HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2025 May 30.

Published in final edited form as:

Cell Rep. 2025 April 22; 44(4): 115471. doi:10.1016/j.celrep.2025.115471.

Comprehensive nucleoside analysis of archaeal RNA modification profiles reveals an m^{7}G in the conserved P loop of 23S rRNA

Yueh-Lin Tsai[1], **Eric J. Wolf**[1], **Kristin A. Fluke**[2], **Ryan T. Fuchs**[1], **Nan Dai**[1], **Sean R. Johnson**[1], **Zhiyi Sun**[1], **Liam Elkins**[2], **Brett W. Burkhart**[2], **Thomas J. Santangelo**[2], **Ivan R. Corrêa Jr.**[1,3,*]

[1]New England Biolabs Inc., Beverly, MA 01915, USA

[2]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523, USA

[3]Lead contact

SUMMARY

Extremophilic Archaea employ diverse RNA modifications for survival. Our understanding of the modified nucleosides and their functions in Archaea is far from complete. Here, we establish an extensive profile of nucleoside modifications in thermophilic and mesophilic Archaea. Through liquid chromatography-tandem mass spectrometry (LC-MS/MS) and rigorous non-coding RNA depletion, we identify four previously unannotated modifications in archaeal mRNA. Nucleoside analysis conducted on total, large, small, and mRNA-enriched subfractions of hyperthermophile *Thermococcus kodakarensis* reveals modifications whose relative abundance is dynamically responsive to growth temperatures. To predict archaeal RNA-modifying enzymes, we leverage open-access databases to compare putative functional domains with previously annotated enzymes. Our approach leads to the discovery of a methyltransferase responsible for the installation of m^{7}G in the P loop of 23S rRNA peptidyl transferase center in *T. kodakarensis*. The methyltransferase activity is confirmed *in vitro* with synthetic substrates and *in vivo* by assessing the presence of the m^{7}G modification in a genetic deletion strain.

Graphical Abstract

*Correspondence: correa@neb.com.

AUTHOR CONTRIBUTIONS

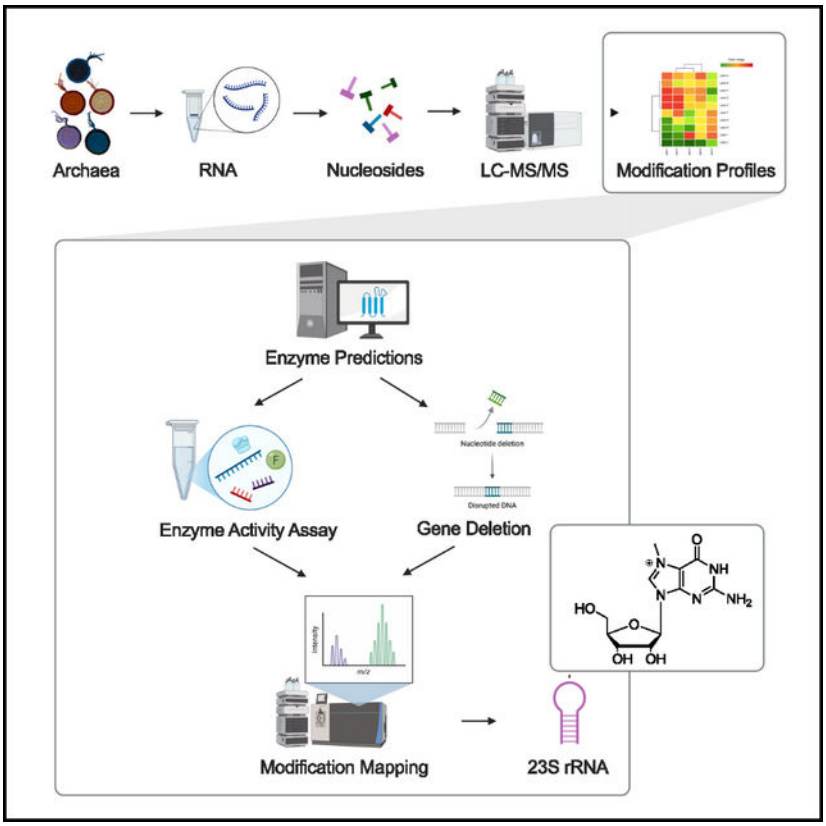
Y.-L.T. and I.R.C. conceptualized the study and wrote the manuscript. I.R.C. and T.J.S. provided input and edited the manuscript. Y.-L.T., E.J.W., K.A.F., R.T.F., N.D., and B.W.B. conducted the experiments. Y.-L.T. performed primary data analysis and interpretation with assistance from E.J.W. and N.D. S.R.J. and Z.S. conducted bioinformatics analysis. L.E. performed protein purification.

DECLARATION OF INTERESTS

Y.-L.T., R.T.F., E.J.W., N.D., S.R.J., Z.S., and I.R.C. are employees of New England Biolabs, Inc. New England Biolabs is a manufacturer and vendor of molecular biology reagents. The authors declare that this affiliation does not affect the authors' impartiality, adherence to journal standards and policies, or availability of data.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2025.115471.



In brief

Tsai et al. describe comprehensive nucleoside profiles of archaeal RNAs and identify temperature-responsive modifications using LC-MS/MS. They establish a workflow for RNA-modifying enzyme discovery by associating enzyme predictions with global nucleoside profiles. This approach reveals a methyltransferase responsible for installing an m^7G in the peptidyl transferase center of 23S rRNA.

INTRODUCTION

Many Archaea thrive in extreme environments on Earth that present very real challenges in maintaining RNA stability and function. A growing body of evidence highlights the importance of chemical modifications introduced post-transcriptionally by specialized RNA-modifying enzymes in conferring thermal stability to RNA in thermophilic Archaea.[1–8] Across the three domains, >170 RNA modifications have been identified, bespeaking to their chemical and functional diversity. Some modifications modulate base-pairing capacity, while others create versatile surfaces for protein interactions.[9–11] The modification landscape is complex with RNA modifications unevenly distributed among RNA classes and many sites of modification occurring substoichiometrically, presenting a formidable challenge for unambiguous detection and quantification. Finally, our incomplete knowledge of the enzymes responsible for installing these modifications further hinders a deeper understanding of their functional roles.

To realize the impact of RNA modifications on the survival of Archaea in extreme environments, it is essential to build a comprehensive catalog of the modified nucleosides generated within these organisms. Nucleoside profiles can serve as roadmaps to reveal the presence or absence of specific RNA-modifying enzymes and to identify modifications that respond to environmental cues. While significant efforts have been made to characterizing nucleosides in archaeal RNA over the past decades,[12–17] the dynamic changes in nucleoside levels in response to varying environmental conditions remain poorly understood. Historically, ribosomal RNA (rRNA) and transfer RNA (tRNA) have been the main subjects for archaeal RNA modification studies owing to their high cellular abundance and dense modification status. In contrast, profiling modifications in mRNA poses additional challenges due to perceived lower abundance, sparse modification density, and the requirement for rigorous non-coding RNA (ncRNA) depletion; to date, only ac^4C and m^5C have been reported in archaeal mRNA.[3,7,18] In light of the diversity of modifications that are present in extremophilic archaeal ncRNAs, a comprehensive survey of the mRNA modification may provide further insights into archaeal adaptive mechanisms.

Next-generation sequencing (NGS) and mass spectrometry (MS) are the primary methods for detecting RNA modifications.[19–21] NGS-based approaches offer several advantages, including low sample input requirements, high-resolution modification mapping, and wide accessibility across various sequencing platforms. However, NGS methods often necessitate nucleoside conversions through chemical or enzymatic treatments to generate distinct readouts at modified positions. Unfortunately, these pretreatments carry the risk of incomplete conversions, potentially leading to false-positive discoveries.[22] Some pretreatment protocols subject RNA to high pH and temperatures, resulting in significant RNA degradation and skewed data distribution.[23,24] Moreover, most pretreatments are tailored for specific modifications, making it difficult to detect multiple modifications simultaneously.[25] Nanopore direct RNA sequencing (RNA-seq) is an emerging technology that has been applied to RNA modification detection by distinguishing electrical signals derived from modified and unmodified nucleosides.[26] Modified nucleosides often induce discernible shifts in ion current signals that lead to base-calling errors. Although computational pipelines have been developed to analyze direct RNA-seq data,[26,27] effectively differentiating base-calling errors arising from modifications versus inherent system noise remains a challenge.[26] Modification detection based on ion current shifts requires extensive training of base-calling models and is highly dependent on the quality of the training data.[27] Liquid chromatography-tandem MS (LC-MS/MS), in contrast, enables direct detection of virtually any RNA modification without cDNA amplification or nucleoside conversion.[28] In one approach, RNA samples are completely digested to nucleosides and subjected to highly sensitive triple-quadrupole MS for nucleoside identification and quantification.[29] A limitation of nucleoside analysis is the complete loss of sequence contexts at the modified positions. The combination of MS-based identification of modified nucleosides while retaining some sequence context is provided by the complementary technique of oligonucleotide LC-MS/MS analysis, which allows inference of the modified positions by analyzing the masses of short RNA fragments derived from partial digestion of longer RNA sequences.[29–32]

Here, we comprehensively investigated RNA modification profiles across five archaeal species using LC-MS/MS. Our analysis confidently identified six RNA modifications (m^1A, m^1I, ac^4C, m^5C, Cm, and Gm) within mRNA-enriched RNA samples. Two-thirds of these modifications, namely m^1A, m^1I, Cm, and Gm, have not been previously reported in archaeal mRNA. Total RNA nucleoside analysis revealed that the patterns of modification in these archaeal species correlate well with their phylogenetic distances calculated through 16S comparisons. Additionally, we uncovered a subgroup of modifications that exhibit temperature-dependent prevalence in the model hyperthermophilic archaeon *Thermococcus kodakarensis*. Finally, we utilized nucleoside profiling to guide bioinformatic prediction of putative RNA-modifying enzymes, leading to the discovery and confirmation of a methyltransferase-encoded gene responsible for installing the m^7G modification in 23S rRNA. Taken together, our results establish the totality of the modification landscape in different archaeal species and delineates a roadmap for future RNA modification enzyme discovery.

RESULTS

Global nucleoside profiling of Archaea and non-archaeal species

Ultra-high-performance liquid chromatography coupled with triple-quadrupole MS (UHPLC-QqQ MS) is a useful technique for detecting diverse set of modified nucleosides in biological samples.[33–36] Each modified nucleoside can be uniquely identified by its specific combination of retention time and mass transition. To determine the presence of modified nucleosides, we extracted total RNA from five different archaeal species (*T. kodakarensis*, *Thermococcus* sp. AM4, *Sulfolobus islandicus* M16.4, *Sulfolobus acidocaldarius*, and *Methanococcus maripaludis*), digested total RNA into nucleosides, and subsequently conducted UHPLC-QqQ analysis. We examined the retention times and mass transitions of the digested RNA samples against 76 nucleoside standards (see Data S1, Table S9, and STAR Methods), and observed 22–32 modified nucleosides with unique modified nucleoside patterns for each species (Figure 1A). 16 modified nucleosides (Cm, Gm, m^1A, m^1G, m^1I, $m^1\Psi$, $m^2{}_2G$, m^2G, m^5C, m^5U, m^6 A, m^6A, I, s^4U, Um, and Ψ) were identified within all tested archaeal species, while another five modifications (Am, ac^4C, ac^4Cm, m^3U, and m^7G) were shared by the thermophilic species (*Sulfolobus* and *Thermococcus* species) to the exclusion of *M. maripaludis*. Of the remaining modifications, four were found exclusively in the *Thermococcus* species (m^2Gm, $m^{2,7}G$, $m^4{}_2C$, and m^5s^2U), six in the *Sulfolobus* species (m^3C, $mcm5s^2U$, mcm5 U, mcm5 Um, s^2Um, and t^6A), and two in the mesophilic archaeon *M. maripaludis* ($mnm5s^2U$ and ncm5 U) (Table 1).

Species within the same genus shared the most similar nucleoside profiles, resulting in closer branch distances when we conducted a hierarchical clustering analysis based on the presence of modified nucleosides within each species (Figure 1B). Beyond the genus level, the branch of *M. maripaludis* was found to be closer to that of *Thermococcus* species than to that of *Sulfolobus* species, aligning well with their evolutionary distances as determined by 16S rRNA sequence variations (Figure 1C). We further investigated whether the tree topology of modified nucleoside profiles also correlated with species outside the Archaea. Along with the five archaeal species, we analyzed the total RNA nucleosides of human,

Saccharomyces cerevisiae, and *Escherichia coli* and reperformed a hierarchical clustering analysis (Figures 1A and 1B). Not surprisingly, non-archaeal species formed a separate cluster, with human and *S. cerevisiae* grouped into the same clade (Figure 1B). These findings suggest that RNA modification profiles can reflect relationships between species[14] and further support the notion that Archaea are chemotaxonomically distinct from Bacteria and Eukarya.

Nucleoside analysis of RNA modifications in mRNA-enriched fractions

rRNA and tRNA constitute more than 95% of the total RNA mass. To specifically detect modified nucleosides in mRNA, we employed a stringent workflow to deplete the majority of the ncRNA species (Figure S1A). Customized DNA probes were designed for each of the archaeal species to deplete their rRNA content by RNase H digestion. Short cellular RNAs, such as tRNAs and small regulatory RNAs (sRNAs) as well as RNA fragments less than 200 nucleotides (nt), were removed with size-exclusion spin columns. Illumina sequencing of the cDNAs resulting mRNA-enriched samples revealed that over 98% of reads in each archaeon mapped to protein-coding genes (Figure 2A; Table S1). Bioanalyzer traces confirmed the exclusion of the small RNA fraction from archaeal mRNA preparations (Figure S1B).

Given the unknown distribution of modified nucleoside in stable (rRNA and tRNA) compared to mRNA, we next sought to determine the identity and relative abundance of the nucleoside modifications retained in the mRNA-enriched fraction. Given that the mRNA modification density is generally lower than that of ncRNAs, we conducted a scaled-up rRNA depletion starting with 50 μg of the large RNA fraction (>200 nt). The resulting mRNA-enriched samples were digested to single nucleosides and ≥100 ng of digested mRNA was injected per analysis. Since UHPLC-QqQ MS reliably detects nucleosides above 1-fmol level, we set an arbitrary threshold for qualitative mRNA modification screening using the mass responses of 1 fmol of each canonical nucleoside, inosine, and pseudouridine standards. Accordingly, 1 fmol of a modified nucleoside in 100 ng of digested RNA corresponds to one modification occurring at a single position at ~13-ppm frequency (assuming an average archaeal mRNA length of 1000 nt with an equal proportion of A, U, G, and C). Our qualitative screening narrowed the initial 22–32 modifications identified within total RNA down to 12–20 modifications within the mRNA-enriched fractions of the five archaeal species (Table S2).

Despite the stringent ncRNA depletion, modified nucleosides originating from residual rRNA and tRNA could still be detected due to the high input sample injection. For quantitative analysis, 1–5 ng of digested mRNA was injected onto the UHPLC-QqQ MS, which falls well within the quantitative range of the mass detector. To confidently select mRNA modifications, we established a cutoff criterion based on the mass response over background monitored at the same mass transition window for each nucleoside (see STAR Methods section). Our analysis identified with high confidence six modifications in the archaeal mRNA: m^1A, m^1I, ac^4C, m^5C, Cm, and Gm (Figures 2B and 2C; Table S3). The relative abundance of each modified nucleoside ranged from 0.02% to 0.76% (expressed differently, ~1/5,000 to ~1/130 of each canonical nucleoside was observed as a modified nucleoside). The mesophilic archaeon *M. maripaludis* exhibited the lowest mRNA

modification level with only three modifications detected (m^1A, m^5C, and Cm) and all below 0.1%. Interestingly, m^1A, m^5C, and Cm were consistently present in mRNA-enriched samples of all tested archaeal species. A low, but high-confidence, level of m^1I (0.083% ± 0.002%) was observed for *S. acidocaldarius*. m^1I is a common tRNA modification in Archaea and eukaryotes.[15] To our knowledge, this is the first time m^1I has been identified in the mRNA of any species. Overall, we found that archaeal mRNA modifications overlap with those of eukaryotic mRNAs, with the exception of m^1I, and are markedly distinct from bacterial mRNA modifications (Figure 2D).[35,37–41] Although our report sheds new light on archaeal modifications, it is important to note that our cumulative knowledge of RNA modification across the three domains is continually evolving with the development of RNA isolation techniques and more sensitive and accurate detection methods.

Identification of temperature-sensitive modifications in *T. kodakarensis*

Dynamic changes in the abundance of several RNA modifications in response to temperature appear to represent fitness-relevant adaptation in some hyperthermophilic Archaea.[3,13,42] We hypothesized that temperature-responsive RNA modifications might display differential abundances at different growth temperatures. Here, we sought to identify RNA modifications whose bulk abundance responds positively or negatively to temperature changes. *T. kodakarensis*, a hyperthermophilic archaeon with a known complex RNA modification profile, thrives within a growth temperature range of ~50°C–~98°C, with a temperature optimum at 85°C. We cultivated *T. kodakarensis* in three biological replicates at temperatures of 65°C, 75°C, 85°C, and 95°C and harvested cells during the exponential growth phase. Total RNA was subsequently extracted and digested to nucleosides for UHPLC-QqQ analysis. Modification levels were assessed by normalizing mass response of each modified nucleoside to the mass response of the corresponding unmodified nucleoside (e.g., m^5C/C). The normalized value represented the mass response ratio. Temperature-dependent fold changes for each modification were calculated by comparing the mass response ratio at 65°C, 75°C, or 95°C to that at 85°C.

Among the 27 nucleoside modifications identified in *T. kodakarensis* total RNA from cultures grown at 85°C, more than a third (~37% or 10 modifications) displayed significant changes in relative abundance as the growth temperature deviated from the optimal 85°C (Figures 3A and 3B; Table S4). Of those, three modifications, ac^4C, ac^4Cm, and m^5U, exceeded a 2-fold change when *T. kodakarensis* was cultured at suboptimal temperatures (Figure 3B). ac^4C and ac^4Cm increased significantly with rising temperatures, consistent with previous observations in *T. kodakarensis* and in another hyperthermophilic archaeon, *Pyrococcus furiosus*.[3,42] Interestingly, ac^4C levels continued to rise at 95°C, whereas ac^4Cm levels remained stable after 85°C (Figure 3B). In contrast, m^5U levels substantially decreased with the increase in temperatures, showing strong accumulation at 65°C. The negative correlation between growth temperature and m^5U level was also observed in *P. furiosus* cultured at 70°C, 85°C, and 100°C[42]

To examine whether the differential regulation of RNA modifications was specific to any RNA subgroup(s), we separated total RNA recovered from *T. kodakarensis* cultures grown at each temperature into large (>200 nt), small (<200 nt), and mRNA-enriched fractions

prior to quantifying the abundance of modified nucleosides (Figures 3A, 3C–3E). Am, Gm, $m^1\Psi$, $m^{2,7}G$, m^3U, $m^3\Psi$, m^4 C, m^7G, and Um displayed higher modification fold changes in the large RNA fraction than in the small RNA fraction (Figure S2A), suggesting that these modifications occur at higher frequency in large RNAs. Cm, m^5C, m^5U, $m^6{}_2A$, and Ψ were also detected in the large RNA fraction despite being predominantly found in small RNAs. Of all modifications detected in large RNA, seven showed significant changes in response to the growth temperature, but only ac^4C exceeded a 2-fold difference; this was not only in the large RNA but also in all other RNA fractions (Figures 3C–3E). Changes in m^5U levels were primarily seen in the small RNA fraction (Figure 3D), supporting the likely presence in m^5U predominately in tRNAs. Several modifications that were barely detectable in total RNA, including 5-methyl-2′-*O*-methylcytidine (m^5Cm), isowyosine (imG2), and 4-demethylwyosine (imG14), displayed a stronger mass response signal in the small RNA fraction (Figure S2B), again supportive of these modifications being enriched in tRNAs. In addition, despite m^5Cm being present at low abundance in the small RNA fraction, its mass response ratio exceeded a 2-fold difference in response to varying growth temperature (Figure 3D). To further explore the responsiveness of *T. kodakarensis* to temperature changes with respect to mRNA modifications, we performed rRNA depletion as described previously and examined the levels of m^1A, Gm, ac^4C, Cm, and m^5C from 65°C to 95°C (Figures 3A and 3E). We found that ac^4C was the only modification in the mRNA-enriched fraction that underwent significant temperature-dependent changes. While total levels of m^1A, Cm, Gm, and m^5C within the mRNA-enriched fraction did not change in nucleoside analysis, it remains plausible that the sites of modification are differentially controlled at different temperatures.

Interestingly, m^5U was depleted when *T. kodakarensis* was cultivated at temperatures ≥85°C. Previous studies have identified m^5U as a precursor to the tRNA modification m^5s^2U in thermophilic bacteria.[43,44] The substantial accumulation of m^5U in the small RNA fraction at 65°C prompted us to hypothesize that a slower m^5s^2U synthesis may occur at this temperature leading to an increase in the m^5U levels. To test this hypothesis, we quantified the nucleosides m^5U, m^5s^2U, and U in the small RNA fractions of *T. kodakarensis* at temperatures ranging from 65°C to 95°C revealing that the m^5U/U molar ratio increased from below 0.1% at 85°C to ~2% (1.98% ± 0.15%; a ~20-fold increase) at 65°C, whereas the m^5s^2U/U molar ratio only modestly decreased from 13.87% ± 0.49% at 85°C to 11.50% ± 0.46% at 65°C (Figure 3F). The observed ~2% increase in m^5U abundance from 85°C to 65°C seemingly correlates with the ~2% decrease in m^5s^2U. Given that m^5U is the precursor of m^5s^2U,[43,44] the accumulation of m^5U at lower temperature may suggest a potential slowdown in the synthesis of m^5s^2U. Positive correlation between m^5s^2U level and growth temperature has also been observed in both thermophilic Archaea and Bacteria.[6,42,45]

Predicting archaeal RNA-modifying enzymes

Although a multitude of enzyme function predictive tools exist,[46–50] *de novo* prediction of RNA-modifying enzymes in individual organisms remains challenging. To infer the RNA-modifying enzymes responsible for the RNA modifications observed in this study, we conducted a sequence homology based HMMER hmmscan search against the Pfam database,[51,52] annotating >75% of protein-coding genes from the five archaeal reference

genomes (Figures S3A–S3B). We then compared the predicted protein domain composition of each annotated archaeal gene with those of known RNA-modifying enzymes in the Modomics database using the Jaccard index.[53,54] An archaeal gene was assumed to be an RNA-modifying enzyme if its Jaccard index was ≥0.5 (Figures S3A; Table S5). Finally, we associated the enzyme candidates with the corresponding modified nucleosides identified in each species as determined by UHPLC-QqQ MS (see Table S5 for a complete list of domain annotations and the associated RNA modifications).

The associations between RNA modifications and RNA-modifying enzymes were categorized into two groups (Figure 4). In group I, a modification was identified through nucleoside analysis, and at least one putative enzyme plausibly involved in the synthesis of this modified nucleoside was predicted, while group II consisted of modifications that were identified but no associated enzyme was found by domain comparison. At least half (~60%) of the identified modifications in each species fall within group I and are predominantly a product of predicted mono-methylation or acetylation. As one example, an archaeal homolog of the eukaryotic ac^{4}C acetyltransferase NAT10, which is a multi-domain enzyme consisting of two Gcn5-related N-acetyltransferase (GNAT acetyltransferase 2) domains; a helicase (RecD) domain; a tRNA (Met) cytidine acetyltransferase, N-terminal (TmcA N) domain; and a tRNA-binding (tRNA bind 2) domain, is included in this group (Table S5). Proteins with similar domain compositions to the eukaryotic ac^{4}C acetyltransferase NAT10 were identified in the *Thermococcus* and *Sulfolobus* species, but not in *M. maripaludis*, in which ac^{4}C modification was not identified (Figure 4). Group II contains 15 modifications that were observed in many of the tested Archaea but without any immediately associated enzyme(s) presumed responsible for generating the modified nucleoside. Among these 15 modifications, 10 are either doubly modified or hypermodified nucleosides and thus likely the product of multiple RNA modification enzymes, and the classification of these modifications into group II is likely due to an insufficient understanding of the synthesis pathways for these complex nucleosides.

Archaeal domain-specific enzymes could not be readily identified by an HMMER search heavily reliant on matching conserved amino acid residues. To address these limitations, we performed a Foldseek search to compare AlphaFold structure similarities among all putative proteins in these five Archaea and 37 known RNA-modifying enzymes reported in Modomics, which are collectively responsible for eight of the group II modifications (Figure S3C).[55–57] The structure-based search revealed that most (28 out of 37; ~75%) known RNA-modifying enzymes share structure similarities (cutoff e-value $<10^{-5}$) with at least one protein from the five tested Archaea (Table S6). Using this approach, we were able to identify 26–45 enzyme candidates from each archaeon that could potentially establish group II modifications.

Identification of TK0008 as a guanine-*N7*-methyltransferase in *T. kodakarensis*

Our Foldseek search showed that a variety of archaeal proteins share structural similarities with bacterial or eukaryotic guanine-*N*7-methyltransferases that modify rRNA or tRNA at internal positions (Table S6). Despite the reports of m^{7}G in the tRNA of thermophilic Archaea,[12,17,58] no archaeal guanine-*N*7-methyltransferases have been reported so far.

Intriguingly, nucleoside analysis of the large and small RNA fractions from *T. kodakarensis* indicated that m^7G is primarily located in the large RNA fraction (Figure S2A). We found no evidence of m^7G in the mRNA-enriched fraction of *T. kodakarensis*, leading us to hypothesize that m^7G may be present in 16S and/or 23S rRNA. To identify the primary RNA harboring m^7G modifications, we individually depleted 5S, 16S, and 23S rRNA and estimated the levels of m^7G co-depletion in these rRNA-depleted subfractions. The post-rRNA digestion m^7G mass response was normalized to that of an mo^5U spike-in control (Figures 5A, 5B, S4A, and S4B). Our results showed that the depletion of 23S rRNA resulted in a strong co-depletion of m^7G compared to 5S and 16S rRNA depletion (Figure S4B). This suggested that the m^7G modification is primarily found in the 23S rRNA of *T. kodakarensis*.

Next, we sought to identify the genes responsible for installing the m^7G modification in the *T. kodakarensis* 23S rRNA. In *E. coli*, the m^7G modification in 23S rRNA is introduced by the ribosome large subunit methyltransferase K/L (RLMKL). RLMKL installs m^2G at position 2445 and m^7G at position 2069 using two distinct methyltransferase domains.[59,60] Our Foldseek search retrieved four genes (*TK0008*, *TK0981*, *TK1785*, *TK1863*) in *T. kodakarensis* encoding methyltransferases that aligned well (e-values $<10^{-9}$) with the AlphaFold-predicted structure of *E. coli* RLMKL (Table S6). Of those, the enzymes encoded by *TK0981* and *TK1863* have been shown to be $m^2{}_2G$ and m^2G tRNA methyltransferases, respectively.[61] The *TK1785*-encoded protein shares the closest sequence and predicted protein structure with the bacterial ribosome large subunit methyltransferase I (RLMI), which is an RNA m^5C methyltransferase (Table S7; Figure S5). Thus, the remaining unannotated gene *TK0008* emerged as a promising candidate for establishing the m^7G modification in the 23S rRNA of *T. kodakarensis*.

To investigate the activity of the protein encoded by *TK0008*, we conducted a comprehensive nucleoside analysis comparing the total RNA from the TS559 wild-type strain with that of the *TK0008*-deletion strain (ΔTK0008 strain). Our analysis revealed a >95% depletion in the m^7G/G mass response ratio in the ΔTK0008 strain, with m^7G modification being indistinguishable from the background (Figure 5C). To confirm m^7G was specifically depleted in the ΔTK0008 strain, we quantified the relative abundance of the other methylated guanosine (m^1G, m^2G, m^2 G, m^2Gm, and Gm) in both total RNA and the RNA fraction in which 5S and 16S rRNA were depleted (23S rRNA-enriched fraction). As expected, none of these modifications underwent significant reduction in the ΔTK0008 strain (Figure 5D), suggesting that m^7G is likely the primary product of the *TK0008*-encoded methyltransferase. To confirm that the *TK0008* gene encodes a guanine-*N7*-methyltransferase, we treated *in vitro* transcribed (IVT) 23S rRNA with purified recombinant TK0008 protein (rTK0008) at 85°C for 30 min. Indeed, rTK0008-treated IVT 23S displayed an enzyme-concentration-dependent increase of m^7G levels as determined by nucleoside analysis (Figure 5E).

To further investigate the sequence context of m^7G within endogenous 23S rRNA, we performed independent digestions of the 23S rRNA-enriched fraction with either RNase T1 or RNase A and analyzed the resulting RNA oligonucleotides using nanoLC-MS/MS (Figures 5F and 5G; Figures S6A–S6B). Through annotation using

the NucleicAcidSearchEngine (NASE),[62] we identified several detectable sequences that contain a monomethylated guanosine (mG) base (Figure S5A). An mG group is defined here as a guanosine harboring a methyl group, regardless of the substitution pattern. Of particular significance were the high spectrum counts of the methylated oligonucleotides with sequence [mG] [mG]G in the RNase T1 digestion and G[mG][mG]GC in the RNase A digestion, which were abundant in TS559 but depleted in ΔTK0008 (Figure 5F). Subsequent analysis of these oligonucleotides using parallel reaction monitoring (PRM)-based targeted MS revealed that the mG base anion resides in the second position of the [mG][mG]G trinucleotide from RNase T1 digestion (Figure 5G). We also detected the abasic a-B fragment ions corresponding to the loss of a monomethylated guanine in the second position of the [mG][mG]G trinucleotide (RNase T1 digestion) and in the third position of the G[mG][mG]GC pentanucleotide (RNase A digestion) in the TS559 parental strain (Figure 5G), whereas such fragment ions were not detected in the ΔTK0008 strain due to the absence of the target precursors. Altogether, these data provide compelling evidence that the TK0008 is responsible for an m^7G modification, which predominantly resides within the G[mG][mG]GC sequence context of 23S rRNA.

Next, we determined whether TK0008 establishes m^7G at specific site(s) within 23S rRNA. The IVT 23S was treated with four concentrations of rTK0008 (0.1, 0.2, 1, 4 μM) at 85°C. After treatment with the methyltransferase, the RNA was digested with RNase 4 to produce oligonucleotide fragments for modification mapping by LC-MS/MS. Greater than 80% methylation was observed at position 2354 across all rTK0008-treated samples (Figure S7). We also detected minor secondary methylation sites along the IVT 23S, but most of these did not appear at low rTK0008 concentration (<1 μM) (Figure S7). The position 2354 resides in the third guanosine of a GGGGC sequence, which is consistent with G[mG][mG]GC oligonucleotide detected in the TS559 parental strain. Together, these results provide strong evidence supporting that position 2354 is the primary TK0008 methylation site *in vitro*.

Identification of m^7G modification in the conserved P loop of 23S rRNA

To identify the precise location of m^7G in the endogenous 23S rRNA, we examined the cryoelectron microscopy (cryo-EM) structure available for *T. kodakarensis* 70S ribosome (PDB: 6SKF). Gm (2′-*O*-methylguanosine) was the only guanosine modification annotated in the 23S rRNA substructure.[3] According to the reported cryo-EM data, Gm was found in many sequence contexts, including within a G[Gm]GGC at position 2353, wherein G2354 was annotated as an unmodified guanosine. In the same study, LC-MS data indicated a Gm[mG]G trinucleotide assigned to 23S rRNA.[3] We reasoned that the m^7G site could be located at the 3′ adjacent position of Gm, thus yielding a [Gm][m^7G]G sequence. To confirm the presence of an m^7G modification at position 2354, we designed a biotinylated DNA probe complementary to the region of the 23S predicted to harbor the m^7G and independently performed fragment depletion with RNase H digestion of the target sequence and enrichment experiments using RNase 4 prior to nucleoside analysis (Figure 6A). RNase H digestion of the probe-hybridized RNA led to a near-complete depletion of m^7G as shown by UHPLC-QqQ analysis (Figure 6A) without any significant change in the abundance of m^5C, ac^4C, or Cm. Conversely, a nuclease protection assay using RNase 4 to remove

regions of the 23S rRNA not protected by binding of a complementary probe, followed by streptavidin-based pull-down of the probe-hybridized RNA, resulted in a >10-fold enrichment of m^7G relative to its level in total RNA (Figure 6A). Along the same lines, Gm, which is found at G2353, among many other positions, was shown to be at least in part depleted in the RNase H treatment and enriched in the RNase 4 treatment. Finally, careful inspection of the cryo-EM density map at G2354 of the *T. kodakarensis* 70S ribosome revealed a clear exocyclic density above position 7 of the purine ring in support of the methyl group attachment at the guanine-*N*7 position (Figure 6B). G2354 is positioned within the P loop of the peptidyl transferase center (PTC) (Figures 6C and 6D) and forms a universally conserved Watson-Crick base pair not internally within the 23S rRNA but instead with cytidine 74 (C74) of the 3′ end CCA sequence of tRNAs.[63] Inspections of the published cryo-EM structures of *E. coli* 70S ribosome (PDB: 7K00) and human 80S ribosome (PDB: 8QOI) revealed no evidence of modifications at the corresponding positions in *E. coli* 23S rRNA (G2252) and in human 28S rRNA (G4197) (Figures 6E and 6F),[64,65] suggesting that m^7G may play a unique role in archaeal or thermophilic translation by impacting this critical contact between tRNAs and rRNAs.

DISCUSSION

Archaea are a rich reservoir for the discovery of unique RNA modifications and RNA-modifying enzymes. In this study, we established comprehensive nucleoside profiles across three different archaeal genera revealing temperature-sensitive modifications and previously unannotated mRNA modifications. Interestingly, a phylogenetic tree based on these total nucleoside modification profiles closely resembles that of 16S rRNA-seq for these archaeal species. Analysis of the mRNA-enriched fractions from five archaeal species revealed the presence of six distinct nucleotide modifications. Remarkably, we found no evidence of a m^6A modification in these species, in contrast to Eukarya and Bacteria, where m^6A can dominate the landscape of mRNA modifications.[37,68] This intriguing difference in nucleoside chemistry may be attributed to the differential deployment of RNA-modifying enzymes in Archaea to survive in challenging and dynamic environments. 2′-*O*-methylation has been generally implicated in modulating RNA conformation and enhancing nuclease resistance,[69] and our analysis determined that 2′-*O*-methylation is an abundant modification for each of the species we investigated and abundant in rRNA and tRNA.[69–72] Perhaps surprisingly, we also found that 2′-*O*-methylation, specifically Gm and Cm, was abundant in the mRNA of most archaeal species investigated here, likely contributing to their hydrolytic stability.

LC-MS/MS enables simultaneous and global profiling of nucleoside modifications without prior knowledge of the prevalence of each modification. We showed that combining MS analysis for modification discovery with data mining from multiple open-access databases facilitates a rapid selection of initial candidates for enzyme validation. A critical step of this approach is to compare the functional domains of predicted coding sequences with known RNA-modifying enzymes.[53] Modomics provides a comprehensive open-access database for RNA modifications, yet most of the annotated RNA-modifying enzymes are derived from Eukarya and Bacteria. As such, it is likely the predicted archaeal RNA-modifying enzymes resultant from using Modomics will result in hits that preserve most of the functional

domain features that are shared by common ancestors with bacterial and eukaryotic RNA-modification enzymes. Enzymes identified here through Modomics searches are associated with group I modifications, where the associated RNA-modifying enzyme was predicted based on Pfam domain annotations from eukaryotic and bacterial proteins. In contrast, group II modifications largely implicate a subset of archaeal RNA modifications may be established by enzymes without obvious orthologs in other domains. Archaeal enzymes with significant sequence divergence (or even no ancestral linkage) from Eukarya and Bacteria are thus very challenging to predict with domain or protein sequence comparisons. Here, we demonstrate that the combination of LC-MS with enzyme-predicting tools and genetic deletions can provide evidence for the existence or lack of certain RNA-modifying enzymes in the selected archaeal species.

Despite these challenges, our analyses predicted, and we subsequently confirmed, that the methyltransferase-encoded gene *TK0008* is responsible for m^7G modification in the 23S rRNA of *T. kodakarensis*. To our knowledge, this is the first report of an m^7G modification in ribosomal RNA and its associated gene in the Archaea. The *TK0008*-encoded protein harbors a UPF0020 domain and an N6/N4 DNA methylase domain, which resembles the Trm14 tRNA m^2G methyltransferase of *Methanocaldococcus jannaschii*.[73] Initially, we classified the TK0008 protein as an m^2G methyltransferase. However, further structural analysis using Foldseek revealed similarities between the *E. coli* m^7G methyltransferase RLMKL and TK0008. Interestingly, the predicted structure of TK0008 exhibits greater structural similarities with the m^2G (2445) methyltransferase domain of RLMKL rather than the domain responsible for installing m^7G (2069).[60] Nevertheless, our nucleoside analysis revealed guanine-*N7* methyltransferase activity of purified TK0008, as well as a significant reduction of m^7G levels in a *TK0008*-deletion strain, providing compelling evidence for the putative role of this enzyme as an m^7G methyltransferase. Oligonucleotide LC-MS/MS analysis revealed an m^7G modification located at the P loop of the PTC in the 23S rRNA. This finding widens our knowledge of the m^7G modification from previous studies, which have primarily identified this modification in archaeal tRNAs, including the first and tenth positions of tRNA-Glu as well as the tenth position of tRNA-Met^i.[12,17,58] The differential distribution of m^7G in different species and RNA subtypes suggests that m^7G may have divergent functions in Archaea. These findings underscore the combination of LC-MS/MS and comparative genomics as a powerful approach for uncovering new RNA modifications and the associated RNA-modifying enzymes.

Recent technological advancements have enabled single-base-resolution mapping of new mRNA modifications beyond m^6A. Several studies have identified the presence of m^1A and m^5C in mRNA at sub-stoichiometric level across different organisms.[7,18,74,75] Unlike the m^6A mRNA modification, which in higher eukaryotes is deposited by a nuclear RNA-specific methyltransferase,[76] many other mRNA modifications share biosynthetic machineries with tRNAs.[75,77–79] This raises the question of whether mRNA modifications are specifically regulated or are off-target activities from tRNA-modifying enzymes.[80] Our enzyme predictions revealed that all six modifications found in archaeal mRNA are linked to potential rRNA-/tRNA-modifying enzymes. It is conceivable that changes in mRNA modification levels under certain conditions can be attributed to the overall changes in tRNA modifications. In fact, the co-regulation of mRNA and ncRNA modifications by

shared enzymes implies that deletion of writer enzymes may not be sufficient to dissect the functional roles of mRNA modifications. Therefore, complementary approaches may be needed to fully elucidate the biological consequences of RNA modifications.

Limitations of the study

This study specifically investigates RNA modifications that can be validated by nucleoside standards. Therefore, the reported archaeal nucleoside profiles do not cover the entire spectrum of RNA modifications. Secondly, putative RNA-modifying enzymes listed in this study were predicted by bioinformatics analysis, predominantly based on known enzymes reported in Modomics.[53] The discovery of TK0008 as a guanine-*N7*-methyltransferase is one example of the approach employed to validate enzyme predictions. While we find our approach is a powerful tool for enzyme discovery, validation of each new RNA-modifying enzyme activities is still required. Lastly, the *N7*-methylation identified at G2354 in 23S rRNA may not be the only m^7G site within the *T. kodakarensis* epitranscriptome. Combining our approach with other direct RNA-seq technologies may be helpful to identify other potential modification sites.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ivan R. Corrêa Jr. (correa@neb.com).

Materials availability

All unique/stable reagents generated in this study are available from the lead contact without restriction.

Data and code availability

- Raw RNA-seq data have been deposited at Sequence Read Archive as Bioproject PRJNA1219756 and are publicly available. Oligonucleotide analysis data have been deposited at Proteomics Identifications Database with identifier PXD060644 and are publicly available. Nucleoside analysis data have been deposited at the NIH Common Fund's National Metabolomics Data Repository (NMDR), the Metabolomics Workbench,[81] as project PR002314 and are publicly available. The database is supported by NIH grant U2C-DK119886 and OT2-OD030544 grants.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

EXPERIMENTAL MODELS AND STUDY PARTICIPANT DETAILS

Wild-type *T. kodakarensis* were grown at 65°C, 75°C, 85°C, or 95°C in anaerobic artificial sea water supplemented with yeast extract and tryptone to mid-exponential phase (O.D. ~0.3) in triplicates before harvest, as previously described. The regular maintenance condition for all *T. kodakarensis* strains (wild-type and Δ *TK0008*) was 85°C. Sources of wild-type *Thermococcus* sp. AM4, *Sulfolobus islandicus* M16.4, *Sulfolobus acidocaldarius*, and *Methanococcus maripaludis* biomasses are listed in Data S2. Universal human reference RNA (Agilent Technologies, Cat #740000) and *Escherichia coli* total RNAs (Thermo Fisher Scientific, Cat #AM7940) were purchased from commercially available sources. *Saccharomyces cerevisiae* (strain S288C) total RNA was obtained as described in Mulroney et al.[82]

METHOD DETAILS

Δ *TK0008* strain construction: The TK0008 open reading frame (ORF) spans positions 7655 to 8755 on the *T. kodakarensis* genome. PCR was used to generate an amplicon including the TK0008 ORF flanked by ~700 bp both up- and down-stream that was cloned into the unique SwaI site of pTS700 (reference Hileman and Santangelo, PMID: 22701112),[83] resulting in pCSU-TK0008A. The coding sequence of TK0008 was excised from pCSU-TK0008A using site-directed QuikChange mutagenesis (Agilent Cat# 200516) to generate pCSU-TK0008B. All plasmid sequences were confirmed by Sanger sequencing.

pCSU-TK0008B was transformed into *T. kodakarensis* strain TS559 and plated onto agmatine-free rich media plates as previously described (reference PMID: 34752288).[86] Transformants were allowed to grow for 4 days anaerobically at 85°C and subsequent colonies were picked into rich liquid media lacking agmatine and grown overnight. Genomic DNA extracted from 1 mL of cells was used as a template for PCRs using two sets of primer pairs to determine the site and directionality of plasmid integration into the genome. One primer from each pair has homology to the genome with the other having homology to the plasmid to ensure PCR amplification involves a region flanking TK0008 on the genome. Colonies confirmed to contain proper integrations at the TK0008 locus were plated on minimal media containing 1 mM agmatine and 100 μM 6-methyl purine and grown anaerobically at 85°C for 2–5 days. Genomic DNA from resultant colonies was used in diagnostic PCRs to identify strains deleted for TK0008 preliminarily. The deletion of TK0008 was ultimately confirmed by Sanger sequencing.

Cloning and recombinant expression of TK0008: The native *T. kodakarensis* sequence encoding RNA methyltransferase TK0008, appended with an N-terminal six histidine encoding sequence, was cloned into a pQE80 expression vector using Infusion cloning (Takara Bio., Cat #638910), transformed into the Nico21 *E coli* cell line (NEB, Cat #C2529H), and colonies were allowed to form at 37°C on LB plates containing 100 μg/mL ampicillin and 25 μg/mL chloramphenicol. A single colony was picked into 10 mL liquid LB medium containing 100 μg/mL ampicillin and 25 μg/mL chloramphenicol and grown at 37°C with shaking (~200–250 rpm) overnight before being used to inoculate a 1 L LB medium culture containing 100 μg/mL ampicillin and 25 μg/mL chloramphenicol that was allowed to grow at 37°C to an optical density ($OD_{600\ nm}$) of ~0.4 prior to addition of

isopropyl β-*d*-1-thiogalactopyranoside (IPTG, 400 μM final concentration) and D-sorbitol (3% final concentration by weight). Cultures were allowed to grow for 3 h at 37°C with agitation before harvesting via centrifugation (10825 × *g*). Cell pellets were stored at −20°C prior to thawing on ice and resuspended in buffer A (25 mM Tris HCl pH 8.0, 500 mM NaCl, 10% glycerol by weight). For each gram of pelleted cells, 3 mL of buffer A was added. The cell suspension was sonicated on ice for repeated cycles of 30 s on, 30 s rest over 30 min prior to centrifugation (48384 × g for 20 min at 4°C). The supernatant was reserved and a second round of resuspension, sonication, and centrifugation performed. The clarified supernatants were combined prior to heating the lysate to 65°C for 20 min before clearing the bulk of non-thermostable proteins via centrifugation (48384 × g for 20 min at 4°C). The heat-treated clarified supernatant was filtered through a 0.45 μM filter immediately prior to loading a 5 mL HiTrap Chelating HP column (Cytiva Life Sciences, Cat # 17040901) that was charged with $NiSO_4$ and equilibrated with buffer A. Following extensive washing with buffer A, bound proteins were eluted with a linear gradient to buffer B (25 mM Tris HCl pH 8.0, 100 mM NaCl, 200 mM imidazole, 10% glycerol by volume) collecting 2 mL fractions. Fractions containing recombinant TK0008 protein (judged to be >95% pure) were identified by SDS-PAGE and protein staining prior to pooling, dialysis into storage buffer (25 mM Tris HCl pH 8.0, 100 mM NaCl, 1 mM DTT, 50% glycerol by volume), and −20°C storage.

Total RNA extraction and rRNA depletion: Harvested archaeal cell pellets were resuspended in 10 mL of TRI reagent (Molecular Research Center, Inc., Cat #TR118). The resuspended cells were homogenized using a beads beater at 4.0 m/s for 20 s × 2 cycles (MP Biomedicals, FastPrep-24TM). Subsequently, the mixture was centrifuged at 14000 g for 5 min to precipitate any cell debris. Supernatants were collected post-centrifugation and treated with 50 μL of BAN reagent (Molecular Research Center, Inc., Cat #BN191) per mL of supernatant for aqueous-organic phase separation. RNA from the aqueous phase was isolated by isopropanol precipitation and subjected to DNase I treatment (NEB, Cat #M0303S) to remove genomic DNA contamination. To further purify the DNase I-treated RNA, an equal volume of acid phenol-chloroform with isoamyl alcohol (125:24:1, Thermo Fisher Scientific, Cat #AM9722) was added to the reaction and centrifuged at 21300 g for 2 min to separate the aqueous phase from the organic phase. The aqueous phase containing RNA was then precipitated with 1.5 volumes of isopropanol and 0.1 volume of sodium acetate (Sigma Aldrich, Cat #S7899) at −20°C overnight. Finally, the precipitated RNA pellets were washed with 75% ethanol and dissolved in nuclease-free water.

To remove rRNA and tRNA, total RNA was separated into large (>200 nt) and small RNA (<200 nt) fractions using the RNA Clean and Concentrator Kit (Zymo Research, Cat #R1017). Subsequently, 50 μg of the large RNA fraction were subjected to rRNA depletion using the NEBNext rRNA Depletion Kit (NEB, Cat #E7850X) with the following changes: The NEBNext rRNA depletion solutions provided in the kit were substituted for customized DNA probe mixtures (at 1 μM for each probe) fully complementary to rRNA sequences corresponding to each archaeal species; All volumes for the probe hybridization, RNase H and DNase I digestion reactions were scaled up by 5-fold in 10 parallel reactions. Following the enzymatic treatment, the reactions were cleaned up using RNA Clean and Concentrator Kit (Zymo Research, Cat #R1017), the mRNA-enriched fractions were eluted in 10 μL water

and combined. The specific probe sequences used for rRNA depletion are detailed in Table S8.

For individual depletion of *T. kodakarensis* 5S, 16S, and 23S rRNA (Figures 5A and 5B; Figures S4A and S4B), 5 μg of *in vitro* transcribed (IVT) erythropoietin (EPO) mRNA modified with 5-methoxyuridine (mo^5U) was spiked into the starting total RNA (50 μg) before rRNA depletion. Since mo^5U was not identified in the endogenous *T. kodakarensis* RNA (Figure 1A), it served as an internal control standard for assessing m^7G levels. The total RNA was then split evenly into five tubes for rRNA depletion using the NEBNext rRNA Depletion Kit (NEB, Cat #E7850X). The NEBNext rRNA depletion solutions provided in the kit were substituted for customized DNA probes (at 1 μM for each probe) fully complementary to individual rRNA sequence. The enzymatic reactions were cleaned up following the procedure described above and eluted in nuclease-free water.

RNA-seq library preparation and data analysis: RNA-seq libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, Cat #E7760S). Briefly, RNA samples with or without rRNA depletion were primed using random primers, fragmented, and followed by first and second strand cDNA synthesis. Adaptors were ligated onto the cDNA, and the second-strand cDNA was removed using the Uracil-Specific Excision Reagent provided in the kit. The adaptor-ligated cDNA libraries underwent PCR enrichment (7–12 cycles) with unique dual index primers (NEB, Cat #E6440S). Enriched libraries were purified using Sample Purification Beads (NEB, Cat #E7767S) and eluted in 0.1x TE buffer. Subsequently, libraries were normalized, pooled, and quality-checked using Bioanalyzer High Sensitivity DNA Analysis (Agilent Technologies, Part #5067–4626). Finally, paired-end sequencing experiments were performed on Illumina NextSeq 500/550 or MiSeq platforms.

The output FASTQ files from Illumina sequencing underwent trimming using the fastx trimmer to eliminate low-quality bases. Trimmed reads were subsequently mapped to archaeal reference genomes using Bowtie2, and the mapped reads were sorted by coordinates using Samtools. Next, the sorted reads were counted based on the annotated exon features using HTSeq in union mode, with a minimum alignment quality threshold of 10. Counts overlapping with mRNA features were then normalized to the total exon read counts to obtain the percentage of mRNA reads in RNA samples with or without rRNA depletion. The accession numbers of the archaeal reference genomes used for mapping were NC_006624, CP002952, NC_012726, NZ_CP020364, and NC_015847.

Selective rRNA fragment enrichment and depletion: To enrich m^7G-modified rRNA fragment, 50 μg of total RNA from *T. kodakarensis* TS559 wild-type strain was mixed with a biotinylated DNA probe (1 μM) fully complementary to the target sequence in 10 mM Tris buffer (pH 7.5) and heated to 95°C for 2 min followed by a ramp down to 22°C. Biotinylated DNA probe sequence: 5′-Biotin-ACCTTCGAGGTGTGCGCCAGCAACTGCCACT-3’. The unhybridized RNAs were digested with 5 μL of RNase 4 (NEB, Cat #M1284L) in 1x NEB buffer r1.1 at 37°C for 1 h. The digestion was stopped by adding 10 μL of human placental RNase inhibitor (NEB, Cat #M0307S) to the reaction mixture and incubated at room temperature for 10 min.

Selective enrichment of the probe-hybridized rRNA region was performed using streptavidin magnetic beads (NEB, Cat #S1421S). Briefly, 250 μL of streptavidin magnetic beads were washed twice with equal volume of low-salt buffer (100 mM Tris pH 7.5, 10 mM EDTA, 50 mM NaCl) and resuspended in 50 μL of high-salt buffer (100 mM Tris pH 7.5, 10 mM EDTA, 250 mM NaCl). The high-salt buffer was removed, and the beads were incubated with the reaction mixture for 15 min at room temperature with occasional agitation. The bead-captured DNA-RNA hybrid was washed and eluted by heating to 80°C for 5 min in 30 μL of nuclease-free water.

For m^7G-modified rRNA fragment depletion, the same biotinylated DNA probe (1 μM) described above was mixed with 5 μg of total RNA from TS559 strain in the probe hybridization buffer provided in NEBNext rRNA Depletion Kit (NEB, Cat #E7850X). The probe-hybridized rRNA mixture was subjected to rRNA depletion according to the NEBNext rRNA Depletion Kit protocol. Following the enzymatic treatment, the reactions were cleaned up using RNA Clean and Concentrator Kit (Zymo Research, Cat #R1017). The eluants from rRNA fragment enrichment and depletion reactions were collected and subjected to UHPLC-QqQ nucleoside analysis (Figure 6A).

Genetic distance tree construction: A genetic distance tree for five archaeal species was constructed using a 16S rRNA multiple sequence alignment. Genetic distances were calculated based on the percentage of identical residues between the aligned sequences. Briefly, FASTA files containing 16S rRNA sequences from the five archaea were used as input for Clustal Omega multiple sequence alignment. The resulting alignment was then employed for distance tree construction using the unweighted pair group method with arithmetic mean (UPGMA) algorithm within Geneious Prime software. The archaeal reference genomes utilized in the 16S rRNA alignment can be identified by the following RefSeq accession numbers: NC_006624, NC_016051, NC_012726, NZ_CP020364, and NC_015847.

Bioinformatics prediction of RNA-modifying enzymes: Domain annotations of archaeal proteins and known RNA-modifying enzymes in Modomics database was carried out by HMMER hmmscan to search against Pfam database with a cutoff e-value at 0.001. Next, we compared the similarity of domain compositions between the tested archaeal proteins and the known RNA-modifying enzymes using the Jaccard index:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

For instance, an archaeal protein comprising a SAM-dependent methyltransferase (Methyltrans SAM) domain and a PUA-like (PUA3) domain will have a Jaccard index of 1 when compared with the ribosomal RNA large subunit methyltransferase I of *E. coli*, which also contains both a Methyltrans SAM domain and a PUA3 domain. Conversely, it will have a Jaccard index of 0.5 if the compared protein only consists of a PUA-like domain. All tested archaeal proteins with a Jaccard index ≥0.5 were selected as RNA-modifying enzyme candidates. Duplicated domains were not considered in the comparison.

To further select RNA-modifying enzyme candidates responsible for Group II modifications in the tested Archaea, AlphaFold2-predicted protein structures of RNA-modifying enzymes in Modomics reported to install Group II modifications were used to search against AlphaFold2-predicted structures of all annotated proteins of the tested Archaea. Cutoff e-value was set at 10^{-5}.

Methyltransferase activity assay: *In vitro* transcribed (IVT) 23S rRNA was prepared using a linearized plasmid template (pBluescript II KS(−)) and HiScribe T7 High Yield RNA Synthesis Kit (NEB, Cat #E2040S). The synthesized transcripts were cleaned up using Monarch Spin RNA Cleanup Kit (NEB, Cat #T2050L). To perform *in vitro* methylation reactions, IVT 23S rRNA was heat-denatured at 95°C for 30 s and cooled to 22°C prior to enzyme treatment. Purified recombinant TK0008 (0.1–4 μM) was incubated with 1 μM of IVT 23S rRNA in 1x methyltransferase buffer (25 mM Tris-HCl pH 7.5, 100 mM NaCl, 1 mM DTT) supplemented with 160 μM of S-adenosyl-methionine (SAM) (NEB, Cat #B9003S) at 85°C for 30 min. The reactions were cleaned up using Monarch Spin RNA Cleanup Kit (NEB, Cat #T2040L). Lastly, the eluted RNA substrates were used for nucleoside or oligonucleotide LC-MS/MS analysis.

LC-MS/MS: For modified nucleoside detection, total RNA, RNA subfractions, and mRNA-enriched samples were digested to nucleosides at 37 ° C overnight using a Nucleoside Digestion Mix (NEB, Cat #M0649S). The digested RNAs were subsequently injected without prior purification on an Agilent 1290 Infinity II UHPLC equipped with a G7117 diode array detector and an Agilent 6495C Triple-Quadrupole Mass Spectrometer operating in positive electrospray ionization (+ESI) mode. UHPLC was conducted on a Waters XSelect HSS T3 XP column (2.1 × 100 mm, 2.5 μm) containing methanol and 10 mM ammonium acetate (pH 4.5) gradient mobile phase. Mass spectrometric data were acquired using dynamic multiple reaction monitoring (DMRM) mode. Identification of each nucleoside species was based on the associated retention time and mass transition in the extracted chromatogram. The *m/z* of precursor and product ions and retention times for nucleoside profiling (Figure 1) are listed in Table S9. Isomeric nucleosides of methylated adenosine, guanosine, cytidine and uridine were either differentiated by retention times or unique mass transition (Figure S8). Methylated adenosines (m^1A, m^2A, m^6A, and m^8A), methylated guanosines (m^1G, m^2G, m^6G, and m^7G), and methylated uridines (m^3U and m^5U) were distinguished by their retention times. Methylated cytidines were differentiated based on our previous publication.[8] In brief, the retention time of m^3C is distinct from m^4C and m^5C. Although retention time of m^4C is nearly indistinguishable from m^5C, m^4C has a unique 258.1 -> 95 mass transition which is not observed in that of m^5C.

For endogenous m^7G-modified oligonucleotide analysis, the 23S rRNA-enriched fraction was digested with either RNase T1 (Thermo Fisher Scientific, Cat #EN0542) or RNase A (NEB, Cat #T3018L). m^7G mapping in the rTK0008-treated IVT 23S rRNA was carried out by digesting ~10 μg of RNA samples with 1 μL of RNase 4 (NEB, Cat #M1284L). All RNase treatments were conducted at 37°C for 1 h and subjected to analysis on an Eclipse Fusion Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with a Vanquish UHPLC (Thermo Fisher Scientific). The RNase-digested endogenous RNA was resolved in

a mobile phase gradient consists of buffer A (1% hexafluoroisopropanol (HFIP), 0.1% N, N-diisopropylethylamine (DIEA), 1 μM EDTA) and buffer B (90% Methanol, 10% water, 0.075% HFIP, 0.0375% DIEA, 1 μM EDTA) on a Waters nanoEase M/Z Peptide BEH C18 column (1.7 μm, 100 μm × 100 mm). The obtained chromatogram peaks were plotted by a Freestyle software package (Thermo Fisher Scientific, USA). Oligonucleotides derived from RNase 4-digested IVT RNA were resolved in the mobile phase described above on a Waters ACQUITY Premier Oligonucleotide BEH C18 Column (1.7 μm, 2.1 × 100 mm).

Nucleoside LC-MS/MS data analysis: Raw data acquired from UHPLC-QqQ MS were processed using Agilent MassHunter WorkStation Quantitative Analysis (QQQ). Nucleoside mass responses were obtained by integrating the underlying areas of chromatographic peaks. For nucleoside quantification, calibration curves were constructed based on the mass responses of nucleoside standards with known concentrations, as determined by a UV spectrometer (Thermo Fisher Scientific, Evolution 220). Relative abundances were calculated by normalizing the amount of modified nucleoside to the corresponding unmodified nucleoside (see Figures 2B, 3F, and 5C). The mass response ratios of modified nucleosides relative to corresponding unmodified nucleosides were calculated as depicted as follows, using $m^3\Psi$ as an example:

$$\text{Mass response ratio of } m^3\Psi = \frac{\text{Mass response of } m^3\Psi}{\text{Mass response of uridine}}$$

The fold change in mass response ratios were employed to assess the level differences of modified nucleosides in *Thermococcus kodakarensis* grown at different temperatures (Figures 3A–3E) as well as in nucleoside profile comparisons between TS559 wild-type strain and ΔTK0008 strain (Figure 5B). This approach can be illustrated as below using $m^3\Psi$ as an example:

$$\text{Mass response ratio fold change of } m^3\Psi \text{ at } 65\,^\circ C = \frac{\text{Mass response ratio of } m^3\Psi \text{ at } 65\,^\circ C}{\text{Mass response ratio of } m^3\Psi \text{ at } 85\,^\circ C}$$

To calculate the modification fold change between large and small RNA fractions (Figure S2A), the mass response ratio of modified nucleosides in the RNA subfraction was normalized to that of total RNA. The equation can be expressed as follows, using m^5C in the small RNA fraction as an example:

$$\text{Modification fold change of } m^5C \text{ (Small)} = \frac{\text{Mass response ratio of } m^5C \text{ in small RNA fraction}}{\text{Mass response ratio of } m^5C \text{ in total RNA}}$$

A qualitative screening for modified nucleosides in the mRNA-enriched samples was conducted by injecting at least 100 ng of digested RNA. The presence of a modified nucleoside was determined using an arbitrary threshold based on the mass responses of ~1 fmol of the corresponding unmodified nucleoside, inosine, and pseudouridine. Nucleosides detectable in total RNA but falling below the arbitrary threshold in mRNA-enriched samples were designated as ncRNA-specific modifications (Table S2). For quantitative analysis, only

detectable nucleosides with known UV extinction coefficients in mRNA-enriched samples were considered as potential mRNA modifications.

Quantitative analysis was performed by injecting 1–5 ng of digested mRNA samples, which was determined as the amount of both modified and unmodified nucleosides falling within the quantitative range of UHPLC-QqQ MS. Subsequently, the mass response of each nucleoside within the mRNA sample was compared to the mass response of the same mass transition window in the blank control. The resultant ratios represent the signal-to-background for each individual nucleoside. To confidently assign modifications originating from mRNA, a threshold of 2 standard deviations above the average signal-to-background ratio of ncRNA-specific modifications was set.

Oligonucleotide LC-MS/MS data analysis: Oligonucleotide MS/MS data analysis on endogenous RNA was performed using the NucleicAcidSearchEngine (NASE) in Open-MS (version: 2.7.0).[62] Precursor and fragment ion mass cutoffs of 10 ppm were utilized, and 0 to +4 precursor isotopes between the charge states −1 to −20 were considered for analysis. Fragment ions as defined by[87] (a-B, a, b, c, d, w, x, y and z) were considered for analysis of tandem mass spectra. MS/MS data were searched against theoretical digests of the appropriate RNA sequence with RNase T1 or RNase An up to two missed cleavages. Only 3′-phosphorylated or cleavage products were considered. Data analysis on IVT 23S rRNA was performed using the Biopharma Finder software (version: 5.1; Thermo Fisher Scientific, U.S.A.). The threshold for differences in precursor ion masses was set at < 5 ppm. The confidence score for oligonucleotide annotations was set at > 90.

QUANTIFICATION AND STATISTICAL ANALYSIS

For temperature-sensitive modification search experiment described in Figures 3B–3F: Statistical significance was tested with one-way *ANOVA* followed by Bonferroni's multiple comparisons test using Prism. All quantifications are displayed as mean ± S.E.M. from three biological replicates. Statistical significance was defined by $p < 0.05$. Domain annotations of proteins were carried out by HMMER hmmscan with a cutoff e-value set at 0.001. Domain composition similarities between unknown archaeal proteins and known RNA-modifying enzymes were assessed by Jaccard index using an R script. Only archaeal proteins with Jaccard index ≥ 0.5 were considered as potential RNA-modifying enzymes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank Peter Weigele, Yan-Jiun Lee, Robert Trachman, Daniel Kneller, and G. Brett Robb for their valuable input. We thank Sonja-Verena Albers, C.S. Raman, Rachel Whitaker, and Barney Whitman for generously providing archaeal biomasses. We thank Donald Comb, Jim Ellard, Salvatore Russello, and Richard J. Roberts for the continuous support of research at New England Biolabs. This work was funded by New England Biolabs, Inc., NSF award 2022065 (to T.J.S.), NIH award GM143963 (to T.J.S.), Dept. of Energy award DE-SC0014597 (to T.J.S.), and NASA award 80NSSC23K1354 (to T.J.S.).

REFERENCES

1. Ushida C, Muramatsu T, Mizushima H, Ueda T, Watanabe K, Stetter KO, Crain PF, McCloskey JA, and Kuchino Y (1996). Structural feature of the initiator tRNA gene from Pyrodictium occultum and the thermal stability of its gene product, tRNA(imet). Biochimie 78, 847–855. 10.1016/s0300-9084(97)84337-4. [PubMed: 9116054]
2. Noon KR, Guymon R, Crain PF, McCloskey JA, Thomm M, Lim J, and 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. 10.1128/JB.185.18.5483-5490.2003. [PubMed: 12949100]
3. 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. 10.1038/s41586-020-2418-2. [PubMed: 32555463]
4. Turner B, Burkhart BW, Weidenbach K, Ross R, Limbach PA, Schmitz RA, de Crécy-Lagard V, Stedman KM, Santangelo TJ, and Iwata-Reuyl D (2020). Archaeosine Modification of Archaeal tRNA: Role in Structural Stabilization. J. Bacteriol. 202, e00748–19. 10.1128/JB.00748-19. [PubMed: 32041795]
5. Ohira T, Minowa K, Sugiyama K, Yamashita S, Sakaguchi Y, Miyauchi K, Noguchi R, Kaneko A, Orita I, Fukui T, et al. (2022). Reversible RNA phosphorylation stabilizes tRNA for cellular thermotolerance. Nature 605, 372–379. 10.1038/s41586-022-04677-2. [PubMed: 35477761]
6. Ohira T, and Suzuki T (2024). Transfer RNA modifications and cellular thermotolerance. Mol. Cell 84, 94–106. 10.1016/j.molcel.2023.11.041. [PubMed: 38181765]
7. Fluke KA, Fuchs RT, Tsai YL, Talbott V, Elkins L, Febvre HP, Dai N, Wolf EJ, Burkhart BW, Schiltz J, et al. (2024). The extensive m(5)C epitranscriptome of Thermococcus kodakarensis is generated by a suite of RNA methyltransferases that support thermophily. Nat. Commun. 15, 7272. 10.1038/s41467-024-51410-w. [PubMed: 39179532]
8. Fluke KA, Dai N, Wolf EJ, Fuchs RT, Ho PS, Talbott V, Elkins L, Tsai YL, Schiltz J, Febvre HP, et al. (2024). A novel N4,N4-dimethylcytidine in the archaeal ribosome enhances hyperthermophily. Proc. Natl. Acad. Sci. USA 121, e2405999121. 10.1073/pnas.2405999121. [PubMed: 39471227]
9. Agris PF (2015). The importance of being modified: an unrealized code to RNA structure and function. RNA 21, 552–554. 10.1261/rna.050575.115. [PubMed: 25780134]
10. Frye M, Harada BT, Behm M, and He C (2018). RNA modifications modulate gene expression during development. Science 361, 1346–1349. 10.1126/science.aau1646. [PubMed: 30262497]
11. Fagre C, and Gilbert W (2024). Beyond reader proteins: RNA binding proteins and RNA modifications in conversation to regulate gene expression. Wiley Interdiscip. Rev. RNA 15, e1834. 10.1002/wrna.1834. [PubMed: 38444048]
12. Edmonds CG, Crain PF, Gupta R, Hashizume T, Hocart CH, Kowalak JA, Pomerantz SC, Stetter KO, and McCloskey JA (1991). Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria). J. Bacteriol. 173, 3138–3148. 10.1128/jb.173.10.3138-3148.1991. [PubMed: 1708763]
13. Noon KR, Bruenger E, and McCloskey JA (1998). Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile Sulfolobus solfataricus. J. Bacteriol. 180, 2883–2888. 10.1128/JB.180.11.2883-2888.1998. [PubMed: 9603876]
14. McCloskey JA, Graham DE, Zhou S, Crain PF, Ibba M, Konisky J, Söll D, and 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. 10.1093/nar/29.22.4699. [PubMed: 11713320]
15. Grosjean H, Gaspin C, Marck C, Decatur WA, and de Crécy-Lagard V (2008). RNomics and Modomics in the halophilic archaea Haloferax volcanii: identification of RNA modification genes. BMC Genom. 9, 470. 10.1186/1471-2164-9-470.
16. Yu N, Jora M, Solivio B, Thakur P, Acevedo-Rocha CG, Randau L, de Crécy-Lagard V, Addepalli B, and Limbach PA (2019). tRNA Modification Profiles and Codon-Decoding Strategies in Methanocaldococcus jannaschii. J. Bacteriol. 201, e00690–18. 10.1128/JB.00690-18. [PubMed: 30745370]

17. Wolff P, Villette C, Zumsteg J, Heintz D, Antoine L, Chane-Woon-Ming B, Droogmans L, Grosjean H, and Westhof E (2020). Comparative patterns of modified nucleotides in individual tRNA species from a mesophilic and two thermophilic archaea. RNA 26, 1957–1975. 10.1261/rna.077537.120. [PubMed: 32994183]

18. Edelheit S, Schwartz S, Mumbach MR, Wurtzel O, and Sorek R (2013). Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. PLoS Genet. 9, e1003602. 10.1371/journal.pgen.1003602. [PubMed: 23825970]

19. Schwartz S, and Motorin Y (2017). Next-generation sequencing technologies for detection of modified nucleotides in RNAs. RNA Biol. 14, 1124–1137. 10.1080/15476286.2016.1251543. [PubMed: 27791472]

20. Jora M, Lobue PA, Ross RL, Williams B, and Addepalli B (2019). Detection of ribonucleoside modifications by liquid chromatography coupled with mass spectrometry. Biochim. Biophys. Acta. Gene Regul. Mech. 1862, 280–290. 10.1016/j.bbagrm.2018.10.012. [PubMed: 30414470]

21. Zhang Y, Lu L, and Li X (2022). Detection technologies for RNA modifications. Exp. Mol. Med. 54, 1601–1616. 10.1038/s12276-022-00821-0. [PubMed: 36266445]

22. Huang G, Ding Q, Xie D, Cai Z, and Zhao Z (2022). Technical challenges in defining RNA modifications. Semin. Cell Dev. Biol. 127, 155–165. 10.1016/j.semcdb.2021.11.009. [PubMed: 34838434]

23. Schaefer M, Pollex T, Hanna K, and Lyko F (2009). RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res. 37, e12. 10.1093/nar/gkn954. [PubMed: 19059995]

24. Zhao LY, Song J, Liu Y, Song CX, and Yi C (2020). Mapping the epigenetic modifications of DNA and RNA. Protein Cell 11, 792–808. 10.1007/s13238-020-00733-7. [PubMed: 32440736]

25. Li X, Xiong X, and Yi C (2016). Epitranscriptome sequencing technologies: decoding RNA modifications. Nat. Methods 14, 23–31. 10.1038/nmeth.4110. [PubMed: 28032622]

26. Furlan M, Delgado-Tejedor A, Mulroney L, Pelizzola M, Novoa EM, and Leonardi T (2021). Computational methods for RNA modification detection from nanopore direct RNA sequencing data. RNA Biol. 18, 31–40. 10.1080/15476286.2021.1978215. [PubMed: 34559589]

27. Zhao X, Zhang Y, Hang D, Meng J, and Wei Z (2022). Detecting RNA modification using direct RNA sequencing: A systematic review. Comput. Struct. Biotechnol. J. 20, 5740–5749. 10.1016/j.csbj.2022.10.023. [PubMed: 36382183]

28. Yoluc Y, Ammann G, Barraud P, Jora M, Limbach PA, Motorin Y, Marchand V, Tisne C, Borland K, and Kellner S (2021). Instrumental analysis of RNA modifications. Crit. Rev. Biochem. Mol. Biol. 56, 178–204. 10.1080/10409238.2021.1887807. [PubMed: 33618598]

29. Ross RL, Cao X, and Limbach PA (2017). Mapping Post-Transcriptional Modifications onto Transfer Ribonucleic Acid Sequences by Liquid Chromatography Tandem Mass Spectrometry. Biomolecules 7, 21. 10.3390/biom7010021. [PubMed: 28241457]

30. Jiang T, Yu N, Kim J, Murgo JR, Kissai M, Ravichandran K, Miracco EJ, Presnyak V, and Hua S (2019). Oligonucleotide Sequence Mapping of Large Therapeutic mRNAs via Parallel Ribonuclease Digestions and LC-MS/MS. Anal. Chem. 91, 8500–8506. 10.1021/acs.analchem.9b01664. [PubMed: 31129964]

31. Wolf EJ, Grünberg S, Dai N, Chen TH, Roy B, Yigit E, and Corrêa IR (2022). Human RNase 4 improves mRNA sequence characterization by LC-MS/MS. Nucleic Acids Res. 50, e106. 10.1093/nar/gkac632. [PubMed: 35871301]

32. Vanhinsbergh CJ, Criscuolo A, Sutton JN, Murphy K, Williamson AJK, Cook K, and Dickman MJ (2022). Characterization and Sequence Mapping of Large RNA and mRNA Therapeutics Using Mass Spectrometry. Anal. Chem. 94, 7339–7349. 10.1021/acs.analchem.2c00765. [PubMed: 35549087]

33. Su D, Chan CTY, Gu C, Lim KS, Chionh YH, McBee ME, Russell BS, Babu IR, Begley TJ, and Dedon PC (2014). Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. Nat. Protoc. 9, 828–841. 10.1038/nprot.2014.047. [PubMed: 24625781]

34. Thuring K, Schmid K, Keller P, and Helm M (2016). Analysis of RNA modifications by liquid chromatography-tandem mass spectrometry. Methods 107, 48–56. 10.1016/j.ymeth.2016.03.019. [PubMed: 27020891]
35. Tardu M, Jones JD, Kennedy RT, Lin Q, and Koutmou KS (2019). Identification and Quantification of Modified Nucleosides in Saccharomyces cerevisiae mRNAs. ACS Chem. Biol. 14, 1403–1409. 10.1021/acschembio.9b00369.
36. Kimura S, Dedon PC, and Waldor MK (2020). Comparative tRNA sequencing and RNA mass spectrometry for surveying tRNA modifications. Nat. Chem. Biol. 16, 964–972. 10.1038/s41589-020-0558-1. [PubMed: 32514182]
37. Deng X, Chen K, Luo GZ, Weng X, Ji Q, Zhou T, and He C (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. Nucleic Acids Res. 43, 6557–6567. 10.1093/nar/gkv596. [PubMed: 26068471]
38. Bar-Yaacov D, Mordret E, Towers R, Biniashvili T, Soyris C, Schwartz S, Dahan O, and Pilpel Y (2017). RNA editing in bacteria recodes multiple proteins and regulates an evolutionarily conserved toxin-antitoxin system. Genome Res. 27, 1696–1703. 10.1101/gr.222760.117. [PubMed: 28864459]
39. Chu JM, Ye TT, Ma CJ, Lan MD, Liu T, Yuan BF, and Feng YQ (2018). Existence of Internal N7-Methylguanosine Modification in mRNA Determined by Differential Enzyme Treatment Coupled with Mass Spectrometry Analysis. ACS Chem. Biol. 13, 3243–3250. 10.1021/acschembio.7b00906. [PubMed: 29313662]
40. McCown PJ, Ruszkowska A, Kunkler CN, Breger K, Hulewicz JP, Wang MC, Springer NA, and Brown JA (2020). Naturally occurring modified ribonucleosides. Wiley Interdiscip. Rev. RNA 11, e1595. 10.1002/wrna.1595. [PubMed: 32301288]
41. Nie W, Wang S, He R, Xu Q, Wang P, Wu Y, Tian F, Yuan J, Zhu B, and Chen G (2020). A-to-I RNA editing in bacteria increases pathogenicity and tolerance to oxidative stress. PLoS Pathog. 16, e1008740. 10.1371/journal.ppat.1008740. [PubMed: 32822429]
42. Kowalak JA, Dalluge JJ, McCloskey JA, and Stetter KO (1994). The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. Biochemistry 33, 7869–7876. 10.1021/bi00191a014. [PubMed: 7516708]
43. Shigi N, Sakaguchi Y, Asai SI, Suzuki T, and Watanabe K (2008). Common thiolation mechanism in the biosynthesis of tRNA thiouridine and sulphur-containing cofactors. EMBO J. 27, 3267–3278. 10.1038/emboj.2008.246. [PubMed: 19037260]
44. Yamagami R, Yamashita K, Nishimasu H, Tomikawa C, Ochi A, Iwashita C, Hirata A, Ishitani R, Nureki O, and Hori H (2012). The tRNA recognition mechanism of folate/FAD-dependent tRNA methyltransferase (TrmFO). J. Biol. Chem. 287, 42480–42494. 10.1074/jbc.M112.390112. [PubMed: 23095745]
45. Watanabe K, Shinma M, Oshima T, and Nishimura S (1976). Heat-induced stability of tRNA from an extreme thermophile, Thermus thermophilus. Biochem. Biophys. Res. Commun. 72, 1137–1144. 10.1016/s0006-291x(76)80250-1. [PubMed: 985514]
46. Dalkiran A, Rifaioglu AS, Martin MJ, Cetin-Atalay R, Atalay V, and Doğan T (2018). ECPred: a tool for the prediction of the enzymatic functions of protein sequences based on the EC nomenclature. BMC Bioinf. 19, 334. 10.1186/s12859-018-2368-y.
47. Hon J, Borko S, Stourac J, Prokop Z, Zendulka J, Bednar D, Martinek T, and Damborsky J (2020). EnzymeMiner: automated mining of soluble enzymes with diverse structures, catalytic properties and stabilities. Nucleic Acids Res. 48, W104–W109. 10.1093/nar/gkaa372. [PubMed: 32392342]
48. Zallot R, Oberg N, and Gerlt JA (2019). The EFI Web Resource for Genomic Enzymology Tools: Leveraging Protein, Genome, and Metagenome Databases to Discover Novel Enzymes and Metabolic Pathways. Biochemistry 58, 4169–4182. 10.1021/acs.biochem.9b00735. [PubMed: 31553576]
49. Oberg N, Zallot R, and Gerlt JA (2023). EFI-EST, EFI-GNT, and EFI-CGFP: Enzyme Function Initiative (EFI) Web Resource for Genomic Enzymology Tools. J. Mol. Biol. 435, 168018. 10.1016/j.jmb.2023.168018. [PubMed: 37356897]
50. Yu T, Cui H, Li JC, Luo Y, Jiang G, and Zhao H (2023). Enzyme function prediction using contrastive learning. Science 379, 1358–1363. 10.1126/science.adf2465. [PubMed: 36996195]

51. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, and Finn RD (2018). HMMER web server: 2018 update. Nucleic Acids Res. 46, W200–W204. 10.1093/nar/gky448. [PubMed: 29905871]
52. Johnson SR, Weigele P, Fomenkov A, Ge A, Vincze A, Roberts RJ, and Sun Z (2024). Domainator, a flexible software suite for domain-based annotation and neighborhood analysis, identifies proteins involved in antiviral systems. Preprint at bioRxiv 2024, 590562. 10.1101/2024.04.23.590562.
53. Cappannini A, Ray A, Purta E, Mukherjee S, Boccaletto P, Moafinejad SN, Lechner A, Barchet C, Klaholz BP, Stefaniak F, and Bujnicki JM (2024). MODOMICS: a database of RNA modifications and related information. 2023 update. Nucleic Acids Res. 52, D239–D244. 10.1093/nar/gkad1083. [PubMed: 38015436]
54. Jeffryes M, and Bateman A (2018). Rapid identification of novel protein families using similarity searches. F1000Res. 7, ISCBCommJ-1975. 10.12688/f1000research.17315.1.
55. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. 10.1038/s41586-021-03819-2. [PubMed: 34265844]
56. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, and Steinegger M (2024). Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. 42, 243–246. 10.1038/s41587-023-01773-0. [PubMed: 37156916]
57. Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, Tsenkov M, Nair S, Mirdita M, Yeo J, et al. (2024). AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. Nucleic Acids Res. 52, D368–D375. 10.1093/nar/gkad1011. [PubMed: 37933859]
58. Tomikawa C, Ohira T, Inoue Y, Kawamura T, Yamagishi A, Suzuki T, and Hori H (2013). Distinct tRNA modifications in the thermo-acidophilic archaeon, Thermoplasma acidophilum. FEBS Lett. 587, 3575–3580. 10.1016/j.febslet.2013.09.021. [PubMed: 24076028]
59. Kimura S, Ikeuchi Y, Kitahara K, Sakaguchi Y, Suzuki T, and Suzuki T (2012). Base methylations in the double-stranded RNA by a fused methyltransferase bearing unwinding activity. Nucleic Acids Res. 40, 4071–4085. 10.1093/nar/gkr1287. [PubMed: 22210896]
60. Wang KT, Desmolaize B, Nan J, Zhang XW, Li LF, Douthwaite S, and Su XD (2012). Structure of the bifunctional methyltransferase YcbY (RlmKL) that adds the m7G2069 and m2G2445 modifications in Escherichia coli 23S rRNA. Nucleic Acids Res. 40, 5138–5148. 10.1093/nar/gks160. [PubMed: 22362734]
61. Hirata A, Suzuki T, Nagano T, Fujii D, Okamoto M, Sora M, Lowe TM, Kanai T, Atomi H, Suzuki T, and Hori H (2019). Distinct Modified Nucleosides in tRNA(Trp) from the Hyperthermophilic Archaeon Thermococcus kodakarensis and Requirement of tRNA m(2)G10/m(2)(2)G10 Methyltransferase (Archaeal Trm11) for Survival at High Temperatures. J. Bacteriol. 201, e00448–19. 10.1128/JB.00448-19.
62. Wein S, Andrews B, Sachsenberg T, Santos-Rosa H, Kohlbacher O, Kouzarides T, Garcia BA, and Weisser H (2020). A computational platform for high-throughput analysis of RNA sequences and modifications by mass spectrometry. Nat. Commun. 11, 926. 10.1038/s41467-020-14665-7. [PubMed: 32066737]
63. Samaha RR, Green R, and Noller HF (1995). A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome. Nature 377, 309–314. 10.1038/377309a0. [PubMed: 7566085]
64. Watson ZL, Ward FR, Meheust R, Ad O, Schepartz A, Banfield JF, and Cate JH (2020). Structure of the bacterial ribosome at 2 A resolution. Elife 9, 60482. 10.7554/eLife.60482.
65. Holvec S, Barchet C, Lechner A, Fréchin L, De Silva SNT, Hazemann I, Wolff P, von Loeffelholz O, and Klaholz BP (2024). The structure of the human 80S ribosome at 1.9 A resolution reveals the molecular role of chemical modifications and ions in RNA. Nat. Struct. Mol. Biol. 31, 1251–1264. 10.1038/s41594-024-01274-x. [PubMed: 38844527]
66. Sweeney BA, Hoksza D, Nawrocki EP, Ribas CE, Madeira F, Cannone JJ, Gutell R, Maddala A, Meade CD, Williams LD, et al. (2021). R2DT is a framework for predicting and visualising RNA secondary structure using templates. Nat. Commun. 12, 3494. 10.1038/s41467-021-23555-5. [PubMed: 34108470]

67. Consortium RNAcentral (2021). RNAcentral 2021: secondary structure integration, improved sequence search and new member databases. Nucleic Acids Res. 49, D212–D220. 10.1093/nar/gkaa921. [PubMed: 33106848]
68. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117–120. 10.1038/nature12730. [PubMed: 24284625]
69. Hofler S, and Carlomagno T (2020). Structural and functional roles of 2'-O-ribose methylations and their enzymatic machinery across multiple classes of RNAs. Curr. Opin. Struct. Biol. 65, 42–50. 10.1016/j.sbi.2020.05.008. [PubMed: 32610226]
70. Picard-Jean F, Brand C, Tremblay-Létourneau M, Allaire A, Beaudoin MC, Boudreault S, Duval C, Rainville-Sirois J, Robert F, Pelletier J, et al. (2018). 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. PLoS One 13, e0193804. 10.1371/journal.pone.0193804. [PubMed: 29601584]
71. Abou Assi H, Rangadurai AK, Shi H, Liu B, Clay MC, Erharter K, Kreutz C, Holley CL, and Al-Hashimi HM (2020). 2'-O-Methylation can increase the abundance and lifetime of alternative RNA conformational states. Nucleic Acids Res. 48, 12365–12379. 10.1093/nar/gkaa928. [PubMed: 33104789]
72. Drazkowska K, Tomecki R, Warminski M, Baran N, Cysewski D, Depaix A, Kasprzyk R, Kowalska J, Jemielity J, and Sikorski PJ (2022). 2'-O-Methylation of the second transcribed nucleotide within the mRNA 5' cap impacts the protein production level in a cell-specific manner and contributes to RNA immune evasion. Nucleic Acids Res. 50, 9051–9071. 10.1093/nar/gkac722. [PubMed: 36018811]
73. Menezes S, Gaston KW, Krivos KL, Apolinario EE, Reich NO, Sowers KR, Limbach PA, and Perona JJ (2011). Formation of m2G6 in Methanocaldococcus jannaschii tRNA catalyzed by the novel methyltransferase Trm14. Nucleic Acids Res. 39, 7641–7655. 10.1093/nar/gkr475. [PubMed: 21693558]
74. Legrand C, Tuorto F, Hartmann M, Liebers R, Jacob D, Helm M, and Lyko F (2017). Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. Genome Res. 27, 1589–1596. 10.1101/gr.210666.116. [PubMed: 28684555]
75. Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, Erlacher M, Rossmanith W, Stern-Ginossar N, and Schwartz S (2017). The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 551, 251–255. 10.1038/nature24456. [PubMed: 29072297]
76. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol. 10, 93–95. 10.1038/nchembio.1432. [PubMed: 24316715]
77. Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, Suter CM, and Preiss T (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res. 40, 5023–5033. 10.1093/nar/gks144. [PubMed: 22344696]
78. Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, Frye M, Helm M, Stoecklin G, and Lyko F (2012). RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat. Struct. Mol. Biol. 19, 900–905. 10.1038/nsmb.2357. [PubMed: 22885326]
79. Liu J, Huang T, Chen W, Ding C, Zhao T, Zhao X, Cai B, Zhang Y, Li S, Zhang L, et al. (2022). Developmental mRNA m(5)C landscape and regulatory innovations of massive m(5)C modification of maternal mRNAs in animals. Nat. Commun. 13, 2484. 10.1038/s41467-022-30210-0. [PubMed: 35513466]
80. Penning A, Jeschke J, and Fuks F (2022). Why novel mRNA modifications are so challenging and what we can do about it. Nat. Rev. Mol. Cell Biol. 23, 385–386. 10.1038/s41580-022-00485-8. [PubMed: 35422515]
81. Sud M, Fahy E, Cotter D, Azam K, Vadivelu I, Burant C, Edison A, Fiehn O, Higashi R, Nair KS, et al. (2016). Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. Nucleic Acids Res. 44, D463–D470. 10.1093/nar/gkv1042. [PubMed: 26467476]

82. Mulroney L, Wulf MG, Schildkraut I, Tzertzinis G, Buswell J, Jain M, Olsen H, Diekhans M, Corrêa IR Jr., Akeson M, and Ettwiller L (2022). Identification of high-confidence human poly(A) RNA isoform scaffolds using nanopore sequencing. RNA 28, 162–176. 10.1261/rna.078703.121. [PubMed: 34728536]

83. Hileman TH, and Santangelo TJ (2012). Genetics Techniques for Thermococcus kodakarensis. Front. Microbiol. 3, 195. 10.3389/fmicb.2012.00195. [PubMed: 22701112]

84. Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. 10.1038/nmeth.1923. [PubMed: 22388286]

85. Anders S, Pyl PT, and Huber W (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169. 10.1093/bioinformatics/btu638. [PubMed: 25260700]

86. Scott KA, Williams SA, and Santangelo TJ (2021). Thermococcus kodakarensis provides a versatile hyperthermophilic archaeal platform for protein expression. Methods Enzymol. 659, 243–273. 10.1016/bs.mie.2021.06.014. [PubMed: 34752288]

87. McLuckey SA, Van Berkel GJ, and Glish GL (1992). Tandem mass spectrometry of small, multiply charged oligonucleotides. J. Am. Soc. Mass Spectrom. 3, 60–70. 10.1016/1044-0305(92)85019-G. [PubMed: 24242838]

Highlights

- LC-MS/MS analysis revealed modified nucleoside profiles of archaeal RNAs
- Temperature-responsive RNA modifications were identified in *T. kodakarensis*
- An m^7G modification in 23S ribosomal RNA P loop and its writer enzyme were discovered

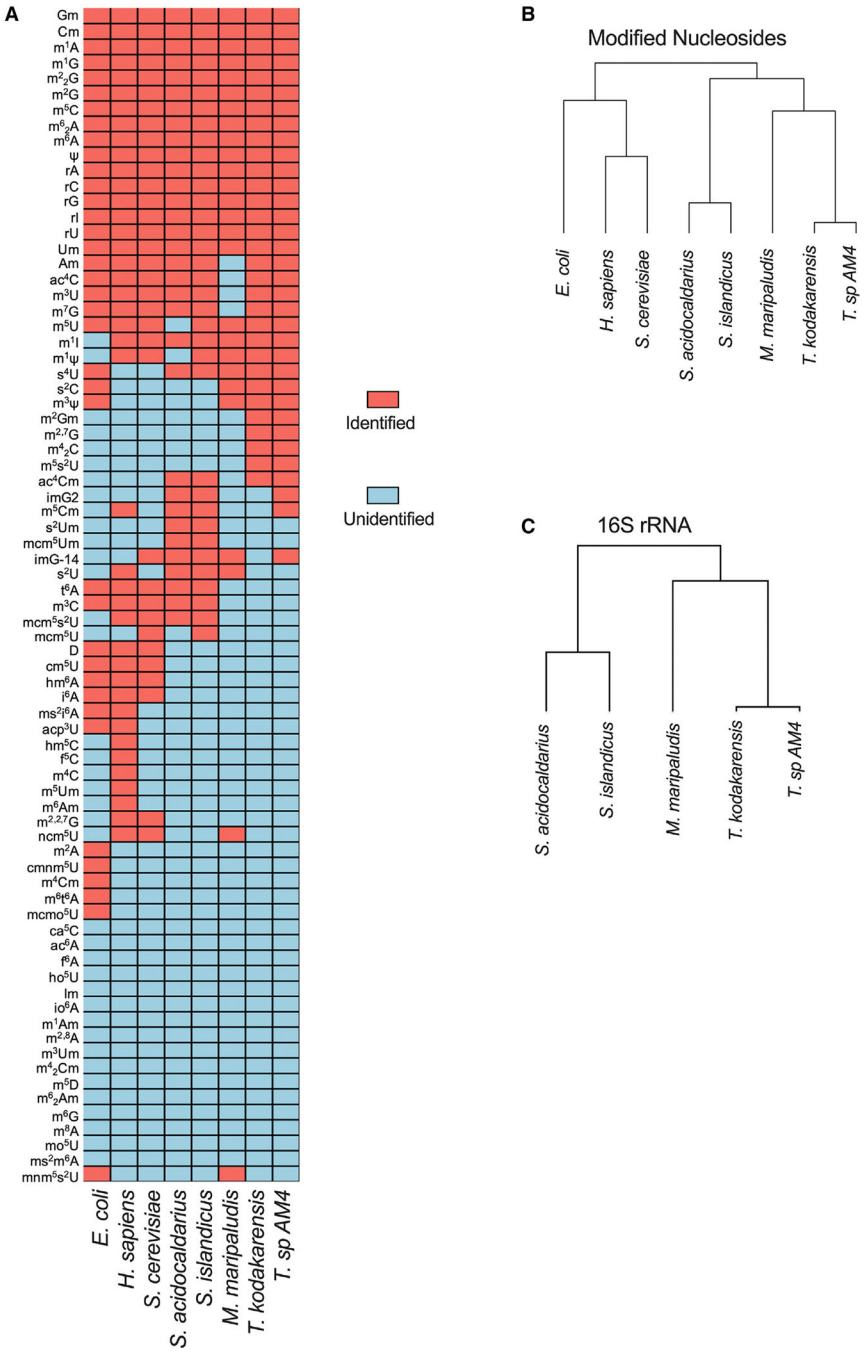


Figure 1. RNA modification profiles in archaeal and non-archaeal species

(A) Total RNA from five Archaea, *E. coli*, *S. cerevisiae* and universal human reference RNA were digested to nucleosides and subjected to UHPLC-QqQ analysis. The colored binary table summarizes the detected nucleosides for each species searched against a panel of 76 nucleoside standards.

(B) Hierarchical clustering dendrogram representing relationship between archaeal and non-archaeal species according to their nucleoside modification profiles.

(C) Unrooted tree built from 16S rRNA multiple sequence alignment of the five archaeal species.

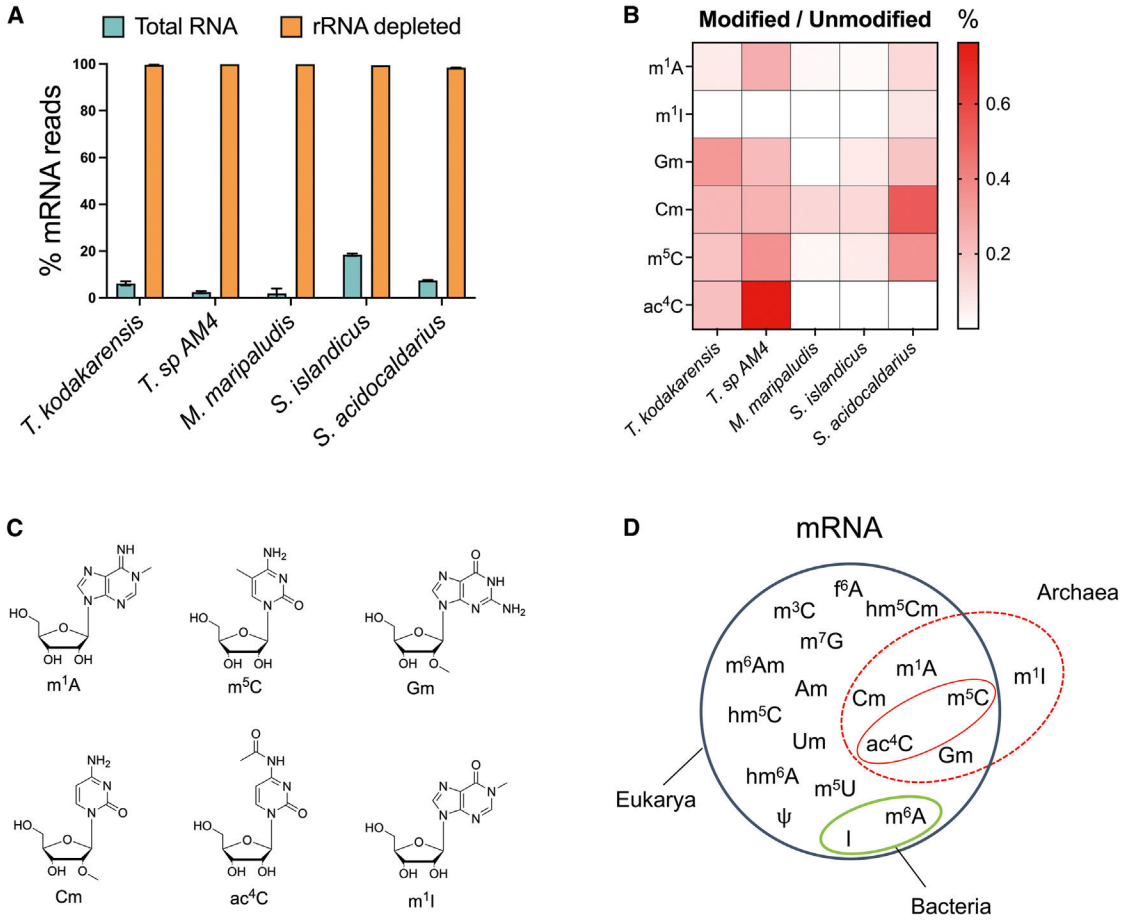


Figure 2. Detection of modified nucleosides in mRNA-enriched archaeal samples

(A) The bar graph indicates the percentage of RNA-seq reads mapped to protein-coding genes annotated in archaeal reference genomes before and after rRNA depletion from two biological replicates. Bars are displayed in mean ± S.D.

(B) UHPLC-QqQ quantification of the modified nucleosides detected in mRNA-enriched archaeal samples. The color scale of the heatmap corresponds to the mean molar ratio of modified nucleosides versus unmodified nucleosides measured from two biological replicates. The molar ratio of m^5C/C in *M. maripaludis* mRNA-enriched fraction was below quantification range in one biological replicate.

(C) Chemical structures of the modified nucleosides identified in archaeal mRNA-enriched samples.

(D) Venn diagram depicting mRNA modifications across the three domains of life. The dark blue, green, and red solid lines represent mRNA modifications previously reported in Eukarya, Bacteria, and Archaea, respectively.[35,37–41] The red dashed line indicates RNA modifications found in archaeal mRNA-enriched samples in this study.

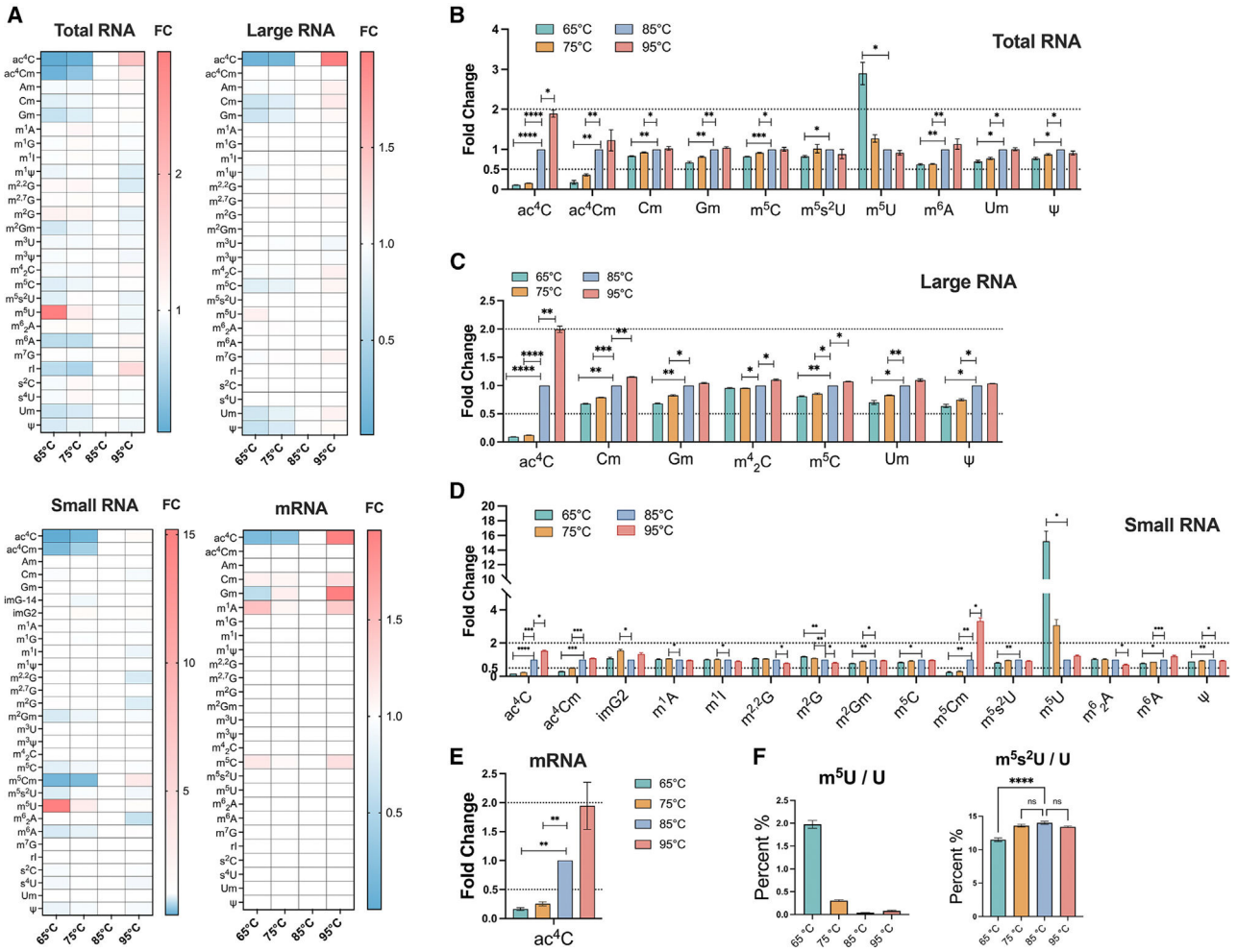


Figure 3. Temperature-dependent changes on modified nucleosides in *T. kodakarensis* RNA

(A) Nucleoside analysis of RNA subfractions from *T. kodakarensis* grown at 65°C, 75°C, 85°C, and 95°C. The heatmaps depict the mass response ratio fold changes (FCs) relative to 85°C in total RNA and different subfractions. The heatmap color scales indicate the mean mass response ratio FC measured from three biological replicates.

(B–E) Relative FCs for modified nucleosides in (B) total RNA, (C) large RNA, (D) small RNA, and (E) mRNA-enriched fractions are illustrated in the bar graphs.

(F) Relative abundance of m^5U and m^5s^2U in the small RNA fractions, as measured in three biological replicates. Statistical significance was tested with one-way ANOVA followed by Bonferroni's multiple comparisons test. All quantifications are displayed as mean ± SEM. Statistical significance: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

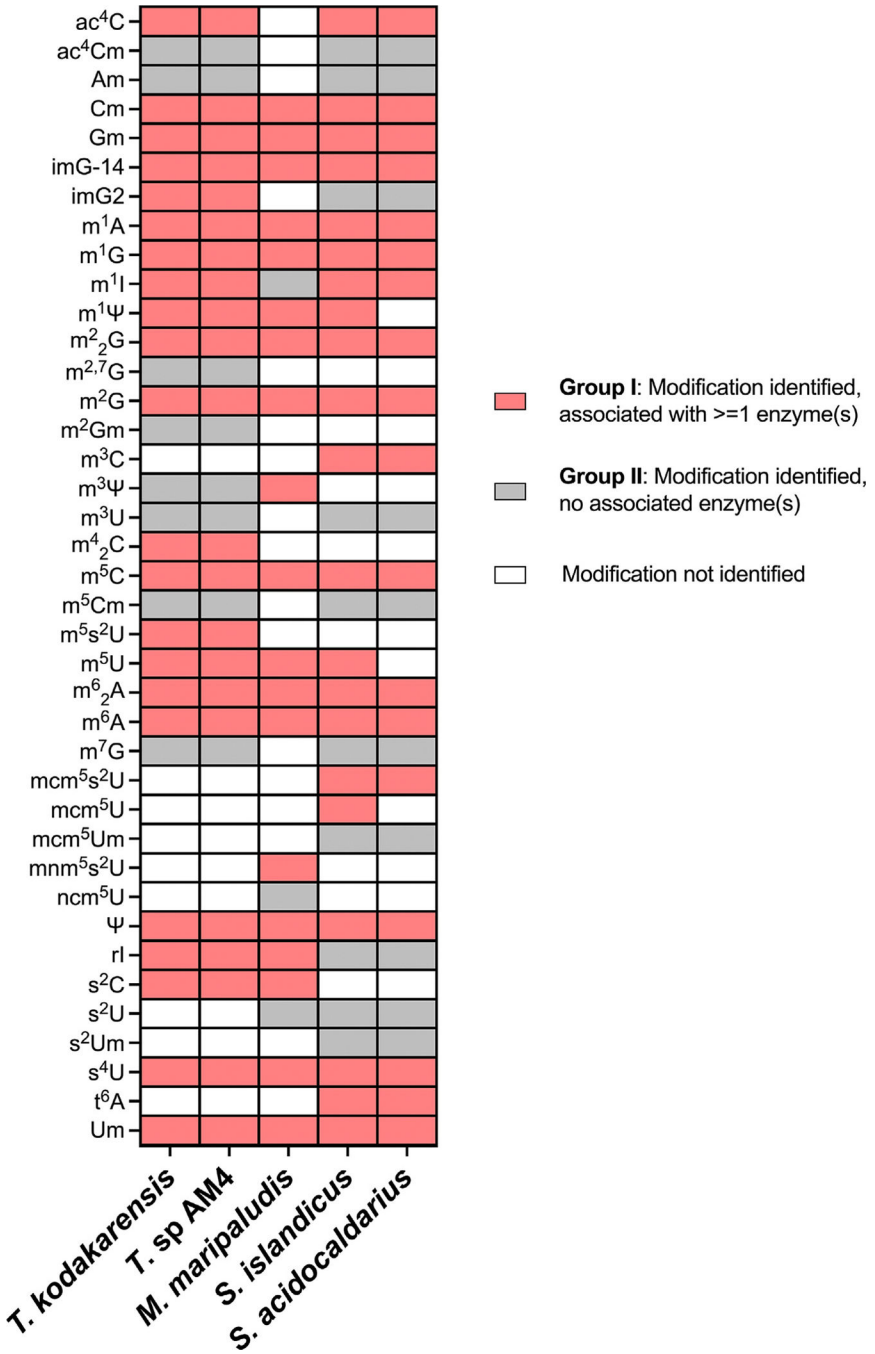


Figure 4. Association of RNA modifications with predicted archaeal RNA-modifying enzymes

Modified nucleosides were categorized into two groups based on their association with putative RNA-modifying enzymes in the reference archaeal genome, as predicted by domain organization comparisons. Group I, modifications associated with one or more RNA-modifying enzymes; group II, modifications detected in archaeal RNA, but the associated enzymes could not be predicted by domain organization comparisons.

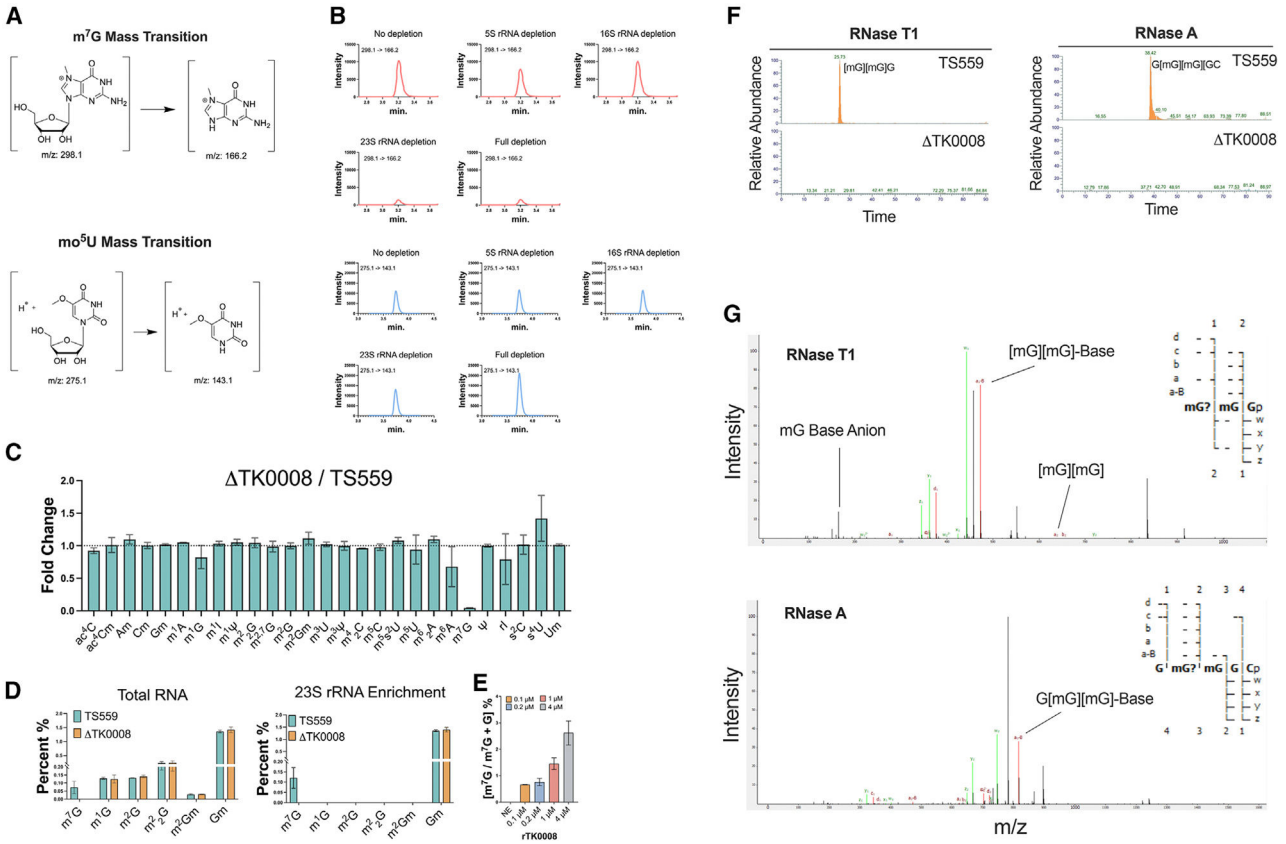


Figure 5. Identification and quantification of an m7G modification in *T. kodakarensis* ribosomal RNA

(A) Proposed precursor and product ion structures of m^7G and mo^5U.

(B) Total RNA from *T. kodakarensis* was selectively depleted of 5S, 16S, 23S, or of all three rRNAs. The resulting RNA samples were subjected to UHPLC-QqQ nucleoside analysis. Representative LC-MS/MS chromatograms of m^7G (298.1 → 166.2, red trace) and mo^5U (275.1 → 143.1, blue trace) in each RNA sample.

(C) The mass response change for all modifications found in total RNA of the ΔTK0008 strain relative to the wild-type TS559. Bar graphs are displayed in mean ± SD from two biological replicates.

(D) Relative abundance of the guanosine modifications m^7G, m^1G, m^2G, m^2_2G, and m^2Gm over unmodified guanosine measured in TS559 and ΔTK0008 strains. Enrichment of 23S rRNA was achieved by depleting 5S and 16S rRNAs. m^1G, m^2G, m^2_2G, and m^2Gm were not detectable after 5S and 16S rRNA depletion. Bar graphs show mean ± SD from two biological replicates.

(E) m^7G nucleoside quantification in IVT 23S rRNA treated with rTK0008 (0.1, 0.2, 1, 4 μM). NE, no-enzyme control. Bar graph shows mean ± SD from two biological replicates.

(F) 23S rRNA-enriched samples of TS559 and ΔTK0008 strains were digested with RNase T1 or RNase A at 37°C for 1 h. The digested RNAs were subjected to nanoLC-MS/MS analysis. The panel displays representative MS1 ion chromatograms of mG-containing oligonucleotides at −2H charge state.

(G) (Upper) Representative chromatograms of mG base anion and a-B ion ([mG][mG]-Base) derived from the [mG][mG]Gp trinucleotide. (Lower) Representative chromatograms of a-B ion (G[mG][mG]-Base) derived from the G[mG][mG]GCp pentanucleotide.

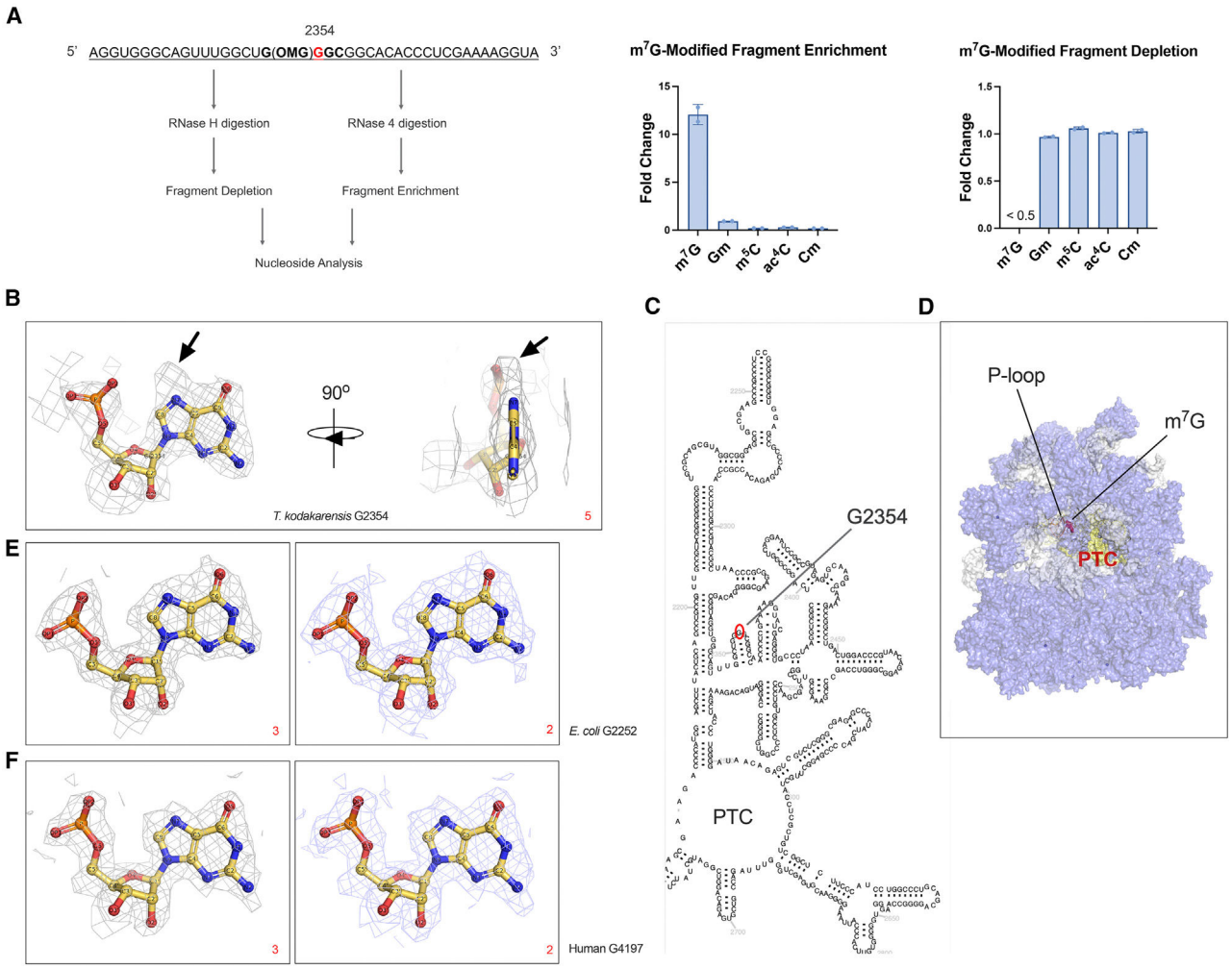


Figure 6. m7G modification resides in the P loop of 23S rRNA

(A) Biotinylated DNA probe-targeted sequence in 23S rRNA containing the prospective m^7G site. The probe-hybridized region was either depleted by RNase H digestion or enriched by RNase 4 protection cleavage. The RNase 4-treated samples were subjected to streptavidin pull-down. UHPLC-QqQ nucleoside analysis was subsequently performed on RNase H- and RNase 4-treated samples to determine the fold change in m^7G abundance relative to total RNA.

(B) Electron microscopy (EM) density of m^7G (2354) in the cryo-EM structure visualized in Pymol. Contour level is indicated in red at the lower-right corner of the panel. The dark arrows indicate the observed density on the *N7* position. EM map: EMD-10223.

(C) Secondary structure of the domain V region of *T. kodakarensis* 23S rRNA.[66,67] The circled guanosine at position 2354 represents the site for m^7G modification. PTC, peptidyl transferase center.

(D) Structural view of m^7G-modified position within *T. kodakarensis* 70S ribosome. Cryo-EM structure PDB: 6SKF.

(E and F) (E) EM density of G2252 in the 23S rRNA of *E. coli* and (F) G4197 in the 28S rRNA of human ribosome. Contour levels are indicated in red at the lower-right corner of each panel. EM maps: EMD-22586, EMD-18539.

Table 1.

Unique and shared RNA modifications found in archaeal samples

All species tested	Thermophiles	*Thermococcus*	*Sulfolobus*	*M. maripaludis*
Cm	Am	m^2Gm	m^3C	mnm^5s^2U
Gm	ac^4C	$m^{2,7}G$	mcm^5s^2U	ncm^5 U
m^1A	ac^4Cm	$m^4{}_2C$	mcm^5 U	-
m^1G	m^3U	m^5s^2U	mcm^5 Um	-
m^1I	m^7G	-	s^2Um	-
m^1 Ψ	-	-	t^6A	-
$m^2{}_2G$	-	-	-	-
m^2G	-	-	-	-
m^5C	-	-	-	-
m^5U	-	-	-	-
$m^6{}_2A$	-	-	-	-
m^6A	-	-	-	-
rI	-	-	-	-
s^4U	-	-	-	-
Um	-	-	-	-
Ψ	-	-	-	-

mcm^5 U was only detected in *S. islandicus*.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Universal human reference RNA	Agilent Technologies	Cat #740000
Escherichia coli total RNA	Thermo Fisher Scientific	Cat #AM7940
Saccharomyces cerevisiae total RNA	Mulroney et al.[82]	N/A
Chemicals, peptides, and recombinant proteins		
Nucleoside standards	Multiple suppliers	Data S1
TK0008	This paper	N/A
Deposited data		
Cryo-EM of *T. kodakarensis* 70S ribosome	Sas-Chen et al., 2020[3]	PDB entry: 6SKF
Cryo-EM of *E. coli* 70S ribosome	Watson et al., 2020[64]	PDB entry: 7K00
Cryo-EM of Human 80S ribosome	Holvec et al., 2024[65]	PDB entry: 8QOI
RNA-Seq data	This study	PRJNA1219756
Nucleoside analysis data	This study	PR002314
Oligonucleotide analysis data	This study	PXD060644
Experimental models: Organisms/strains		
Thermococcus kodakarensis	This paper	N/A
Thermococcus sp. AM4	C.S. Raman, University of Maryland	Data S2
Methanococcus maripaludis	Barney Whitman, University of Georgia	Data S2
Sulfolobus islandicus	Rachel Whitaker, University of Illinois at Urbana-Champaign	Data S2
Sulfolobus acidocaldaius	Sonja-Verena Albers, University of Freiburg	Data S2
NiCo21(DE3) Competent *E. coli*	New England Biolabs	Cat #C2529H
Oligonucleotides		
rRNA depletion probes	Integrated DNA Technologies	Table S8
Biotinylated DNA probe	Integrated DNA Technologies	N/A
Recombinant DNA		
pTS700	Hileman and Santangelo[83]	N/A
pCSU-TK0008A	This paper	N/A
pCSU-TK0008B	This paper	N/A
pQE80	Takara Bio	Cat #638910
Software and algorithms		
Agilent MassHunter WorkStation Quantitative Analysis (QQQ)	Agilent Technologies	https://www.agilent.com/en/product/software-informatics/mass-spectrometry-software/data-analysis/quantitative-analysis

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BioPharma Finder (version 5.1)	Thermo Fisher Scientific	https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/liquid-chromatography-mass-spectrometry-lc-ms/lc-ms-software/multi-omics-data-analysis/biopharma-finder-software.html
NucleicAcidSearchEngine	Wein et al.[62]	N/A
Prism 10	Dotmatics	https://www.dotmatics.com/solutions/prism
Geneious Prime 2024.0.7	Dotmatics	https://www.dotmatics.com/solutions/geneious-prime
R (4.4.2)	The R Foundation	https://www.r-project.org/
PyMOL	Schrödinger, Inc.	https://www.pymol.org/
Bowtie2	Langmead and Salzberg[84]	https://bowtie-bio.sourceforge.net/bowtie2/manual.shtml
HTSeq	Anders et al.[85]	https://htseq.readthedocs.io/en/release_0.11.1/index.html