Comparative patterns of modified nucleotides in individual tRNA species from a mesophilic and two thermophilic archaea

PHILIPPE WOLFF,[1] CLAIRE VILLETTE,[2] JULIE ZUMSTEG,[2] DIMITRI HEINTZ,[2] LAURA ANTOINE,[1] BÉATRICE CHANE-WOON-MING,[1] LOUIS DROOGMANS,[3] HENRI GROSJEAN,[3] and ERIC WESTHOF[1]

[1]Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire du CNRS, Université de Strasbourg, F-67084, Strasbourg, France

[2]Institut de Biologie Moléculaire des Plantes du CNRS, Université de Strasbourg, F-67084, Strasbourg, France

[3]Laboratoire de Chimie Biologique, Université Libre de Bruxelles, Institut Labiris, B-1070, Belgium

ABSTRACT

To improve and complete our knowledge of archaeal tRNA modification patterns, we have identified and compared the modification pattern (type and location) in tRNAs of three very different archaeal species, *Methanococcus maripaludis* (a mesophilic methanogen), *Pyrococcus furiosus* (a hyperthermophile thermococcale), and *Sulfolobus acidocaldarius* (an acidophilic thermophilic sulfolobale). Most abundant isoacceptor tRNAs (79 in total) for each of the 20 amino acids were isolated by two-dimensional gel electrophoresis followed by in-gel RNase digestions. The resulting oligonucleotide fragments were separated by nanoLC and their nucleotide content analyzed by mass spectrometry (MS/MS). Analysis of total modified nucleosides obtained from complete digestion of bulk tRNAs was also performed. Distinct base- and/or ribose-methylations, cytidine acetylations, and thiolated pyrimidines were identified, some at new positions in tRNAs. Novel, some tentatively identified, modifications were also found. The least diversified modification landscape is observed in the mesophilic *Methanococcus maripaludis* and the most complex one in *Sulfolobus acidocaldarius*. Notable observations are the frequent occurrence of ac^{4}C nucleotides in thermophilic archaeal tRNAs, the presence of m^{7}G at positions 1 and 10 in *Pyrococcus furiosus* tRNAs, and the use of wyosine derivatives at position 37 of tRNAs, especially those decoding U1- and C1-starting codons. These results complete those already obtained by others with sets of archaeal tRNAs from *Methanocaldococcus jannaschii* and *Haloferax volcanii*.

Keywords: mass spectrometry; Archaea; tRNA; modifications; hyperthermophiles

INTRODUCTION

Transfer RNAs are the most modified RNA molecules in terms of number of modified positions and diversity of chemical modifications. A whole gamut of modification enzymes had to differentially evolve in the three domains of life to mature properly tRNAs (Helm and Alfonso 2014; Boccaletto et al. 2018). In short, tRNA modifications play two central roles. The first one is to guarantee the maintenance of the uniqueness and stability of the tRNA architectural fold (Helm 2006; Motorin and Helm 2010, 2011), a requirement for proper recognition by key factors like aminoacyl-tRNA synthetases (Giégé and Springer 2016), the Elongator complex (Karlsborn et al. 2014), and ribosomal subunits (Selmer et al. 2006). The modifications occur in the whole body of the tRNA but especially in the elbow created by the intricate contacts formed between the D- and T-loops (Machnicka et al. 2014). The second main role played by tRNA modifications is to guarantee fidelity and efficiency during ribosomal translation at the decoding site, thereby participating in the regulation of the translational activity and the control of proteostasis (Pollo-Oliveira and de Crécy-Lagard 2019). These modifications occur mainly in the extended anticodon loop of tRNAs (Yarus 1982). Their roles are (i) to maintain a conformation of the anticodon loop preorganized for pairing with the mRNA codon in the A site (Vendeix et al. 2012); (ii) to stabilize the weak AU-rich codon/anticodon pairs (Grosjean and Westhof 2016); (iii) to avoid miscoding (e.g., Met/Ile or Trp/Stop) (Cantara et al. 2013); (iv) to allow

Corresponding authors: e.westhof@ibmc-cnrs.unistra.fr, p.wolff@ibmc-cnrs.unistra.fr

Article is online at http://www.rnajournal.org/cgi/doi/10.1261/rna.077537.120. Freely available online through the *RNA* Open Access option.

© 2020 Wolff et al. This article, published in *RNA*, is available under a Creative Commons License (Attribution 4.0 International), as described at http://creativecommons.org/licenses/by/4.0/.

the decoding of purine-ending codons in split codon boxes by promoting unusual base-pairings that fit within the decoding ribosomal grip (Rozov et al. 2016). Although the nature and number of the modifications vary considerably between the three domains of life, the great majority of the tRNA positions that are modified are highly conserved throughout phylogeny.

Unfortunately, the complete patterns of tRNA modifications are known for only a limited number of species (e.g., *E. coli* in Bacteria or *S. cerevisiae* in Eukarya). In Archaea, although studied for a long time, the landscape of tRNA modifications remained disperse, and known only for a few particular tRNA isoacceptors or bulk cellular tRNAs (see for examples: Kuchino et al. 1982; Edmonds et al. 1991; Tomikawa et al. 2013). Only in the cases of *Haloferax volcanii* (Gupta 1984, 1986; Grosjean et al. 2008a) and very recently in the case of *Methanocaldococcus jannaschii* (Yu et al. 2019), the modification landscape of a complete set of cellular tRNAs has been elucidated. These data reveal that a few modifications are unique to certain archaea, while others are present in most, if not all archaeal species studied so far. Among the archetypal ones are a N1-methylated pseudouridine at position 54 instead of a thymine (Pang et al. 1982; Gupta 1984), the presence of archaeosine, or 7-formamidino-7-deazaguanosine, at position 15 (Watanabe et al. 1997), certain wyosine derivatives like imG, imG2 and mimG at position 37 (de Crécy-Lagard et al. 2010) and the presence at position 34 of tRNA-Ile of agmatidine, a modified C where the carbonyl group is replaced by decarboxy-arginine (Ikeuchi et al. 2010; Mandal et al. 2010).

Here, we examined the landscapes of tRNA modifications in three archaeal species with very different evolutionary history. Two belong to the Kingdom *Euryarchaeota*: *Methanococcus maripaludis*, a methane-producing anaerobic mesophilic archaeon belonging to the same Methanococcales clade as *M. jannaschii* (but the latter is thermophilic) and *Pyrococcus furiosus*, an anaerobic hyperthermophilic archaeon belonging to the Thermococcales clade (Forterre 2015). The third species studied here is *Sulfolobus acidocaldarius* (an obligate aerobic, acidophilic, sulfur-oxidizing thermophile that belongs to the Sulfolobales clade of the *Crenarchaeota* Kingdom). The last two archaea therefore belong to families not yet systematically analyzed and *S. acidocaldarius* is the first crenoarchaeon for which the modification landscape is reported. Further, all euryarchaeal tRNAs analyzed so far are from anaerobic microorganisms (with thermophilic or hyperthermophilic character), while *S. acidocaldarius* is an aerobic moderate thermophile.

As performed earlier in the case of *H. volcanii* (Gupta 1984, 1986), two-dimensional gel electrophoresis was used to separate individual tRNA isoacceptors from purified bulk cellular tRNAs of each of the three archaea, allowing the analysis of 79 cellular tRNAs out of a total of 116 theoretical corresponding tDNA genes (Chan and Lowe 2009, 2016). In-gel digestion by specific nucleases followed. The resulting tRNA digests, for which sequences were deduced from their known genomes, were then separated by chromatography and analyzed by mass spectrometry. The analysis of modified nucleosides of bulk tRNAs was also performed and the results compared with those obtained from the analysis of oligonucleotide sequences. Altogether, our results support the idea that, despite the fact that Archaea share typical modified nucleotides present in both Bacteria and/or Eukarya, they also display unique and specific modifications.

RESULTS

tRNA genes and their cellular mature products

In the GtRNAdb (Chan and Lowe 2009, 2016), there are 37 genes coding for tRNAs in *M. maripaludis* and 46 in both *P. furiosus* and *S. acidocaldarius*, all predicted with excellent scores using tRNAscan-SE 2.0 (Lowe and Eddy 1997). In *S. acidocaldarius*, 50 tRNA genes are predicted but four have scores below 45.0 and would be considered as pseudogenes (Chan and Lowe 2009, 2016). Based on genomic information, in *M. maripaludis*, there are only two tRNA genes corresponding to 4-codon boxes and only one gene corresponding to 2-codon boxes, except for Asp, Glu, and Lys where there are two isoacceptors. In *P. furiosus* and *S. acidocaldarius,* there are three genes coding for isoacceptors in 4-codon boxes and only one in 2-codon boxes, except again for Glu, Lys, but also Gln, Ile, Arg(AGR), and Leu(UUR) (all codons of 2/3-codon boxes ending with R3 and thus decoded by Y34-containing tRNAs). In the three species, there is one gene coding for tRNA-Met^{i} and another one for tRNA-Met, except in *M. maripaludis*, where there are two copies of tRNA-Met^{i}. As in all other archaeal genomes sequenced so far, there is no gene coding for A34-containing tRNA. Remarkably, in *M. maripaludis*, the C34-containing tRNAs are absent in both 4- and 2-codon boxes except of course in tRNA-Ile(CAU), tRNA-Met and tRNA-Met^{i}(CAU), and tRNA-Trp(CCA). Thus, besides these five particular tRNAs, all the other tRNA anticodon triplets start with either G34 or U34. In contrast, in the two thermophilic archaea, *P. furiosus* and *S. acidocaldarius,* the C34-containing tRNAs are present, which explains the increase from 37 to 46 naturally occurring tRNA genes. These distributions follow the sparing strategies in Archaea described by Grosjean et al. (2010). The percentage of GC-content increases from *M. maripaludis* (34%) to *S. acidocaldarius* (37.5%) and to *P. furiosus* (41.1%). The codon usage is such that U3- and A3-ending codons are highly preferred (Nakamura et al. 2000; Emery and Sharp 2011; Nayak 2013). This is particularly striking in the mesophilic *M. maripaludis* where U3- and A3-ending codons are decoded solely by G34-tRNAs

and U34-tRNAs, respectively. In *M. maripaludis*, the percentage of GC-content at the third position is 25.6%, while it is 28.8% in *P. furiosus* and as high as 39.3% in *S. acidocaldarius*. Such differences in decoding strategies have consequences on the modification identity found at position 34 and in the extended anticodon stem–loop of individual isoacceptor tRNA (see below and Grosjean and Westhof 2016). The knowledge of tRNA genes and copy numbers does not allow yet to predict which identified genes correspond to lowly or highly expressed cellular tRNAs (minor/major species). The low abundance of certain cellular tRNAs can be below the detection mapping method, which explains why some tRNA isoacceptors are missing in our analyses of bulk tRNAs. However, for each archaeon, representative tRNAs corresponding to each of the 20 amino acids could be analyzed. They probably correspond to the most abundant naturally occurring species. Lastly, in *Archaea*, especially the hyperthermophilic ones, tRNA genes often contain introns (Sugahara et al. 2008). Among the three archaeal species studied, the *S. acidocaldarius* contains the highest number of tRNAs with introns (21 out of 46) (Chan and Lowe 2009, 2016). Such intron-containing tRNAs are often the targets of site-specific 2′-O-ribose methylations or uridine isomerization into pseudouridine via the sRNA-guided FlpA C/D box or the Cbf5 H/ACA enzymatic machinery, respectively (see below).

Purification and sequencing of isolated tRNA species

Each tRNA isoacceptor was isolated using two-dimensional PAGE. The process leads to a series of tRNA spots. Each spot contains generally one isoacceptor, while some contain two or rarely three isoacceptors (Supplemental Fig. S1). For each spot, all CID spectra were manually inspected and sequenced. The tRNAs were identified by unique sequences in RNase digestion products (see Materials and Methods). From such 2D-gel electrophoresis, we could purify and analyze 27 post-transcriptionally matured tRNAs corresponding to obviously major naturally occurring species from *M. maripaludis* and 30 tRNAs from *P. furiosus* (over 37 and 46, respectively) but only 22 (over 46) from *S. acidocaldarius*. In this latter case, of the 21 intron-containing primary transcripts predicted from the genome sequence only six were identified as matured species in the purified bulk matured tRNAs, while 16 over 25 of the predicted intron-less tRNAs were detected. In *P. furiosus* and *M. maripaludis*, only two tRNAs (specific for Met [CAU] and Trp[CCA]) are transcribed as intron-containing species and each of them was obtained and analyzed as the matured species.

A critical point in the present experimental strategy is to assign a modified position to the tRNA it belongs to. Indeed, because archaeal tRNAs are highly GC-rich (especially in stems), there is a high level of similarities between tRNA sequences especially for the highly conserved D- and TΨC-loops. To be sure to assign unambiguously an MS/MS sequencing spectrum, each tRNA isoacceptor was analyzed separately (see Materials and Methods). All CID (Collision Induced Decay) MS/MS spectra were manually examined and sequenced. Although LC MS/MS spectra of digestion products allow the localization of a modified nucleotide within the sequence, some of the modifications unfortunately share the same (m/z) mass. For example, methylation can be detected but, solely on the basis of the mass spectra, it is not possible to localize the methyl group (either on the ribose or on the base). The same situation occurs between uridine and pseudouridine. However, knowing the presence of a given modified nucleotide at the same position in a homogous tRNA of a closely related archaeon, together with the existence of corresponding modification enzymes (and its ORF in the genome), sometimes allows to assign the most probable chemical modification after verification that a homolog gene exists in the genome of the archaea studied (Supplemental Table S2). The ambiguous cases that are left are discussed in the text or legends. Analyses of total modified nucleosides of bulk tRNAs were also performed allowing to identify the presence of modified nucleotides that escape the above analysis of sufficiently long oligonucleotides (Supplemental Figs. S2, S3). The code for modified nucleosides used throughout this paper, except when specifically mentioned, follows either the one used in MODOMICS (Boccaletto et al. 2018) or the chemically based nomenclature (Motorin and Helm 2011; Helm and Alfonso 2014).

Structural tRNA alignments

Alignments of the sequenced archaeal tRNAs studied in this work, including the modified nucleotides we have been able to detect at the oligonucleotide digest products, are shown in Figure 1 (see also Table 1; Supplemental Fig. S10). These alignments follow the usual nomenclature of tRNA structure, and they are structural in the sense that equivalent positions in the three-dimensional structure (assumed by homology with known crystal structures) are vertically aligned. They show that the tRNA sequences follow the expected patterns of covariations and conservations as observed in the majority of cellular tRNAs of all three domains of life, demonstrating that they do conform to the known three-dimensional structure of tRNAs. The numbers of Watson–Crick base pairs (often GC-rich and with a small number of GU pairs, especially in the thermophiles) in the stems are as usual: seven in the acceptor stem (AA-stem), four in the dihydrouridine stem (D-stem), five in the anticodon stem (AC-stem) and the thymine stem (TΨC-stem). Also, a long variable region is always present in the long-arm tRNAs specific for leucine and serine. The tRNA-Leu (anticodons YAG, Y = U/C) in *M. maripaludis* and *P. furiosus*

Amino Acid	anticodon	Acc-stem		D-stem	D-loop	D-stem		Ac-stem	Ac-loop	Ac-stem	V-region	T-stem	T-loop	T-stem	Acc-stem		CCA
		1	8	10	14	22	26	27	32	39	44	49	54	61	66		74
M. maripaludis																	
Ala	UGC	GGGCCCG	UA	GCUC	A[G+]UUGGGA	[Gm]AGC	G	CCGCC	CUUGC/[ac4C]AA	GGCGG	AGCUC	GUGGG	[m1Y]U[Cm][m1I][m1A]AU	CCCGC	CUAGGCU	-	
Arg	UCU	GGGCCCG	UG	GCCU	AGUCUGGAUA	CGGC	A	CCGGC	CUUCU[t6A]A	GCCGG	GGAUC	GGGGG	[m1Y]UCG[m1A]AU	CCCUC	CGGGUCC	G	CC-
Asn	GUU	GCCUCCU	UA	GCUC	A[G+]UAGGUAG	CAGC	[m2,2G]	ACGGA	CUGUU[hn6A]A	UCCGU	AGGUC	GCAGG	UUCGAGC	CCUGC	AGGAGGC	G	CCA
Asp	GUC	GCUCUGG	UG	GUGU	A[G+]CUCGGCCUA	UCAU	A	CAGGA	CUGUC[f6A]C	UCCUG	UGAC	UCGGG	[m1Y]U[Cm][m1I][m1A]AU	CCCGG	CCAGAGC	G	CCA
Glu	UUC	GCUCCAG	UG	[m2,2G]UGU	AGUCCGGCCAA	UCAU	C	CGGCC	CU[mnm5U]UC[m1G]A	GGCCG	GGAC	UCGGG	[m1Y]U[Cm][m1I][m1A]AU	CCCGG	CUGGAGC	A	CCA
Gly	GCC	GCGGCUU	UG	AUGU	A[G+]ACUGGUA	UCAU	A	CGGCC	CUGCCAC	GGCCG	ACAC	CCGGG	UUCAAAU	CCCGG	AGGCCGC	-	
Gly	UCC	GCGCCAG	UG	GUGU	AGCCUGGUA	UCAC	U	UUGGC	CU[cnm5U]CCAA	GCCAG	UAAC	UCGGG	[m1Y]U[Cm][m1I][m1A]AU	CCCGA	CUGGCGC	A	CCA
His	GUG	GCCGUGG	UA	GGGU	AGUGGCUA	UCCU	G	AGGGA	CUGUG[m1G]A	UCCCU	CGAC	CCGGG	UUCGAUU	CCCGG	UCUCGGC	C	CC-
Ile	GAU	AGGGCAG	UG	GCUC	A[G+]C[s2C]UGGUUA	GAGU	[m2,2G]	CUCGG	CUGAU[hn6A]A	CCGAG	UGGUC	CGGGG	UU[Cm]G[m1A]AU	CCCCG	CUGCCCU	A	CCA
Ile	CAU	GGGCCCG	UA	GCUC	AGGCUGGUU-A	GAGU	[m2,2G]	CUCGG	CU[C+]AU[hn6A]A	CCGAG	UGGUC	AUGGG	[m1Y]U[Cm][m1I][m1A]AU	CCCAU	CGGGCCC	A	
Ini	CAU	AGCGAGG	UA	GGGU	A[G+]CCAGGCCUA	UCCC	G	CCGGG	CUCAUAA	CCCGG	AGAUC	AGAAG	UUCAAAU	CUUCU	CCUCGCU	A	CCA
Leu	GAG	GCGAGGG	UU	GCCA	A[G+]C[s2C]UGGUUAA	AGGC	[m2,2G]	CUAGG	UUGAGGG	CCUAG	UCUCAUAGGAGUUC	GAGGG	UUCG[m1A]AU	CCCUU	CCCUCGC	-	
Leu	UAA	GGAGGGA	UU	GCCA	AGCCUGGUCAA	AGGC	G	UCGGA	CUUAAGA	UCCGA	UCCGGUAGCGGUUC	GAGGG	UUCAAAU	CCCUU	UCCCUCC	A	
Leu	UAG	GCAGGGG	UU	GUCG	A[G+]CCUGGCCAA	AGAU	[m2,2G]	CAGGA	C[s2U][xU]AGAA	UCCUG	UCCAGUAGUGGUUC	CAGGG	UUCAAAU	CCCUG	CCCCUGC	A	
Lys	UUU	GGGCCCG	UA	[m2G]CUU	A[G+]UCUGGUA	GAGC	[m2,2G]	CCUGA	CU[cmnm5s2U]UU[t6A]A	UCAGG	CGGUC	GAGGG	UUCG[m1A]AU	CCCUU	CGGGCCC	G	CCA
Met	CAU	GCCGAGG	UG	GCUU	A[G+]GCUGGUUA	UAGC	[m2,2G]	CUCGG	CU[Cm]AU[hn6A]A	CCGAG	AGGUC	GGGGG	[m1Y]U[Cm][m1I]AGU	CCCCC	CCUCGGC	-	
Phe	GAA	GCCAAGG	UA	GUUC	AGCCUGGGA	GAAC	[m2,2G]	CUGGA	[Cm]UGAA[imG-14]A	UCCAG	UUGUC	GGGUG	UU[Cm][m1I][m1A]AU	CACCC	CCUUGGC	A	CCA
Pro	UGG	GGGCCUG	UG	GGGU	A[G+]CCUGGUCA	UCCU	U	UGGGA	UUUGG[m1G]A	UCCUG	AAAC	CCCAG	UUCG[m1A]AU	CUGGG	CAGGCCC	A	CCA
Thr	GGU	GCUUCAG	UG	GCUC	A[G+]CCUGGUA	GAGC	[m2,2G]	CCUGC	UUGGU[hn6A]A	GCAGG	AGGCC	GCGGG	UUCAAAC	CCCGC	CUGAAGC	U	
Thr	UGU	GCCUCGG	UG	GCUC	AGCCUGGUA	GAGC	G	CCUGA	CUUGU[hn6A]A	UCAGG	UGGUC	GGGGG	UUCGAAU	CCCCC	CCGGGGC	U	
Trp	CCA	GGGAGUA	UG	GUGU	A[G+]U[s2C]UGGUCUA	UCAU	C	GGGGA	CUCCAAA	UCCCU	GGAC	CUGGG	UUCAAAU	CCCAG	UACUUCC	A	CCA
Tyr	GUA	CCCGCGA	UA	GUUC	A[G+]ACUGGUA	GAAC	[m2,2G]	GCGGA	CmUGUAUxG]A	UCCGC	AUGUC	GCUGG	[m1Y]U[Cm][m1I][m1A]AU	CCGGC	UCGCGGG	A	CCA
Ser	UGA	GCAGAGG	UG	GUUG	A[G+]CUUGGCCAA	AGGC	[m2,2G]	CCGGA	[s2C]UUGAAA	[s2U]CCGG	UUCUCCACUGGGGAGC	GGGGG	[m1Y]U[Cm][m1I][m1A]AU	CCCUC	CCUCUGC	-	
Ser	GCU	GCGGAGA	UG	GCCC	A[G+]UUGGGUA	CGGC	G	CGGGA	CUGCU[t6A]A	UCCCG	UUACGCUCUGCGUAAC	CCGAG	UUCAAAU	CUCGG	UCUCCGC	-	
Ser	GGA	GCAGAGA	UA	GUCU	A[G+]CCUGGGA	AGGC	[m2,2G]	UACGG	CUGGAAA	CCGUA	UGAGGCUUUGCCUCCC	GGGGG	UUCAAAU	CCCCC	UCUCUGC	G	CC-
Val	GAC	GAGUUCG	UG	GUCU	A[G+]UUGGCUA	UGAU	A	CCGCC	CUGACAC	GGCGG	UGAUC	GGGAG	UUCG[m1A]AU	CUCCC	CGGACUC	A	CC-
Val	UAC	GGACCU[ac6A]	UG	GUCU	A[G+]UUGGCUA	UGAC	A	UCGCC	CUUACAA	GGCGA	GGGUC	GCCGG	UUCGAAU	CCGGC	UAGGUCC	A	CCA
P. furiosus																	
Ala	GGC	GGGCCGG	UA	GCUC	AGCCUGGUAU	GAGC	G	CCGCC	UUGGCAA	GGCGG	AGGC	CCGGG	UUCAAAU	CCCGG	CCGGUCC	A	CCA
Ala	UGC	GGGC[ac4C][m2,2G]G	UA	GCUC	A[G+]CCUGGUAU	GAGC	[m2,2Gm]	CCGCC	CU[cnm5U]GCAA	GGCGG	AGGCC	CCGGG	[m5U]U[Cm][m1I][m1A]AU	CC[ac4C]GG	CCGGUCC	A	CCA
Arg	CCU	GGAC[ac4C][m2,2G]G	UA	GCCU	A[G+]CCAGGAUA	GGGC	[m2,2Gm]	GCGGC	[s2C]UCCU[t6A]A	GCCGC	AGGUC	CGGGG	UUCAAAU	CCCCG	CCGGUCC	G	CCA
Arg	UCU	GGGCCCG	UA	GCCU	A[G+]CAGGAUA	GGGC	G	CCGGC	CU[cnm5U]CUAA	GCCGG	AGGUC	CCGGG	UUCG[m1A]AU	CC[ac4C]GG	CGGGCCC	G	CCA
Arg	GCG	GCCCCGG	UG	GCCU	AGCCUGGAU-A	GGGC	[m2G]	CGAGG	CUGCG[mimG]A	CCU[Cm]G	AGGUC	CGGGG	UUCAAAU	CCCCG	CCGGGCG	C	CCA
Asp	GUC	GCCCGGG	UG	GUGU	AGCCCGGCCUA	UCAU	[m2G]	CGGGA	CUGUCAC	UCCCG	CGAC	CCGGG	UUCAAAU	CCCGG	CCCGGGC	G	CCA
Cys	GCA	GCCGG[m2,2G]A	U[m1A]	GCCU	A[G+]AGGCC	AGGC	[m2,2G]/[m2,2Gm]	GGGGA	CUGCAGA	UCCCC	UUUAC	CCGGG	[m5s2U]U[Cm][m1I][m1A]AU	CCCGG	UCCCGGC	U	CCA
Glu	CUC	p[2H-m7G]CCC[ac4C][m2,2G]G	UG	[2H-m2,7Gm]UGU	A[G+]CCCGGUCAA	UCAU	[m2G]	CGGGA	CUCUC[m1G]A	UCCCG	CGAC	CCGGG	UUCAAAU	CCCGG	CCGGGGC	A	CCA
Gln	CUG	AGCCCCG	UG	GUGU	A[G+]C[Gm]GCCAA	GCAU	G	CGGGA	CU[ac4C]UGGA	UCCCG	CGA[m5C]	[m5C][Gm]GGG	UUCGAAU	CCCCG	CGGGGCU	A	CCA
Gln	UUG	AGCCCCG	UG	GUGU	A[G+]C[Gm]GCCAA	GCAU	G	CGGGA	CU[cnm5U]UGGA	UCCCG	CGA[m5C]	[m5C]CGGG	UUCGAAU	CC[ac4C]GG	CGGGGCU	A	CCA
Gly	CCC	GCGGUGG	UA	GUCU	A[G+]C[s2C]UGGUCUA	GGAC	[m2,2G]	CCGGC	CUCCCAA	GCCGG	CGA[m5C]	[m5C]CGGG	UUCAAAU	CCCGG	CCACCGC	A	CCA
Gly	GCC	GCGGU[m2,2G]G	UA	GUCU	AGCCUGGUCUA	GGAC	G	CCACC	CUGCCAA	GGUGG	AGA[m5C]	CCGGG	UU[Cm][m1I][m1A]AU	CCCGG	CCACCGC	A	CCA
Gly	UCC	GCGGUGG	UA	GUCU	A[G+]CCUGGUCUA	GGAC	[m2G]	CCGGC	CU[cnm5U]CC[m1G]A	GCCGG	CGAC	CCGGG	UUCAAAU	CCCGG	CCACCGC	A	CC
Ile	GAU	GGGCCCG	UG	GCUC	A[G+]C[s2C]UGGUCA	GAGC	[m2,2Gm]	CCCGC	CUGAU[hn6A]A	GCGGG	AGGUC	CGGGG	UUCGAAG	CCC[ac4C]G	CGGGCCC	A	CCA
Ini	CAU	AGCGGG[m2G]G	U[Gm]	[m7G]GGC/[Cm]	A[G+]CUAGGAG[Um]	[Gm]CCC	[m2,2G]	CCGGG	CUCAUAA	CCCGG	AGG[xU]C	CGAGG	[m5s2U]U[Cm][m1I][m1A]AU	CUCG	GCCCGCU	A	CCA
Leu	CAG	GCGGGGG	UU	GCCG	A[G+]C[s2C]UGGUCAA	AGGC	[m2,2Gm]	CGGGA	UUCAGGG	UCCCG	UCCCGUAGGGGUUC	CGGGG	UUCAAAU	CCCCG	CCCCCGC	A	CCA
Leu	UAA	GCGGGGG	UG	CCCG	AGCCUGGCCAA	AGGG	G	CCGGA	CUUAAGA	UCCGG	UGCCGUAGGGCUGC	GCGGG	UUCGAAU	CCCGC	CCCCCGC	A	CCA
Leu	UAG	GCGGGGG	UU	GCCG	A[G+]CCUGGUCAA	AGGC	[m2,2Gm]	CGGGA	UU[cnm5U]AGGG	UCCCG	UCCCGUAGGGGUUC	CGGGG	UUCAAAU	CCCCG	CCCCCGC	A	CCA
Lys	CUU	GGGCCGG	UA	GCUU	AGCCUGGUU-A	GAGC	[m2,2Gm]	GCGGA	CUCUU[ms2t6A]A	UCCGC	AGGUC	GGGGG	[m5U]U[Cm][m1I][m1A]AU	CCCCC	CCGGCCC	G	CC
Met	CAU	GCCGG[m2,2G]G	UA	GCUU	AGCCUGGUCA	AAGC	[m2,2Gm]	[m5C]CCGG	CU[Cm]AU[hn6A]A	CCGGG	AGA[xU]C	CGGGG	[m5s2U]U[Cm]G[m1A]AG	CCC[ac4C]G	CCCCGGC	A	CCA
Phe	GAA	GGGGCGG	UA	GCUC	A[G+]C[s2C]UGGGA	GAGC	A	CCGGA	CUGAA[mimG]A/[Am]	UCCGG	GUGUC	GGGGG	UU[Cm][m1I][m1A]AU	CCCCC	CCGCCCC	A	CCA
Pro	UGG	GGGGCCG	UA	GGGU	AGCUUGGCCCA	UCCU	[m2G]	CGGGC	UU[cnm5U]GG[m1G]GA	GCC[ac4C]G	UGAC	CCGGG	UUCAAAU	CCCGG	CGGCCCC	A	CCA
Ser	GCU	GCCGCGG	UA	GCCU	A[G+]CCUGGGUA	GGGC	[m2,2Gm]	CCGGC	CUGCU[t^{6}A]A/[Am]	GCCGG	UGGGCGUGGUCCCGC	CGGGG	UUCAAAU	CCCCG	CCGCGGC	G	CCA
Ser	GGA	GCCGGGG	UA	GCCU	AGCCAGGG--A	AGGC	G	CGGGC	CUGGA[mimG]A	G/[m2,2G]CCCG	UGGGCGUUCGCCCAC	CGGGG	UUCAAAU	CCCCG	CCCCGGC	G	CC-
Ser	UGA	GCCGG[m2G]/[m22G]A	U/[Um][m1A]	GCCU	AGCCUGGU--A	GGGC	[m2,2Gm]	CGGGC	CUUGA[imG2]/[mimG]A	GCCCG	[Um]GGGCGUUUGCCCGC	CGGGG	[m5U]/[m5s2U]UCUAAAU	CCCCG	UCCCGGC	G	CC-
Thr	GGU	GCCCCGG	UG	GCUC	A[G+]CC/[s2C]UGGUA	GAGC	[m2,2Gm]	GCCGC	UUGGU[t6A]/[hn6A]A	GCGGC	AGG[xU]C	CCGGG	[m5U]/[m5s2U]UCUm][m1I][m1A]AG	CCCGG	CCG/[Gm]GGGC	U	CCA
Thr	UGU	GCCCGGG	UA	GCCU	AGCCUGGUG	GGGC	[m2,2Gm]	GGGGA	CU[cnm5U]GU[hn6A]A	UCCCC	AGGUC	GCGGG	UUCAAAU	CCCGC	CCCGGGC	U	CCA
Trp	CCA	GGGGGCG	UG	GUGU	AGCCUGGUCCA	UCAU	C	GCGGG	CUCCAG/[mimG]A	[Cm]C[Cm][Gm]C	GGA[m5C]	CGGGG	UU[Cm][m1I][m1A]AU	CCCCG	CGCCCCC	A	CCA
Val	GAC	GGGCCCG	UG	GUCU	A[G+]ACUGGUUA	UGAC	[m2,2G]	CCACC	CUGACAA	GGUGG	AGGUC	CGGGG	UUCAAAU	CCCCG	CGGGCCC	A	CCA
Val	UAC	GGGCCCG	UG	GUCU	A[G+]CUGGUUA	UGAC	G	CCGCC	CUUACGA	GGCGG	AGGUC	CGGGG	UUCG[m1A]AU	CCC[ac4C]G	CGGGCCC	A	CCA
S. acidocaldarius																	
Ala	GGC	GGGCCGG	UA	GCUC	A[G+]C[Cm]UGGAA	GAGU	[m2,2G]	CGUGG	UUGGCAU	CCACG	AGGUC	CCGGG	UUCAAAU	CCCGG	CCGGUCC	A	CCA
Ala	UGC	GGGCCGG	UA	GCUC	A[G+]C[Cm]UGGA--A	GAGU	[m2,2G]	CUUGG	[Cm]U[Um]GCAA	C[ac4C]GAG	AGGUC	CCGGG	UUCAAAU	CCCGG	CCGGUCC	A	CCA
Asn	GUU	GCCGCGG	UA	GCUC	AGUUGGCA	GAGC	G	CCGGG	CUGUUAA	CCCGG	UGGUC	GGUGG	UCCGAGU	CCACC	CCGCGGC	G	
Asp	GUC	GC[ac4C]GCGG	U[m1A]	[m2,2G]UCU	A[G+]CGGUCUA	GGA[Um]	[m2,2G]	GGGGC	CUGUC[m1G]/[imG]/[mimG]A	GCCCU	UGA[m5C]	[m5C][Cm][Gm]GG	[s2U]U[Cm][m1I][m1A]AU	CCCGG	CCG[Cm]GG[m5C]	G	
Gln	UUG	AGCCGGG	UC	GUCU	A[G+]C[Gm]GUCAA	G[Gm]AU	[Cm]	CAGGG	CU[Um]/[s2Um]UG[m1G]C	CCCUG	GGA[m5C]	[m5C]AGGG	UUCGAGU	CCC[Um]G	CCCGGC[Um]	A	
Glu	CUC	GCCGCGG	UA	[m22G]GUAU	A[G+]CCC[Gm]GUC-A	GUAU	[m2,2G]	[m5C]GGGC	[Cm]UCUCAA	GCCCG	UGA[m5C]	[m5C]C[Gm]GG	[s2U]U[Cm][m1I][m1A]AU	CCCGG	CCGCGGC	A	
Glu	UUC	GCCGCGG	UA	[m2,2G]UAU	A[G+]CCC[Gm]GUCAA	GUAU	[m2,2G]	[m5C]GGGC	CUUUCGA	GCCCG	UGA[m5C]	[m5C][Cm]GGG	UUCAAAU	CCCGG	CCGCGGC		
Gly	GCC	GCGGCCG	UA	GUCU	A[G+]CCUGG[Am]UUA	GGAC	[m2,2G]	CCUGC	CUGCCAC	GCAGG	AGG[xU]C	C[Cm]GGG	[s2U]U[Cm][m1I][m1A]AU	CCCGG	CGGUCGC	A	
Leu	CAA	GCGGGGG	UG	CCCG	A[G+]CUGGUCAA	AGGG	G	GCGGA	CUCAAGA	UCCGC	UGGCGAAGGCCU[Am][Cm]	GCGGG	UUCAAAU	CCCGU	CCCCCGC	A	
Leu	CAG	GCGGGGG	UG	CCCG	AGC[Am]AGGUCAA	AGGG	[m2G]/[m2,2G]	UCGGG	[Cm]U[Cm]AG[m1G]C	CCCGA	UGGUGUAGGCCUGC	G[Um]GGG	[s2U]UCAAAU	CCCAC	CCCCCGC	A	
Leu	GAG	GCGGGGG	UG	CCCG	A[G+]C[Am]AGGUCAA	AGGG	[m2,2G]	GCGGA	[Cm]UGAG[m1G]C	UCCGC	UGGUGUAGGCCUGC	G[Um]GGG	[s2U]U[Cm][m1I][m1A]GU	CCCAC	CCCCCGC	A	
Leu	UAA	GCGGGGG	UG	CCCG	A[G+]CAGGUCAA	AGGG	G	GCGGA	CUUAAGA	UCCGC	UGGCGUAGGCCmU[m2,2G]C	GUGGG	UU[Cm][m1I][m1A]AU	CCCAC	CCCCCGC	A	
Leu	UAG	GCGGGGG	UG	CCCG	AGCAAGGUCAA	AGGG	G	UCGGG	[Cm]U[s2Um]AG[m1G]C	CC[ac4C]GA	UGGUGUAGGCCUGC	GUGGG	UUCAAAU	CCCAC	CCCCCGC	A	
Ini	CAU	AGCGGCG	U[m1A]	[m2G]GGA	A[G+]CCUGGUA	UCCC	[m2,2Gm]	CA[Gm]GG	[Cm]UCAUAA	CCCUG	AGGUC	CC[Um]GG	[s2U]U[Cm][m1I][m1A]AU	C[Cm]AGG	CGCGGCU	A	
Pro	UGG	GGGGCCG	UA	GUCU	AGCUUGG[Am]CUA	GGAU	[m2,2G]	CCAGG	CUUGG[m1G]C	CCUGG	UGGUC	CCGGG	UUCAAAU	CCCGG	CGGCCCC	A	
Phe	GAA	GCCGCCG	UA	GCUC	AGCCCGGGA	GAG[Cm]	[m2,2G]	CCCGG	CUGAA[mimG]A	CCGGG	UUGUC	CGGGG	[s2U]U[Cm][m1I][m1A]GU	CCCCG	CGGCGGC	A	
Thr	GGU	GCCGCGG	UA	GCUC	AGCCUGGCU-A	GAGC	[m2,2G]	UUGCC	CUGGU[t6A]A	GGCAA	AGGUC	CCGGG	UUCAAAU	CCCGG	CCGCGGC	U	
Trp	CCA	GGGCCCA	U[m1A]	[m2G]CUC	A[G+]CCAGGUAA	GAGU	[m2,2Gm]	CCGGG	[Cm]UCCG[mimG]A	[Cm]CCGG	UGGUC	CGGGG	UUCAAGU	CCCCG	UGGGCCC	A	
Tyr	GUA	CCCGCCG	UA	GCUC	A[G+]CCCGGUG	GAGC	[m2,2Gm]	CCCGG	CUGUA[mimG]A	CCCGG	UGGUC	CGGGG	[s2U]U[Cm][m1I][m1A]GU	CCCCG	CGGCGGG	A	
Ser	UGA	GCCGGGG	UG	CCCG	AGCGGUCCA	AGGG	[m2,2G]	[m5C]UGGC	CU[mchm5Um]GA[mimG]A	GCCAG	UAGGGUUAUCCCU[m2,2G]C	G[Um]GGG	[s2U]UCAAAU	CCCAC	CCCCGGC	G	
Ser	GGA	GCCGGGG	UG	CCCG	AGUGGACUA	AGGG	G	CUGGC	CUGGAGA	GCCAG	UGUGGAUCUUCCACGC	GCGGG	UUCAAAU	CCCGC	CCCCGGC	G	
Val	GAC	GGGCCCG	UC	[m2G]UCU	A[G+]C[Cm]UGGUU-A	GGA[Cm]	[m2,2G]	CUGCC	[Cm]UGAC[m1G]C	GGCAG	AAA[xU]C	[m5C][Um]GGG	[s2U]/[m^{5}s^{2}U]U[Cm][m1I][m1A]GU	CCCAG	CGGGCCC	A	

FIGURE 1. Compilation of modified tRNA sequences from *S. acidocaldarius*, *M. maripaludis*, and *P. furiosus*. Red nucleotides indicate fragments obtained by RNase T1, and/or RNase A, and/or RNase U2 digestion, while black nucleotides represent regions that could not be analyzed. Nucleotides in gray are modified nucleotides with a mass corresponding to several possible modifications.

both contain, after the conserved U8, U9 instead of the very common R9 (R = A/G). This is also the case in *Thermoplasma acidophilum*, a thermo-acidic Euryarchaeon where s^4U8 and s^4U9 have been detected (Tomikawa et al. 2013). This occurrence of two consecutive Us at positions 8 and 9 is surprisingly accompanied by an unusual G38 opposite to U32 in the anticodon loop (especially with a CAG anticodon). Although we could not detect s^4U at either position 8 or 9, the nucleoside analysis (Supplemental Figs. S2, S3) confirms the presence of s^4U in both *M. maripaludis* and *P. furiosus*. An analysis of the GtRNAdb shows that these correlations stand out in *Euryarchaeota* and *Thaumarchaeota* (Chan and Lowe 2009, 2016). The D-stem presents also some particularities like a preference for G10-Y25 and Y13-G22 with the central two base pairs maintained as Watson–Crick. It is likely that some, if not all, of the uridines at positions 13, 22, and 25 of the D-stem and position 39 of the anticodon stem (AC-stem) are pseudouridines (Ψ) which cannot be differentiated from U by mass spectrometry. Moreover, the AC-stem has a pronounced preference for a G30–C40 pair, while in other tRNAs, especially of mesophilic organisms, a A30–U40 pair is present (Marck and Grosjean 2002). There are seven residues in the AC- and T-loops with the conserved residues U33 and the favored C32/A38 opposition in the anticodon loop. In the Sulfolobales clade of the *Crenarchaeota*, the tRNA-Cys(GCA) has an unusual C33 (together with G27oU43), but unfortunately we could not isolate

TABLE 1. List of modified positions in different Archaea

	Halophile	Methano	Coccales	Pyro	Coccales	Sulfolobales
Position	*Haloferax volcanii*	*Methanococcus maripaludis*	*Methanocaldococcus jannaschii*	*Thermococcus kodakarensis*	*Pyrococcus furiosus*	*Sulfolobus acidocaldarius*
1					**2H-m^{7}G**	
3						**ac^{4}C**
5					**ac^{4}C**	
6			m^{2}G	Cm	m^{2}G/m$^2{}_2$G	
7		**ac^{6}A**				
8		*(s^{4}U)**	s^{4}U	s^{4}U	**Um**/*(s^{4}U)**	
9				(m^{1}A/m^{1}G)*	m^{1}A/m^{1}G	m^{1}A
10>>	m^{2}G/m$^2{}_2$G	m^{2}G/m2$_2$G	m^{2}G/m$^2{}_2$G	m$^2{}_2$G	m^{7}G/2H-m^{2,7}Gm/(m$^2{}_2$G*)	m^{2}G/m$^2{}_2$G
15>>	G+	G+	G+	G+	G+	G+
17		s^{2}C		s^{2}C	s^{2}C	Cm/Am
18					**Gm**	**Gm**
20				Um		Am
21					**Um**	
22		**Gm**			**Gm**	
23						**Gm**
25						**Cm/Um**
26>>	m^{2}G/m$^2{}_2$G	m$^2{}_2$G	m$^2{}_2$G/m$^2{}_2$Gm	*(m$^2{}_2$G)**	m^{2}G/m$^2{}_2$G/m$^2{}_2$Gm	m^{2}G/m$^2{}_2$G/m$^2{}_2$Gm
27					**m^{5}C**	**m^{5}C**
29						**Gm**
32	Cm	*Cm/(s^{2}C)**	Cm/Um/s^{2}C	m^{5}Cm/*(s2C**	s^{2}C	Cm
33		**s^{2}U**				
34>37	Anticodon and its 5′ flanking purine-37 (see Table 2)					
38					**Am**	
39	Um/m^{5}C	s^{2}U		Cm	m$^2{}_2$G/Cm	Cm
40	m^{5}C					ac^{4}C
41					**Cm**	**ac^{4}C**
42				Gm	Cm/ac^{4}C/Gm	
44					**Um**	
47					xU	Am/m$^2{}_2$G/xU
48>>	m^{5}C	*(m^{5}C)**	m^{5}C	m^{5}C	m^{5}C	m^{5}C
49>>	m^{5}C	*(m^{5}C)**	m^{5}C	m^{5}C	m^{5}C	m^{5}C
50					Gm	**Um/Cm**
51						**Gm**
54	m^{1}Y	m^{1}Y	m^{1}Y	m^{5}s^{2}U	m^{5}U/m^{5}s^{2}U	s^{2}U/s^{x}m^{x}U

Continued

TABLE 1. *Continued*

	Halophile	Methano	Coccales	Pyro	Coccales	Sulfolobales
Position	*Haloferax volcanii*	*Methanococcus maripaludis*	*Methanocaldococcus jannaschii*	*Thermococcus kodakarensis*	*Pyrococcus furiosus*	*Sulfolobus acidocaldarius*
55						
56>>	Cm	Cm	Cm	Cm	Cm	Cm
57>>	m^1I	m^1I	m^1I	m^1I	m^1I	m^1I
58		m^1A	m^1A	m^1A	m^1A	m^1A
62						**Cm**
63					**ac^4C**	
64					**ac^4C**	**Um**
67			m^2G	m^2G		
68					**Gm**	
69						**Cm**
72				*(m^5C)**	*(m^5C)**	**Um/m^5C**

See Table 2 for positions 34–37 in the anticodon loop. Note that, in the case of *T. kodakarensis*, the information corresponds to the single tRNA-Trp (Hirata et al. 2019), with the supplemental information indicated by an asterisk that should exist in other tRNAs of that archaeon (Krishnamohan et al. 2019 and Supplemental Table S2). New modified positions found in the three Archaea of this study are in bold. Pseudouridines are not indicated, except at position 54 (Y). In the case of methylation, our technique does not always allow to place unambiguously the methyl group (either on the base or on the ribose). However, knowing the type of chemical modification at equivalent positions in tRNAs of other closely related archaea, it is often possible to guess for putative types of modification. For example, in both *S. sulfolobus* and *T. kodakarensis*, A9 was demonstrated to be catalyzed by N1-methylating enzyme Trm10 (Kempenaers et al. 2010; Krishnamohan et al. 2019), an enzyme for which the corresponding gene (PF0678) also exists in *P. furiosus* (Supplemental Table S2). Likewise, the detected monomethylation of C56 in *M. maripaludis*, *P. furiosus* and *S. acidocaldarius* was interpreted as Cm56, again based on the presence in their genomes of the gene coding for a well characterized almost ubiquitous Trm56 enzyme that is present in all archaea sequenced so far (Supplemental Table S2). All these putative modifications are indicated as (XXX)*. When uncertain, the methylation is noted as a 2′O-methyl and underlined. Likewise, the modifications not completely identified are underlined. Data for *H. volcanii* are from Gupta (1984, 1986) and Grosjean et al. (2008a), data for *M. jannaschii* are from Yu et al. (2019).

those tRNAs. The T-loop is closed by an invariant G53–C61 pair with always the possibility to form a U54/A58 *trans* Watson–Crick/Hoogsteen stacking against it. The *trans* Watson–Crick/Hoogsteen between U8 and A14 is always present and triple formation with A21 is possible in the large majority of sequences. Interestingly, 15–48 is always a G15–C48 pair (Supplemental Fig. S11A) that stacks with the invariant A60 (from the T-loop). In three tRNA sequences (Ile, elongator Met, and Thr) of *P. furiosus*, residue 60 is G instead of the usual U, and is preceded by A59, which forms a rare combination. C48, as expected, is often methylated on C5 (m^5C); in which case, the methyl group would be in the tRNA core and in the neighborhood of the charged formamidino group of G15 (G^+). The positive charge on G15 points in a cavity surrounded by three negative phosphate groups from residues 7, 14, and 15. Altogether, these peculiarities correspond to tRNAs well stabilized, even in the mesophilic *M. maripaludis*.

New modified positions

Table 1, Figure 1, Supplemental Figure S10, and the corresponding Supplemental Table S2 list all the modified nucleotides we have detected, among them several are present at newly identified positions (mostly methylations and acetylations in the two thermophilic archaea). In eukaryotic tRNAs, the $m^2{}_2G$ modification is exclusively found at position 26 (exceptionally at position 27) at the interface between the D- and AC-stems (Machnicka et al. 2014), while in archaeal tRNAs, it is present at many other locations. The $m^2{}_2G10$ is observed when residue 25 is a U, in which case the $m^2{}_2G$ modification stabilizes the G10oU25 wobble pair (discussed in Urbonavicius et al. 2006). Similarly, the 26-44 pair displays a frequent G26 modification (m^2G, $m^2{}_2G$ but also $m^2{}_2Gm$), forming either G26oU44 or G26–A44 pairs (Supplemental Fig. S11D), rarely G26–C44 (with a slight bias toward $m^2{}_2G$ at position 26 when pairing with U44, especially for the long-arm tRNAs Leu and Ser). Surprisingly, in tRNA-Ser(GGA) of *P. furiosus*, MS/MS data and T1 cleavage analysis (Supplemental Table S1) indicate the presence of $m^2{}_2G$ at positions 39, which would induce a pronounced propeller-twist of the C31 = $m^2{}_2G39$ pair (as seen in Supplemental Fig. S11D). A similar situation exists for $m^2{}_2G6$–C67 in the amino acid stem. Only thermophilic archaea present the triple methylation ($m^2{}_2Gm$) on both the base and the ribose. Such "doubly" modified nucleosides (xNm) are hallmarks of hyperthermophilic tRNAs (McCloskey et al. 2001; Grosjean et al. 2008b; Hori et al. 2018).

In tRNAs of both Bacteria and Eukarya, the positively charged m^7G is often found at position 46 in the variable region (Machnicka et al. 2014). m^7G has been identified in unfractionated bulk tRNAs of two thermophilic archaea, *Thermoproteus neutrophilium* and *Thermoplasma acidophilium* (Edmonds et al. 1991). Recently m^7G was located at position 49 of *T. acidophilum* tRNA-Leu (UAG) (Tomikawa et al. 2013). Here, we report the presence of m^7G at position 10 of tRNA-Met^i (CAU) of *P. furiosus*. We suggest also the presence of a reduced neutral form of 2H-m^7G at positions 1 and 10 of tRNA-Glu (CUC) (for position 10 we suggest an unexpected and unique case of tri-methylated 2H-$m^2{}_7Gm$). These suggestions are based on the following observations. The monomethylated nucleotide at position 10 of tRNA-Met^i has a mass of 359 Da with a neutral loss of 165 Da that is characteristic of a normal m^7G, while at position 1 of tRNA-Glu, the methylated nucleotide has a mass of 361 Da (Supplemental Fig. S4), which could correspond to the reduced form of m^7G (2H-m^7G) (Supplemental Fig. S4D; Wintermeyer and Zachau 1975). The same reduced form exists for 2H-$m^2{}_7Gm$ at position 10 of tRNA-Glu. The accuracy is less than 0.05 Da; however, in the absence of standard substance to compare against, these suggestions are still tentative. The presence of an m^7G derivative at position 1 of an archaeal tRNA would indeed be remarkable. Usually, such guanosine derivatives are found in the form of a positively charged cap-like structure protecting the tRNA against 5′-exonucleolytic degradation (see for example Ohira and Suzuki 2016). In m^7G cap, the 5′ extremity is a 5′OH, while our MS/MS data show a classical 5′P extremity on the *P. furiosus* tRNA-Glu (see Supplemental Fig. S4B, which corresponds to the MS/MS spectrum of the 5′-end of tRNA-Glu where, in the ion series c, c1 corresponds to p2H-m^7G). A putative protecting role of such terminal m^7Gp against specific exonucleolytic degradation processes remains to be demonstrated in Archaea.

Very recently, Sas-Chen et al. (2020) published a thorough analysis of ac^4C modifications in rRNAs and tRNAs across phylogeny including several archaeal species. They found high concentrations of ac^4C in the hyperthermophiles *Thermococcus kodakarensis, P. furiosus*, *Thermococcus sp.* AM4, and *S. solfataricus*, with a preference for CCG sites, the modified C being the middle C. Residues ac^4C most certainly contribute to the stabilization of base pairs (Kawai et al. 1992). In this work we identified 13 ac^4C modifications in the isolated tRNAs (Table 1; Fig. 1) and all, except the one at the wobble position 34 of tRNA-Gln (ac^4CmUG—see below), are found in the middle position of a CCG motif within a stem. For example, in *S. acidocaldarius*, ac^4C is found at position 41 of the AC-stem of tRNA-Leu(UAG) and at position 40 of tRNA-Ala(UGC). m5C40 is found in tRNA-Ile of *H. volcanii* (Grosjean et al. 2008a). In position 40, the modification cannot be a ribose methylation (Cm40) because the O2′ (C40) is functionally engaged in H-bonds with the conserved A1339 of the 16S rRNA during the P state of translation (Selmer et al. 2006; Watson et al. 2020). Indeed, the base pairs of the AC-stem 30–40 and 29–41 are both involved in contacts with, respectively, G1338 and A1339 of the 16S rRNA in the P state (Supplemental Fig. S12A,

B). In tRNA-Met^{i}(CAU) of *S. acidocaldarius,* residue G29 pairing with C41 is methylated but in this case the 2′-hydroxyl group is far enough from G1338 and therefore can accommodate a methyl group, leading to a final assignment of 2′O-methyl G29 (Gm). In *S. acidocaldarius*, we found a m5C residue at position 72 in the acceptor stem of tRNA-Asp(GUC), while at the same position in tRNA-Gln(UUG), an undetermined monomethylated U72 (indicated as Um) was found. Residue m5C72 has been already reported in *S. solfataricus* tRNA-Glu/-Gly/-Met (Wagner et al. 2004), as well as in tRNA-Cys and tRNA-Thr of humans (Haag et al. 2015). Residue m5C is quite common at positions 40 (see above), 48, 49 of tRNAs in both Eukarya and Archaea (Machnicka et al. 2014). Worth to note is that, in tRNAs of *S. acidocaldarius* (and to a lower extent of *P. furiosus),* many residues are monomethylated all over the molecules. Some are catalyzed by site-specific protein-only methyltransferases while others are catalyzed by Fibrillarin-C/D box sRNP guide machinery acting on specific intron-containing pre-tRNAs (Gaspin et al. 2000; Omer et al. 2000; Clouet d'Orval et al. 2001; Dennis et al. 2001). From the database maintained by Todd Lowe (provided at http://lowelab.ucsc.edu/snoRNAdb/), this enzymatic machinery in *S. acidocaldarius* could target site-specific methylations to tRNA-Gly(CCC) at C50 in the D-arm (Tang et al. 2005; Zago et al. 2005) and to tRNA-Gln (UUG) at U34 and G18 of the AC- and D-loops, respectively (see also Ziesche et al. 2004), a situation indeed observed in the present study. However, in *S. acidocaldarius*, a few other observed monomethylated nucleotides remain orphan and await corresponding snoRNAs or stand-alone specific methyltransferases to be identified (Table 1; Supplemental Table S2, they are indicated as Xm). Likewise, in tRNAs of *P. furiosus*, several positions were predicted to be methylated via the Fibrillarin-C/D box machinery: tRNA-Trp(CCA) at C32, C34, C39, tRNA-Leu(UAA) at G47, tRNA-Leu(CAA) at C34, tRNA-Gln at C16, tRNA-Val and tRNA-Gly at G26 and tRNA-Asp(GUC) at U35 (http://lowelab.ucsc.edu/snoRNAdb/). However, although we do observe Cm34, Am38, and $m^2{}_2$Gm26 in some tRNAs, the other putative methylated positions listed above were presently not detected in the set of tRNAs we analyzed.

In *M. maripaludis*, most modified nucleotides were expected, except a methylated G22 (Gm) in tRNA-Ala (UGC) and an unidentified adenine derivative in the acceptor stem of tRNA-Val(GAC). The corresponding nucleoside A*7 displays a mass of 309.1 Da corresponding to either hypermodified $m^6{}_2$Am or monomodified N6-acetylA. We tend to favor the latter (ac^{6}A), since such acetylated adenosine has been discovered in another methanogen (Sauerwald et al. 2005) with its location, however, tentatively assigned to residue 37. In *M. maripaludis*, a mass corresponding to f^{6}A37 was found instead in tRNA-Asp (GUC) (see below). Intriguingly, in tRNA-Leu(UAG), the C32 is unmodified but U33 of the AC-loop, a very rarely modified conserved residue, is thiolated (s^{2}U) (Supplemental Fig. S6). An s^{2}U33 has already been found in trypanosomid mitochondrial tRNA-Trp(CCA) (Crain et al. 2002), where the stop codon is translated as Trp (Alfonso et al. 1999). The probability to find site-specific 2′O-methylations guided by snoRNA in *M. maripaludis*, is meager as the frequency of occurrence of potential intron-containing tRNA targets in mesophilic archaea is much lower than in the hyperthermophilic archaea (Sugahara et al. 2008).

Modifications in D-loops

In mesophilic bacterial and eukaryotic tRNAs, U17 as well as a few other Us at positions 16, 20, and 20a of the D-loop, are usually modified to dihydrouridine (D). In thermophilic bacteria and hyperthermophilic archaea, this thermolabile dihydrouridine D, is rare or even totally absent (Edmonds et al. 1991), while in psychrophilic bacteria D is abundant (Dalluge et al. 1997). From these observations, Dalluge and coworkers (1977) suggested an interesting functional role for D in the maintenance of a certain degree of conformational flexibility in tRNAs, especially important to organisms growing at low temperatures where the dynamics of thermal motions of tRNAs are severely compromised. Unexpectedly, in tRNAs of *M. maripaludis* and of *P. furiosus* we found 2-thiocytidine at position 17 (Table 1; Fig. 1), a situation found also in the hyperthermophilic tRNA-Trp of *T. kodakarensis* (Hirata et al. 2019). However, in *S. acidocaldarius,* position 17 contains instead a methyl group on either C or A, most probably a 2′-O-methyl (discussed above) (Table 1). Usually, s^{2}C and Cm are found at position 32 of anticodon loops (Jühling et al. 2009; Boccaletto et al. 2018). Thiolation of pyrimidines and methylation of 2′-O-ribose of nucleotides are known to favor stacking, thus limiting local flexibility of the RNA (Plesiewicz et al. 1976; Larsen et al. 2015), a property that is obviously important in organisms thriving at high temperatures.

In most crystal structures, residue 17 bulges out of the tRNA core structure and is exposed to solvent. In addition, residues 16 and 17 are in the vicinity of two other bulging residues from the T-loop, 59 and 60, and modifications in the D-loop may limit the tendency of residues 16 or 17 to bulge out of the loop. The locations of residues 16 or 17, either within or outside the tRNA core, may influence interactions between other residues in the T-D environment. When residue 16 is a pyrimidine, it is often observed that it forms a pair with residue 59, especially when 59 is also a pyrimidine. However, in the present archaeal tRNAs, residue 59 is always A and residue 60 is mostly U. In that case, a couple of crystal structures show A59 stacked on G15–C48 and U60 forming an H-bond between N3(U60) and the phosphate group between the two invariant G18–G19 of the D-loop (see for example PDB entry 2DU3, Supplemental Fig. S12C; Fukunaga and Yokoyama

2007). As stated above the modified G bearing a positively charged formamidino group (G^+, archaeosine), a hallmark of archaeal tRNAs, also contributes to the global stability of the 3D-core of tRNA. In Bacteria and Eukarya, nucleotide 15 is never modified.

Modifications in TΨC-loops

Nucleotide U54 is nearly always modified to thymine or 5-methyluridine (m^5U) in Eukaryotes and Bacteria (Machnicka et al. 2014). However, in most Archaea (mainly Euryarchaeota), a 1-methylpseudouridine ($m^1\Psi$) is generally found instead (Pang et al. 1982; Gupta 1984; McCloskey et al. 2001; Chatterjee et al. 2012; Yu et al. 2019). Such a modification adds a methyl group to the pseudouridine (Ψ) at the free N1 atom, a position structurally equivalent to the C5 atom of T54. In the case of *Ignococcus hospitalis*, a crenarchaeon thriving at temperatures up to 100°C, $m^1\Psi54$ was shown to be further hypermodified into $s^4m^1\Psi$ (Rose et al. 2020). The sulfur atom at position 4 in Ψ54 is structurally equivalent to the sulfur atom at position 2 of U54 (see Fig. 6 of Rose et al. 2020). In a few other archaea belonging to the thermococcales clade among the Euryarchaeota, as *P. abyssi*, *P. furiosus*, and *T. kodakarensis*, the bacterial-like m^5s^2U has been identified instead (Kowalak et al. 1994; Urbonavicius et al. 2008; Hirata et al. 2019; for review, see Hori et al. 2018).

In this work, we confirm the presence of $m^1\Psi54$ in the mesophilic *M. maripaludis*, and in the hyperthermophilic *P. furiosus* a mix of m^5U54 and m^5s^2U54 (Supplemental Figs. S2 and S9). For *P. furiosus*, the result is consistent with the fact that the 2-thiolation process occurs after 5-methylation of U54 and 1-methylation of A58 (Shigi et al. 2006). For *S. acidocaldarius*, the MS/MS sequencing spectra show the presence of thiolated U/Ψ (Supplemental Fig. S9A) and, only in tRNA-Val, the presence of a methylthiolated U/Ψ (Supplemental Fig. S9B). In the latter case a neutral loss of 142 was observed, which corresponds to a modified U with a methyl group and a sulfur atom, a situation that was not observed in the case of m^5s^2U54-containing tRNAs of *P. furiosus*. The total absence of m^5s^2U is also evident in the chromatogram profile of nucleoside digests from *S. acidocaldarius*, in comparison to *P. furiosus* (Supplemental Fig. S2). Altogether, these facts strongly suggest the presence of a methylthiolated derivative like $s^4m^1\Psi$ as in *I. hospitalis* (Rose et al. 2020). This situation is however perplexing. Indeed, examination of the *S. acidocaldarius* genome reveals the lack of genes coding for both Pus10 (catalyzing formation of Ψ55 and Ψ54) and TrmY (catalyzing $m^1\Psi$), while genes coding for TtuA and TtuB responsible for thiolation of U/Ψ (into s^2U or $s^4\Psi$) are present (Supplemental Table S2). It might be that in *S. acidocaldarius* a new type of s^2U54/or $s^4\Psi54$-methylating enzyme exists. Therefore, we prefer to indicate that position as $s^xm^xU/\Psi54$.

In all three archaea analyzed, A58 is methylated at position N1 (m^1A58), adding a positive charge on the base while still allowing for the formation of the usual *trans* Watson–Crick/Hoogsteen 54–58 pair (Supplemental Fig. S11B). In sum, thiolation of m^5U, $m^1\Psi$ and U/Um at position 54 of tRNA appears as a hallmark of (hyper)thermophilic archaea. In thermophilic bacteria, such as *Thermus thermophilus*, the thiolation process was demonstrated to be thermo-inducible, as the level of 2-thiolation of m^5U54 increases with the cultivation temperature (Shigi et al. 2006). The van der Waals radius of the sulfur atom is 0.3 Å larger than that of the oxygen atom and its presence may fill the cavity present around that position in the overall compact T-loop, thereby excluding solvent molecules and promoting stacking.

Residue 56 is a conserved C, usually 2′-O-methylated on its ribose in almost all archaeal tRNAs analyzed so far (Clouet-d'Orval et al. 2005; Renalier et al. 2005). The ribose of C56 is highly accessible in the turn of the TΨC-loop and methylation allows protection against hydrolysis, especially at high temperatures. Cm56 forms a Watson–Crick pair with the conserved G19 of the D-loop (G19–C56) and therefore cross-bridges the two parts of the tRNA core. Residue 57 is either G or A where A57 is often doubly modified first into m^1A, positively charged, and then into m^1I neutral (Grosjean et al. 1995). Residue 57 intercalates between the invariant G18 and the conserved G19–C56 pair. With G57, there is an H-bond between N2(G57) or N1(G57) and one anionic oxygen atom of the phosphodiester bond between 18 and 19 and with m^1I57 the methyl group will be on the solvent exposed surface of the T-loop. It is noteworthy that nucleotide 17 is probably 2′-O-methylated on the ribose in *S. acidocaldarius* and *P. furiosus*, and 2-thiolated in *P. furiosus* and *M. maripaludis* (Fig. 3). The 2′-O-methyl group of Cm56 could protect the bent backbone from cleavage. Both Cm56 and m^1I57 are unique and frequent in archaeal tRNAs (Table 1), while m^1A58 is also present in many bacterial and eukaryotic tRNAs (Jühling et al. 2009).

Modifications at position 34 (Table 2)

The majority of the modifications observed at position 34 are the same as those identified in several tRNAs, mostly from Bacteria, such as $cmnm^5s^2U$, mnm^5U, cnm^5U, $mchm^5U$, Cm, Um, ac^4C, except that, in the present thermophilic archaeal tRNAs, some of them are doubly modified with an extra 2′-O-methylribose, $mchm^5Um$, ac^4Cm, s^2Um (Grosjean et al. 2008a,b, 2010; Jühling et al. 2009). In tRNA-Met^i of *M. maripaludis* and *P. furiosus* (recognizable by the invariant last three G = C pairs of the AC-stem [Kuchino et al. 1982]) and tRNA-Trp of *S. acidocaldarius*, there is a modified Cm34 (Table 2). The methyl group in that position occupies a tight space in the decoding site locking C34 for pairing only with G3 (Met codon AUG or

TABLE 2. List of the archaeal tRNAs for which we could fully characterize the tRNA anticodon regions

Amino acid	Anticodon	AC-loop
Methanococcus maripaludis		32–38
Asn	GUU	CUGUU[hn^{6}A]A
Asp	GUC	CUGUC[f^{6}A]C
Glu	UUC	CU[mnm^{5}U]UC[m^{1}G]A
Gly	UUC	CU[cnm^{5}U]CCAA
Ile	GAU	CUGAU[hn^{6}A]A
Ile	CAU	CU[C+]AU[hn^{6}A]A
Lni	CAU	CUCAUAA
Leu	UAG	C[s^{2}U][xU]AGAA
Lys	UUU	CU[cmnm^{5}s^{2}U]UU[t^{6}A]A
Met	CAU	CU[Cm]AU[hn^{6}A]A
Phe	GAA	[Cm]UGAA[imG-14]A
Tyr	GUA	[Cm]UGUA[xG]A
Pyrococcus furiosus		32–38
Ala	GGC	UUGGCAA
Ala	UGC	CU[cnm^{5}U]GCAA
Arg	CCU	[s^{2}C]UCCU[t^{6}A])A
Arg	UCU	CU[cnm^{5}U]CUAA
Gln	UUG	CU[cnm^{5}U]UGGA
Gln	CUG	CU[ac^{4}Cm]UGGA
Glu	CUC	CUCUC[m^{1}G]A
Gly	CCC	CUCCCAA
Gly	GCC	CUGCCAA
Gly	UCC	CU[cnm^{5}U]CC[m^{1}G]A
Ile	GAU	CUGAU[hn^{6}A]A
Lni	CAU	CUCAUAA
Leu	CAG	UUCAGGG
Leu	UAG	UU[cnm^{5}U]AGGG
Lys	CUU	CUCUU[ms^{2}t^{6}A]A
Met	CAU	CU[Cm]AU[hn^{6}A]A
Phe	GAA	CUGAA[mimG]A/[Am]
Pro	UGG	UU[cnm^{5}U]GG[m^{1}G]GA
Thr	GGU	UUGGU[t^{6}A]/[hn^{6}A]A
Thr	UGU	CU[cnm^{5}U]GU[hn^{6}A]A
Ser	GCU	CUGCU[t^{6}A]A
Val	GAC	CUGACAA
Sulfolobus acidocaldarius		32–38
Ala	UGC	[Cm]U[Um]GCAA
Asp	GUC	CUGUC[m^{1}G]/[imG2]/[mimG]A
Gln	UUG	CU[Um]/[s^{2}Um]UG[m^{1}G]C
Glu	CUC	[Cm]UCUCAA
Gly	GCC	CUGCCAC
Leu	CAG	[Cm]U[Cm]AG[m^{1}G]C
Leu	UAG	[Cm]U[s^{2}Um]AG[m^{1}G]C
Lni	CAU	[Cm]UCAUAA
Trp	CCA	[Cm]U[Cm]CA[mimG]A
Phe	GAA	CUGAA[mimG]A
Tyr	GUA	CUGUA[mimG]A

Continued

TABLE 2. *Continued*

Amino acid	Anticodon	AC-loop
Val	GAC	[Cm]UGAC[m^1G]C
Ser	UGA	CU[mchm^{5}Um]GA[mimG]A

This list is not as complete as those in the other figures or tables. Red nucleotides indicate fragments obtained by RNase T1 and/or RNase A digestion, while black nucleotides represent regions that could not be analyzed. Nucleotides in gray are modified nucleotides with a mass corresponding to several chemistries for modifications. Positions 34, 35, and 36 are underlined.

Trp codon UGG) and preventing mispairing of the C34 residue with A3 (Ile codon AUA or stop codon UGA). The Cm34 modified nucleotide found in tRNA-Trp in *S. acidocaldarius* was previously observed in *H. volcanii* (Gupta 1984) and in *T. kodakarensis* (Hirata et al. 2019). In tRNA-Ile, the anticodon CAU has to read exclusively the Ile codon AUA and the C34-tRNA is modified into agmatidine (C^+) (Ikeuchi et al. 2010; Mandal et al. 2010). The corresponding mass spectrum for tRNA-Ile from *M. maripaludis* is shown on Supplemental Figure S8. In tRNA-Gln(CUG) of *P. furiosus*, the modified residue ac^{4}Cm34, the same as in the homolog tRNA-Gln of *H. volcanii*, was found (Gupta 1984). NMR studies have shown that ac^{4}Cm is exceptionally rigid in conformation owing to the additive nature of the acetylation and methylation modifications which stabilize the 3′-*endo* sugar conformation (Kawai et al. 1992). The same remark probably applies to Um34 and s^{2}Um34 that occur in tRNA-Gln(UUG) and tRNA-Leu(UAG) of *S. acidocaldarius*. Interestingly, ac^{4}C residues were also found in the acceptor stems of *P. furiosus* and *S. acidocaldarius*, as well as in the AC-stem of *S. acidocaldarius* (see above and Table 1). Modification of U34 is necessary for decoding G3 ending codons. The modification in U34 (U34*) changes the chemical structure of the U34 so that a pair U34*–G3, with the U displaced into the minor groove, and not into the major groove, can be stabilized (Rozov et al. 2016; Westhof et al. 2019). Several U34-containing tRNAs were observed modified either at position C5 and/or at position C2. The case of xU34 in tRNA-Leu of *M. maripaludis* is explained below.

Wyosine and modifications at position 37

Residue 37 is commonly a purine, often (hyper)modified, that stacks on the first 1–36 bp formed between the codon and the anticodon during translation on the ribosome. Residue 37 should not be 2′-O-methylated because it forms an H-bond with N6(A1913) of helix H69 from the large subunit in the A state in known crystal structures of ribosomes (Supplemental Fig. S12D; Selmer et al. 2006). The type of modification at base 37 usually correlates with the rest of the so-called extended anticodon stem–loop, especially with the adjacent nucleotide 36 of the anticodon (Yarus 1982; Grosjean and Westhof 2016). In *E. coli*, for example, m^1G exclusively occurs in tRNAs recognizing codons CCN (Pro), CGN (Arg), CUG (Leu), that is, in tRNAs decoding in the codon quadrant starting with C1 (Fig. 2; Supplemental Fig. S10). All the other tRNAs, belonging to the three remaining decoding quadrants, harbor either an unmodified A37 or a modified A37 (m^2A, t^6A, m^6t^6A, ms^2i^6A) with large modifications in the codon quadrants starting with U1 or A1 (Fig. 2; Supplemental Fig. S10). In the halophilic mesophilic *H. volcanii*, m^1G37 occurs in tRNAs decoding codons starting with C1 (G-ending codons) and U1 (A-ending codons, with the exceptions of tRNA-Ser(CGA), tRNA-Ser(GGA) where A38 is found and tRNA-Glu(CUU) with m^1G; Gupta 1986). In all the other tRNAs of *H. volcanii*, A37 or a modified A37 is used (Grosjean et al. 2008a).

In *M. jannaschii*, *P. furiosus, M. maripaludis*, and *S. acidocaldarius*, the landscape is striking. Indeed, the tRNAs for the codon quadrants starting with C1 and U1 contain m^1G but also imG-14/imG2, a wyosine derivative of m^1G37 that is reminiscent of the yW37 found exclusively in tRNA-Phe of Eukaryotes (de Crécy-Lagard et al. 2010). One finds indeed imG-14 or mimG at position 37 of tRNA-Phe(GAA) in *M. maripaludis* and *P. furiosus* with a wyosine-like (xG) (as discussed below), in tRNA-Tyr(GUA) of *M. maripaludis* and mimG37 in tRNA-Arg(GCG) of *P. furiosus* (Fig. 2; Supplemental Fig. S10). In *S. acidocaldarius*, wyosine derivatives are present in tRNA-Asp(GUC), tRNA-Ser(UGA), tRNA-Phe(GAA), tRNA-Trp(CCA), and tRNA-Tyr(GUA). This surprising result was noted in the case of *M. jannaschii* (Yu et al. 2019) where besides tRNA-Phe, the tRNA-Arg (UCG), tRNA-Cys(GCA), tRNA-Leu(UAA), tRNA-Ser(GGA), and tRNA-Tyr(GUA) contain wyosine derivatives (Fig. 2). In *T. kodakarensis* (Hirata et al. 2019), mimG was found at position 37 of tRNA-Trp. Obviously, the presence of wyosine derivatives is more prevalent in Archaea than in Eukarya, especially in the U1-quadrant.

The tRNAs corresponding to the codon quadrants starting with A1 still prefer large modifications on A37 (t^6A, ms^2t^6A, hn^6A), while the G1-quadrant prefers A and in a few isolated cases also m^1G (Fig. 2; Supplemental Fig. S10). In the case of tRNA-Asp of *M. maripaludis*, a modified A37, which has the same nucleoside mass (295.1) than either di-methyl-A ($m^6{}_2$A) or N6-formyl-A (f^6A) was found. The latter f^6A derivative is the most probable modification. The modified f^6A derivative has been identified in mammalian mRNAs (Fu et al. 2013). Interestingly, as discussed below,

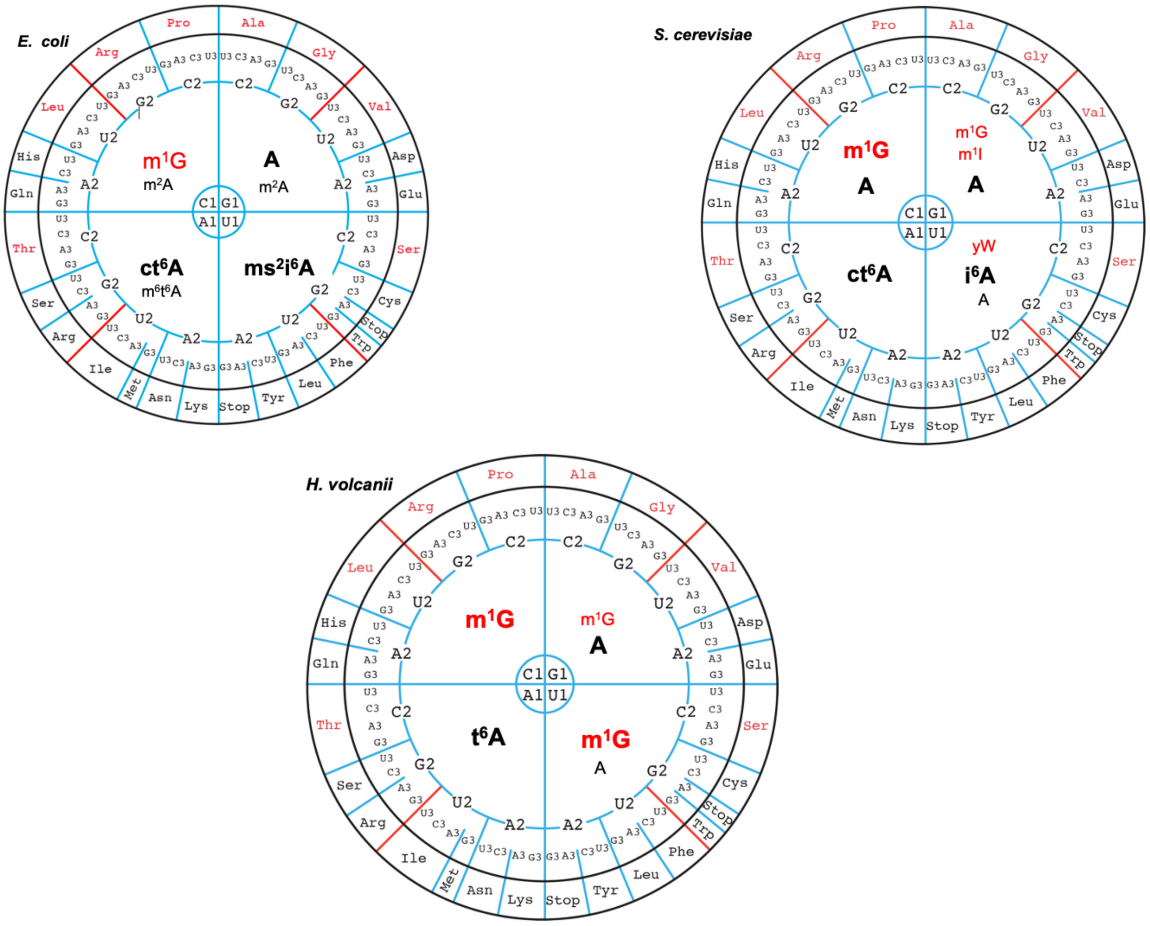

FIGURE 2. Patterns of distribution of modified nucleotides at position 37 in *E. coli*, *S. cerevisiae* and five archaeal species on the wheels of the genetic code (Grosjean and Westhof 2016). In that representation, GC-rich codons are at the *top* and AU-rich at the *bottom* of the wheel. The codon positions are numbered 1, 2, 3 so that G1A2C3 gives an Asp codon. The four quadrants are characterized by the first codon position. Data for *E. coli* and *S. cerevisiae* are from MODOMICS (Boccaletto et al. 2018). Data for *H. volcanii* are from Gupta (1984, 1986) and Grosjean et al. (2008a), data for *M. jannaschii* are from Yu et al. (2019). The wheels for *E. coli*, *S. cerevisiae*, and *H. volcanii* are adapted from Grosjean and Westhof (2016). In each quadrant, the observed modified nucleotides at position 37 are shown (red for G derivatives and black for A derivatives; the very unusual U37 of tRNA-Ala(GGC) of *M. jannaschii* is also shown in black). The modifications observed frequently have a large font and are bold, while those observed only once or in a single tRNA species have a smaller font and are not bold.

in *M. maripaludis*, a mass corresponding to ac^{6}A, a modified nucleotide in the same biochemical pathway as f^{6}A, was observed for residue 7 of the acceptor stem. The homologous tRNA in *M. jannaschii* harbors an unmodified A. There is also one report of the presence of m$^6{}_2$A in *Mycobacterium bovis*, but without identification of either tRNA species or tRNA position (Chan et al. 2011). Also, in *E. coli*, tRNA-Val(UAC) contains m^{6}A37 (Golovina et al. 2009).

Thus, as a rule, in *Archaea*, tRNAs decoding the codon quadrant starting with C1 and U1 harbor either unmodified A37 or mostly modified G37. The G1-quadrant has a preference for unmodified A or slightly modified A with some occurrences of modified G37. While tRNAs decoding the codon quadrant starting with A1 seem to harbor mostly hypermodified A37. In short, the stacking power of a G37 derivative (modified or not) on the first codon–anticodon base pair (Y1-R36) is obviously preferred for efficient decoding on the ribosome, while for R1-Y36 another type of stacked hypermodified A37 is favored. In addition to residue 37, other elements of the extended anticodon stem–loop, including generally simpler chemical modifications on the base and/or the ribose also contribute to the global efficiency and accuracy of the translation process, whatever the temperature at which the archaeon is growing. These rules are more restrictive than those observed in other organisms like *E. coli*, *S. cerevisiae* or *H. volcanii* (Fig. 2; see Grosjean and Westhof 2016).

Tentative identification of three novel modified nucleosides

The sequence analysis of each tRNA allowed the detection of three possibly novel chemical modifications (designated xG and xU). The first one was found at position G37 of

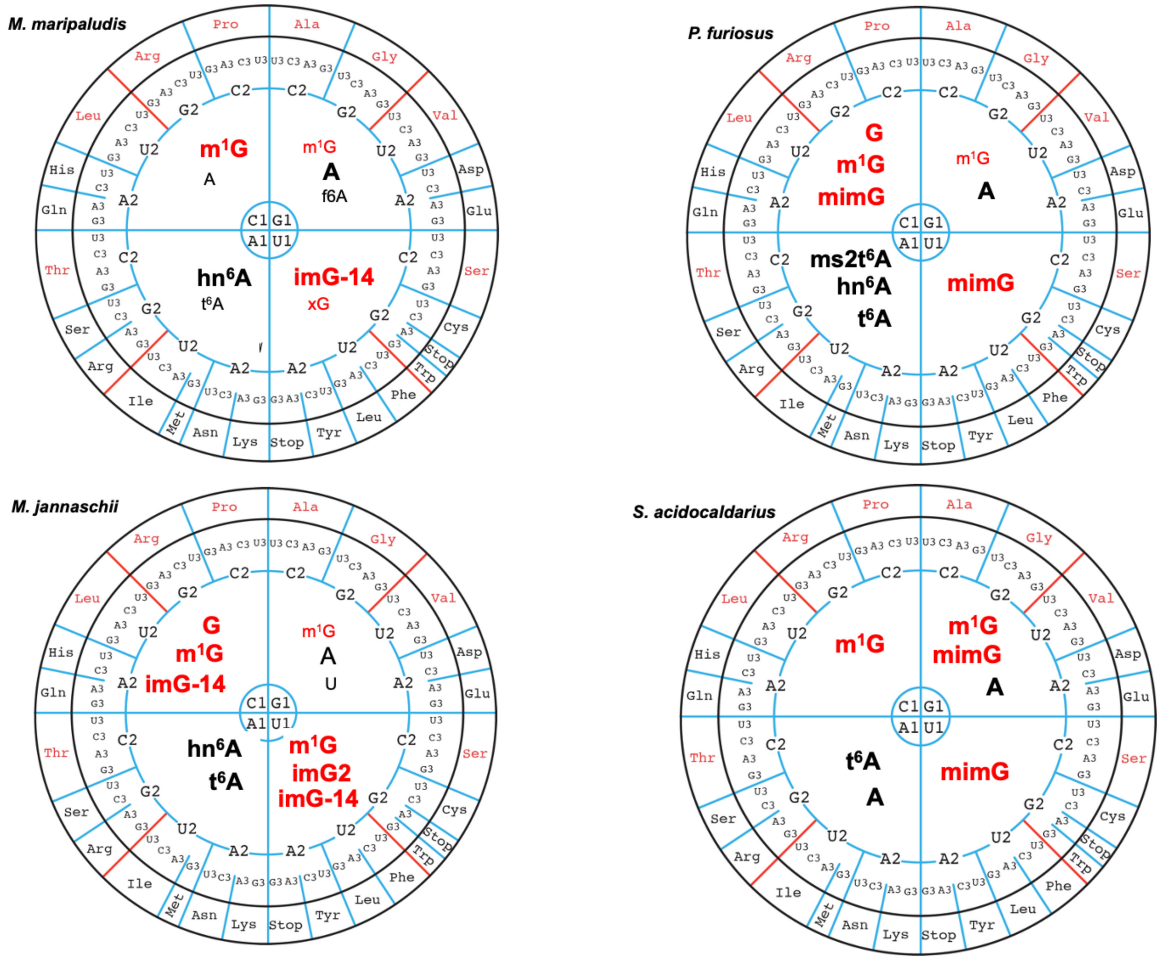

FIGURE 2. *Continued.*

tRNA-Tyr(GUA) of *M. maripaludis* (Supplemental Fig. S5) with a nucleoside mass of 392 Da, much higher than for the expected m^1G37. In the homolog tRNA-Tyr of *M. jannaschii*, G37 is unexpectedly imG-14 of the wyosine metabolism (Yu et al. 2019). If this is also the case for *M. maripaludis* tRNA-Tyr, on the basis of previous work, one would therefore expect to find either imG, yW-86 or yW-72 (see G37 pathway #5 in Fig. 4 of de Crécy-Lagard et al. 2010). Taking into account that CID fragmentation of xG37 occurs between guanine and the modification, the MS/MS spectrum shows the complete mass of guanosine (345 Da) and a neutral loss of 109 Da representing the mass of the adduct (Supplemental Fig. S5). The nucleoside mass of xG (392 Da) could therefore correspond to yW-72 (436.17 Da) with the loss (natural or accidental) of the carboxyl group (44.17 Da).

The second novel modification is xU at position 34 of tRNA-Leu(xUAG) of *M. maripaludis* with a nucleoside mass of 269 Da. Again, in several tRNAs of *M. jannaschii* (Yu et al. 2019), either 5-cyanomethyl-U (cnm^5U) or 2-thiolated-5-cyanomethyl-U (cnm^5s^2U) have been found, as in other archaeal tRNAs (Mandal et al. 2014). We propose that xU34 in *M. maripaludis* is the simpler 5-cyano-U (cn^5U), with the cyano group directly linked to the C5 atom of uracil. Such a derivative is known from organic chemistry (Mao et al. 2018) but was never identified in tRNA so far.

A third unidentified modification, with a nucleoside mass of 338 Da, was found at position U47 of two tRNAs of *S. acidocaldarius*, tRNA-Val(GAC) (Supplemental Fig. S7A) and tRNA-Gly(GCC) (Supplemental Fig. S7B), two tRNAs of *S. acidocaldarius*, tRNA-Met elongator(CAU) (Supplemental Fig. S7C) and tRNA-Thr(UGU) of *P. furiosus*. In Bacteria and Eukarya, 3-(3-amino-3-carboxypropyl)-uridine (acp^3U, nucleoside mass 345.1 Da) is widely conserved in the D- and variable loops (Takakura et al. 2019). It is likely that in hyperthermophilic archaea U47 is modified differently.

DISCUSSION

All the tRNA modifications identified in this work are compiled in several figures and tables. Figure 1 and

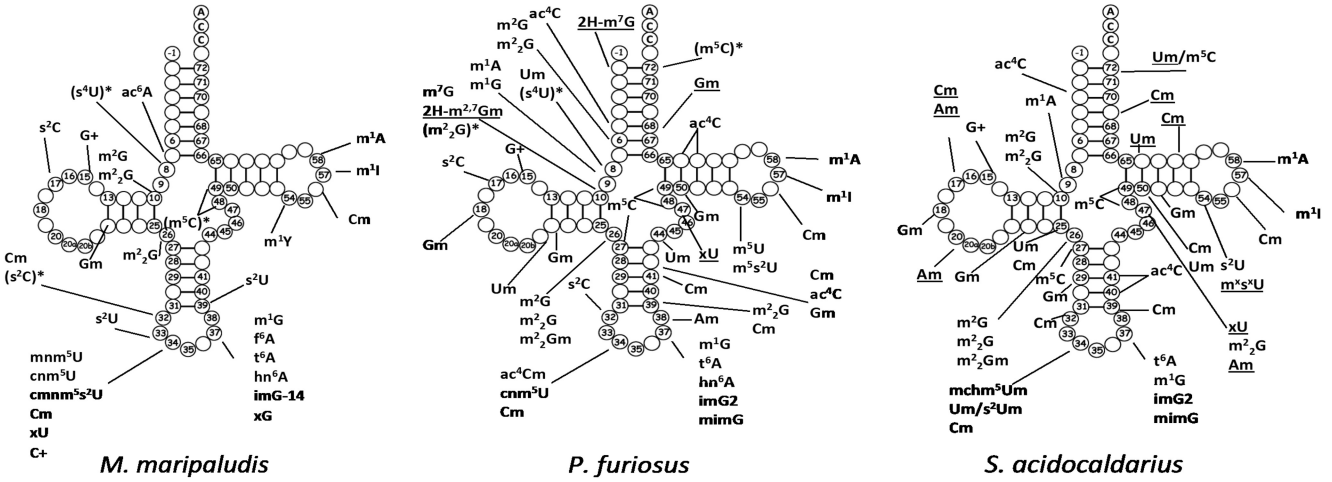

FIGURE 3. Nucleoside modification patterns in archaeal tRNAs determined by experimental MS/MS sequencing. The different modifications observed are reported on a typical cloverleaf two-dimensional structure. Underlined nucleotides are modified nucleotides with a mass corresponding to several possible chemistries for modifications, most of them corresponding to simple monomethylation products (see legend to Table 1). A "/" between two notations means a mixture of modifications. An "x" before a nucleotide, for example, xG, or after a modification symbol, for example, m^x, means that the modification has not been formally identified (see text). Modified nucleotides indicated in brackets followed by an asterisk correspond to modifications deduced from the convergence of information from both the analysis of total tRNA and the presence of the gene coding for the corresponding modification enzyme in the genome of the particular archaeon. The modifications, uncertain or not completely identified, are underlined. Pseudouridines are not represented, except in the case of N1-methylpseudouridine (m^1Y) because of the methyl group.

Supplemental Figure S10 align the intron-less tRNA sequences as deduced from the genomes on which the modified nucleotides are indicated; Figure 3 summarizes the data in cloverleaf representations; Supplemental Figure S3 displays the relative amounts of modified nucleosides identified in bulk tRNAs; Supplemental Tables S2 and S3 list the oligonucleotide fragments obtained after RNase digests. Supplemental Table S2 lists all the archaeal tRNA modification enzymes (and their corresponding coding genes) that have been experimentally validated so far in independent works or deduced from their close similarities with genuine modification enzymes in each of the Methanococcales, Pyrococcales and Sulfolobales groups of Archaea. Altogether, 79 naturally occurring fully matured isoacceptor tRNAs coding for 20 amino acids have been analyzed out of a theoretical total set of 116 species. Thus, the possibility still exists that a few modifications that are specific to the missing tRNAs (especially in the anticodon loop including the wobble position 34) still escaped our analysis.

Nevertheless, the major general conclusions that came out from this comparative analysis of modified nucleotide patterns in three very different archaeal species are the following: (i) the least diversified chemical modification pattern is observed in the mesophilic *M. maripaludis*; (ii) a larger diversity of modifications is found in the two hyperthermophiles; (iii) the largest amount of 2′-hydroxyl ribose methylations occurs in the acidophilic hyperthermophile *S. acidocaldarius*, most of them appear being catalyzed by Fibrillarin-C/D box sRNP guide machinery; (iv) depending on the tRNA species, nucleotides at positions 32, 38, and 39 of the extended anticodon loop are frequently but diversely modified; (v) beyond those found at positions 34 and 37 in the anticodon loop, a few characteristic modifications are found in the body of most if not all isoacceptor species of the three archaea (G^+15 in the D-loop, m^2G/$m^2{}_2G$ at positions 10 and 26 at the beginning and the end of the D-arm, m^5C at positions 48/49 of the variable loop, Cm56, m^1I57 and m^1A58 in the TΨC-loop), some of them are hallmarks of archaeal tRNAs; (vi) except for ac^4C at the wobble position 34, this modified residue is present in the amino-acceptor and anticodon stems of only thermophilic archaeal tRNAs; (vii) at variance with the situation in bacterial and eukaryal tRNAs, m^2G/$m^2{}_2G$ are found at positions other than 10 and 26; (viii) remarkably, the chemical adducts on the conserved U54 of the T-loop depends on archaeon analyzed ($m^1\Psi$, m^5U/m^5s^2U, or $s^4m^1\Psi$); (ix) the suggested presence of m^7G modifications at the nucleotides 1 and 10 in *P. furiosus* would also be remarkable.

All these observations complete and reinforce similar conclusions made by others about the importance of certain post-transcriptional modifications for correct tRNA folding and on final cellular stability of the already G/C-rich tRNAs in thermophiles (Edmonds et al. 1991; Kowalak et al. 1994; McCloskey et al. 2001; Noon et al. 2003; for reviews, see Machnicka et al. 2014; Lorenz et al. 2017; Hori et al. 2018). As a rule, methylations

promote precise H-bonded pairs (e.g., m^1A favors Hoogsteen pairs or $m^2{}_2G$ favors GoU or GoΨ pairs) and electrostatic charges introduced by the chemical adducts are localized in shielded pockets of the tRNA fold (m^1A, m^7G, archaeosine G^+). Modifications in the tRNA core, although not close in sequence, tend to form clusters of modifications filling empty space on the surface of the compact tRNA core. In addition, methylations on the base and/or the ribose affect the hydration shells in complex ways (Auffinger and Westhof 2001). Thiolation of U or C, acetylation of C, and isomerization of U to Ψ, can stabilize the 3′-*endo* sugar conformation, fill in space and enhance stacking power or base-pairing (Kawai et al. 1992; Davis 1995; Larsen et al. 2015; Sas-Chen et al. 2020). High-resolution crystallographic structures would be necessary to apprehend the effects of such complex modification scaffolds.

MATERIALS AND METHODS

Culture and tRNA isolation

Total tRNA of *M. maripaludis* and *P. furiosus* were prepared as described in de Crécy-Lagard et al. (2010). Total tRNA of *S. acidocaldarius* was obtained from a 12 L culture using the procedure described in Buck et al. (1983).

Individual tRNA purification by two-dimensional PAGE

tRNA isoacceptors were isolated using two-dimensional gel electrophoresis as previously described (Antoine et al. 2019; Antoine and Wolff 2020). Briefly, the total tRNA of each organism was separated in a first dimension gel under denaturing conditions using 12.5% polyacrylamide gel and 8 M urea, followed by a second dimension under semi-denaturing conditions using 20% polyacrylamide gel and 4 M urea at room temperature (Supplemental Fig. S1). Gel are staining with an ethidium bromide solution (10 µg.l–1) for 10 min. Spots containing tRNAs are visualized and excised under UV light (302 nm).

In-gel RNase digestion

Gel spots containing tRNAs were dried and rehydrated by 20 µL of 0.1 U/µL of RNase T1 (ThermoFisher Scientific) or by 20 µL of 0.01 U/µL of RNase A (Thermo Fisher Scientific) in 100 mM ammonium acetate (pH is not adjusted). For a few selected samples, spots were digested by RNase U2, by using 50 µL of RNase U2 (homemade) at 0.3 ng.µL^{-1} in 100 mM ammonium acetate (pH is not adjusted). The spots were incubated 4 h at 50°C. Using ZipTip C18 (Milllipore) samples were desalted by several washes with 200 mM ammonium acetate and eluted with 50% acetonitrile in milliQ water and dried under vacuum.

NanoLC-MS/MS of RNA oligonucleotides

Pellet containing RNase digestion products is resuspended in 3 µL of milliQ water and separated on an Acquity peptide BEH C18 column (130 Å, 1.7 µm, 75 µm × 200 mm) using a nanoAcquity system (Waters). The column was equilibrated in buffer A containing 7.5 mM TEAA (Triethylammonium acetate), 7.0 mM TEA (Triethylammonium) and 200 mM HFIP (Hexafluoroisopropanol) at a flow rate of 300 nL/min. Oligonucleotides were eluted using a gradient from 15% to 35% of buffer B (100% methanol) for 2 min followed by elution with an increase of buffer B to 50% in 20 min. MS and MS/MS analyses were performed using SYNAPT G2-S (quadrupole time-of-flight mass spectrometer) from Waters Corporation. All experiments were performed in negative mode with a capillary voltage set at 2.6 kV and a sample cone voltage set at 30 V. Source was heated to 130°C. The samples were analyzed over an *m/z* range from 500 to 1500 for the full scan, followed by fast data direct acquisition scan (Fast DDA).

Data analysis

All CID were deconvoluted using MassLynx software from Waters and manually sequenced by following the y and/or c series (w ions were also useful when sequencing was difficult or in order to confirm the sequence). Experimental masses of parents and fragments were compared to the theoretical masses obtained by the Mongo Oligo Mass Calculator (https://mods.rna.albany.edu/masspec/Mongo-Oligo; Cantara et al. 2011). tRNA identification was done by comparisons with the genomic sequences obtained from GtRNAdb (http://gtrnadb.ucsc.edu/; Chan and Lowe 2009, 2016). Data about nucleoside modification were obtained from Modomics (Boccaletto et al. 2018).

Data analysis by MassSpec-Toolkit for RNAs

The identification and characterization of modified tRNAs by LC-MS/MS spectrum analysis is difficult and time-consuming. To help in this process, we implemented MassSpec-Toolkit for RNAs, a Python web application (http://labex-ibmc.u-strasbg.fr/MassSpec-Toolkit/, accessible upon request) linked to a local MongoDB database that stores user-provided RNA genomic sequences and their theoretical digestion products obtained by specific ribonucleases. For each studied species, mature tRNA sequences retrieved from GtRNAdb (Chan and Lowe 2009, 2016) were submitted to the application and digested in silico with RNases T1, A and/or U2, in the "RNA Digestion" module. Genomic tRNA sequences longer than 100 nt were discarded and for the remaining ones with a nonambiguous anticodon position, fragments with common U34 or G34 and/or A37 or G37 modifications were generated when appropriate. In addition, fragment variants containing up to five additional methylations were also computed for each digestion product. Experimental data such as parent ion masses or manually reconstructed subsequences could then be compared to the theoretical ones in the "RNA Search" module. Additional criteria, like the species of interest, the ribonuclease used, as well as the presence of expected methylations or modifications at specific positions, can be specified to reduce the search space in the database. Candidate tRNAs are given a score comprised between 0 and 1 depending on the number of matching masses or subsequences they present with the list provided by the user. Besides these two main modules, the application gathers a set of "Additional Tools" under a third module that includes some of the tools present in the Mongo

Oligo Mass Calculator (Cantara et al. 2011) and "Total Mass Decipherer." The latter program can be very useful to identify modifications in case of incomplete MS/MS series, since it computes all combinations of a chosen set of (modified) nucleotides matching the mass of an RNA fragment obtained after cleavage by RNase T1.

RNase cleavage

LC MS/MS of digestion products allows the localization of methylation in the correct nucleotide but does not allow the localization on the ribose or on the base (Supplemental Table S1). To correctly assign the type of methylation, we used known modified tRNA sequences and the presence of modification enzymes in the species genome (Supplemental Table S2). To confirm the type of methylation, we also used RNase T1 and RNase A cleavage profiles. Indeed, a methyl group on the 2′ ribose protects RNA against RNase cleavages. With the methyl group on the base, the interpretation of RNase cleavages is not straightforward. Our data show that in the case of RNase T1, m^2G and $m^2{}_2G$ can be cleaved (one example for each, m^2G10 in *S. acidocaldarius* tRNA-Trp and $m^2{}_2G39$ in *P. furiosus* tRNA-Ser[GGA]) but m^7G, m^1G are never cleaved. With RNase A, m^5C and m^1Y are cleaved and Um, m^5U and m^5s^2U are not cleaved. It is interesting to note that, for tRNA-Cys and tRNA-Met of *P. furiosus* tRNAs, RNase A does not cleave m^5U54. A previous work shows that m^5U54 is always cleaved (Antoine and Wolff 2020). A possible explanation is the presence of a stretch of three to four Gs 5′ preceding the modified m^5U54 that could prevent access and binding of RNase A.

LC-MS/MS of nucleosides

Total tRNA was desalted by ethanolic precipitation with 200 mM ammonium acetate (Supplemental Figs. S2, S3). For nucleoside analysis, tRNAs are diluted to a concentration of 5 µg/µL in H_2O. Digestion was carried out in the following order: 14 µL H_2O; 2 µL buffer P1 10× (2 mM $ZnCl_2$, 250 mM NH_4OAc, pH 5.0); 21 µL of tRNA and 2 µL of P1 (0.5 U/µL). The mixture is incubated at 37°C for 2 h followed by addition of 2 µL of snake venom phosphodiesterase (0.1 U/µL) for 4 h at 37°C. After digestion, 20 µL of BAP (1.5 U/µL in 100 mM NH_4OAc) were added to the mixture. The latter was then incubated at 37°C for 2 h, dried under vacuum SpeedVac and resuspended with 100 µl of methanol. Nucleosides were analyzed by liquid chromatography coupled to mass spectrometry using an Ultimate 3000 (Thermo) chromatography coupled to an EvoQ triple quadrupole (Bruker). Separation was performed on an Acquity UPLC HSST3 column (1.8 µm, 2.1 × 100 mm, Waters) equipped with an Acquity UPLC HSST3 precolumn (1.8 µm, 2.1 × 5 mm, Waters). A gradient of solvent A (H_2O, 0.1% formic acid [Sigma Aldrich]) and solvent B (methanol [Fisher Chemicals], 0.1% formic acid [Sigma Aldrich]) was used as follows: 2% B during 2 min, 7% B at 4 min, 100% B at 12 min, hold during 1.5 min and back to 2% of B at 13.5 min, hold during 1.5 min for a total run time of 15 min. The column was operated at 35°C with a flow rate of 0.32 mL/min; 10 µL of samples were injected for each run. The triple quadrupole was used in positive ion mode, the spray voltage was set at 3500 V and cone temperature at 350°C. Nucleosides were identified using multiple reaction monitoring (MRM) with one to three transitions per nucleotide. The identifications were based on the retention time, *m/z* of the parent ion and *m/z* of the daughter ions in MS Data Review software (Bruker), with a signal-to-noise (S/N) ratio set at 10 and a search window of ±0.2 min.

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

We wish to thank the three reviewers for their constructive comments. E.W. wishes to thank G. and D. Crahay for their hospitality during the coronavirus lockdown. The project was supported by the French National Program Investissement d'Avenir (Labex NetRNA) administered by the Agence Nationale de la Recherche (ANR-10-LABX-0036_NETRNA) and ANR-15-CE11-0021-01 [to E.W.]) and by the EpiRNA funding from the Region Grand Est.

Received August 1, 2020; accepted September 9, 2020.

REFERENCES

Alfonzo JD, Blanc V, Estevez AM, Rubio MA, Simpson L. 1999. C to U editing of the anticodon of imported mitochondrial tRNATrp allows decoding of the UGA stop codon in *Leishmania tarentolae*. *EMBO J* **18:** 7056–7062. doi:10.1093/emboj/18.24.7056

Antoine L, Wolff P. 2020. Mapping of posttranscriptional tRNA modifications by two-dimensional gel electrophoresis mass spectrometry. *Methods Mol Biol* **2113:** 101–110. doi:10.1007/978-1-0716-0278-2_8

Antoine L, Wolff P, Westhof E, Romby P, Marzi S. 2019. Mapping post-transcriptional modifications in *Staphylococcus aureus* tRNAs by nanoLC/MSMS. *Biochimie* **164:** 60–69. doi:10.1016/j.biochi.2019.07.003

Auffinger P, Westhof E. 2001. Hydrophobic groups stabilize the hydration shell of 2′-*O*-methylated RNA duplexes. *Angew Chem Int Ed Engl* **40:** 4648–4650. doi:10.1002/1521-3773(20011217)40:24<4648::AID-ANIE4648>3.0.CO;2-U

Boccaletto P, Machnicka MA, Purta E, Piątkowski P, Bagiński B, Wirecki TK, de Crécy-Lagard V, Ross R, Limbach PA, Kotter A, et al. 2018. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* **46:** D303–D307. doi:10.1093/nar/gkx1030

Buck M, Connick M, Ames BN. 1983. Complete analysis of tRNA-modified nucleosides by high-performance liquid chromatography: the 29 modified nucleosides of *Salmonella typhimurium* and *Escherichia coli*. *Anal Biochem* **129:** 1–13. doi:10.1016/0003-2697(83)90044-1

Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FA, Fabris D, Agris PF. 2011. The RNA Modification Database, RNAMDB: 2011 update. *Nucleic Acids Res* **39:** D195–D201. doi:10.1093/nar/gkq1028

Cantara WA, Murphy F, Demirci H, Agris PF. 2013. Expanded use of sense codons is regulated by modified cytidines in tRNA. *Proc Natl Acad Sci* **110:** 10964–10969. doi:10.1073/pnas.1222641110

Chan PP, Lowe TM. 2009. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res* **37:** D93–D97. doi:10.1093/nar/gkn787

Chan PP, Lowe TM. 2016. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res* **44:** D184–D189. doi:10.1093/nar/gkv1309

Chan CT, Chionh YH, Ho CH, Lim KS, Babu IR, Ang E, Wenwei L, Alonso S, Dedon PC. 2011. Identification of N^6,N^6-dimethyladenosine in transfer RNA from *Mycobacterium bovis* Bacille Calmette-Guérin. *Molecules* **16:** 5168–5181. doi:10.3390/molecules16065168

Chatterjee K, Blaby IK, Thiaville PC, Majumder M, Grosjean H, Yuan YA, Gupta R, de Crécy-Lagard V. 2012. The archaeal COG1901/DUF358 SPOUT-methyltransferase members, together with pseudouridine synthase Pus10, catalyze the formation of 1-methylpseudouridine at position 54 of tRNA. *RNA* **18:** 421–433. doi:10.1261/rna.030841.111

Clouet d'Orval B, Bortolin ML, Gaspin C, Bachellerie JP. 2001. Box C/D RNA guides for the ribose methylation of archaeal tRNAs. The tRNATrp intron guides the formation of two ribose-methylated nucleosides in the mature tRNATrp. *Nucleic Acids Res* **29:** 4518–4529. doi:10.1093/nar/29.22.4518

Clouet-d'Orval B, Gaspin C, Mougin A. 2005. Two different mechanisms for tRNA ribose methylation in Archaea: a short survey. *Biochimie* **87:** 889–895. doi:10.1016/j.biochi.2005.02.004

Crain PF, Alfonzo JD, Rozenski J, Kapushoc ST, Mccloskey JA, Simpson L. 2002. Modification of the universally unmodified uridine-33 in a mitochondria-imported edited tRNA and the role of the anticodon arm structure on editing efficiency. *RNA* **8:** 752–776. doi:10.1017/S1355838202022045

Dalluge JJ, Hamamoto T, Horikoshi K, Morita RY, Stetter KO, McCloskey JA. 1997. Posttranscriptional modification of tRNA in psychrophilic bacteria. *J Bacteriol* **179:** 1918–1923. doi:10.1128/JB.179.6.1918-1923.1997

Davis DR. 1995. Stabilization of RNA stacking by pseudouridine. *Nucleic Acids Res* **23:** 5020–5026. doi:10.1093/nar/23.24.5020

de Crécy-Lagard V, Brochier-Armanet C, Urbonavicius J, Fernandez B, Phillips G, Lyons B, Noma A, Alvarez S, Droogmans L, Armengaud J, et al. 2010. Biosynthesis of wyosine derivatives in tRNA: an ancient and highly diverse pathway in archaea. *Mol Biol Evol* **27:** 2062–2077. doi:10.1093/molbev/msq096

Dennis PP, Omer A, Lowe T. 2001. A guided tour: small RNA function in Archaea. *Mol Microbiol* **40:** 509–519. doi:10.1046/j.1365-2958.2001.02381.x

Edmonds CG, Crain PF, Gupta R, Hashizume T, Hocart CH, Kowalak JA, Pomerantz SC, Stetter KO, McCloskey JA. 1991. Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria). *J Bacteriol* **173:** 3138–3148. doi:10.1128/JB.173.10.3138-3148.1991

Emery LR, Sharp PM. 2011. Impact of translational selection on codon usage bias in the archaeon *Methanococcus maripaludis*. *Biol Lett* **7:** 131–135. doi:10.1098/rsbl.2010.0620

Forterre P. 2015. The universal tree of life: an update. *Front Microbiol* **6:** 717. doi:10.3389/fmicb.2015.00717

Fu Y, Jia G, Pang X, Wang RN, Wang X, Li CJ, Smemo S, Dai Q, Bailey KA, Nobrega MA, et al. 2013. FTO-mediated formation of N^6-hydroxymethyladenosine and N^6-formyladenosine in mammalian RNA. *Nat Commun* **4:** 1798. doi:10.1038/ncomms2822

Fukunaga R, Yokoyama S. 2007. Structural insights into the first step of RNA-dependent cysteine biosynthesis in archaea. *Nat Struct Mol Biol* **14:** 272–279. doi:10.1038/nsmb1219

Gaspin C, Cavaillé J, Erauso G, Bachellerie JP. 2000. Archaeal homologs of eukaryotic methylation guide small nucleolar RNAs: lessons from the *Pyrococcus* genomes. *J Mol Biol* **297:** 895–906. doi:10.1006/jmbi.2000.3593

Giegé R, Springer M. 2016. Aminoacyl-tRNA synthetases in the bacterial world. *EcoSal Plus* **7**. doi:10.1128/ecosalplus.ESP-0002-2016

Golovina AY, Sergiev PV, Golovin AV, Serebryakova MV, Demina I, Govorun VM, Dontsova OA. 2009. The *yfiC* gene of *E. coli* encodes an adenine-N^6 methyltransferase that specifically modifies A37 of tRNA1Val(cmo5UAC). *RNA* **15:** 1134–1141. doi:10.1261/rna.1494409

Grosjean H, Westhof E. 2016. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res* **44:** 8020–8040. doi:10.1093/nar/gkw608

Grosjean H, Constantinesco F, Foiret D, Benachenhou N. 1995. A novel enzymatic pathway leading to 1-methylinosine modification in *Haloferax volcanii* tRNA. *Nucleic Acids Res* **23:** 4312–4319. doi:10.1093/nar/23.21.4312

Grosjean H, Gaspin C, Marck C, Decatur WA, de Crécy-Lagard V. 2008a. RNomics and Modomics in the halophilic archaea *Haloferax volcanii*: identification of RNA modification genes. *BMC Genomics* **9:** 470–496. doi:10.1186/1471-2164-9-470

Grosjean H, Gupta R, Maxwell ES. 2008b. Modified nucleotides in Archaeal RNAs. In *Archaea, new models for prokaryotic biology* (ed. Blum P), pp. 171–196. Caister Academic Press, Norfolk, UK.

Grosjean H, de Crécy-Lagard V, Marck C. 2010. Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. *FEBS Lett* **584:** 252–264. doi:10.1016/j.febslet.2009.11.052

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.

Gupta R. 1986. Transfer RNAs of *Halobacterium volcanii*: sequences of five leucine and three serine tRNAs. *Syst Appl Microbiol* **7:** 102–105. doi:10.1016/S0723-2020(86)80131-X

Haag S, Warda AS, Kretschmer J, Günnigmann MA, Höbartner C, Bohnsack MT. 2015. NSUN6 is a human RNA methyltransferase that catalyzes formation of m^5C72 in specific tRNAs. *RNA* **21:** 1532–1543. doi:10.1261/rna.051524.115

Helm M. 2006. Post-transcriptional nucleotide modification and alternative folding of RNA. *Nucleic Acids Res* **34:** 721–733. doi:10.1093/nar/gkj471

Helm M, Alfonso JD. 2014. Posttranscriptional RNA modifications: playing metabolic games in a cell's chemical legoland. *Chem Biol* **21:** 174–185. doi:10.1016/j.chembiol.2013.10.015

Hirata A, Suzuki T, Nagano T, Fujii D, Okamoto M, Sora M, Lowe TM, Kanai T, Atomi H, Suzuki T, et al. 2019. Distinct modified nucleosides in tRNATrp from the hyperthermophilic archaeon *Thermococcus kodakarensis* and requirement of tRNA m2G10/m22G10 methyltransferase (Archaeal Trm11) for survival at high temperatures. *J Bacteriol* **201:** e00448-19. doi:10.1128/JB.00448-19

Hori H, Kawamura T, Awai T, Ochi A, Yamagami R, Tomikawa C, Hirata A. 2018. Transfer RNA modification enzymes from thermophiles and their modified nucleosides in tRNA. *Microorganisms* **6:** 110. doi:10.3390/microorganisms6040110

Ikeuchi Y, Kimura S, Numata T, Nakamura D, Yokogawa T, Ogata T, Wada T, Suzuki T, Suzuki T. 2010. Agmatine-conjugated cytidine in a tRNA anticodon is essential for AUA decoding in archaea. *Nat Chem Biol* **6:** 277–282. doi:10.1038/nchembio.323

Jühling F, Mörl M, Hartmann RK, Sprinzl M, Stadler PF, Pütz J. 2009. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res* **37:** D159–D162. doi:10.1093/nar/gkn772

Karlsborn T, Tükenmez H, Mahmud AKMF, Xu F, Xu H, Byström AS. 2014. Elongator, a conserved complex required for wobble uridine modifications in eukaryotes. *RNA Biol* **11:** 1519–1528. doi:10.4161/15476286.2014.992276

Kawai G, Yamamoto Y, Kamimura T, Masegi T, Sekine M, Hata T, Iimori T, Watanabe T, Miyazawa T, Yokoyama S. 1992. Conformational rigidity of specific pyrimidine residues in tRNA arises from posttranscriptional modifications that enhance steric interaction between the base and the 2′-hydroxyl group. *Biochemistry* **31:** 1040–1046. doi:10.1021/bi00119a012

Kempenaers M, Roovers M, Oudjama Y, Tkaczuk KL, Bujnicki JM, Droogmans L. 2010. New archaeal methyltransferases forming 1-methyladenosine or 1-methyladenosine and 1-methylguanosine at position 9 of tRNA. *Nucleic Acids Res* **38:** 6533–6543. doi:10.1093/nar/gkq451

Kowalak JA, Dalluge JJ, McCloskey JA, Stetter KO. 1994. The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry* **33:** 7869–7876. doi:10.1021/bi00191a014

Krishnamohan A, Samantha Dodbele S, Jackman JE. 2019. Insights into catalytic and tRNA recognition mechanism of the dual-specific tRNA methyltransferase from *Thermococcus kodakarensis*. *Genes (Basel)* **10:** 100. doi:10.3390/genes10020100

Kuchino Y, Ihara M, Yabusaki Y, Nishimura S. 1982. Initiator tRNAs from archaebacteria show common unique sequence characteristics. *Nature* **298:** 684–685. doi:10.1038/298684a0

Larsen AT, Fahrenbach AC, Sheng J, Pian J, Szostak JW. 2015. Thermodynamic insights into 2-thiouridine-enhanced RNA hybridization. *Nucleic Acids Res* **43:** 7675–7687. doi:10.1093/nar/gkv761

Lorenz C, Lünse CE, Mörl M. 2017. tRNA modifications: impact on structure and thermal adaptation. *Biomolecules* **7:** 35. doi:10.3390/biom7020035

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25:** 955–964. doi:10.1093/nar/25.5.955

Machnicka MA, Olchowik A, Grosjean H, Bujnicki JM. 2014. Distribution and frequencies of post-transcriptional modifications in tRNAs. *RNA Biol* **11:** 1619–1629. doi:10.4161/15476286.2014.992273

Mandal D, Köhrer C, Su D, Russell SP, Krivos K, Castleberry CM, Blum P, Limbach PA, Söll D, RajBhandary UL. 2010. Agmatidine, a modified cytidine in the anticodon of archaeal $tRNA^{Ile}$, base pairs with adenosine but not with guanosine. *Proc Natl Acad Sci* **107:** 2872–2877. doi:10.1073/pnas.0914869107

Mandal D, Köhrer C, Su D, Babu IR, Chan CT, Liu Y, Söll D, Blum P, Kuwahara M, Dedon PC, et al. 2014. Identification and codon reading properties of 5-cyanomethyl uridine, a new modified nucleoside found in the anticodon wobble position of mutant haloarchaeal isoleucine tRNAs. *RNA* **20:** 177–188. doi:10.1261/rna.042358.113

Mao S, Ranganathan SV, Tsai HC, Haruehanroengra P, Shen F, Valsangkar VA, Han B, Hassan AEA, Chen A, Sheng J. 2018. Cyano modification on uridine decreases base-pairing stability and specificity through neighboring disruption in RNA duplex. *Chembiochem* **19:** 2558–2565. doi:10.1002/cbic.201800399

Marck C, Grosjean H. 2002. tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* **8:** 1189–1232. doi:10.1017/S1355838202022021

McCloskey JA, Graham DE, Zhou S, Crain PF, Ibba M, Konisky J, Söll D, Olsen GJ. 2001. Post-transcriptional modification in archaeal tRNAs: identities and phylogenetic relations of nucleotides from mesophilic and hyperthermophilic Methanococcales. *Nucleic Acids Res* **29:** 4699–4706. doi:10.1093/nar/29.22.4699

Motorin Y, Helm M. 2010. tRNA stabilization by modified nucleotides. *Biochemistry* **49:** 4934–4944. doi:10.1021/bi100408z

Motorin Y, Helm M. 2011. RNA nucleotide methylation. *WIREs RNA* **2:** 611–631. doi:10.1002/wrna.79

Nakamura Y, Gojobori T, Ikemura T. 2000. Codon usage tabulated from the international DNA sequence databases: status for the year 2000. *Nucleic Acids Res* **28:** 292. doi:10.1093/nar/28.1.292

Nayak KC. 2013. Comparative genome sequence analysis of *Sulfolobus acidocaldarius* and 9 other isolates of its genus for factors influencing codon and amino acid usage. *Gene* **513:** 163–173. doi:10.1016/j.gene.2012.10.024

Noon KR, Guymon R, Crain PF, McCloskey JA, Thomm M, Lim J, Cavicchioli R. 2003. Influence of temperature on tRNA modification in archaea: *Methanococcoides burtonii* (optimum growth temperature [Topt], 23 degrees C) and Stetteria hydrogenophila (Topt, 95 degrees C). *J. Bacteriol* **185:** 5483–5490. doi:10.1128/JB.185.18.5483-5490.2003

Ohira T, Suzuki T. 2016. Precursors of tRNAs are stabilized by methylguanosine cap structures. *Nat Chem Biol* **12:** 648–655. doi:10.1038/nchembio.2117

Omer AD, Lowe TM, Russell AG, Ebhardt H, Eddy SR, Dennis PP. 2000. Homologs of small nucleolar RNAs in Archaea. *Science* **288:** 517–522. doi:10.1126/science.288.5465.517

Pang H, Ihara M, Kuchino Y, Nishimura S, Gupta R, Woese CR, McCloskey JA. 1982. Structure of a modified nucleoside in archaebacterial tRNA which replaces ribosylthymine 1-methylpseudouridine. *J Biol Chem* **257:** 3589–3592.

Plesiewicz E, Stepień E, Bolewska K, Wierzchowski KL. 1976. Stacking self-association of pyrimidine nucleosides and of cytosines: effects of methylation and thiolation. *Nucleic Acids Res* **3:** 1295–1306. doi:10.1093/nar/3.5.1295

Pollo-Oliveira L, de Crécy-Lagard V. 2019. Can protein expression be regulated by modulation of tRNA modifications profiles? *Biochemistry* **58:** 355–362. doi:10.1021/acs.biochem.8b01035

Renalier MH, Joseph N, Gaspin C, Thebault P, Mougin A. 2005. The Cm56 tRNA modification is catalyzed either by a specific 2′-*O*-methylase, or a C/D sRNP in Archaea. *RNA* **11:** 1051–1063. doi:10.1261/rna.2110805

Rose S, Auxilien S, Havelund JF, Kirpekar F, Huber H, Grosjean H, Douthwaite S. 2020. The hyperthermophilic partners *Nanoarchaeum* and *Ignicoccus* stabilize their tRNA T-loops via different but structurally equivalent modifications. *Nucleic Acids Res* **48:** 6906–6918. doi:10.1093/nar/gkaa411

Rozov A, Demeshkina N, Khusainov I, Westhof E, Yusupov M, Yusupova G. 2016. Novel base-pairing interactions at the tRNA wobble position crucial for accurate reading of the genetic code. *Nat Commun* **7:** 10457. doi:10.1038/ncomms10457

Sas-Chen A, Thomas JM, Matzov D, Taoka M, Nance KD, Nir R, Bryson KM, Shachar R, Liman GLS, Burkhart BW, et al. 2020. Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. *Nature* **583:** 638–643. doi:10.1038/s41586-020-2418-2

Sauerwald A, Sitaramaiah D, McCloskey JA, Söll D, Crain PF. 2005. N^6-acetyladenosine: a new modified nucleoside from *Methanopyrus kandleri* tRNA. *FEBS Lett* **579:** 2807–2810. doi:10.1016/j.febslet.2005.04.014

Selmer M, Dunham CM, Murphy FV, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313:** 1935–1942. doi:10.1126/science.1131127

Shigi N, Suzuki T, Terada T, Shirouzu M, Yokoyama S, Watanabe K. 2006. Temperature-dependent biosynthesis of 2-thioribothymidine of *Thermus thermophilus* tRNA. *J Biol Chem* **281:** 2104–2113. doi:10.1074/jbc.M510771200

Sugahara J, Kikuta K, Fujishima K, Yachie N, Tomita M, Kanai A. 2008. Comprehensive analysis of archaeal tRNA genes reveals rapid increase of tRNA introns in the order thermoproteales. *Mol Biol Evol* **25:** 2709–2716. doi:10.1093/molbev/msn216

Takakura M, Ishiguro K, Akichika S, Miyauchi K, Suzuki T. 2019. Biogenesis and functions of aminocarboxypropyluridine in tRNA. *Nat Commun* **10:** 5542. doi:10.1038/s41467-019-13525-3

Tang TH, Polacek N, Zywicki M, Huber H, Brugger K, Garrett R, Bachellerie JP, Hüttenhofer A. 2005. Identification of novel non-coding RNAs as potential antisense regulators in the archaeon

Sulfolobus solfataricus. *Mol Microbiol* **55:** 469–481. doi:10.1111/j.1365-2958.2004.04428.x

Tomikawa C, Ohira T, Inoue Y, Kawamura T, Yamagishi A, Suzuki T, Hori H. 2013. Distinct tRNA modifications in the thermo-acidophilic archaeon, *Thermoplasma acidophilum*. *FEBS Lett* **587:** 3575–3580. doi:10.1016/j.febslet.2013.09.021

Urbonavicius J, Armengaud J, 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. doi:10.1016/j.jmb.2005.12.087

Urbonavicius J, Auxilien S, Walbott H, Trachana K, Golinelli-Pimpaneau B, Brochier-Armanet C, Grosjean H. 2008. Acquisition of a bacterial RumA-type tRNA(uracil-54, C5)-methyltransferase by Archaea through an ancient horizontal gene transfer. *Mol Microbiol* **67:** 323–335. doi:10.1111/j.1365-2958.2007.06047.x

Vendeix FA, Murphy F, Cantara WA, Leszczynska G, Gustilo EM, Sproat B, Malkiewicz A, Agris PF. 2012. Human tRNA(Lys3)(UUU) is pre-structured by natural modifications for cognate and wobble codon binding through keto-enol tautomerism. *J Mol Biol* **416:** 467–485. doi:10.1016/j.jmb.2011.12.048

Wagner TM, Nair V, Guymon R, Pomerantz SC, Crain PF, Davis DR, McCloskey JA. 2004. A novel method for sequence placement of modified nucleotides in mixtures of transfer RNA. *Nucleic Acids Symp Ser* **48:** 263–264. doi:10.1093/nass/48.1.263

Watanabe M, Matsuo M, Tanaka S, Akimoto H, Asahi S, Nishimura S, Katze JR, Hashizume T, Crain PF, McCloskey JA, et al. 1997. Biosynthesis of archaeosine, a novel derivative of 7-deazaguanosine specific to archaeal tRNA, proceeds via a pathway involving base replacement on the tRNA polynucleotide chain. *J Biol Chem* **272:** 20146–20151. doi:10.1074/jbc.272.32.20146

Watson ZL, Ward FR, Méheust R, Ad O, Schepartz A, Banfield JF, Cate JHD. 2020. Structure of the bacterial ribosome at 2 Å resolution. *eLife* **9:** e60482. doi:10.7554/eLife.60482

Westhof E, Yusupov M, Yusupova G. 2019. The multiple flavors of GoU pairs in RNA. *J Mol Recognit* **32:** e2782. doi:10.1002/jmr.2782

Wintermeyer W, Zachau HG. 1975. Tertiary structure interactions of 7-methylguanosine in yeast $tRNA^{Phe}$ as studied by borohydride reduction. *FEBS Lett* **58:** 306–309. doi:10.1016/0014-5793(75)80285-7

Yarus M. 1982. Translational efficiency of transfer RNA's: uses of an extended anticodon. *Science* **218:** 646–652. doi:10.1126/science.6753149

Yu N, Jora M, Solivio B, Thakur P, Acevedo-Rocha CG, Randau L, de Crécy-Lagard V, Addepalli B, Limbach PA. 2019. tRNA modification profiles and codon-decoding strategies in *Methanocaldococcus jannaschii*. *J Bacteriol* **201:** e00690-18. doi:10.1128/JB.00690-18

Zago MA, Dennis PP, Omer AD. 2005. The expanding world of small RNAs in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Mol Microbiol* **55:** 1812–1828. doi:10.1111/j.1365-2958.2005.04505.x

Ziesche SM, Omer AD, Dennis PP. 2004. RNA-guided nucleotide modification of ribosomal and non-ribosomal RNAs in Archaea. *Mol Microbiol* **54:** 980–993. doi:10.1111/j.1365-2958.2004.04319.x