO) is responsible for m^5U1939 formation in 23S RNA. RlmFO has replaced the S-adenosyl-l-methionine (SAM)-enzyme RlmD that adds the same modification in the ancestor and which is still present in mollicutes from the Hominis subgroup. Another paralog of this family, the TrmFO-like protein, has a yet unidentified function that differs from the TrmFO and RlmFO homologs. Despite having evolved towards minimal genomes, the mollicutes possess a repertoire of m^5U modifying enzymes that is highly dynamic and has undergone horizontal transfer. This emphasizes the necessity for combining bioinformatics predictions with empirical testing and structural information to get a reliable functional annotation of these enzymes.

Keywords: base modification, methyltransferases, flavoenzymes, tRNA, rRNA, mycoplasmas, spiroplasmas, acholeplasmas, evolution, minimal cell, moonlighting function

1. Introduction

Methylation reactions are essential for a large number of cellular processes including DNA synthesis, genome protection against restriction systems, gene expression and regulation, and post-transcriptional modification of RNAs[1] [1,2]. Among these reactions, the C5-methylation of uracil yielding 5-methyluracil (m^5U) constitutes one of the most common nucleic acids modifications found

both in DNA and RNA. 5-methyluracil is ubiquitous in DNA as one of the four canonical bases deoxythymidine (dT or deoxym^{5}U), and is formed by the *de novo* methylation of the deoxyuridine monophosphate (dUMP) precursor to deoxythymidine monophosphate (dTMP) [3]. In RNA, 5-methyluracil is present as ribothymidine (rT or m^{5}U) and is synthetized post-transcriptionally [1,2]. So far, m^{5}U has been found at position 54 in the T-loop of tRNAs (m^{5}U54), where it is conserved in most organisms, and at the corresponding position in bacterial tmRNA (formerly 10Sa RNA) [4-6]. However, in some organisms, including a few mycoplasmas, m^{5}U54 is absent from the tRNA molecules [7-9]. This modification also occurs in some bacterial 23S rRNAs at positions m^{5}U747 and m^{5}U1939 [10] [11]. The presence of m^{5}U stabilizes RNA molecules by increasing the intrinsic stacking power and rigidity of the loop containing the modified nucleosides [12-14], allowing also protection against degradation by specific nucleases [15,16]. Furthermore, 5-methyluracil has recently been shown to be a component of the polyoxin antibiotic (PolB) produced by *Streptomyces cacaoi* subsp. *asoensis*, where antibiotic synthesis is initiated by C5-methylation of UMP to m^{5}UMP [17]. This finding indicates that 5-methyluracil can engage in additional pathways other than those specifically linked to DNA and RNA metabolism.

Interestingly, methylation of C5-uracil can occur via four distinct mechanisms involving four structurally unrelated enzyme families with distinct evolutionary origins, highlighting how independent solutions have evolved to catalyze the same reaction [3,18-21] (**Table 1**). The most widespread enzyme involved in dTMP synthesis is the homodimeric thymidylate synthase ThyA (COG0207), encoded in most prokaryotes by the *thyA* gene and by its ortholog in humans [22,23]. This enzyme uses N5, N10-methylenetetrahydrofolate (CH_2THF) as a methylene donor and THF as a reducing agent (**Table 1**). In the reaction pathway, a conserved cysteine in the enzyme activates the C6-position of dUMP via a classical Michael's addition, increasing the nucleophilic character of the C5-position. Direct transfer of CH_2 from CH_2THF then takes place via the formation of a transient enzyme-dUMP-folate covalent intermediate. Removal of the H5 proton promotes the heterolytic cleavage of the C-C bond between CH_2THF and dUMP, yielding an exocyclic methylene, which is ultimately reduced to a methyl group by THF (**Figure 1**).

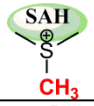
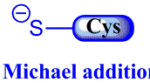
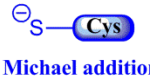
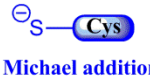
The subsequent discovery of the thymidylate synthase ThyX (a distinct COG1351 as of ThyA), established the existence of an alternate mechanism for dUMP methylation [24]. ThyX is less widespread than its ThyA counterpart and mainly found in prokaryotes and a few eukaryotes. ThyX is a homotetrameric flavoenzyme that employs the flavin adenine dinucleotide (FAD) as a coenzyme, with nicotinamide adenine dinucleotide phosphate (NADPH) as the initial source of hydride and CH_2THF as a methylene donor (**Table 1**). In contrast to ThyA, the CH_2 moiety from CH_2THF is first transferred to the N5 atom of reduced flavin ($FADH^-$), obtained from the preliminary reaction of FAD and NADPH, leading to the previously unseen flavin iminium species FAD=CH_2, which acts as the *bona fide* dUMP methylating agent [25,26] (**Figure 1**). An additional mechanistic feature that distinguishes ThyA from ThyX is that the active site of ThyX has two essential conserved arginine residues that become polarized to activate dUMP [27]. Following the activation step, the electrophilic methylene on FAD=CH_2 is passed to the C5-dUMP via the formation of a transient dUMP-CH_2-FAD adduct, which eventually breaks down after abstraction of the H5 proton. The resulting exocyclic methylene is then reduced by a hydride donated by $FADH^-$, as opposed to the case of ThyA where reduction occurs via THF. This flavin- and folate-dependent mechanism is shared by PolB, a ThyX paralog also belonging to COG1351 that is involved in C5-UMP methylation during polyoxin

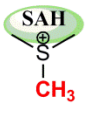
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2020 doi:10.20944/preprints202003.0097.v1

Peer-reviewed version available at Biomolecules 2020, 10, 587; doi:10.3390/biom10040587

biosynthesis [17]. PolB can also methylate dUMP but with a lower catalytic efficacy than its natural substrate (UMP).

Formation of m^{5}U in tRNAs can also be catalyzed by two fundamentally different pathways one of which is S-adenosyl-l-methionine (SAM)-dependent while the other uses flavin and folate [28-30]. The SAM-dependent pathway is the most common and uses a mechanism similar to 5-methylcytosine DNA and RNA methyltransferases. In most characterized organisms, m^{5}U54 in tRNAs and at the corresponding position in bacterial tmRNA is synthesized by a SAM-dependent methyltransferase, such as the *Escherichia coli* enzyme TrmA [31-33]. In addition, m^{5}U1939 that is present in most bacterial 23S rRNAs, and m^{5}U747 which is less common and found mainly in Gram-negative Beta, Epsilon and Gamma Proteobacteria, are respectively catalyzed by the SAM-dependent RlmD (formerly RumA) and RlmC (formerly RumB) methyltransferases [34,35]. Interestingly, both of these rRNA m^{5}U modifications in *Bacillus subtilis* are catalyzed by the same SAM-dependent RNA methyltransferase RlmCD, showing that dual target specificity is possible [36]. All these m^{5}U tRNA or rRNA methyltransferases are member of the same superfamily (COG2265) and thus share a common ancestry, and all use a simple mechanism based on direct transfer of the methyl group from the electrophilic carbon of the SAM cofactor to the activated C5-uracil. As in ThyA, this carbon is activated by a conserved cysteine that plays the role of nucleophile (**Table 1** and **Figure 1**). Another COG2265 family member YfjO is encoded in the *B. subtilis* genome and, although its function remains unknown, this putative enzyme has been included in our screening process.

Enzymes	Substrates	Coenzymes	Carbon donor	Reducing agent	Uracil activation
(I) TrmA RlmCD	tRNA (U54) rRNA(U747/1939)	None		None	
(II) TrmFO RlmFO	tRNA (U54) rRNA(U1939)			NAD(P)H	
(III) ThyA	dUMP	None			
(IV) ThyX PolB	dUMP UMP			NAD(P)H	



S-Adenosyl-Methionine

5,10 methylenetetrahydrofolate

tetrahydrofolate

Flavin Adenine Dinucleotide

Table 1. Different enzyme families and mechanisms of C5-uracil methylation.

Alternative mechanisms of RNA m^{5}U modification, which are analogous to dTMP synthesis, are seen with the two flavoenzyme homologs TrmFO that adds the m^{5}U54 modification in some tRNAs [30] and RlmFO that is responsible for the m^{5}U1939 modification in 23S rRNA [37]. These RNA

methyltransferases belong to a distinct COG1206. To date, only one case of m^5U1939 formation by RlmFO has been described, and this is in the mollicutes *Mycoplasma capricolum* subsp. *capricolum* [37]. TrmFO, while apparently more common and found mainly in Gram-positive bacteria and some mollicutes [30,38], remains much less prevalent than the SAM-dependent pathway. Although the TrmFO/RlmFO enzymes employ a chemical mechanism using FAD=CH_2 as the methylating agent, similar to that of ThyX [18,39,40], they differ from ThyX in their means of substrate activation where TrmFO/RlmFO rely on a conserved cysteine nucleophile in a manner similar to ThyA [41] (**Table 1** and **Figure 1**).

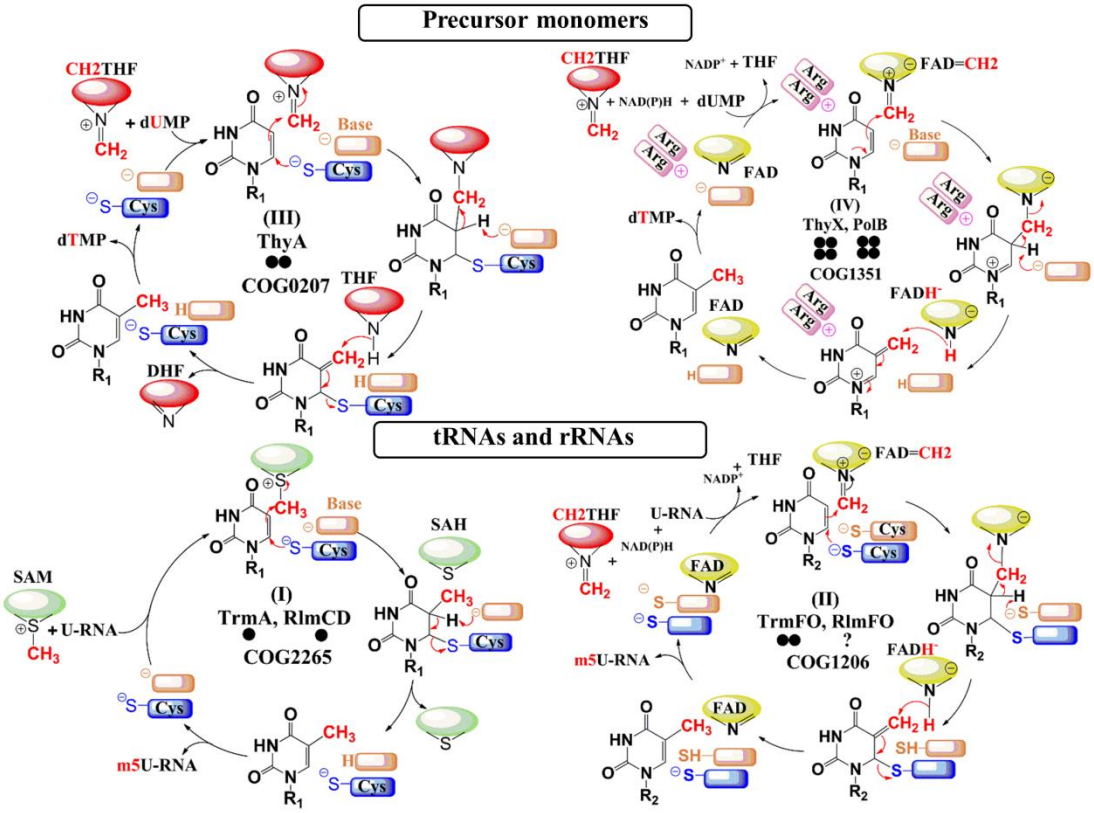


Figure 1 . Enzymatic mechanisms of the C5-uracil methyltransferases. Folate, SAH and FAD derivatives are colored as red, green and yellow, respectively. Residues playing the role of base in the catalytic cycle are in orange whereas the nucleophile are in blue. The bold circles below each protein's name represent the oligomeric state of the enzyme.

These observations show that while fundamentally different types of m^5U-modifying enzyme have evolved convergently to modify the same nucleotide target, other m^5U-modifying enzymes that are structurally similar have diverged in their functions to modify different RNA sites. In the present study, we investigate how such phenomena might have evolved within the Mollicutes. These organisms, with their small genome sizes of ~1 Mbp on average, represent excellent models for defining a minimal set of genes required for life and, more generally, for the studying mechanisms of genome reduction and evolution [42-44]. We have studied a diverse array of mollicutes to systematically explore the distribution and function of two different SAM- and folate-dependent

families of RNA m^5U methyltransferases (COG1206 and COG2265). The complex patterns of acquisition and loss of the genes for these enzymes have been mapped, and predictions of enzyme substrate specificity have been tested empirically.

2. Materials and Methods

2.1. In silico genome and protein analyses

Escherichia coli and *Bacillus subtilis* m^5U modification enzymes were used in blastp searches for mollicutes homologs in the MolliGen (http://molligen.org) database [45]. MolliGen and MBGD (http://mbgd.genome.ad.jp/) [46] databases were used to study the genomic contexts of genes of interest. The phylogenetic tree of mollicutes was generated using the maximum likelihood method from the concatenated multiple sequence alignments of 79 selected orthologous proteins involved in translation [47]. For phylogenetic analyses of TrmFO and RlmD homologs, protein alignments were obtained with MUSCLE and cured from unreliable positions using Gblocks [48]. Phylogenetic trees were then inferred using the maximum likelihood method using the PhyML software implemented at phylogeny.fr [49]. An overview of conserved positions was obtained from protein alignments created using Jalview [50].

2.2. Functional domain analysis and secondary structure prediction

The TrmFO homologs were modelled using the *Swiss model* server [51]. The QSQE score is a number between 0 and 1, reflecting the expected accuracy of the interchain contacts for a model built based a given alignment and template. In general a higher QSQE is "better", while this complements the GMQE score which estimates the accuracy of the tertiary structure of the resulting model. QSQE is only computed for the top ranked templates. Protein electrostatic surfaces were calculated using APBS (v1.4) software [52]. Calculations were performed at 310K with 150 mM NaCl with the same grid size (193,193,161) in all cases showing electrostatic potential within ± 3 kTe-1.

2.3. RNA extraction and HPLC analysis of tRNAs

Mollicutes cells were grown to late log phase and harvested by centrifugation at 10000×*g* for 20 min. Cells (0.5×*g*) were washed twice by resuspending in 100 mL buffer A (50 mM Tris-Cl pH 7.2, 10 mM $MgCl_2$, 100 mM NH_4Cl) and pelleting by centrifugation. Cells were lysed by sonication at 4 °C in 10 mL buffer A. Cell debris containing the chromosomal DNA was removed by centrifugation at 15000×*g* for 10 min. The supernatant was extracted with phenol/chloroform and total RNA was recovered by ethanol precipitation before redissolving in 100 µL H_2O. Half of each sample was kept for rRNA analysis (below), and the remainder was passed through a Nucleobond® RNA/DNA 400 column (Macherey-Nagel) to isolate the tRNA fraction. Bulk tRNAs were digested to completion to form nucleosides [53] before being subjected to reverse-phase chromatography on a Agilent Technologies 1200 series HPLC with a Phenomenex Luna C18 column (2 × 250 mm, 5 µm particles, 100 Å pores). Nucleosides were eluted as described previously [37,53] with 40 mM ammonium acetate pH 6 and a linear gradient of 0 to 40% acetonitrile, detecting eluents at 260 nm.

2.4. Analysis of RNA by Matrix-Assisted Laser Desorption-Ionization Mass Spectrometry (MALDI-MS)

Total RNA extracts from mollicutes cells were analyzed within the 23S rRNA regions previously shown in other organisms to contain m^5U methylations. In each case, 100 pmol of total RNA was hybridized to 500 pmol of the 48-mer deoxyoligonucleotide, 5′-GCCACAAGTCATCCAAAGTCTTTTCAACGAATACTGGTTCGGTCCTCC, complementary to the sequence G725-C772 in domain II of 23S rRNA, or to the 55-mer 5′-CGGGTCAGAATTTACCTGACAAGGAATTTCGCTACCTTAGGACCGTTATAGTTAC, complementary to the sequence G1910-G1964 within domain IV of23S rRNA (Figure 4). The exposed regions within the RNAs were digested away with nucleases, and the sequences protected by hybridization were separated by gel electrophoresis [54,55]. The protected rRNA fragments were extracted and digested with RNases A or T1 in aqueous solution and analyzed by MALDI-MS (Ultraflextreme, Bruker Daltronics). Spectra were recorded in reflector and positive ion mode and processed using Flexanalysis (Bruker Daltonics) [56].

2.5. Complementation tests of ΔthyA::kan E. coli strain.

A transition mutation of A to G was introduced at the 5′-end of MCAP_0613 to create an NcoI site, and the gene was cloned into the NcoI and PstI restriction sites of pBAD24. This change corresponds to a K2E substitution at the N-terminus of the MCAP_0613 protein. *E. coli* strains BW25113 (F- Δ(*araD-araB*)567, *lacZ*4787(Δ)::*rrnB*-3, LAM-*rph*-1, DE(*rhaD-rhaB*)568, *hsd*R514) and its Δ*thyA::kan* derivative were transformed with pBAD24 and pBAD24::MCAP_0613. Cells were grown to OD600 of 0.8 in LB medium supplemented with ampicillin (100 $\mu g.mL^{-1}$) and thymidine (0.3 mM). Cells were washed twice in water and 10 µL of serial dilutions were spotted on LB plates containing ampicillin, arabinose (0.02 %) and, in some cases, supplemented with thymidine at 0.3 mM. Cells were incubated at 37°C for 24 h. Growth rates and yields were measured in liquid cultures following standard procedures.

3. Results

3.1. Distribution of predicted 5-methyluracil synthesis enzymes in mollicutes

Thirty-nine representatives of the main phylogenetic sub-groups of the class *Mollicutes* were selected as a reference set for this study (**Figure S1** and **Table S1**). In order to predict the repertoire of enzymes involved in dyUMP, rRNA and tRNA 5-methyluracil modifications in mollicutes, blastp queries against deduced proteomes of the selected reference set were conducted using the input sequences: ThyA (b2827); TrmA (b3965); RlmC (b0859) and RlmD (b2785) from *E. coli*; TrmFO (BSU16130) and RlmCD and YfjO (BSU08020) from *B. subtilis*; ThyX (P9WG57) from *Mycobacterium tuberculosis*; and PolB (Uniprot: C1IC19) from *S. cacaoi* subsp. *asoensis.* We also searched for the presence of homologs of the deoxyT salvage enzyme thymidine kinase Tdk (b1238, BSU37060) as in the absence of ThyA, Tdk becomes essential to provide dTMP precursors. If no candidate enzymes were found using this first approach, further analyses were performed using mollicutes homologs as queries and tblastn. Examples of paralogy were detected with more than one copy of TrmFO or YjfO homologs per genome, and in these instances, further phylogeny and synteny analyses were carried out to separate the sub-families.

TrmFO homologs were found in twelve of the thirty-nine mollicutes species, with ten of the genomes containing more than one copy (**Table S2**). In order to clarify the evolutionary relationships between TrmFO-related homologs, a phylogenetic tree was constructed adding other TrmFO homologs identified in Gram-positive bacteria and in recently sequenced mollicutes genomes (**Figure S2**). Two main groups supported by 100% statistical values were clearly identified. One of them includes the TrmFO-related homolog (MCAP0476) from *M. capricolum* subsp. *capricolum* that was

previously shown to catalyze the formation of m^5U1939 in 23S rRNA, and subsequently renamed RlmFO [37]. The genomic context region around RlmFO encodes genes that are highly conserved, even among remote species of the Spiroplasma phylogenetic group (**Figure S3**). The other well-defined group of TrmFO-related homologs includes the *M. capricolum* subsp. *capricolum* paralog MCAP0613, the function of which remains unknown, and we have renamed this subgroup 'TrmFO-like'. Analysis of the genomic context of TrmFO-like homologs showed only a moderate conservation between related mycoplasmas from the Mycoides cluster (**Figure S4**). The remaining TrmFO-related homologs were distributed between two other subgroups supported by statistical values of 85% and 87%, separating proteins in the Acholeplasma clade from the rest of mollicutes. Gene synteny was however observed among genomic regions surrounding all other *trmFO*-related homologs (**Figure S5**), showing some conservation of gene order with *trmFO* homologs from other Gram-positive bacteria including *B. subtilis*. This conserved synteny suggests that all these *trmFO*-related genes are true orthologs of the genuine ancestral *trmFO* present in the common ancestor of Gram-positive bacteria and mollicutes, and we now refer to them as the TrmFO subgroup.

Four RlmD homologs were identified in *Acholeplasma laidlawii,* an unexpectedly high number for such a genome-reduced bacterium (1.5 Mpb). Further investigation showed that most Acholeplasma species also have four homologs, and this seems to be a recent expansion in the group (**Figure S6, Table S3**). The full comparative genomic analysis with accession numbers of all identified proteins, and the grouping in the different paralogous subgroups (**Table S2**), is summarized in **Figure 2** . The functional hypotheses derived from this information are presented below.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2020 doi:10.20944/preprints202003.0097.v1

Peer-reviewed version available at Biomolecules 2020, 10, 587; doi:10.3390/biom10040587

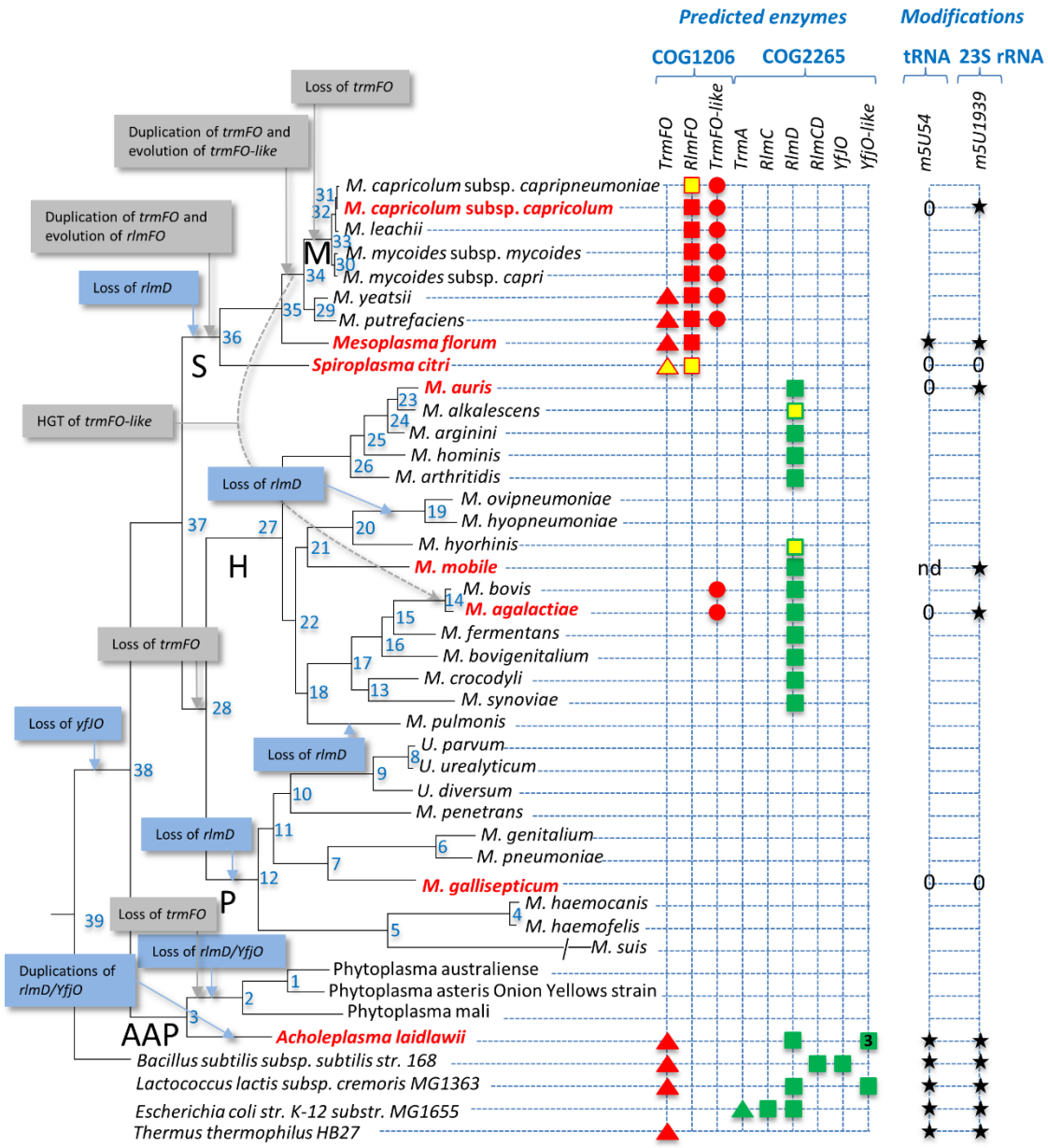


Figure 2. Occurrence of genes involved in rRNA/tRNA modification in mollicutes. The phylogenetic tree was inferred using the maximum likelihood method from the concatenated multiple alignments of 79 proteins encoded by genes present at one copy in each genome. Nodes are numbered. Main phylogenetic groups are indicated, S, Spiroplasma, H, Hominis, P, Pneumoniae, AAP, Acholeplasma/Phytoplasma, M, Mycoides cluster of ruminant mycoplasmas. Non-cultivated species are framed by a red dotted rectangle. Species analysed by MS are indicated in bold red. Targets are indicated as follows: square, rRNA, triangle, tRNA, circle, unknown. Symbols are coloured as follows: red, folate-dependent enzymes; green, SAM-dependent enzymes; yellow, predicted pseudogene. Predicted evolution events are indicated in grey boxes. nd, not determined.

3.2. Synthesis by ThyA homologs and deoxyT salvage are the two main routes to dTMP synthesis in mollicutes.

The phylogenetic distribution analysis of enzyme involved in dTMP synthesis revealed that there are no homologs of ThyX and PolB in mollicutes. In addition, ~40% (16/39) of the mollicutes genomes analyzed encode a ThyA homolog, and these are scattered over different phylogenetic subgroups (**Table S2**). Furthermore, all but one of the mollicutes (*Mycoplasma bovigenitalium* cl-51080)

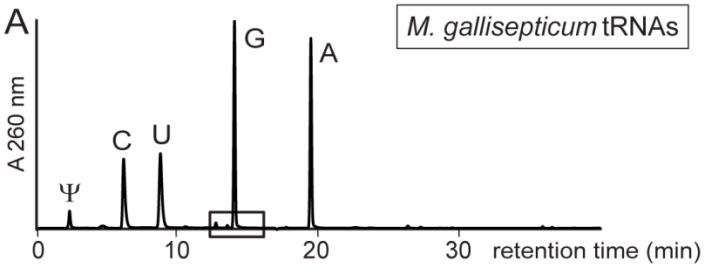
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2020 doi:10.20944/preprints202003.0097.v1

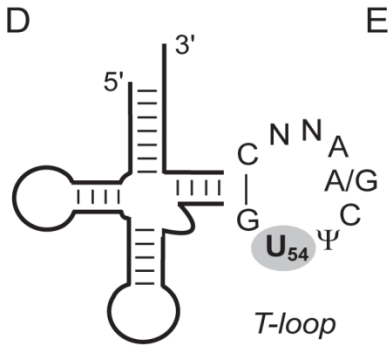
Peer-reviewed version available at Biomolecules 2020, 10, 587; doi:10.3390/biom10040587

encode Tdk homologs (**Table S2**) and, as no other specific enzyme for dTMP synthesis is presently known, this would suggest that thymidine salvage is the major pathway for this process in mollicutes.

3.3. Most mollicutes have lost the m⁵U54 modification in tRNA

No bacterial TrmA homolog was found in any of the mollicutes (**Table S2**) and the majority (34/39) also lack a TrmFO homolog, suggesting that the corresponding tRNA modification has been lost in most members of the class. Putative orthologs of TrmFO are present only in *A. laidlawiii* and in four members of the Spiroplasma group, where in one of these, *Spiroplasma citri*, it is a pseudogene (see below) and thus presumably in the process of being lost (**Figure 2**). In order to test our functional predictions, we analyzed the tRNAs from seven strains for the presence of m⁵U. A combination of strains was chosen to cover the various permutations of absence and presence of TrmFO, RlmFO and TrmFO-like homologs. The tRNA m⁵U modification was found to be present only when a strain encodes what appears to be a functional *trmFO* gene (**Figure 3**).





E

Species	Predicted enzyme	m^5U_{54}
A. laid	ACL_1161	+
Me. flor	Mfl532	+
S. citri	[*TrmFO*][1]	-
M. agal	*none*	-
M. auri	*none*	-
M. mob	*none*	*nd*
M. capr	*none*[2]	-
M. gall	*none*	-

Figure 3. HPLC profiles of tRNA nucleosides from Mollicutes bulk tRNAs. (**A**) The digestion products of total tRNAs from *M. gallicepticum* showing the positions of the unmodified nucleosides and pseudouridine detected by their absorbance at 260 nm (A260 nm). Under the conditions used here, m⁵U elutes in the boxed region after guanosine at approximately 15 min (Lartigue et al 2014). (**B**) Enlargement of the boxed region for *M. gallicepticum* and (**C**) for *Me. florum*, where only the tRNAs of this latter strain contain m⁵U. (**D**) Schematic of the mollicutes tRNAs showing the consensus sequence of the T-loop nucleotides and the position of U54. (**E**) Mollicutes species shown empirically to have (+) or not have (-) m⁵U54 in their tRNAs. This modification correlates with the presence of an active TrmFO enzyme predicted by bioinformatics. Species abbreviations: *A. laid*, *A. laidlawii*; *Me. flor*, *Me florum*; *S. citri*, *S. citri*; *M. agal*, *M. agalactiae*; and the additional *Mycoplasma* species *M. auris (M. auri)*, *M. mobile (M. mob)*, *M. capricolum* subsp. *capricolum (M. capr)*, and *M. gallisepticum (M. gall)*. [1]The *trmFO* sequence in *S. citri* is a nonfunctional pseudogene; [2]*M. capr* has no *trmFO*, but possesses a functional *rlmFO*. nd, not determined.

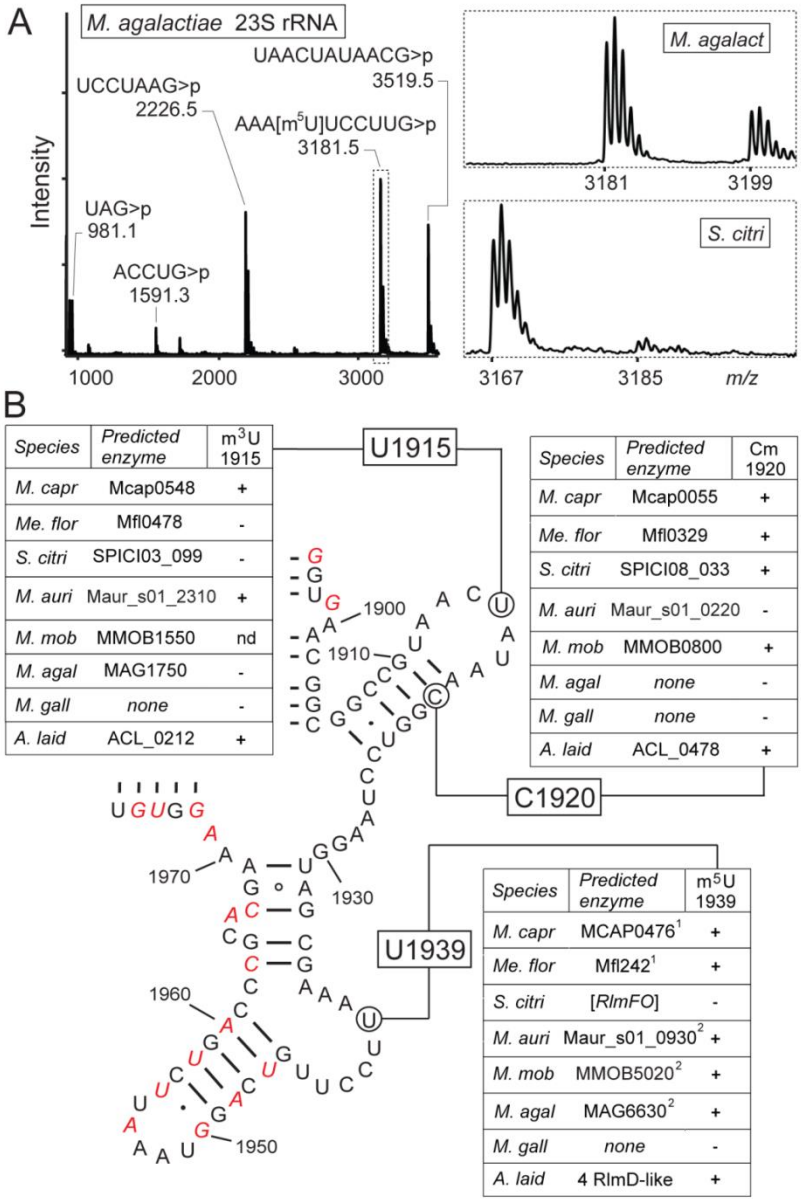
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2020 doi:10.20944/preprints202003.0097.v1

Peer-reviewed version available at Biomolecules 2020, 10, 587; doi:10.3390/biom10040587

This modification is absent in the tRNAs from *M. capricolum* subsp. *capricolum*, which encodes both RlmFO and TrmFO-like genes but no TrmFO, and is also absent in *S. citri* which encodes a TrmFO-pseudogene.

3.4. Mollicutes modify 23S rRNA m⁵U1939 via either RlmD or RlmFO.

Some species of mollicutes possess a homolog of either RlmD or RlmFO that directs the 23S rRNA modification at m⁵U1939 (**Figure 4**). The genes encoding these two enzymes appear mutually exclusive in these bacteria. However, some mollicutes appear to be in the process of losing their rRNA methyltransferase and, for instance, this is seen as a degenerate non-functional version of the *rlmFO* gene in *S. citri* (**Figure 4**) and what appears to be an *rlmD* pseudogene in some of the *Mycoplasma hyorhinis* strains (**Figure 2**). No RlmC homolog was identified in any of the mollicutes, nor were there any homologs of the dual-specific enzyme RlmCD seen in *B. subtilis* [36] and *Streptococcus pneumoniae* [57]. Consistent with this, MS analyses of the G725-C772 in domain II of the mollicutes 23S rRNAs confirmed that there was no modification at U747, or at any other nucleotide in this region (not shown).



Species	Predicted enzyme	m³U 1915
M. capr	Mcap0548	+
Me. flor	Mfl0478	-
S. citri	SPICI03_099	-
M. auri	Maur_s01_2310	+
M. mob	MMOB1550	nd
M. agal	MAG1750	-
M. gall	*none*	-
A. laid	ACL_0212	+

Species	Predicted enzyme	Cm 1920
M. capr	Mcap0055	+
Me. flor	Mfl0329	+
S. citri	SPICI08_033	+
M. auri	Maur_s01_0220	-
M. mob	MMOB0800	+
M. agal	*none*	-
M. gall	*none*	-
A. laid	ACL_0478	+

Species	Predicted enzyme	m⁵U 1939
M. capr	MCAP0476[1]	+
Me. flor	Mfl242[1]	+
S. citri	[*RlmFO*]	-
M. auri	Maur_s01_0930[2]	+
M. mob	MMOB5020[2]	+
M. agal	MAG6630[2]	+
M. gall	*none*	-
A. laid	4 RlmD-like	+

Figure 4. (**A**) MALDI-MS spectrum of RNase T1 fragments from the *M. agalactiae* 23S rRNA region around nucleotide U1939. The fragment sequences are shown above the peaks with their protonated masses (*m/z*); the *m/z* measurements are for fragments with a cyclic 3′-phosphate (>p), and are within 0.1 Da of the theoretical monoisotopic values. The spectral region (boxed) with nucleotide 1936-1945 fragment is expanded, showing that the *M. agalactiae* sequence at *m/z* 3181 contains the m^5U1939 modification. The lower box shows the same spectral region from *S. citri* 23S rRNA, where the corresponding fragment at *m/z* 3167 is unmethylated. Minor peaks of the hydrated linear fragments with the same sequence (+ 18 Da) are visible in both spectra. Also of note here is the *M. agalactiae* nucleotide 1911-1921 containing U1915 and C1920, and with *m/z* 3519 is unmodified. These two nucleotides are analysed in greater detail in **Figure S7**. (**B**) Schematic of the mollicutes 23S rRNA secondary structure around the potential methylation sites at nucleotides U1915, C1920 and U1939. The structure shown here is highly conserved in all bacteria, and varies amongst the mollicutes only at the nucleotides highlighted in red (where the sequence specific for *M. capricolum* subsp. *capricolum* is shown here). The tables list the bioinformatics predictions for the presence of methyltransferases in *A. laidlawii*, *Me. florum*, *S. citri*, *M. agalactiae*; *M. auris*, *M. mobile*, *M. capricolum* subsp. *capricolum* and *M. gallisepticum* (abbreviated as in **Fig. 3**); the presence of the rRNA modifications was ascertained empirically, as above. Where present, the m^5U1939 modification was added either by an RlmD homolog or an RlmFO homolog; *S. citri* appears to contain a non-functional pseudo-RlmFO. *A. laidlawii* possesses four RlmD paralogs, and it is not yet clear which of these has the m^5U1939 modification function. nd, not determined.

Close to the U1939 region of the 23S rRNA, two other nucleotide modifications were identified in some of the mollicutes. The m^3U1915 modification, which usually is dependent on the prior isomerization of U1915 to Ψ [58,59], lies close to Cm1920 within helix 69. Helix 69 is essential for ribosomal subunit interaction, P-site tRNA binding and recycling of the ribosome after translational termination, and these modifications are thought to facilitate these processes [60-62]. Intriguingly, combinations of these modifications in mollicutes ranged from all three (m^3U1915, Cm1920 and m^5U1939) in *M. capricolum* subsp. *capricolum* and *A. laidlawii*; to various pairs of two (Cm1920/m^5U1939 or m^3U1915/m^5U1939) respectively in *Me. florum* and *M. auris*; to different single modifications in *S. citri* (Cm1920) and *M. agalactiae* (m^5U1939); to none at all in *M. gallisepticum* (**Figure 4 and S7**). This array can be compared to bacteria with larger genomes where *B. subtilis* has all three modifications, while *E. coli* and other Enterobacteria make do with $m^3Ψ1915$ and m^5U1939 [63]. Surprisingly, a gene encoding the RlmH enzyme responsible for m^3U1915 modification was predicted in *Me. florum*, *S. citri* and *M. agalactiae* while this site was shown to remain unmodified. This was despite the presence of an *rluD* gene homolog (encoding the enzyme for U1915 isomerization) in each of these three organisms, and would suggest a change or loss of either RlmH or RluD function.

3.5. Binding sites of folate and flavin are conserved in TrmFO-like proteins

The empirical analyses of the mollicutes rRNAs and tRNAs confirmed the functions of the TrmFO and RlmFO enzyme subgroups in respectively targeting U1939 in 23S rRNA and U54 in tRNA. However, the TrmFO-like subgroup modifies neither these sites nor U747, and its function remains unknown. For example, *M. capricolum* subsp. *capricolum* encodes the TrmFO-like protein MCAP0613 in addition to the RlmFO protein MCAP0476 that adds the m^5U1939 rRNA modification, and lacks the m^5U54-tRNA modification (**Figure 3**) [37]. Aligning the nine available TrmFO-like sequences from different *Mycoplasma* spp. with the sequence of canonical m^5U54 flavin- and folate-dependent methyltransferase TrmFO from *Thermus thermophilus* ($TrmFO_{Tt}$) showed that the TrmFO-like proteins are approximately twenty residues shorter than $TrmFO_{Tt}$ (**Figure 5**). Structural information is available for $TrmFO_{Tt}$ (PDB: 3G5S) and this enzyme exhibits roughly 30% sequence identity with the TrmFO-like proteins.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2020 doi:10.20944/preprints202003.0097.v1

Peer-reviewed version available at Biomolecules 2020, 10, 587; doi:10.3390/biom10040587

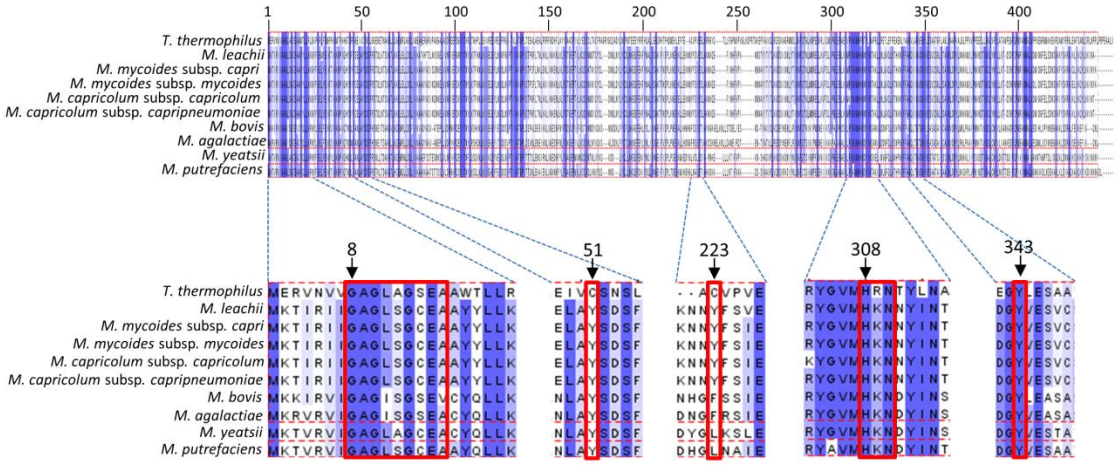


Figure 5. Conserved positions in TrmFO-like homologs. Multiple alignement of TrmFO-like proteins with reference TrmFO from *T. thermophilus* was implemented into Jalview to get an overview of the conserved positions. Amino-acids conserved at >50% are indicated in blue. Important regions described in the main text are enlarged underneath the overview alignment. Positions indicated by arrows correspond to numbering in the reference TrmFO protein of *T. thermophilus*.

We note several key points from the alignment of the TrmFO$_{Tt}$ and TrmFO-like protein sequences (**Figure 5**). First, the GAGx[A/S]GxE[A/V] motif involved in the recognition of the pyrophosphate group of FAD is conserved in TrmFO-like proteins. Second, the residues H308 R309/or K309 and N310 that are specifically involved in the folate binding are also strictly conserved. Here, the TrmFO$_{Tt}$ crystal structure shows that the side chain of H308 is rotated from its original position in the free-form structure to interact with the pteridine moiety [29]. Also, the R309 residue of TrmFO$_{Tt}$ is replaced in the TrmFO-like sequences by lysine, another positively charged residue. Finally, the two catalytic cysteines C51 and C223, which are essential for the U54 methylating activity of TrmFO, are replaced by tyrosine residues in all the TrmFO-like proteins. In TrmFO, C51 acts as a general base and C223 plays the role of a nucleophile that activates the C5-carbon of the uracil target [39,41]. A cysteine residue at position 195-196 of the alignment seems to be conserved in both TrmFO subgroups, but this residue is located far from the active site and is not involved in *B. subtilis* TrmFO catalysis [41].

The substitution of C51 and C223 with tyrosines is illustrated in the structural model of the TrmFO-like protein MCAP_0613 (**Figure 6**) that is based on the TrmFO$_{Tt}$ crystal structure (PDB: 3G5S).

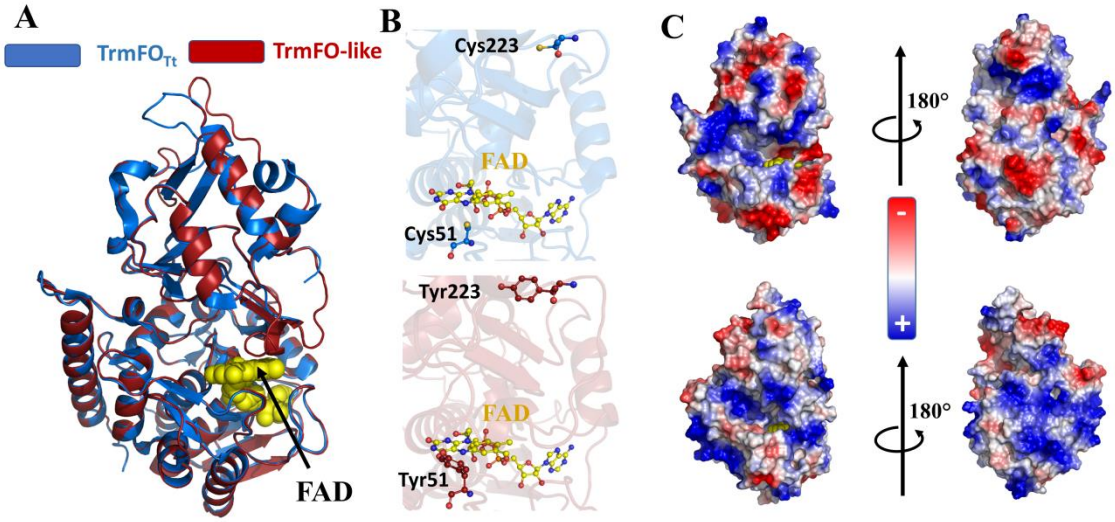


Figure 6. Structural model of TrmFO-like protein. (A) Structural overlay of the crystal structure of *T. thermophilus* TrmFO (blue) with the 3D-model of *M. capricolum* subsp. *capricolum* TrmFO-like protein MCAP_0613 (red). The FAD coenzyme is represented as ball sticks in yellow. (B) Location of the two cysteines (top panel) that are strictly conserved and required for m^5U54-tRNA methylation activity in $TrmFO_{Tt}$. The lower panel shows the two tyrosine residues at the corresponding locations in TrmFO-like proteins. (C) Electrostatic surface of $TrmFO_{Tt}$ (top) and the TrmFO-like protein MCAP_0613 from *M. capricolum* (bottom).

The model was built using the fully automated protein structure homology-modelling server, *SWISS-MODEL*, achieving acceptable reliability values with GMQE and QMEAN scoring functions of nearly 0.7 and -3.28, respectively. Alignment of the TrmFO-like model against the crystal structure of $TrmFO_{Tt}$ does not reveal any major changes (RMSD = 0.353 Å, over 370 atoms), which would indicate that TrmFO-like proteins adopt the same structural topology seen in the TrmFO/RlmFO subgroups. Compared to $TrmFO_{Tt}$, the identities and spatial locations of all of the residues required to bind FAD and THF are fully conserved in TrmFO-like proteins. Notably, the peculiar Y343 residue that stacks against the isoalloxazine ring in $TrmFO_{Tt}$, and plays an essential role in maintaining active redox state of FAD [64-66], is also preserved in TrmFO-like proteins and could feasibly have a similar function. The structural model shows that the two tyrosine residues Y51 and Y223 occupy positions identical to the two cysteines which they replace in $TrmFO_{Tt}$.

Our sequence and structural analyses suggest that TrmFO-like proteins are similar to TrmFO folate- and FAD-binding proteins, although their enzymatic functions differ. Synteny analysis did not give any clear hint about TrmFO-like function, although the gene immediately downstream of *trmFO*-like possibly encodes an EcfS binding component of a folate ECF transporter (**Figure S4**), reinforcing the putative connection in folate metabolism.

3.6. TrmFO-like subfamily members are presumably nucleic acid binding proteins

As a prelude to methylation by ThyX, uracil is activated via polarization of two arginine residues within the enzyme's active site. This differs from that mechanism of ThyA, TrmFO/RlmFO, RlmC and RlmD, which involve the use of a nucleophile (**Table 1**). The replacement of both cysteines by two tyrosines (**Figure 6**) led us to first hypothesize that TrmFO-like proteins could function as 5-methyluracil methyltransferases via a mechanism of uracil activation that differs from ThyX and ThyA (**Figure 1**). This hypothesis was tested genetically by attempting to suppress the dT auxotrophy of an *E. coli thyA* mutant by expressing the MCAP_0613 gene in *trans*. However, no suppression of thymidine auxotrophy was observed, although expressing MCAP0613 both in BW25113 and *ΔthyA* cells in the presence of dT, did lead to higher cell densities both on plates and in liquid cultures density (**Figure 7)**.

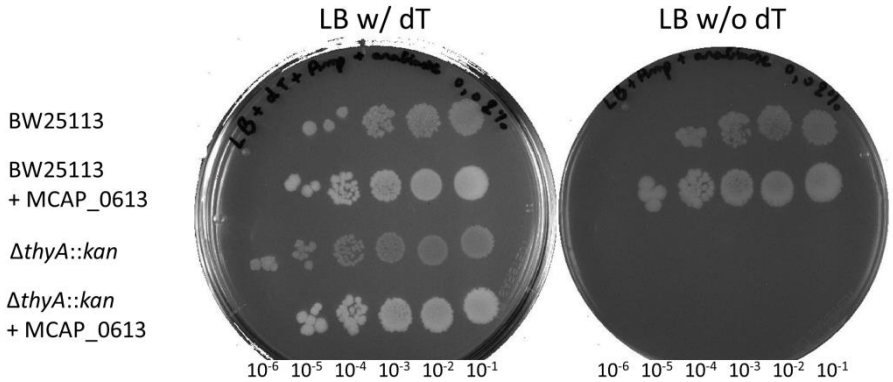


Figure 7. Complementation tests of Δ*thyA*::kan *E. coli* strain. Cells were grown as described in Material and Methods. Serial dilutions were spotted from right to left, as indicated.

The absence of thymidylate synthase activity led us to explore a potential implication of this protein in nucleic acid metabolism. Accordingly, we undertook the calculation of the electrostatic surface of TrmFO-like models. Analysis of the electrostatic surface of the *M. capricolum* subsp. *capricolum* TrmFO-like protein revealed two notable features (**Figure 6C**). First, in a manner similar to $TrmFO_{Tt}$, the TrmFO-like protein harbours electropositive patches that surround its active site. This electropositive surface extends around the TrmFO-like protein and is accessible to the solvent, whereas $TrmFO_{Tt}$ has a negatively charged surface on the face opposite to the active site. This suggests that the substrate for TrmFO-like protein might be a large nucleic acid molecule. Taking this idea further, we generated models to analyze the surface electrostatic potentials of TrmFO-like proteins from *M. mycoides, M. yeatsii, M. putrefaciens, M. bovis* and *M. agalactiae*, obtaining QMQE and QMEAN values that indicate the models are reliable (**Table S4**). The TrmFO-like proteins, with the exception of the *M. bovis* and *M. agalactiae* homologs, exhibit similar electrostatic surfaces with positively charged patches around the FAD binding site **(Figure S8)**. Unexpectedly, the TrmFO-like models for *M. bovis* and *M. agalactiae* have electronegative patches formed by a pair of glutamate residues, E204/E283, located in the inserted domain at the distal side of FAD. This would reduce significantly the electropositive surface of these proteins compared to the other TrmFO-like proteins. Consistently, these results could indicate that the *M. bovis* and *M. agalactiae* TrmFO-like proteins have progressively lost tRNA-binding capability while specializing in another cellular function that requires both flavin and folate.

4. Discussion

Taking into account the phylogenetic data, the genomic context and the experimental validation presented here, we can propose different scenarios for the reductive evolution of the COG2265 and COG1206 families from bacteria with larger genomes to mollicutes. Evolution of the SAM-dependent methyltransferases (COG2265) seems rather simple. We hypothesize that one or more the bacterial genes *rlmC, rlmD, rlmCD* and *YfjO*, but not *trmA*, were present in the ancestor of mollicutes (**Figure 2**). This is in agreement with the origin of mollicutes from Gram-positive ancestors. The gene coding for YfjO was duplicated several times only during the evolution of the Acholeplasma branch and was completely lost in Phytoplasmas (node 2) as well as in all other mollicutes examined (node 38). The function of the different YfjO-like paralogs of *Acholeplasma* spp. (at least three) is not yet understood and, with the exception of the loss of RNA-binding TRAM domains in the A2 group, no major differences were observed in the residues within the active site of these proteins (**Figure S9**). It is of course possible that these YfjO-like paralogs have acquired another function and do not methylate an RNA macromolecule. The other ancestral *rlmC*/*rlmD*/*rlmCD* genes, which encode rRNA-specific methyltransferases, were all lost in the ancestors of the Spiroplasma (node 36) and Pneumoniae groups (node 12), whereas in the Hominis group, *rlmD* was kept, still encoding the m^5U1939-specific methyltransferase (node 27). While this is the most parsimonious scenario explaining the data, a novel acquisition event in the ancestor of the Hominis group cannot be formally excluded. A more recent *rlmD* loss was observed for some species, including *M. pulmonis*, and subgroups of species including *M. ovipneumoniae* and *M. hyopneumoniae* (node 19). A recent gene-essentiality study based on transposon mutagenesis in *M. bovis* indicated that *rlmD* (MBOVJF4278_00748) was not essential [67] reinforcing the dispensability of these rRNA methylases.

The evolution of the folate-dependent methyltransferases is more complex than the SAM-dependent ones. Again, in agreement with the origin of mollicutes from Gram-positive ancestors, the formation of m^5U54 formation in mollicute tRNAs is encoded only by *trmFO* genes and never by *trmA* genes as in Gram-negative bacteria. This *trmFO* gene was probably present in the ancestor of mollicutes (node 38), and in the ancestors of the Acholeplasma/Phytoplasma (AAP, node 3) and Spiroplasma/ Hominis/Pneumoniae (SHP, node 37) groups. During the evolution of the AAP group, *trmFO* was maintained in Acholeplasma species but lost in the ancestor of Phytoplasma (node 2). The general lack of any *trmFO*-related gene in the Hominis and Pneumoniae groups indicates a probable

loss in the common ancestor of the two sub-groups (node 28). By contrast, the distribution of *trmFO*-related genes in the Spiroplasma phylogenetic group (S) suggests a different stepwise evolution. First, there has been a duplication in the subgroup ancestor (node 36) with evolution of the *rlmFO* paralog to encode a folate-dependent methyltransferase responsible for the 23S RNA m^5U1939 modification. This was followed by a second duplication of *trmFO* or *rlmFO* in the ancestor of the mycoplasmas of ruminants (node 34) with subsequent evolution of *trmFO*-like paralogs. Cases of gene degradation are also visible in the *S. citri trmFO* and *rlmFO* genes, and for *rlmFO* of *M. capricolum* subsp. *capripneumoniae*. In addition, a *trmFO*-like gene from the Hominis subgroup was probably transferred by HGT (discussed below) to the ancestor of *M. agalactiae*/*M. bovis*, and finally, there was a loss of the original *trmFO* in the ancestor of the Mycoides cluster (node 33). HGT between the ruminant mycoplasmas from the Mycoides cluster and the *M. agalactiae*/*M. bovis* cluster have been described previously [68-70].

The complex evolution among the Spiroplasma phylogenetic group (node 36) may have favored the parallel diversification of their methyltransferases. Indeed, appearance of the unique *rlmFO* paralog encoding a folate-dependent 23S rRNA methylase could correspond to an evolutionary relay to maintain the m^5U1939 modification within the Spiroplasma group as *rlmD* was being lost [37]. From this point of view, enzymatic activity of RlmD and of RlmFO proteins appears to be mutually exclusive, as described for TrmA and TrmFO above. More striking is the occurrence of the *trmFO*-like paralog in the mycoplasmas of ruminants related to the Mycoides cluster, leading to species with up to three *trmFO*-related genes (i.e. *M. putrefaciens* and *M. yeatsii*). *In silico* analyses suggest that these TrmFO-like proteins might have conserved a folate- and flavin-dependent methylase activity, however their substrate(s) remain unknown (**Figure 6** and **Figure S8**). We have ruled out a potential role in dTMP synthesis (**Figure 7**). Interestingly, the presence of important patches of positive charges around the active site surface of TrmFO-like structures from *M. capricolum* subsp. *capricolum, M. leachi, M. mycoides* subsp. *capri, M. yeatsii* and *M. putrefaciens* suggests that the substrate for these proteins would be a nucleic acid. If TrmFO-like proteins are still methylases, that would require methylation mechanism differing from that of TrmFO/RlmFO because of the replacement of the nucleophilic cysteine with a tyrosine (Figure 6). The role of this tyrosine as nucleophile would be atypical in nucleic acids enzymology but is not unprecedented as several glycosidases utilize a tyrosine believed to be activated by an adjacent base residue as a catalytic nucleophile [71,72].

As mentioned above, genome comparisons indicate that an HGT of a *trmFO*-like gene probably took place from the Mycoides cluster to the phylogenetically remote *M. bovis*/*M. agalactiae* cluster. Among the genes predicted to have been subjected to HGT between these ruminant pathogens, most are related to virulence, metabolism and mobile elements and none were known to be related to the processing or maintenance of genetic information [68-70]. Further analyses on all available complete genomes of *M. agalactiae* (four strains) and *M. bovis* (eleven strains) indicated that a *trmFO-like* gene has been conserved in these species (**not shown**), suggesting that it is biologically significant. It is possible that the transfer and subsequent fixation of trmFO-like genes are not due to a role a in RNA methylation but to fulfil a new moonlighting function (see [73] for examples). Our structural homology models indicate that the TrmFO-like proteins of *M. bovis*/*M. agalactiae* have a smaller electropositive surface than their counterparts in other mollicutes, which would suggest that they may have lost their nucleic acid binding capacity while preserving a biological activity that depends on both folate and flavin.

In relation to possible moonlighting function of certain methyltransferases, a recent paper by [74], has claimed that the TrmFO-like protein (renamed according to our definition here) functions as an adhesin in *M. bovis*. Their conclusion was mainly driven by the demonstration of a fibronectin-binding activity of TrmFO-like using ELISA and direct adhesion assays on embryonic bovine lung (EBL) cells, including inhibition by anti-TrmFO-like polyclonal antibiodies. In another study, a transposon library of *M. bovis* strain JF-4278 was recently shown to include a *trmFO*-like disrupted mutant which ability to bind to primary bovine mammary gland epithelial (bMec) cells was reduced, suggesting this gene is a virulence factor in *M. bovis* [67].

5. Conclusions

In summary, our genetic and structural analyses illustrate evolutionary schemes in which some mollicutes species have kept, while others have lost, the characteristic ancestral Gram-positive C5-uracil methyltransferases. Aa a consequence of such changes, several new enzyme paralogs have evolved in branches of the mollicutes, where it can be seen that RlmFO and TrmFO-like proteins in most Spiroplasma, and YijO-like proteins in Acholeplasma remain unique to the mycoplasma/mollicute clade. The drastic genomic reduction at an earlier stage in the evolution of other mollicutes has led to the loss of their capacity to catalyze C5-uracil methylation, and similar findings have been made for other protein factors and enzymes connected with the protein synthesis machinery [47].

Our present studies on the mollicutes COG2265 and COG1206 methyltransferases support the notion that several unique, evolutionary trajectories have taken place within the different phyla of these minimal fast-evolving bacteria, and this has resulted in both the convergence and divergence of enzyme functions. This adds credence to the idea that mollicutes have displayed a remarkable ability to recycle genes for new functions.

Supplementary Materials: The following are available online, Figure S1: Phylogenetic tree of mollicutes, Figure S2: Phylogenetic analysis of TrmFO-related proteins in mollicutes, Figure S3: Genomic contextof *rlmFO* homologs, Figure S4: Genomic context of *trmFO*-like homologs, Figure S5: Genomic context of *trmFO* homologs, Figure S6: Phylogenetic analysis of RlmD/RlmCD/YfjO- related proteins in mollicutes, Figure S7: MALDI-MS analyses for m^3U1915 and Cm1920 methylations, Figure S8: Electrostatic surface of model of TrmFO-like proteins from different mycoplasma, Figure S9: Conserved cysteins in RlmD-related proteins, Table S1: List of 39 selected mollicutes, Table S2: Homologs of tRNA and rRNA modification enzymes and some other folate and S-AdoMet dependent enzymes in mollicutes, Table S3: Supplementary RLMD/RLMCD/YFJO used for phylogenetic analyses, Table S4: GMQE and QMEAN values for TrmFO-like proteins models.

Author Contributions: Conceptualization, P.S.-P., H.G., S.D., D.H. and V.d.C.-L. ; data production and analyses, P.S.-P., D.B., C.G., L.B., S.R., S.D., D.H. and V.d.C.-L., paper writing, P.S.-P., D.B., L.B., A.B., H.G., S.D., D.H. and V.d.C.-L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Centre National de la Recherche Scientifique, Université Pierre et Marie Curie, Université de Bordeaux as well as program from the French State (IDEX Sorbonne Université ANR-11-IDEX-0004-02 and ANR-15-CE11-0004-01 to D.H.). This work was funded in part by the National Institutes of Health (grant GM70641 to V.d.C.-L.) and support from the Danish Research Agency (FNU-rammebevilling 10-084554) to S.D. is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hori, H. Methylated nucleosides in tRNA and tRNA methyltransferases. *Frontiers in genetics* **2014**, *5*, 144, doi:10.3389/fgene.2014.00144.
2. Motorin, Y.; Helm, M. RNA nucleotide methylation. *Wiley interdisciplinary reviews. RNA* **2011**, *2*, 611-631, doi:10.1002/wrna.79.
3. Myllykallio, H.; Sournia, P.; Heliou, A.; Liebl, U. Unique features and anti-microbial targeting of folate- and flavin-dependent methyltransferases required for accurate maintenance of genetic information. *Frontiers in microbiology* **2018**, *9*, 918, doi:10.3389/fmicb.2018.00918.
4. Juhling, F.; Morl, M.; Hartmann, R.K.; Sprinzl, M.; Stadler, P.F.; Putz, J. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic acids research* **2009**, *37*, D159-162, doi:10.1093/nar/gkn772.
5. Felden, B.; Hanawa, K.; Atkins, J.F.; Himeno, H.; Muto, A.; Gesteland, R.F.; McCloskey, J.A.; Crain, P.F. Presence and location of modified nucleotides in *Escherichia coli* tmRNA: structural mimicry with tRNA acceptor branches. *The EMBO journal* **1998**, *17*, 3188-3196, doi:10.1093/emboj/17.11.3188.
6. Boccaletto, P.; Machnicka, M.A.; Purta, E.; Piatkowski, P.; Baginski, B.; Wirecki, T.K.; de Crecy-Lagard, V.; Ross, R.; Limbach, P.A.; Kotter, A., et al. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic acids research* **2018**, *46*, D303-D307, doi:10.1093/nar/gkx1030.
7. Johnson, L.; Hayashi, H.; Söll, D. Isolation and properties of a transfer ribonucleic acid deficient in ribothymidine. *Biochemistry* **1970**, *9*, 2823-2831, doi:10.1021/bi00816a011.

8. Vani, B.R.; Ramakrishnan, T.; Taya, Y.; Noguchi, S.; Yamaizumi, Z.; Nishimura, S. Occurrence of 1-methyladenosine and absence of ribothymidine in transfer ribonucleic acid of *Mycobacterium smegmatis*. *Journal of bacteriology* **1979**, *137*, 1084-1087.
9. Hsuchen, C.C.; Dubin, D.T. Methylation patterns of mycoplasma transfer and ribosomal ribonucleic acid. *Journal of bacteriology* **1980**, *144*, 991-998.
10. Piekna-Przybylska, D.; Decatur, W.A.; Fournier, M.J. The 3D rRNA modification maps database: with interactive tools for ribosome analysis. *Nucleic acids research* **2008**, *36*, D178-183, doi:10.1093/nar/gkm855.
11. Sergeeva, O.V.; Bogdanov, A.A.; Sergiev, P.V. What do we know about ribosomal RNA methylation in *Escherichia coli*? *Biochimie* **2015**, *117*, 110-118, doi:10.1016/j.biochi.2014.11.019.
12. Motorin, Y.; Helm, M. tRNA stabilization by modified nucleotides. *Biochemistry* **2010**, *49*, 4934-4944, doi:10.1021/bi100408z.
13. Chawla, M.; Oliva, R.; Bujnicki, J.M.; Cavallo, L. An atlas of RNA base pairs involving modified nucleobases with optimal geometries and accurate energies. *Nucleic acids research* **2015**, *43*, 9573, doi:10.1093/nar/gkv925.
14. Lorenz, C.; Lunse, C.E.; Morl, M. tRNA modifications: impact on structure and thermal adaptation. *Biomolecules* **2017**, *7*, doi:10.3390/biom7020035.
15. Whipple, J.M.; Lane, E.A.; Chernyakov, I.; D'Silva, S.; Phizicky, E.M. The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. *Genes & development* **2011**, *25*, 1173-1184, doi:10.1101/gad.2050711.
10.1101/gad.2050711.
16. Hopper, A.K.; Huang, H.Y. Quality control pathways for nucleus-encoded eukaryotic tRNA biosynthesis and subcellular trafficking. *Molecular and cellular biology* **2015**, *35*, 2052-2058, doi:10.1128/MCB.00131-15.
17. Chen, W.; Li, Y.; Li, J.; Wu, L.; Wang, R.; Deng, Z.; Zhou, J. An unusual UMP C-5 methylase in nucleoside antibiotic polyoxin biosynthesis. *Protein & cell* **2016**, *7*, 673-683, doi:10.1007/s13238-016-0289-y.
18. Hamdane, D.; Grosjean, H.; Fontecave, M. Flavin-dependent methylation of RNAs: complex chemistry for a simple modification. *Journal of molecular biology* **2016**, *428*, 4867-4881, doi:10.1016/j.jmb.2016.10.031.
19. Hou, Y.M.; Perona, J.J. Stereochemical mechanisms of tRNA methyltransferases. *FEBS letters* **2010**, *584*, 278-286, doi:10.1016/j.febslet.2009.11.075.
20. Kealey, J.T.; Gu, X.; Santi, D.V. Enzymatic mechanism of tRNA (m^5U54)methyltransferase. *Biochimie* **1994**, *76*, 1133-1142, doi:10.1016/0300-9084(94)90042-6.
21. Swinehart, W.E.; Jackman, J.E. Diversity in mechanism and function of tRNA methyltransferases. *RNA biology* **2015**, *12*, 398-411, doi:10.1080/15476286.2015.1008358.
22. Carreras, C.W.; Santi, D.V. The catalytic mechanism and structure of thymidylate synthase. *Annual review of biochemistry* **1995**, *64*, 721-762, doi:10.1146/annurev.bi.64.070195.003445.
23. Myllykallio, H.; Leduc, D.; Filee, J.; Liebl, U. Life without dihydrofolate reductase FolA. *Trends in microbiology* **2003**, *11*, 220-223.
24. Myllykallio, H.; Lipowski, G.; Leduc, D.; Filee, J.; Forterre, P.; Liebl, U. An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **2002**, *297*, 105-107, doi:10.1126/science.1072113.
25. Mishanina, T.V.; Yu, L.; Karunaratne, K.; Mondal, D.; Corcoran, J.M.; Choi, M.A.; Kohen, A. An unprecedented mechanism of nucleotide methylation in organisms containing t*hyX*. *Science* **2016**, *351*, 507-510, doi:10.1126/science.aad0300.
26. Bou-Nader, C.; Cornu, D.; Guerineau, V.; Fogeron, T.; Fontecave, M.; Hamdane, D. Enzyme activation with a synthetic catalytic co-enzyme intermediate: nucleotide methylation by flavoenzymes. *Angew Chem Int Ed Engl* **2017**, *56*, 12523-12527, doi:10.1002/anie.201706219.
27. Conrad, J.A.; Ortiz-Maldonado, M.; Hoppe, S.W.; Palfey, B.A. Detection of intermediates in the oxidative half-reaction of the FAD-dependent thymidylate synthase from *Thermotoga maritima*: carbon transfer without covalent pyrimidine activation. *Biochemistry* **2014**, *53*, 5199-5207, doi:10.1021/bi500648n.
28. Delk, A.S.; Nagle, D.P., Jr.; Rabinowitz, J.C. Methylenetetrahydrofolate-dependent biosynthesis of ribothymidine in transfer RNA of *Streptococcus faecalis*. Evidence for reduction of the 1-carbon unit by FADH2. *The Journal of biological chemistry* **1980**, *255*, 4387-4390.
29. Nishimasu, H.; Ishitani, R.; Yamashita, K.; Iwashita, C.; Hirata, A.; Hori, H.; Nureki, O. Atomic structure of a folate/FAD-dependent tRNA T54 methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 8180-8185, doi:10.1073/pnas.0901330106.
30. Urbonavicius, J.; Skouloubris, S.; Myllykallio, H.; Grosjean, H. Identification of a novel gene encoding a flavin-dependent tRNA:m^5U methyltransferase in bacteria--evolutionary implications. *Nucleic acids research* **2005**, *33*, 3955-3964, doi:10.1093/nar/gki703.
31. Bjork, G.R. Transductional mapping of gene trmA responsible for the production of 5-methyluridine in transfer ribonucleic acid of *Escherichia coli*. *Journal of bacteriology* **1975**, *124*, 92-98.

32. Ny, T.; Bjork, G.R. Cloning and restriction mapping of the *trmA* gene coding for transfer ribonucleic acid (5-methyluridine)-methyltransferase in *Escherichia coli* K-12. *Journal of bacteriology* **1980**, *142*, 371-379.
33. Ranaei-Siadat, E.; Fabret, C.; Seijo, B.; Dardel, F.; Grosjean, H.; Nonin-Lecomte, S. RNA-methyltransferase TrmA is a dual-specific enzyme responsible for C5-methylation of uridine in both tmRNA and tRNA. *RNA biology* **2013**, *10*, 572-578, doi:10.4161/rna.24327.
34. Agarwalla, S.; Kealey, J.T.; Santi, D.V.; Stroud, R.M. Characterization of the 23 S ribosomal RNA m5U1939 methyltransferase from *Escherichia coli*. *The Journal of biological chemistry* **2002**, *277*, 8835-8840, doi:10.1074/jbc.M111825200.
35. Madsen, C.T.; Mengel-Jorgensen, J.; Kirpekar, F.; Douthwaite, S. Identifying the methyltransferases for m(5)U747 and m(5)U1939 in 23S rRNA using MALDI mass spectrometry. *Nucleic acids research* **2003**, *31*, 4738-4746, doi:10.1093/nar/gkg657.
36. Desmolaize, B.; Fabret, C.; Bregeon, D.; Rose, S.; Grosjean, H.; Douthwaite, S. A single methyltransferase YefA (RlmCD) catalyses both m^5U747 and m^5U1939 modifications in *Bacillus subtilis* 23S rRNA. *Nucleic acids research* **2011**, *39*, 9368-9375, doi:10.1093/nar/gkr626.
37. Lartigue, C.; Lebaudy, A.; Blanchard, A.; El Yacoubi, B.; Rose, S.; Grosjean, H.; Douthwaite, S. The flavoprotein Mcap0476 (RlmFO) catalyzes m^5U1939 modification in *Mycoplasma capricolum* 23S rRNA. *Nucleic acids research* **2014**, *42*, 8073-8082, doi:10.1093/nar/gku518.
38. Urbonavicius, J.; Brochier-Armanet, C.; Skouloubris, S.; Myllykallio, H.; Grosjean, H. In vitro detection of the enzymatic activity of folate-dependent tRNA (Uracil-54,-C5)-methyltransferase: evolutionary implications. *Methods in enzymology* **2007**, *425*, 103-119, doi:10.1016/S0076-6879(07)25004-9.
39. Hamdane, D.; Argentini, M.; Cornu, D.; Golinelli-Pimpaneau, B.; Fontecave, M. FAD/folate-dependent tRNA methyltransferase: flavin as a new methyl-transfer agent. *Journal of the American Chemical Society* **2012**, *134*, 19739-19745, doi:10.1021/ja308145p.
40. Hamdane, D.; Bruch, E.; Un, S.; Field, M.; Fontecave, M. Activation of a unique flavin-dependent tRNA-methylating agent. *Biochemistry* **2013**, *52*, 8949-8956, doi:10.1021/bi4013879.
41. Hamdane, D.; Argentini, M.; Cornu, D.; Myllykallio, H.; Skouloubris, S.; Hui-Bon-Hoa, G.; Golinelli-Pimpaneau, B. Insights into folate/FAD-dependent tRNA methyltransferase mechanism: role of two highly conserved cysteines in catalysis. *The Journal of biological chemistry* **2011**, *286*, 36268-36280, doi:10.1074/jbc.M111.256966.
42. Kamminga, T.; Koehorst, J.J.; Vermeij, P.; Slagman, S.J.; Martins Dos Santos, V.A.; Bijlsma, J.J.; Schaap, P.J. Persistence of functional protein domains in Mycoplasma species and their role in host specificity and synthetic minimal life. *Frontiers in cellular and infection microbiology* **2017**, *7*, 31, doi:10.3389/fcimb.2017.00031.
43. Hutchison, C.A., 3rd; Chuang, R.Y.; Noskov, V.N.; Assad-Garcia, N.; Deerinck, T.J.; Ellisman, M.H.; Gill, J.; Kannan, K.; Karas, B.J.; Ma, L., et al. Design and synthesis of a minimal bacterial genome. *Science* **2016**, *351*, aad6253, doi:10.1126/science.aad6253.
44. Chen, L.L.; Chung, W.C.; Lin, C.P.; Kuo, C.H. Comparative analysis of gene content evolution in phytoplasmas and mycoplasmas. *PloS one* **2012**, *7*, e34407, doi:10.1371/journal.pone.0034407.
45. Barre, A.; de Daruvar, A.; Blanchard, A. MolliGen, a database dedicated to the comparative genomics of Mollicutes. *Nucleic acids research* **2004**, *32*, D307-310, doi:10.1093/nar/gkh114.
46. Uchiyama, I.; Mihara, M.; Nishide, H.; Chiba, H. MBGD update 2015: microbial genome database for flexible ortholog analysis utilizing a diverse set of genomic data. *Nucleic acids research* **2015**, *43*, D270-276, doi:10.1093/nar/gku1152.
47. Grosjean, H.; Breton, M.; Sirand-Pugnet, P.; Tardy, F.; Thiaucourt, F.; Citti, C.; Barre, A.; Yoshizawa, S.; Fourmy, D.; de Crecy-Lagard, V., et al. Predicting the minimal translation apparatus: lessons from the reductive evolution of mollicutes. *PLoS genetics* **2014**, *10*, e1004363, doi:10.1371/journal.pgen.1004363.
48. Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular biology and evolution* **2000**, *17*, 540-552, doi:10.1093/oxfordjournals.molbev.a026334.
49. Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.F.; Guindon, S.; Lefort, V.; Lescot, M., et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic acids research* **2008**, *36*, W465-469, doi:10.1093/nar/gkn180.
50. Waterhouse, A.M.; Procter, J.B.; Martin, D.M.; Clamp, M.; Barton, G.J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189-1191, doi:10.1093/bioinformatics/btp033.
51. Bienert, S.; Waterhouse, A.; de Beer, T.A.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository-new features and functionality. *Nucleic acids research* **2017**, *45*, D313-D319, doi:10.1093/nar/gkw1132.

52. Baker, N.A.; Sept, D.; Joseph, S.; Holst, M.J.; McCammon, J.A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 10037-10041, doi:10.1073/pnas.181342398.
53. Giessing, A.M.; Jensen, S.S.; Rasmussen, A.; Hansen, L.H.; Gondela, A.; Long, K.; Vester, B.; Kirpekar, F. Identification of 8-methyladenosine as the modification catalyzed by the radical SAM methyltransferase Cfr that confers antibiotic resistance in bacteria. *RNA* **2009**, *15*, 327-336, doi:10.1261/rna.1371409.
54. Andersen, T.E.; Porse, B.T.; Kirpekar, F. A novel partial modification at C2501 in *Escherichia coli* 23S ribosomal RNA. *RNA* **2004**, *10*, 907-913, doi:10.1261/rna.5259404.
55. Douthwaite, S.; Kirpekar, F. Identifying modifications in RNA by MALDI mass spectrometry. *Methods in enzymology* **2007**, *425*, 3-20, doi:10.1016/S0076-6879(07)25001-3.
56. Kirpekar, F.; Douthwaite, S.; Roepstorff, P. Mapping posttranscriptional modifications in 5S ribosomal RNA by MALDI mass spectrometry. *RNA* **2000**, *6*, 296-306, doi:10.1017/s1355838200992148.
57. Shoji, T.; Takaya, A.; Sato, Y.; Kimura, S.; Suzuki, T.; Yamamoto, T. RlmCD-mediated U747 methylation promotes efficient G748 methylation by methyltransferase RlmAII in 23S rRNA in *Streptococcus pneumoniae*; interplay between two rRNA methylations responsible for telithromycin susceptibility. *Nucleic acids research* **2015**, *43*, 8964-8972, doi:10.1093/nar/gkv609.
58. Ero, R.; Peil, L.; Liiv, A.; Remme, J. Identification of pseudouridine methyltransferase in *Escherichia coli*. *RNA* **2008**, *14*, 2223-2233, doi:10.1261/rna.1186608.
59. Purta, E.; Kaminska, K.H.; Kasprzak, J.M.; Bujnicki, J.M.; Douthwaite, S. YbeA is the m^3Psi methyltransferase RlmH that targets nucleotide 1915 in 23S rRNA. *RNA* **2008**, *14*, 2234-2244, doi:10.1261/rna.1198108.
60. Ali, I.K.; Lancaster, L.; Feinberg, J.; Joseph, S.; Noller, H.F. Deletion of a conserved, central ribosomal intersubunit RNA bridge. *Molecular cell* **2006**, *23*, 865-874, doi:10.1016/j.molcel.2006.08.011.
61. Ero, R.; Leppik, M.; Liiv, A.; Remme, J. Specificity and kinetics of 23S rRNA modification enzymes RlmH and RluD. *RNA* **2010**, *16*, 2075-2084, doi:10.1261/rna.2234310.
62. Liu, Q.; Fredrick, K. Intersubunit bridges of the bacterial ribosome. *Journal of molecular biology* **2016**, *428*, 2146-2164, doi:10.1016/j.jmb.2016.02.009.
63. Purta, E.; O'Connor, M.; Bujnicki, J.M.; Douthwaite, S. YgdE is the 2'-O-ribose methyltransferase RlmM specific for nucleotide C2498 in bacterial 23S rRNA. *Molecular microbiology* **2009**, *72*, 1147-1158, doi:10.1111/j.1365-2958.2009.06709.x.
64. Hamdane, D.; Bou-Nader, C.; Cornu, D.; Hui-Bon-Hoa, G.; Fontecave, M. Flavin-Protein Complexes: aromatic stacking assisted by a hydrogen bond. *Biochemistry* **2015**, *54*, 4354-4364, doi:10.1021/acs.biochem.5b00501.
65. Nag, L.; Sournia, P.; Myllykallio, H.; Liebl, U.; Vos, M.H. Identification of the TyrOH(*+) Radical Cation in the Flavoenzyme TrmFO. *Journal of the American Chemical Society* **2017**, *139*, 11500-11505, doi:10.1021/jacs.7b04586.
66. Dozova, N.; Lacombat, F.; Bou-Nader, C.; Hamdane, D.; Plaza, P. Ultrafast photoinduced flavin dynamics in the unusual active site of the tRNA methyltransferase TrmFO. *Physical chemistry chemical physics : PCCP* **2019**, *21*, 8743-8756, doi:10.1039/c8cp06072j.
67. Josi, C.; Burki, S.; Vidal, S.; Dordet-Frisoni, E.; Citti, C.; Falquet, L.; Pilo, P. Large-scale analysis of the *Mycoplasma bovis* genome identified non-essential, adhesion- and virulence-related genes. *Frontiers in microbiology* **2019**, *10*, 2085, doi:10.3389/fmicb.2019.02085.
68. Thomas, A.; Linden, A.; Mainil, J.; Bischof, D.F.; Frey, J.; Vilei, E.M. *Mycoplasma bovis* shares insertion sequences with *Mycoplasma agalactiae* and *Mycoplasma mycoides* subsp. mycoides SC: Evolutionary and developmental aspects. *FEMS microbiology letters* **2005**, *245*, 249-255, doi:10.1016/j.femsle.2005.03.013.
69. Sirand-Pugnet, P.; Lartigue, C.; Marenda, M.; Jacob, D.; Barre, A.; Barbe, V.; Schenowitz, C.; Mangenot, S.; Couloux, A.; Segurens, B., et al. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. *PLoS genetics* **2007**, *3*, e75, doi:10.1371/journal.pgen.0030075.
70. Lo, W.S.; Gasparich, G.E.; Kuo, C.H. Convergent evolution among ruminant-pathogenic Mycoplasma involved extensive gene content changes. *Genome biology and evolution* **2018**, *10*, 2130-2139, doi:10.1093/gbe/evy172.
71. Watts, A.G.; Damager, I.; Amaya, M.L.; Buschiazzo, A.; Alzari, P.; Frasch, A.C.; Withers, S.G. *Trypanosoma cruzi* trans-sialidase operates through a covalent sialyl-enzyme intermediate: tyrosine is the catalytic nucleophile. *Journal of the American Chemical Society* **2003**, *125*, 7532-7533, doi:10.1021/ja0344967.
72. Amaya, M.F.; Watts, A.G.; Damager, I.; Wehenkel, A.; Nguyen, T.; Buschiazzo, A.; Paris, G.; Frasch, A.C.; Withers, S.G.; Alzari, P.M. Structural insights into the catalytic mechanism of *Trypanosoma cruzi* trans-sialidase. *Structure* **2004**, *12*, 775-784, doi:10.1016/j.str.2004.02.036.
73. Koliadenko, V.; Wilanowski, T. Additional functions of selected proteins involved in DNA repair. *Free radical biology & medicine* **2020**, *146*, 1-15, doi:10.1016/j.freeradbiomed.2019.10.010.

74. Guo, Y.; Zhu, H.; Wang, J.; Huang, J.; Khan, F.A.; Zhang, J.; Guo, A.; Chen, X. TrmFO, a fibronectin-binding adhesin of *Mycoplasma bovis*. *International journal of molecular sciences* **2017**, *18*, doi:10.3390/ijms18081732.