

Site-Specific Profiling of 4-Thiouridine Across Transfer RNA Genes in *Escherichia coli*

Praneeth Bommiseti and Vahe Bandarian*

Cite This: *ACS Omega* 2022, 7, 4011–4025

Read Online

ACCESS |



Metrics & More

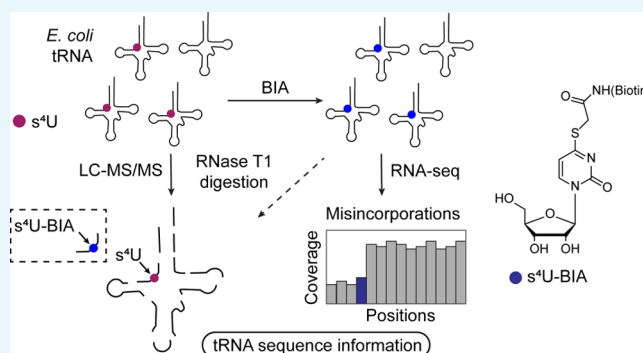


Article Recommendations



Supporting Information

ABSTRACT: The transfer RNA (tRNA) modification 4-thiouridine (s^4U) acts as a near-ultraviolet (UVA) radiation sensor in *Escherichia coli* (*E. coli*), where it induces a growth delay upon exposure to the UVA radiation (~ 310 – 400 nm). Herein, we report sequencing methodology for site-specific profiling of s^4U modification in *E. coli* tRNAs. Upon the addition of iodoacetamide (IA) or iodoacetyl-PEG2-biotin (BIA), the nucleophilic sulfur of s^4U forms a reaction product that is extensively characterized by liquid chromatography–mass spectrometry (LC–MS/MS) analysis. This method is readily applied to the alkylation of natively occurring s^4U on *E. coli* tRNA. Next-generation sequencing of BIA-treated tRNA from *E. coli* revealed misincorporations at position 8 in 19 of the 20 amino acid tRNA species. Alternatively, tRNA from the $\Delta thil$ strain, which cannot introduce the s^4U modification, does not exhibit any misincorporation at the corresponding positions, directly linking the base transitions and the tRNA modification. Independently, the s^4U modification on *E. coli* tRNA was further validated by LC–MS/MS sequencing. Nuclease digestion of wild-type and deletion strains *E. coli* tRNA with RNase T1 generated smaller s^4U/U containing fragments that could be analyzed by MS/MS analysis for modification assignment. Furthermore, RNase T1 digestion of tRNAs treated either with IA or BIA showed the specificity of iodoacetamide reagents toward s^4U in the context of complex tRNA modifications. Overall, these results demonstrate the utility of the alkylation of s^4U in the site-specific profiling of the modified base in native cellular tRNA.



INTRODUCTION

A common characteristic among cellular ribonucleic acids, such as ribosomal RNA (rRNA), transfer RNA (tRNA), and messenger RNA (mRNA), is the presence of post-transcriptional modifications. While these RNA species possess numerous post-transcriptional modifications, the numbers and diversity of modifications are the most significant in tRNA.¹ To date, a total of 108 modifications have been reported in tRNA, and the presence of these modifications across all domains of life highlights their importance.² The anticodon stem-loop region on tRNA contains most of these modifications, presumably because of its role in codon–anticodon interactions during protein synthesis, with modifications at positions 34 and 37 being most prevalent.^{1,3}

The discovery of many nucleic acid modifications coincided with the efforts in the late 1960s and early 1970s to determine the nucleic acid sequence, the roles of quite a few remaining enigmatic to this day. The emerging consensus is that many are not limited to RNA. For example, 7-deazapurine-based modified bases, which were initially discovered in the wobble positions of tRNA Asp, Asn, Tyr, and His, have recently been identified in bacteriophage DNA. Their presence is thought to protect the phage DNA from the host restriction endonuclease response.^{4,5} There is also cross-talk between DNA and RNA

modifications in eukaryotes, where a class of DNA methylating enzymes alkylate certain tRNA.⁶ The dynamics of cellular insertion, detection, and removal of these modifications represent an important dimension in understanding their physiological roles. While detecting the modified bases is relatively trivial using high-resolution analytical separation and mass spectrometry, site-specific profiling methods remain challenging.

The advent of commercially available high-throughput next-generation sequencing (NGS) technologies during the last 2 decades has opened new research avenues for profiling post-transcriptional RNA modifications.⁷ In these experiments, the conversion of RNA to complementary DNA (cDNA) in the reverse transcription (RT) step serves as a readout where a non-canonical nucleoside in the RNA template can result in the misincorporation of deoxynucleotides, premature termi-

Received: September 13, 2021

Accepted: January 11, 2022

Published: January 27, 2022



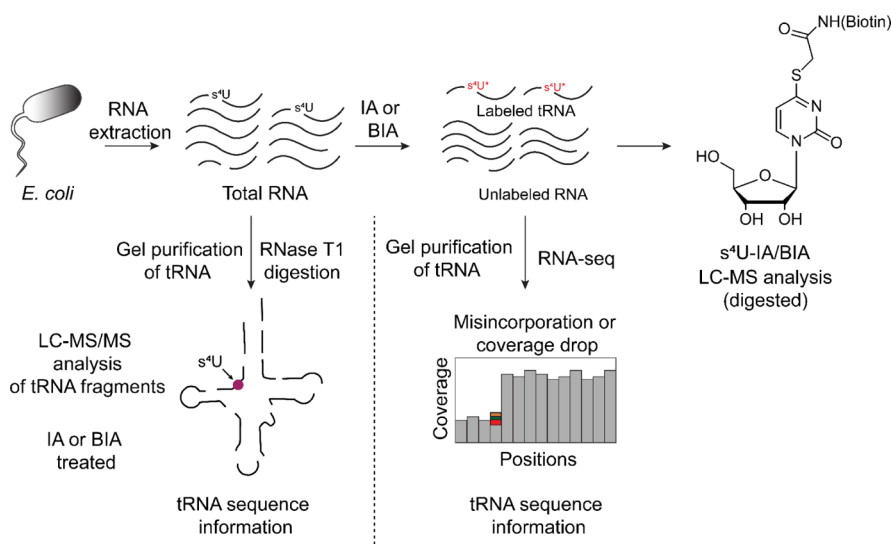


Figure 1. Overview of the workflow for the detection of s^4U using RNA-Seq and mass spectrometry. In this workflow, extracted RNA is treated with IA or BIA prior to RNA-Seq to identify misincorporation or coverage changes. Likewise, tRNA was nuclease digested with RNase T1 followed by LC–MS/MS analysis of tRNA fragments which independently establish the assignment. RNase T1 digestion followed by LC–MS/MS analysis of either IA- or BIA-treated tRNA establish the specificity of iodoacetamide reagents toward s^4U in the context of complex tRNA modifications.

nation, or both during the RT.⁷ While some non-canonical bases such as inosine (I) may induce these directly, others require treatment with a chemical reagent before they lead to a distinct readout.⁷ For example, the use of chemical reagents specific to pseudouridine (ψ) and 5-methylcytosine (m^5C) facilitate transcriptome-wide detection of these modifications.^{8–10} Studies from different laboratories also show that NGS conditions are responsive to many tRNA modifications, where empirical analysis of misincorporations and terminations (collectively known as RT events) can predict the presence of base modifications in uncharacterized bacterial tRNA genes.^{11,12} However, in many cases, these require extensive statistical analysis to uncover as the probabilities for the misincorporation or terminations are low. Ideally, chemical reagents that target specific modifications and increase RT event probabilities could overcome these limitations. However, the diversity of the modifications requires the development of a new procedure that leverages the unique reactivity of the modified base in each case.

Mass spectrometry could be used as an alternative technique for detecting RNA modifications, presenting a more straightforward approach to the sequence-specific analysis of RNAs.¹³ The principles used in such methods are adapted from proteomics, where upon protease specific digestion of a protein generates a smaller peptide fragment library that could subsequently be analyzed using liquid chromatography (LC)–mass spectrometry (MS/MS) analysis.⁷ Similarly, the cleavage of RNA to smaller fragments using specific nuclease generates a smaller RNA library that LC could separate. The MS/MS fragmentation of eluting peaks can be compared against the theoretically generated fragmentation data from the input sequences to detect the modifications on the RNA.^{14–17} While these methods are much more direct than NGS, they are limited by the need for specialized instrumentation and the lack of robust tools for analyzing MS fragmentation data. In the current report, using the tRNA as a model system for 4-thiouridine (s^4U) containing RNA species, we developed and validated a methodology to site specifically profile the base in

the stationary growth phase of *Escherichia coli* (see Figure 2 for the s^4U structure).

Native profiling of tRNA modifications shows that RNA-Seq conditions are sensitive to the presence of the s^4U , where it causes low rates of mismatches.^{11,12,18} The s^4U alone displays empirical mismatch rates between 0.1 and 0.3, which may not be conclusive when applied in a context where the existence of s^4U is unknown. For example, while s^4U is commonly present at the 8th or 9th position of bacterial tRNA, in the case of archaeal tRNA or uncharacterized bacterial tRNA, limitations may occur in confidently assigning the s^4U position on the gene based on the statistical probabilities alone.

The nucleophilicity of the thiouridine base and its reactivity with thiol modification reagents is a viable method for introducing a bulky modification that could potentially report its presence in NGS sequencing experiments.^{19–22} In the SLAM-seq method, s^4U is metabolically incorporated into the eukaryotic mRNA pool in pulse-chase experiments, which is subjected to NGS after iodoacetamide (IA) treatment. The dynamics are monitored by exploring the frequency of mismatches observed in 3′-untranslated regions of mRNA. However, under these conditions, nearly $\sim 0.5\%$ of the total cellular RNA is labeled with s^4U after 6 h of incubation with 0.1 mM s^4U .²³

We envisioned a methodology, starting from *E. coli* cells using RNA-Seq, to detect s^4U modification in a site-specific manner (Figure 1). To achieve this, the documented susceptibility of s^4U to electrophilic reagents, such as IA, was explored.^{21,22} We reasoned that the presence of a large adduct, such as biotin-IA, would lead to a robust mismatch response in NGS, allowing single-nucleotide resolution s^4U detection. This article reports the characterization of s^4U -IA or s^4U -iodoacetyl-PEG2-biotin (BIA) products by high-resolution MS methods to establish conditions that maximized the yield of labeled s^4U . Next, these methods were extended to bulk RNA from *E. coli* and NGS experiments using size-selected RNA to enrich tRNA. Remarkably, the base transitions observed in the NGS data clearly show that most detectable tRNAs in *E. coli* have the modification at position 8 or 9 with robust probabilities.

These base transitions are absent in identically treated samples from cells where a key tRNA sulfur insertion enzyme, ThiI, was absent. Finally, LC–MS/MS analysis of untreated, IA or BIA-treated tRNA from wild type and deletion *E. coli* strains digested with RNase T1 nuclease corroborate the positional modification data from the high-throughput NGS studies. Together, these data provide a workflow for single-base resolution sequencing for $s^4\text{U}$ on bacterial or archaeal tRNAs.

RESULTS AND DISCUSSION

The overall outline of the studies presented in this article is shown in Figure 1. To ensure that the sequencing analyses' data reflect the $s^4\text{U}$ modification, we first carried out several control experiments to characterize the BIA modification reaction.

Characterization of Reaction between $s^4\text{U}$ and IA (or BIA). When combined with IA, $s^4\text{U}$ carries out a nucleophilic attack via the sulfur atom at the C-2 of IA (or BIA), as shown in Figure 2 (or Figure S1). Therefore, $s^4\text{U}$ was incubated with

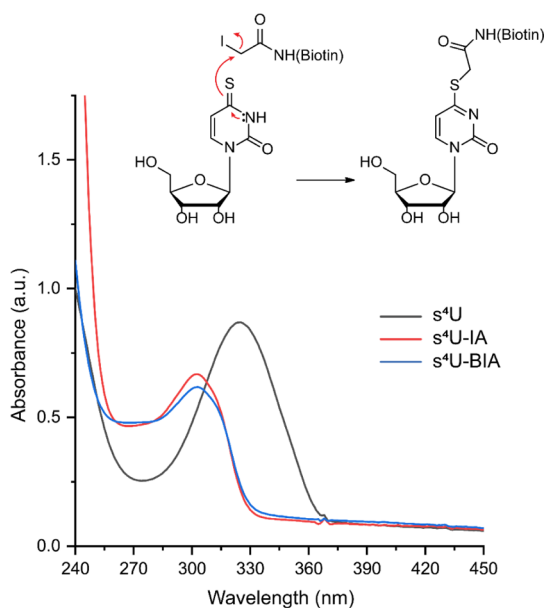


Figure 2. UV–visible spectrophotometric analysis IA-modified $s^4\text{U}$. The spectra are of 20 μM authentic $s^4\text{U}$ and modified $s^4\text{U}$ after treatment with IA or BIA. The spectra were obtained in 0.05 M NaPi (pH 8.0).

excess IA or BIA, and the reaction mixture was subsequently analyzed by UV–visible spectroscopy. Figure 2 shows the absorbance spectra of $s^4\text{U}$ (black trace) with an absorption maximum at ~ 330 nm, $s^4\text{U}$ treated with IA (red trace), and $s^4\text{U}$ treated with BIA (blue trace), both of which have an absorption maximum at ~ 303 nm. The ~ 27 nm blue shift is consistent with a previous report.²³

To confirm that the incubation of $s^4\text{U}$ with IA leads to the formation of the expected product, reactions were analyzed by LC–MS (Figure S2). Under these conditions, $s^4\text{U}$ elutes at 11.8 min. The mass spectrum of the species eluting at 11.8 min contains two species, corresponding to the m/z values of [$s^4\text{U} + \text{Na}^+$] and [$s^4\text{U} + \text{K}^+$], within six ppm of the theoretical m/z values (Figure S2C). In the presence of IA, a new peak at 12.3 min appears. The mass spectrum of the species eluting in this peak is consistent with the m/z of [$s^4\text{U-IA} + \text{H}^+$], [$s^4\text{U-IA} +$

Na^+], and [$s^4\text{U-IA} + \text{K}^+$], which are within seven ppm of the theoretical values (Figure S2B).

To assess the specificity of IA for $s^4\text{U}$ versus the four canonical nucleobases, $s^4\text{U}$ was treated with IA in the presence of equimolar quantities of A, U, G, and C. The reaction mixtures were subsequently analyzed by HPLC–MS (Figure S3). Under these conditions, in the presence of IA, the $s^4\text{U}$ peak at 10.9 min is replaced with one at 11.5 min. The canonical nucleosides remained unaffected. The UV–visible spectra corresponding to each of $s^4\text{U}$ and $s^4\text{U-IA}$ peaks (Figure S3B,C) are consistent with those shown in Figure 2. As is apparent, no other changes are visible in the chromatograms. In passing, we note that a new peak at 16 min is also observed in these experiments, which has the same UV–visible features as $s^4\text{U-IA}$. We do not know its identity, and because it was a minor component, it was not investigated further.

To determine if BIA reacted with $s^4\text{U}$ in a similar manner, $s^4\text{U}$ was reacted with BIA and analyzed via LC–MS experiments. The extracted ion chromatogram (EIC) shows that $s^4\text{U}$ (m/z —261.05) and BIA (m/z —542.10) elute at 13.5 and 34.3 min, respectively (Figure 3A). The mass spectrum corresponding to the species eluting at 13.5 min is a mixture of [$s^4\text{U} + \text{H}^+$], [$s^4\text{U} + \text{Na}^+$], and [$s^4\text{U} + \text{K}^+$], as observed previously (compare Figures 3C and S2C). The observed m/z values for these species are within three ppm of theoretical. The reaction product elutes at 33.03 min, and the mass spectrum of the peak exhibits a [$s^4\text{U-BIA} + \text{H}^+$] species with m/z of 675.25 (Figure 3B), which is within 0.3 ppm of the expected mass.

The peaks corresponding to the $s^4\text{U}$, $s^4\text{U-IA}$, and $s^4\text{U-BIA}$ were subjected to high-resolution MS/MS analysis to characterize the species further. A representative example MS–MS spectrum for the $s^4\text{U-BIA}$ adduct is shown in Figure S4. Fragmentation of $s^4\text{U-BIA}$ via CID at 12–15 eV results in daughter ions with an m/z of 565.19 arising from the loss of ribose sugar (peak (a), Figure S4). Further fragmentation of the peak (a) leads to a loss of the nucleobase with an m/z of 397.19 [peak (b), Figure S4]. Experimental m/z values for the daughter ions are within eight ppm of theoretical m/z values calculated for those ions.

To further characterize the modification reaction of $s^4\text{U}$ to $s^4\text{U-BIA}$ unambiguously, large-scale modification of $s^4\text{U}$ with BIA was carried out, purified by preparative HPLC conditions, and subjected to ^1H NMR experiments (see additional methods in the Supporting Information). The proton NMR spectra of commercially sourced $s^4\text{U}$ and purified $s^4\text{U-BIA}$ are shown in Figures S5 and S6, respectively. The spectral assignments for $s^4\text{U}$ and $s^4\text{U-BIA}$ were made by reference to uridine, biotin, and the linker ethylene glycol ^1H NMR spectra found on publicly available biological magnetic resonance data bank (BMRB).²⁴ As seen from the $s^4\text{U-BIA}$ proton spectrum (Figure S6), resonance peaks for uracil aromatic protons, ribose protons, biotin protons, and polyethylene glycol protons can be seen clearly. These data and MS/MS fragmentation data collectively establish that the incubation of $s^4\text{U}$ with IA and BIA leads to the covalent modification of the nucleoside, establishing the feasibility of using these as chemical biology tools.

IA or BIA Modification of RNA. Encouraged by the efficiency of the in vitro modification data with standards, we attempted to extend the studies to RNA. In these experiments, total RNA extract from wild-type *E. coli* was treated either with IA or BIA, and the resulting sample was digested with

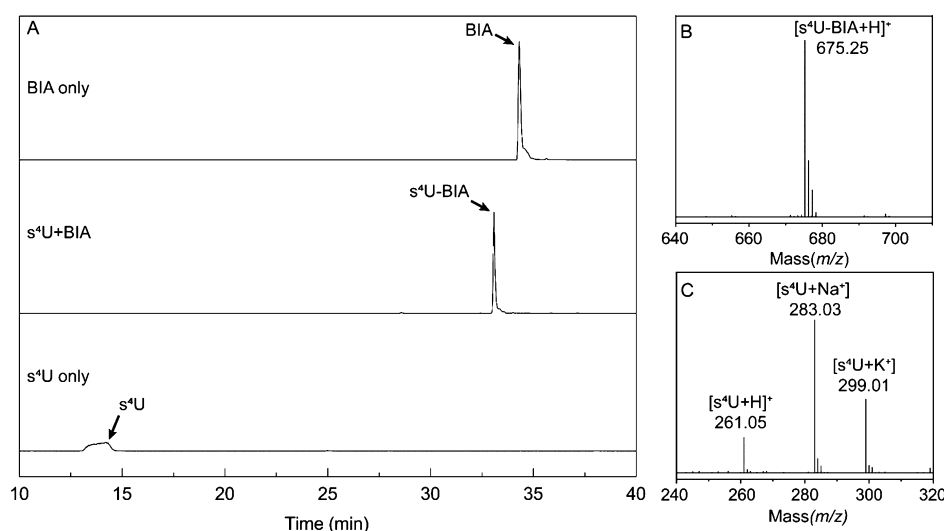


Figure 3. LC–MS analysis of the in vitro modification of s⁴U with BIA. (A) EIC of s⁴U (m/z range 260.55–261.55), s⁴U-BIA (m/z range 674.75–675.75), and BIA (m/z range 541.60–542.60) samples show that upon the reaction with BIA, a new peak at 33 min appears in the chromatogram, which is distinct from that observed for unreacted s⁴U or BIA. (B) Mass spectrum of s⁴U-BIA. (C) Mass spectrum observed for s⁴U.

nucleosides by subsequent actions of P1 nuclease, phosphodiesterase, and alkaline phosphatase enzymes.²⁵ The UV–visible traces from the HPLC analysis of wild-type *E. coli* RNA with or without IA treatment are shown in Figure S7. In addition to the C, U, G, and A, the traces reveal a clear peak for s⁴U at 10.5 min, with the expected UV–visible spectrum for the nucleoside. A time course for the modification of s⁴U to s⁴U-IA from total RNA is shown in Figure S8. After 2 h of incubation with IA, the signal for s⁴U is replaced with one at 11.3 min, corresponding to the s⁴U-IA adduct. The data show that the modification is essentially complete by 4 h. Therefore, the reactions were allowed to proceed overnight (12–14 h). The identity of the s⁴U-IA adduct is further confirmed by examining the mass spectrum of the species eluting in the 11.3 min peak (Figure S7B). The observed m/z values are within 10–11 ppm of theoretical and identical to those observed with the standards (compare mass spectrum insets from Figures S2 and S7). The studies described for the modification of s⁴U with IA were extended to BIA. As with the above, the BIA-treated RNA was analyzed by LC–MS analysis. The UV–vis (190–800 nm) absorbance traces show a peak for s⁴U-BIA [trace (a), Figure 4A], though the peak for its precursor s⁴U co-elutes with G [trace (b), Figure 4A]. Because of the hydrophobic nature of BIA, a modified elution method was used, which did not separate G and s⁴U as well. However, when we examine the data at 330 nm, which corresponds to the absorbance maximum for s⁴U [Figure 4B, trace (f), see Figure 2], a peak corresponding to the nucleoside is visible at 13.4 min in the untreated samples. By contrast, when the traces are examined at 305 nm, which corresponds to the absorbance maximum for the alkylated nucleoside [Figure 4B, trace (c), see Figure 2], the peak corresponding to the BIA modified nucleoside at 33 min is prominent in treated samples. We note that the retention times observed for the adduct and unmodified nucleoside are consistent with those observed with the commercially obtained s⁴U (Figure 3A). Finally, the mass spectrum of the s⁴U-BIA corresponds (within 3 ppm) to that expected for the adduct (Figure 4C).

The s⁴U and s⁴U-BIA originating from biological RNA samples showed similar retention times to commercially

sourced s⁴U and purified s⁴U-BIA characterized by ¹H NMR (Figures S9, see S5 and S6). Additionally, the degradation of RNA treated by this protocol was examined using denaturing poly-acrylamide gel electrophoresis (PAGE) as described in additional methods in the Supporting Information. RNA samples taken at various time points were analyzed by PAGE and represented in Figure S10. The gel image shows minimal (if any) degradation of the total RNA treated in the protocol. Together these results validate the use of the mentioned conditions in an NGS workflow.

BIA Modification of RNA from Control Strains. Before applying the IA methodology in sequencing experiments, we conducted experiments to specifically implicate the cellular sulfur insertion machinery in the modification (Figure S11).²⁶ In the biosynthetic pathway of s⁴U, the PLP-dependent enzyme IscS mobilizes S from Cys as a persulfide attached to a Cys residue in the protein. The persulfide donates the sulfur to ThiI, which in an ATP-dependent reaction incorporates the sulfur into tRNA. In these experiments, total RNA from wild type and Δ iscS and Δ thiI strains of *E. coli* was isolated, reacted with BIA, and digested to nucleosides. The resulting mixtures were analyzed by HPLC, as described in Materials and Methods (Figure 5). Interestingly, the sample from the Δ iscS variant appears to contain s⁴U [trace (c), Figure 5], which elutes similarly to the nucleoside as in the wild-type samples [trace (a), Figure 5] and exhibits UV–visible features that are characteristic of s⁴U. Perhaps this is not surprising, as *E. coli* have several overlapping systems for sulfur mobilization from cysteine.²⁷ By contrast, the Δ thiI variant lacks s⁴U [trace (e), Figure 5]. When compared to wild-type-treated samples [trace (b), Figure 5], the corresponding BIA-treated samples from the Δ iscS variant clearly show the appearance of s⁴U-BIA [trace (d), Figure 5], whereas the Δ thiI sample shows no evidence for the presence of s⁴U-BIA [trace (f), Figure 5]. These data establish that the BIA-modification specifically highlights a ThiI-dependent process, which previous studies have shown to be the incorporation of S to form s⁴U.²⁶ Finally, the data shown in Figure 5 allow one to estimate that under the conditions of the experiments, the modification is nearly quantitative ($95 \pm 4\%$) in several biological and technical

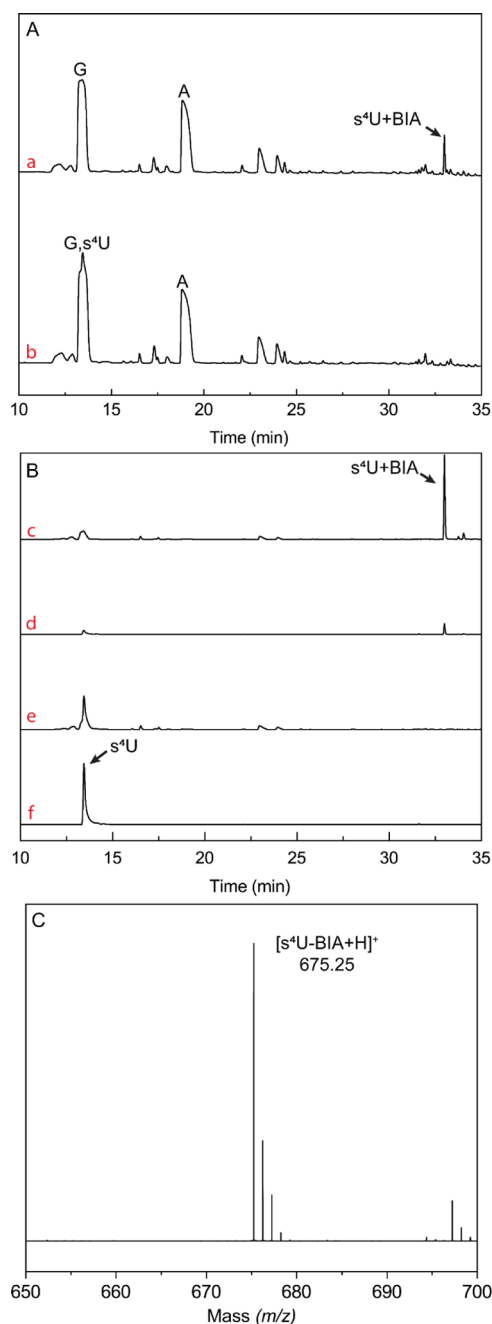


Figure 4. LC–MS analysis of nucleosides in WT *E. coli* total RNA treated with BIA. *E. coli* RNA was digested to nucleosides and analyzed by LC–MS. (A) UV–visible traces (190–800 nm) of nucleosides from BIA-treated *E. coli* RNA [trace (a)] and from untreated *E. coli* RNA [trace (b)]. Under these conditions, s^4U -BIA elutes at 33 min while its precursor s^4U elutes with G at 13 min. (B) UV–visible traces were examined at 305 nm for s^4U -BIA and at 330 nm for s^4U in *E. coli* RNA treated with BIA [traces (c) and (d), respectively]. The chromatograms show a very diminished peak for s^4U , as compared to untreated *E. coli* RNA samples at 305 and 330 nm [traces (e) and (f), respectively]. The y-axis scale for traces (a) and (b) differs from that of the scale for traces (c–f) and the RNA inputs were identical. (C) Mass spectrum observed for s^4U -BIA.

replicates (representative data are shown in Figure 5). Therefore, the reaction is sufficiently robust for s^4U sequencing.

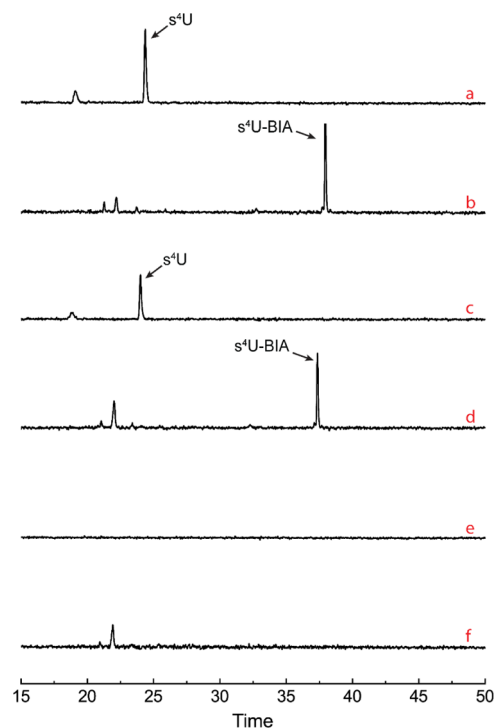


Figure 5. Representative analysis of RNA from $\Delta thiI$ and $\Delta iscS$ *E. coli* deletion strains for the presence of s^4U . The HPLC traces extracted at 330 nm (corresponding to the maximum absorbance for s^4U) of samples containing total RNA from WT and $\Delta iscS$ strains [traces (a), (b), and (c), respectively] show that the deletion of *iscS* does not eliminate s^4U but that the deletion of *thiI* does [trace (e)]. The traces at 305 nm (corresponding to the maximum absorbance for s^4U -BIA) of samples containing total RNA WT and $\Delta iscS$ strains, [traces (b) and (d), respectively] clearly show the presence of s^4U -BIA. By contrast, RNA from the *thiI* deletion strain does not show any s^4U -BIA [trace (f)], which is consistent with the loss of s^4U in the corresponding control [see trace (e)]. To facilitate a direct comparison of intensities at 330 and 305 nm, all the y-axes are on the same scale.

We considered using the BIA to enrich samples for the s^4U -containing species before the sequencing (discussed below) using affinity pull-down (see additional methods in the Supporting Information for experimental details). To this end, BIA-treated RNA from the wild-type, $\Delta iscS$, and $\Delta thiI$ strains were incubated with streptavidin beads and eluted with excess biotin under denaturing conditions, as described in the Materials and Methods section. The resulting samples were then analyzed by denaturing PAGE (Figure S12). Samples lacking the BIA treatment served as controls. In each set of samples, we see an RNA fragment (~ 80 bp) being enriched only in the BIA-containing samples. However, we see a similar band in the $\Delta thiI$ sample as well. At this point, it is not known whether the sample is identical to that observed with the wild type and $\Delta iscS$ RNA. We cannot rule out the possibility of a hitherto unidentified nucleoside, which is sensitive to treatment with IA-based reagents. Considering these results, we did not attempt any further enrichment experiments and moved forward with sequencing both unenriched and enriched samples resulting from the treatment of RNA with BIA, as described below.

Small RNA Sequencing. RNA sequencing reactions were set up with the sample groups shown in Table S1. In these experiments, untreated RNA served as a control for false positives. The small RNA samples for sequencing were isolated

by preparative denaturing PAGE experiments²⁸ after the indicated treatments described in the **Materials and Methods** section (see **Table S2** for yields). In each case, a duplicate sample was included resulting in a total of 12 samples. Sequencing libraries were prepared with a NEBNext multiplex small RNA library prep set for Illumina from 1.5 μg of purified RNA and sequenced on an Illumina NextSeq instrument in a pair-ended manner (2×75 bp).

Data Analysis. The **Materials and Methods** section describes the detailed data analysis workflow. Initially, raw reads were analyzed by FastQC. Adaptor contents were trimmed by Trimmomatic,²⁹ and the resulting reads were mapped to the *E. coli* K12 genome by STAR.³⁰ Next, the cufflinks suite of tools was used to count the transcripts and analyze gene expression across the sample groups.³¹ The alignments were visually assessed for the presence of mismatches, and misincorporation analysis was performed using the data generated by the BCFtools program.³² The results of the analyses are described below.

Gene Expression Analysis. The cufflinks suite of tools was employed to assemble, merge, and perform several RNA-Seq data analyses. An in-depth analysis of each sample group was performed, and the gene expression profiles were compared against each other. Traditionally, differential expression analysis is performed to identify the upregulated or downregulated genes across different sample groups. However, in our case, such an analysis was only used to look at the libraries' profiles and ensure no significant biological variation between the sample groups.

The aggregated read count for genes obtained from the expression analysis was grouped according to the gene biotypes, and percentages for each biotype were calculated (**Table 1**). Unexpectedly, the profiles for each sample group

Table 1. Read Composition Percentages by Gene Biotype^a

class	WTU	WTT	WTE	dTU	dTT	dTE
ncRNA	14.1	28.3	26.4	28.9	31.1	48.0
mRNA	76.3	62.8	57.3	63.8	61.0	39.3
pseudo-gene	1.0	0.9	1.0	1.7	1.6	1.5
rRNA	4.5	4.1	4.1	1.6	1.7	3.8
tmRNA	0.4	0.3	0.8	0.6	0.6	0.5
tRNA	3.8	3.7	10.4	3.6	4.0	6.9

^ancRNA: non-coding RNA, mRNA: messenger RNA, rRNA: ribosomal RNA, tmRNA: transfer messenger RNA, tRNA: transfer RNA. WTU: wild-type untreated, WTT: wild-type treated, WTE: wild-type treated and enriched, dTU: $\Delta thiI$ untreated, dTT: $\Delta thiI$ treated, dTE: $\Delta thiI$ treated and enriched.

(except dTE) are dominated by protein-coding mRNAs. Because we prepared libraries from 60 to 100 nt fragments isolated from PAGE, this may indicate that larger RNA transcripts are cleaved to shorter ones while labeling. Though **Figure S10** depicts a minimal degradation of RNA under the incubation conditions, RNA-Seq is a sensitive experiment that picks up minute quantities of RNA that may still be present and co-migrated with the tRNA. Nevertheless, because we are interested in tRNAs in this analysis, subsequent visualization and variant analysis were limited to tRNA transcripts only.

Visualization. In the next stage of analysis, the aligned reads are visualized by an integrative genomic viewer (IGV)³³ to identify any changes to a sequence readout that would result from the presence of BIA. **Figure S13** depicts a representative

example of tRNA transcripts, highlighting *hisR* (sense strand) and *glnV* (antisense strand). In both cases, the reads are identical to the genome sequence, except at a single position in the BIA-treated samples. In the *glnV* data, a misincorporation occurs at position 8. Indeed, *glnV* has been shown previously to contain s⁴U in this position.³⁴ By contrast, the misincorporation in the *hisR* data occurs at position 9 of the tRNA (which aligns with position 8 of the sequence of all other tRNA species) and has been demonstrated to contain s⁴U at this position.³⁵ A close inspection of similarly constructed alignments of all the reads corresponding to tRNA genes shows misincorporations at position 8. We note that the misincorporations only occurs in the BIA-treated samples with wild-type RNA but not in the *thiI* deletion strain, which our studies show lacks s⁴U (see **Figure S11**). Analysis of the misincorporations and mapping to the DNA shows that the BIA-modified s⁴U is decoded as C. Depending on whether the gene occurs on the sense or antisense strand, the base observed in the alignments is either C or G, respectively. This transition is not observed in samples that were not treated with BIA (wild type untreated), nor is it observed in RNA from the *thiI* deletion strain ($\Delta thiI$ untreated and treated). This latter observation further validates the assignment of this transition to the presence of s⁴U.

Misincorporation Analysis. The BCFtools program was used to carry out a detailed analysis of all the sequences obtained in these studies incorporating the corresponding biological replicates in each case. In this analysis, genotype likelihoods are determined initially by estimating the most likely base at a position by examining all the reads that have been aligned at that position. Possible misincorporations are identified by comparing the most likely base at a given position to the reference genome. The frequency of misincorporation was calculated from the number of correct and misincorporated bases at each position as described in the **Materials and Methods** section.

The misincorporation frequency heat maps for wild-type untreated (WTU) and wild-type-treated (WTT) sample groups are depicted in **Figure 6**. As seen in previous studies,^{11,12} the presence of s⁴U modification by itself at position 8 led to the misincorporation of deoxynucleotides during the RT reaction but at a lower frequency in the untreated sample (**Figure 6A**). The values observed for WTU at position 8 are in the range 0.003–0.35 with a median of 0.04 in contrast to the WTT sample (**Figure 6B**), which displayed much higher values in the range of 0.002–0.98 with a median of 0.75. Likewise, the wild-type-treated and enriched (WTE) sample group displayed similar values at position 8 in the range of 0.02–0.94, with a median of 0.69 (**Figure S14C**). At this position, either T > C or A > G transitions are seen depending on whether the gene is on the sense strand or the antisense strand. The control $\Delta thiI$ untreated (dTU), $\Delta thiI$ treated (dTT), $\Delta thiI$ treated, and enriched (dTE) samples depicted in **Figure S14** show that position 8 does not display misincorporations compared to wild-type samples. The median values of misincorporation at position 8 for dTU, dTT, and dTE are in the range of 0.00–0.01. Alternatively, the tRNAs corresponding to selenocysteine and isoleucine (Ile) have values close to zero at position 8. The misincorporation frequency values are shown in **Table S3**. It is also notable that the misincorporation frequencies were not uniform across all genes, resulting from the incomplete conversion of s⁴U to s⁴U-BIA or non-uniform cellular s⁴U incorporation. Nevertheless, it is evident from the heat maps that the treatment of BIA leads

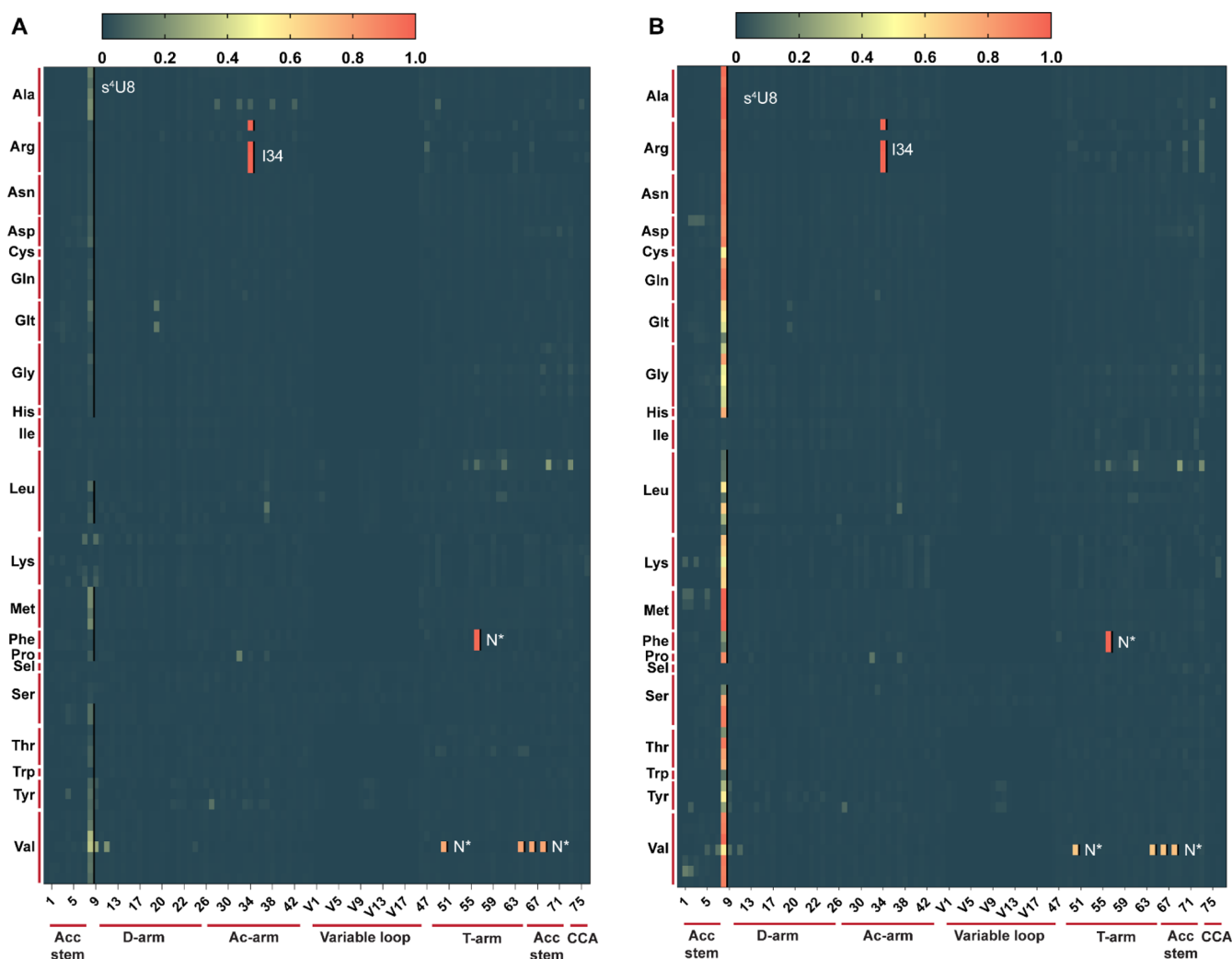


Figure 6. Misincorporation profiles of tRNAs derived from WT *E. coli*. Heat maps summarizing misincorporation frequency for tRNAs from (A) wild-type untreated (WTU) and (B) wild-type-treated (WTT) sample groups. The modifications I and s^4U (before treatment in WTU, after treatment in WTT) are indicated by a black line and labeled in white. Positions labeled as N* contain unknown post-transcriptional modifications. Alignment of tRNA genes of *E. coli* is carried out based on the alignments derived from tRNAdb. For clarity, the x-axis tick marks denote every fourth base. The y-axis shows tRNA species corresponding to each amino acid. The height of the bar correlates to the individual tRNA coding the amino acids. Gapped regions within the alignments were set to background color. The secondary structure attributes below the x-axis are labeled based on the alignments listed on tRNAdb. Acc-stem: acceptor stem, Ac-arm: anticodon arm, CCA: terminal nucleosides of tRNA genes. Each of the images are generated by considering both replicates in the respective sample groups. The raw data used to generate the heat map is reproduced in Figure S14.

to increased misincorporation at position 8, which is diagnostic for the presence of s^4U .

Close inspection of all sample groups (Figures 6 and S14) reveals the presence of additional misincorporation hotspots. However, unlike with s^4U , these are present in all samples, including the *thiI* deletion strain. For example, the I present at position 34 on the anticodon loop of tRNA^{Arg} (ACG) consistently leads to a misincorporation of C instead of G.⁷ Intriguingly, additional misincorporations observed in all sample groups (denoted by N* in Figures 6 and S14) are, to our knowledge, not correlated with any known modifications and may be novel. Because all of these are present $\pm thiI$, if they contain an S, *ThiI* is not the source. We note that the bulky post-transcriptional modification [3-(3-amino-3-carboxypropyl)uridine] (*acp*^{3U}) has been shown to induce misincorporation previously^{11,12} was not seen in our studies. However, this difference may be due to the selectivity of the polymerase employed.

Sequencing by MS. The RNA-Seq experiments clearly show a measurable increase in the transition frequencies related to IA modification and the presence of the *thiI* gene. However, a direct demonstration was sought to independently verify the presence of s^4U on *E. coli* tRNAs to validate the RNA-Seq workflow. The presence of s^4U on various *E. coli* tRNAs was obtained by digesting *E. coli* total tRNA with the RNase T1 enzyme and analyzing the resulting fragments by MS/MS. The nuclease cleaves single-stranded RNA adjacent to G on the 3' end to generate smaller fragments. Typical amounts of size-selected tRNA used for RNase T1 experiments are shown in Table S4. Figure 7 shows HPLC chromatograms of RNase T1-digested tRNA samples from wild-type and deletion *E. coli* strains. As shown in Figure 2, s^4U absorbs at 330 nm, which allows the UV detection of s^4U containing fragments. The chromatographs at 330 nm show numerous peaks, which are consistent with the presence of s^4U . Control experiments with tRNA from $\Delta iscS$ also show peaks, whereas

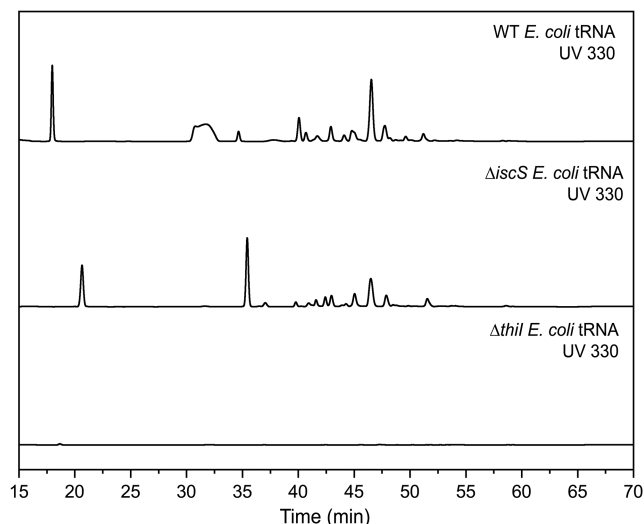


Figure 7. Representative LC analysis of RNaseT1 digest of tRNA from WT and deletion *E. coli* strains. The HPLC traces extracted at 330 nm (corresponding to the maximum absorbance for s^4U) of RNase T1 digestion samples are shown. Fragments containing s^4U display absorbance at 330 nm and distinct peaks are visible in the samples. The corresponding strains are labeled. To facilitate comparison, the intensities on y -axes are normalized to per μg of the sample and depicted to similar scales.

the $\Delta thiI$ chromatograms show no features, consistent with the loss of the modification (Figure 7). These results are in harmony with the nucleoside analysis of tRNA from wild-type and deletion *E. coli* strains (see Figures 4 and 5) that show a requirement of *thiI* in the modification pathway.

The samples were further analyzed by subjecting the eluent to MS/MS analysis in the negative ion mode. In these experiments, the five highest intensity species at each time were subjected to MS fragmentation as described in the Materials and Methods section. The resulting data were analyzed by the NucleicAcidSearchEngine (NASE)¹⁷ using a database containing all the fully modified *E. coli* tRNA sequences in *E. coli*. The resulting output was examined for the s^4U containing fragments and visualization in TOPPView from an OpenMS toolset.³⁶ Some of the tRNA species generated smaller fragments such as s^4UGp and s^4UAGp , which could not be assigned unambiguously to a specific tRNA, while others generated unique fragments that were readily mapped. Likewise, tRNA from $\Delta thiI$ *E. coli* was used as a control sample where the corresponding unmodified U8 fragment for each tRNA was identified to assign the position of s^4U via mass spectrometry confidently. A representative MS/MS analysis for s^4U containing fragment $CUA[s^4U]AGp$ from tRNA^{Ala} is shown in Figure S15. Data with the $\Delta thiI$ tRNA sample shows that the corresponding $CUAUAGp$ is not modified. The results of these analyses are summarized in supplementary extended Table.

The data from LC–MS/MS analyses are compared to sequencing analyses, tRNAdb³⁷ and MODOMICS³⁸ databases in Table 2. The tRNA sequences used for the analysis are provided in supplementary extended Table. In few cases, there are minor differences between the reported modification data at position 8 in MODOMICS and/or tRNAdb and assignments from the RNA-Seq and MS/MS analyses. However, there is a general agreement between the RNA-Seq and mass spectrometry analyses. For example, the observed enhance-

ment of the misincorporation frequency for Ala tRNA genes is corroborated with the presence of s^4U in the MS/MS analysis of all five Ala tRNA genes (see Tables S3 and 2). In the MODOMICS and tRNAdb databases, both modified and unmodified U is observed (Table 2).

The MS/MS data do not just confirm sequencing results but reveal additional modification data. The MS/MS analysis, for example, reveals the presence of s^4U at two adjacent positions in Tyr tRNA. The MS/MS data with the wild-type sample show $[s^4U][s^4U]CCCGp$, which span positions 8 and 9 in the RNA. The corresponding positions in the $\Delta thiI$ tRNA sample contain U. By contrast to this data, the enhanced misincorporation frequency is only reflected at position 8 of Tyr tRNA genes in the RNA-Seq analysis (see Figure 6).

In few cases, the data are consistent with the presence of both modified and unmodified U, which could be attributed to the incomplete sulfuration of U8 in those tRNAs. This observation is in line with the RNA-Seq analysis whereupon BIA treatment, few genes showed close to 80–90% misincorporation at position 8, while others were in 30–70% range (Table S3). Therefore, there are differential modification frequencies of certain tRNA species in the stationary phase of growth, though the data preclude the quantification of dynamics of s^4U installation on individual tRNA species.

In addition to the untreated samples, RNase T1 digestion was carried out on IA, and BIA-treated tRNA samples to supplement the sequencing analysis. The NASE utilized in this study allows the use of custom modifications by utilizing the monoisotopic mass of the modified nucleoside. Therefore, the data from IA or BIA-treated tRNA samples were analyzed against sequences containing either s^4U -IA or s^4U -BIA in place of s^4U on *E. coli* tRNA sequences as described in the Materials and Methods section. Treatment of the sample with IA or BIA leads to loss of the UV–vis detection of peaks in LC, also consistent with the change of the UV–visible absorbance of treated s^4U (Figures S16 and S17). A representative MS/MS analysis of $CUA[s^4U$ -IA]AGp and $CUA[s^4U$ -BIA]AGp fragment originating from tRNA^{Ala} after IA and BIA treatments, respectively, is depicted in Figure S18. The observed dissociation pattern for those fragments allows for the confident assignment of these sequences to a precursor ion with observed mass. Interestingly, in Tyr tRNA, fragments containing s^4U -IA and s^4U -BIA at positions 8 and 9 are found, corroborating the presence of s^4U at both positions 8 and 9.

Overall, all the MS/MS analysis data of RNase T1-digested samples by NASE is tabulated in extended supplementary Table. These observations indicate that the iodoacetamide reaction toward s^4U is very specific in the context of complex tRNA modifications present on *E. coli* tRNA. In conclusion, the iodoacetamide reaction largely modifies the s^4U both specifically and efficiently in the context of tRNA. Overall, our method demonstrates the utility of the chemical treatment of s^4U with iodoacetamide in site-specific profiling of natively occurring s^4U on tRNAs.

Discussion. The first nucleoside preparation of s^4U was prepared in 1958,³⁹ long before it was discovered in a soluble RNA extract from *E. coli* in 1965.⁴⁰ The presence of 2-thiouridine derivatives in RNA was also uncovered around the same time.⁴¹ These revelations sparked considerable interest in the roles of these non-canonical nucleosides. Analytical studies by Favre et al. revealed that the irradiation of *E. coli* tRNA^{val} with 334 nm light led to adduct formation between s^4U at position 8 and C at position 13.⁴² Efforts were also made to

Table 2. Comparison of tRNA Modification at Position 8 from Databases against the RNA-Seq and MS/MS Analyses^a

gene	local DB	anticodon	MODOMICS	tRNAdb	RNA-Seq	RNase T1-MS	RNaseT1—theoretical fragment
<i>alaT, alaU, alaV</i>	Ala1	VGC	U	s ⁴ U/U	s ⁴ U	s ⁴ U	CUA[s ⁴ U]AGp
<i>alaW, alaX</i>	Ala2	GGC	U	N	s ⁴ U	s ⁴ U	CUA[s ⁴ U]AGp
<i>argQ, argV, argY, argZ</i>	Arg1	ICG	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]AGp
<i>argX</i>	Arg2	CCG	U	U	n.d.	Amb	[s ⁴ U]AGp
<i>argU</i>	Arg3	{CU	U	U	s ⁴ U	n.d.	CCCU[s ⁴ U]AGp
<i>argW</i>	Arg4_NA [#]	{CU			n.d.	n.d.	UCCUCU[s ⁴ U]AGp
<i>asnT, asnU, asnV, asnW</i>	Asn	QUU	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]AGp
<i>aspT, aspU, aspV</i>	Asp	QUC	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]AGp
<i>cysT</i>	Cys	GCA	s ⁴ U	s ⁴ U	s ⁴ U	U/s ⁴ U	U[s ⁴ U]AACAAAGp
<i>glnU, glnW</i>	Gln1	NUG	s ⁴ U	s ⁴ U	s ⁴ U	s ⁴ U	UA[s ⁴ U]CGp
<i>glnV, glnX</i>	Gln2	CUG	s ⁴ U	s ⁴ U	s ⁴ U	s ⁴ U	UA[s ⁴ U]CGp
<i>gluT, gluU, gluV, gluW</i>	Glu	SUC	U	U	s ⁴ U	U	UCCCCUUCGp
<i>glyU</i>	Gly1	CCC	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]AGp
<i>glyT</i>	Gly2	NCC	U	U	s ⁴ U	U/s ⁴ U	CA[s ⁴ U]CGp
<i>glyV, glyW, glyX, glyY</i>	Gly3	GCC	s ⁴ U	U	s ⁴ U	U/s ⁴ U	AA[s ⁴ U]AGp
<i>hisR</i>	His	QUG	s ⁴ U	s ⁴ U	s ⁴ U	s ⁴ U	CUA[s ⁴ U]AGp
<i>ileT, ileU, ileV</i>	Ile1	GAU	U	U	U	Amb	UAGp
<i>ileX</i>	Ile2	}AU	s ⁴ U	s ⁴ U	n.d.	n.d.	CCCCU[s ⁴ U]AGp
<i>ileY[#]</i>	Ile3_NA	}AU			n.d.	n.d.	CCCCU[s ⁴ U]AGp/CCCUUAGp
<i>metV, metW, metZ*</i>	Ini1	CAU	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]Gp
<i>metY*</i>	Ini2	CAU	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]Gp
<i>leuP, leuQ, leuT, leuV</i>	Leu1	CAG	U	U	s ⁴ U	Amb	UGp
<i>leuZ</i>	Leu3)AA	s ⁴ U	s ⁴ U	s ⁴ U	Amb	A[s ⁴ U]Gp
<i>leuU</i>	Leu2	GAG	U	U	s ⁴ U	Amb	[s ⁴ U]Gp
<i>leuX</i>	Leu4	BAA	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]Gp
<i>leuW[#]</i>	Leu5_NA	VAG			s ⁴ U	Amb	[s ⁴ U]Gp
<i>lysQ, lysT, lysV, lysW, lysY, lysZ</i>	Lys	SUU	U	s ⁴ U	s ⁴ U	U/s ⁴ U	U[s ⁴ U]AGp
<i>metT, metU</i>	Met	MAU	s ⁴ U	s ⁴ U	n.d.	Amb	[s ⁴ U]AGp
<i>pheU, pheV</i>	Phe	GAA	s ⁴ U	s ⁴ U	s ⁴ U	s ⁴ U	A[s ⁴ U]AGp
<i>proK</i>	Pro	CGG	U	s ⁴ U	n.d.	s ⁴ U/U	AU[s ⁴ U]Gp
<i>proL[#]</i>	Pro2_NA				n.d.	Amb	[s ⁴ U]AGp/UAAGp
<i>proM[#]</i>	Pro3_NA				s ⁴ U	Amb	[s ⁴ U]AGp/UAAGp
<i>selC</i>	Sec	UCA	s ⁴ U	U	U	U	UCGp
<i>serT</i>	Ser1	VGA	s ⁴ U	s ⁴ U	U	Amb	[s ⁴ U]Gp/UGp
<i>serU</i>	Ser2	CGA	U	U	s ⁴ U	s ⁴ U	A[s ⁴ U]Gp
<i>serV</i>	Ser3	GCU	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]Gp
<i>serW, serX</i>	Ser4	GGA	s ⁴ U	s ⁴ U/U	s ⁴ U	Amb	[s ⁴ U]Gp
<i>thrV</i>	Thr1	GGU	U	U	s ⁴ U	s ⁴ U	AUA[s ⁴ U]Gp
<i>thrT</i>	Thr2	GGU	U	U	s ⁴ U	U/s ⁴ U	AUA[s ⁴ U]AGp
<i>thrU</i>	Thr3_NA	VGU	NA	NA	s ⁴ U	s ⁴ U	ACU[s ⁴ U]AGp
<i>thrW</i>	Thr4_NA	NA	NA	NA	s ⁴ U	s ⁴ U	AUA[s ⁴ U]AGp
<i>trpT</i>	Trp	CCA	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]AGp
<i>tyrT, tyrU</i>	Tyr1	QUA	[s ⁴ U8][s ⁴ U9]	U8[s ⁴ U9]	s ⁴ U8	[s ⁴ U8][s ⁴ U9]	[s ⁴ U][s ⁴ U]CCCCGp
<i>tyrV</i>	Tyr2	QUA	[s ⁴ U8][s ⁴ U9]	U8[s ⁴ U9]	s ⁴ U8	[s ⁴ U8][s ⁴ U9]	[s ⁴ U][s ⁴ U]CCCCGp
<i>valT, valU, valX, valY, valZ</i>	Val1	VAC	s ⁴ U	s ⁴ U	s ⁴ U	s ⁴ U	AU[s ⁴ U]AGp
<i>valW</i>	Val2	GAC	s ⁴ U	s ⁴ U	s ⁴ U	Amb	[s ⁴ U]AGp
<i>valV</i>	Val3	GAC	s ⁴ U	s ⁴ U	s ⁴ U	U/s ⁴ U	UUCA[s ⁴ U]AGp

^aLocal DB: name of tRNA in the local sequence database; refer to [supplementary extended Table](#). Anticodon: anticodon on the tRNA. MODOMICS: the modification at position 8 on the MODOMICS³⁸ website. tRNAdb: the modification at position 8 on tRNAdb.³⁷ RNA-Seq: the modification at position 8 using RNA-Seq experiments. RNase T1-MS: the modification at position 8 using LC-MS/MS analysis. RNaseT1—theoretical fragment: theoretical oligonucleotide generated upon RNase T1 digestion of a tRNA. Amb: ambiguous; the theoretical fragment generated is detected but could not be assigned to one tRNA species unambiguously. n.d.: not detected; N: unknown modified uridine; V: uridine 5-oxyacetic acid; I: inosine; { : 5-methylaminomethyluridine; Q: queuosine; \$: 5-carboxymethylaminomethyl-2-thiouridine; S: 5-methylaminomethyl-2-thiouridine; B: 2'-O-methylcytidine; 5-carboxymethylaminomethyl-2'-O-methyluridine; *initiator methionine tRNAs. [#]These tRNAs have a gene, but the RNA sequences are not represented in the MODOMICS and/or tRNAdb databases.

understand the structure of the photoadduct.^{43,44} Additional studies^{45,46} showed similar reactivity in vivo, leading to the suggestion that the modified base acts as a UVA radiation sensor to induce the growth delay and promote photo-protection during the stress.^{47,48} Experiments focused on

understanding the mechanism of growth delay indicated the dependency on the levels of the molecule ppGpp, which inhibits rRNA synthesis similarly to that observed in amino acid starvation.^{48–50} The hypothesis of growth delay by s⁴U was also supported by the observation of a minimal growth

delay in *nuv* *E. coli* variants lacking s^4U following the irradiation.^{47,49} Favre and colleagues tried to examine the fate of cross-linked tRNA in *E. coli* after removing the stress.^{51,52} Intriguingly, the adduct appears to be resolved upon restoring favorable growth conditions, but there is no established mechanism for such a process.^{51,52} While analytical studies based on fluorescence measurements indicate a s^4U presence in 70% *E. coli* tRNA,⁴⁷ understanding the role of the modification and its dynamics under other common stress conditions has not been possible because of the lack of site-specific profiling methods.

The reactivity and photochemistry of s^4U have also been leveraged into chemical biology tools to explore RNA dynamics.^{18,53} For example, incorporating externally supplied s^4U during growth allows labeling of newly synthesized RNA.^{23,54–56} Additionally, the nucleophilicity of the sulfur atom has been exploited to attach desired reagents/fluorescent probes for analytical or drug development purposes.^{19,20} While these studies underscore the unique reactivity of the modified base, to date, it has not been exploited for the site-specific detection of natively occurring s^4U in *E. coli*.

While methylation has received substantial contemporary attention, it is not the only modification that is known to occur in RNA. Indeed, pioneering work from several groups starting from the 1970s has brought to light many chemically diverse bases.^{2,38} Current estimates are that there are 105 numbers of unique modifications in tRNA.² Simple one-step routes that make some but many hyper-modifications require an army of enzymes and cofactors to introduce. In this regard, Q ,^{57–60} Y ,^{61,62} and s^4U ^{26,63,64} are notable for their biosynthetic complexity.

In this article, we have leveraged the nucleophilic reactivity of the s^4U to carry out single nucleotide level profiling of tRNA of *E. coli*. The data show that the modification is present in 19 of 20 amino acid coding tRNAs under these conditions. In all but one case, the modification occurs at position 8 of the tRNA. In *hisR*, the modification occurs at position 9, which is consistent with previously reported sequencing results.³⁵ It is notable that under the conditions used in these experiments, the “readout” is a base transition in all cases, providing a concise method for identifying the modification in tRNA. It is also important to point out that while tRNA is known to undergo many modifications, our data show increased misincorporation frequency for s^4U -BIA in the treated samples, highlighting the fact that the s^4U modification can be observed in the background of all other possible modifications. The extensive analytical studies that underpin the work provide confidence that the reactivity of IA or BIA is unique to the nucleoside, as we observed no modification of the canonical nucleosides, even under conditions, where high concentrations of both nucleoside and the reagent were present. Finally, the observation that the transition is not observed in the *thiI* deletion strain provides confidence in the assignment. Therefore, the method could directly be applied to the profiling of natively occurring s^4U on tRNA, for example, under different physiological stress conditions or from bacterial species where the presence of s^4U is not known.

MATERIALS AND METHODS

Reaction of IA or BIA with s^4U . Commercially available s^4U was treated with a tenfold excess of either iodoacetamide (IA, TCI Chemicals) or iodoacetyl-PEG₂-biotin (BIA, Thermo Scientific) as follows. s^4U (0.1 mM) was incubated with 1 mM

of either IA or BIA in 0.05 M sodium phosphate (NaPi) buffer (pH 8) at 50 °C overnight. The reaction mixtures were quenched with a tenfold excess of DTT (10 mM). UV–visible absorption spectra were recorded on an Agilent 8454 diode-array spectrophotometer between wavelengths from 190 to 800 nm. The reaction mixtures were analyzed further on an Ultimate 3000 HPLC instrument with a photodiode array detector coupled to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) as follows. The mixtures were injected onto a Hypersil Gold C-18 column [particle size—1.9 μ M, dimensions—2.1 mm (D) \times 150 mm (L), Thermo Fisher] pre-equilibrated with 0.05 M ammonium acetate (pH 5.3) at a flow rate of 0.2 mL/min. Separation was carried out with a gradient of 40% acetonitrile (buffer B, optima grade from Thermo Fisher) in water (optima grade from Thermo Fisher) over 17 min for IA-treated samples (time: 0–3.5 min, % B: 0–0.8; time: 3.5–3.75 min, % B: 0.8–3.2; time: 3.75–4.0, % B: 3.2–5.0; time: 4–12 min, % B: 5.0–25.0; time: 12–15 min, % B: 25–50; time: 15–17 min, % B: 50–75; time: 17–17.1 min, % B: 75–100; time: 17.1–20 min, % B: 100) and over 45 min for the BIA-treated samples (time: 0–4.4 min, % B: 0–0.2; time: 4.4–5.8 min, % B: 0.2–0.8; time: 5.8–7.2, % B: 0.8–1.8; time: 7.2–8.6 min, % B: 1.8–3.2; time: 8.6–10 min, % B: 3.2–5.0; time: 10–25 min, % B: 5–25; time: 25–30 min, % B: 25–50; time: 30–34 min, % B: 50–75; time: 34–37 min, % B: 75; time: 37–45 min, % B: 75–100; time: 45–48 min, % B: 100). Variations to the above LC method have been used to analyze few samples. Those experimental details are listed in additional methods in the [Supporting Information](#).

The method employed for BIA-treated samples was longer to accommodate the hydrophobicity of the BIA moiety. All MS data on a LTQ Orbitrap XL mass spectrometer was recorded in the positive ion mode with an FT analyzer at a resolution setting of 100,000 and m/z range of 50–1400. The instrument used for recording the mass spectra was maintained at a capillary temperature of 275 °C, with sheath gas flow 35, auxiliary gas flow 12, and a source voltage of 3 kV. MS/MS fragmentation was achieved by collision-induced dissociation at energy settings, as noted in the results. Xcalibur software was used to analyze all the LC–MS data (Thermo Fisher), and the deconvoluted spectra were generated from the raw data by exporting with the use of Xcalibur (Thermo Fisher).

Total RNA Extraction from *E. coli*. Total RNA was extracted from either wild-type or deletion strains of *E. coli* strains by employing guanidine-based methods with some modifications as detailed below.⁶⁵ The cell pellet from 1 L of the growth was resuspended in 6 mL of denaturation buffer consisting of 4 M guanidine thiocyanate, 0.025 M sodium citrate, 0.5% sarkosyl, and 0.1 M 2-mercaptoethanol. The resulting suspension was equally distributed into 10–12 two mL Eppendorf tubes, each of which contained 0.7 mL of the suspension. Next, 2 M sodium acetate at pH 4 (0.1 mL), phenol (1 mL), and 1-bromo-3-chloropropane (0.2 mL) were added in succession, vortexed, and incubated on ice for 15 min. The tubes were centrifuged at 15,000g, and the resulting top transparent layer (0.3–0.4 mL) was transferred to a new 1.5 mL Eppendorf tube. An equivalent volume of isopropanol was added before placing the tubes at –20 °C for at least 30 min. The precipitated RNA was collected by centrifugation, and the supernatant was discarded. The RNA pellet obtained was further purified using a Qiagen miRNeasy mini kit following the manufacturer's directions for the purification of RNA >18 nt as follows.

In brief, the pellet was resuspended in 0.05 mL of RNase free water (Invitrogen) and vortexed. The Qiazol reagent (0.25 mL) and chloroform (0.05 mL) were added in succession, vortexed, and centrifuged at 21,000g for 1 min. The top transparent layer was transferred to a new 1.5 mL Eppendorf tube containing 1.5 volumes of cold ethanol and mixed by pipetting. The entire content of the tube was loaded onto an Omega BioTek HiBind RNA spin-columns fitted with collection tubes. The column was then washed once with 0.7 mL RWT (Qiagen) and twice with 0.5 mL RPE (Qiagen) buffers in succession, with centrifugation to permit the removal of the wash solutions before eluting the RNA with 40 μ L of RNase free water (Invitrogen). RNA prepared from the same starting cell pellets were pooled and stored at -80 $^{\circ}$ C until further use in smaller aliquots.

Reaction of IA or BIA with Total RNA from *E. coli*. The RNA extracted was treated with either IA or BIA as follows. Purified RNA (0.04 mL) was incubated with 1 mM of either IA or BIA in 0.05 mM NaPi buffer (pH 8) overnight at 50 $^{\circ}$ C. The reaction mixtures were quenched with a tenfold excess of DTT (10 mM) over the reagent, and RNA was purified using the Qiagen miRNeasy mini kit (see above). The eluted RNA was digested to the nucleoside level by the successive action of P1 nuclease, phosphodiesterase, and alkaline phosphatase, as reported previously.²⁵ The reaction mixtures were filtered on VWR centrifugal filters (PES membrane, MWCO 10 kDa) to remove the protein components prior to analysis. LC–MS analyses of these mixtures were carried out as described above for the nucleoside samples.

Purification of RNA for Sequencing Experiments. The desired small RNA (\sim 60–100 nt) for sequencing experiments was purified by employing preparative PAGE (28 cm H \times 16.5 cm W) prepared with 6% acrylamide, 8 M urea, 1 \times TBE buffer, TEMED, and 30% APS.²⁸ The RNA samples were mixed with an equivalent volume of 2 \times loading buffer.²⁸ The RNA was separated at 25 W for 2 h or until the lower dye band was two-thirds through the gel. The bands were detected by UV-shadowing and excised from the gel. RNA in the bands was eluted into solution by incubating the gel pieces overnight at 4 $^{\circ}$ C in the crush soak buffer, which was prepared by mixing 1 mL of 1 M Tris-HCl pH 7.5, 4 mL of 5 M NaCl, and 0.2 mL of 0.5 M EDTA pH 8 (diluted to 100 mL by adding water and autoclaved). Following centrifugation at 21,000g for 5 min, the supernatant was transferred to a new tube containing 0.7 mL of cold ethanol. RNA was precipitated by placing the tubes at -20 $^{\circ}$ C for 2 h. The RNA pellet was collected by centrifugation and dried under vacuum conditions in a Savant speed-vac at room temperature. The resulting white residue was resuspended in 0.02 mL of Tris-EDTA buffer (made by mixing 0.1 mL of Tris-HCl pH 7.5 and 0.02 mL of 0.5 M EDTA pH 8.0, made to 10 mL by adding water and sterile filtered) and quantified using a Nanodrop. The amount of tRNA obtained for sequencing experiments is listed in Table S2.

RNase T1 Digestion and LC–MS/MS Analysis. The RNA samples used for RNase T1 digestion are obtained with a slight variation as described above. RNA from either WT or deletion strains was obtained by phenol extraction²⁵ followed by gel purification as described above, prior to digestion with the RNase T1 enzyme. The cell pellet from 1 L of the growth was resuspended in 1 mL of resuspension buffer (10 mM Tris-HCl (pH 8), 10 mM MgCl₂ and 0.15 M NaCl) per mg of the cell pellet, and mixed by rotation at 4 $^{\circ}$ C. After

resuspension, 1 mL of saturated phenol buffered at pH 4.3 (Thermo Scientific) was added to 1 mg of the starting cell pellet and mixed by inversion for 1 h at 4 $^{\circ}$ C. The suspension was centrifuged at 8000 rpm at 4 $^{\circ}$ C for 30 min, and the top semi-transparent layer was removed, mixed with saturated phenol (\sim 1 mL/mg of starting cell pellet), and centrifuged at 8000 rpm at 4 $^{\circ}$ C for 30 min. After centrifugation, the upper transparent layer was removed and mixed with 0.1 volumes of 3 M sodium acetate pH 5.5 and 3 volumes of ethanol. The sample was incubated for 2 h at -20 $^{\circ}$ C, and the precipitated RNA was collected by centrifugation for 30 min at 12,000 rpm and at 4 $^{\circ}$ C. The resulting pellet was redissolved in water, and total tRNA was isolated using the preparative PAGE purification method as described above.

The RNase T1 digestion reaction was set up as follows. The purified tRNA (49 μ L) was incubated with 1000 U of RNase T1 (1000 U/ μ L, Thermo Scientific) in buffer containing 0.05 M Tris-HCl, 10 mM EDTA (pH 7.5) for 2 h at 37 $^{\circ}$ C. For IA-treated samples, the purified tRNA was first incubated with IA as described above and digested with RNase T1 after the reaction cleanup with a Qiagen miRNeasy mini kit as described in the Total RNA extraction section. The concentrations of tRNA used for each of the replicates in RNase T1 digestion are mentioned in Table S4. Likewise, \sim 50 μ g of each of the sequenced WTU and WTT samples are similarly treated with RNase T1. The reaction mixtures were analyzed on a Vanquish UHPLC instrument (Thermo Fisher Scientific) with a photodiode array detector and a Q-exactive mass spectrometer. The mixture was injected onto a Hypersil Gold C-18 column [particle size—1.9 μ M, 2.1 mm (D) \times 150 mm (L), Thermo Fisher] pre-equilibrated with 97.5% of buffer A containing 0.2 M hexafluoroisopropanol (HFIP) and 0.085 M TEA in water (optima grade from Thermo Fisher) 2.5% of buffer B containing 0.1 M HFIP, 0.042 M triethylamine (TEA) in methanol (optima grade from Thermo Fisher) at a flow rate of 0.05 mL/min. Separation was carried out with a gradient of buffer B over 70 min (time: 0–13.1 min, % B: 2.5; time 13.1–52.4 min, % B: 2.5–35; time: 52.5–71 min, % B: 100). The UV-vis data are plotted as described in the above sections. However, the UV-vis chromatograms depicted under RNase T1 analysis are normalized to per μ g of the input sample.

Data-dependent MS/MS analysis in the negative ion mode was carried out on the eluent from the column from above. The acquisition of full MS at a resolution of 140,000 followed by data-dependent fragmentation of the 5 most intense peaks at a resolution of 70,000. An intensity threshold of 1×10^5 and normalized collision energy of 20 was employed. The acquired Thermo RAW files were converted to mzML using MSConvert with vendor peak picking and MS subset filters.^{66,67} The mzML files were processed with nucleic acids search engine (NASE) software which generates theoretical fragments and compares them against the experimental data.¹⁷ A FASTA file containing modified tRNA sequences was used as the input. Precursor and fragment mass tolerances were set to 5 ppm. The program was set to consider fragments of 2 or more nucleotides, and the number of missed cleavages by RNase T1 was set to zero. The tRNA sequence used in the analysis, along with the output from each sample, is tabulated in supplementary extended Table.

Library Preparation and Data Accumulation. cDNA libraries were prepared from 1.5 μ g of isolated RNA for each of the sample conditions (Table S1) using a NEBNext Multiplex Small RNA library prep kit for Illumina. The libraries were

made following the manufacturer's instructions. Size selection of libraries was carried out using the native PAGE conditions described in the manufacturer's protocol. The amount of size-selected libraries submitted to the sequencing facility is listed in Table S2.

RNA-Seq Data Collection and Analysis. The sequencing was carried out on an Illumina NextSeq instrument in a pair ended manner with a read length of 75 bp by the Advanced Genomics Core at the University of Michigan. The data obtained were analyzed in house. All the analyses are performed using open-source Linux tools following the guidelines of each program.

Quality Assessment and Trimming. Raw reads were analyzed by the program FastQC⁶⁸ which looks for adaptor contents, duplication levels, and GC content. Following the quality check, Trimmomatic²⁹ was used to remove the adaptor content (provided with truseq3 adaptors FASTA file) and to filter out reads below 35 nt in length. The resulting sequences were then mapped to the *E. coli* K-12 MG1655 genome.

Mapping. FASTA sequence file and annotation files required for mapping reads to the genome are obtained from the ensemble website (<ftp://ftp.ensemblgenomes.org/pub/bacteria/release-44>). Spliced transcript alignment to a reference (STAR 2.7), a short-read alignment program was used for mapping reads to the genome.⁵⁰ The alignments are carried out in a two-step process where the genome is indexed first by the program, which would later be used in the mapping process (default parameters were used except for the sjdb-overhang parameter, which was set to 75 following the instructions). The alignments were outputted to a BAM format file (sorted by coordinate).

Gene Expression Analysis. Gene expression analysis was carried out using cufflinks, cuffmerge, cuffdiff, and cummeRbund. The cufflink program within the suite takes GTF annotation file and BAM files as an input (library-type = fr-first-strand, read length = 76 nt) and generates a set of tracking files for each sample group.³¹ Transcripts (information in tracking files) from each assembly were merged into the single master-transcriptome assembly by the program cuffmerge, taking the biological/technical replicates into account. Cuffdiff (input = BAM files, reference FASTA sequence & master assembly GTF) processes the master transcript assembly to perform rigorous statistical analysis on the count values obtained. The data produced by cuffdiff was navigated with the use of an R library package called cummeRbund. A count matrix for all the genes was extracted from the expression analysis. The counts were summed according to the gene biotype in R by matching the cufflink ids (obtained from master transcriptome assembly) with gene information present in the GFF3 annotation file (obtained from the ensemble website).

Visualization. The read coverages for the tRNA genes are visualized in the program IGV.³³ GTF, GFF3, and BED files corresponding to the *E. coli* K-12 genome and the alignment files are loaded in the program. Screenshots corresponding to the respective alignments are generated in scalable vector graphics (SVG) and modified appropriately in an Adobe Illustrator CC 2019.

Misincorporation Frequency. The misincorporation frequency was calculated from the genotype likelihoods generated by the bcftools program³² as follows. In the first step, BCFTools "mpileup" argument [maximum depth at a position 100,000, FASTA file, ignored duplicates, output mpileup in variant call

file (VCF) format] calculates the genotype likelihoods at each genomic position along with coverage. Replicates in the same sample group are considered together for obtaining the VCF files. The output VCF files are then processed with BCFTools "call" argument to discover the mutations (multi-allelic caller, $p = 0.99$, BED feature file containing tRNA information). Several variant analysis attributes corresponding to each position was written in VCF files. The number of high-quality reference and alternate bases (DP4 attribute) at each position was extracted from the VCF file using VariantsToTable utility from the GATK suite of tools.⁶⁹ The misincorporation frequency at each of the tRNA positions is calculated in a Microsoft Excel spreadsheet as follows

$$\begin{aligned} & \text{misincorporation frequency} \\ &= [\text{no. of high quality alternate bases}] \\ & / [\text{no. of high quality alternate bases} \\ & + \text{no. of high quality reference bases}] \end{aligned}$$

The obtained data file was then processed in R software for statistical computing and graphics to generate a matrix file suitable for plotting heat maps. The obtained matrix file was then manually adjusted according to the alignments listed on the transfer RNA database (tRNAdb) and summarized in [supplementary extended Table](#).³⁷ Additionally, the MODO-MICS database was referred to as well.³⁸ The heat maps are generated with GraphPad Prism plotting software. The positions with no values were given the same color as the background of the plot. The genes *argX*, *argW*, *ileX*, *ileY*, *lysV*, *metT*, *metU*, *proK*, and *proL* were ignored while plotting heat maps owing to the low coverage.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.1c05071>.

LC-MS/MS analysis of standard nucleoside, tRNA digested to nucleosides, tRNA digested to small oligo nucleotides, ¹H NMR spectra of s⁴U and s⁴U-BIA, misincorporations in RNA-Seq, experimental setup, and concentrations and misincorporation frequencies observed in RNA-Seq ([PDF](#))

tRNA sequence, tRNA gene alignment, and summary of RNase T1 digested LC-MS analysis ([XLSX](#))

■ AUTHOR INFORMATION

Corresponding Author

Vahe Bandarian – Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States;

📧 orcid.org/0000-0003-2302-0277; Email: vahe@chem.utah.edu

Author

Praneeth Bommiseti – Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.1c05071>

Funding

The research reported in this publication was supported by the National Institutes of General Medical Sciences of the National

Institutes of Health under the award number R35 GM126956. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors also acknowledge start-up funds from the University of Utah.

Notes

The authors declare no competing financial interest.

Data Availability: the raw sequencing data referenced in this manuscript have been deposited at Sequence Read Archive at NCBI. The accession numbers are SAMN15420428, SAMN15420429, SAMN15420430, SAMN15420431, SAMN15420432, SAMN15420433, SAMN15420434, SAMN15420435, SAMN15420436, SAMN15420437, SAMN15420438, and SAMN154204. The mass spectrometry transcriptomic data have been deposited to the ProteomeX-change Consortium via the PRIDE⁷⁰ partner repository with the data set identifier PXD028405.

ACKNOWLEDGMENTS

P.B. thanks Dr. Anthony Young for fruitful discussions. P.B. also acknowledges the assistance from Andrew Jochimsen in NMR data acquisition and interpretation. P.B. thanks Dr. Buck-Koehntop, Dr. Burrows, and Dr. Hammond labs for fruitful discussions and resources. We are also grateful for Dr. Olivia Koues and her team at the Advanced Genomics core of the University of Michigan for helpful suggestions. The NMR measurements were recorded at the David M. Grant NMR Center, which is a University of Utah Core Facility. The funding for the construction of the Center and the helium recovery system were obtained from the University of Utah and the National Institutes of Health awards 1C06RR017539-01A1 and 3R01GM063540-17W1. The NMR instruments were purchased with the support of the University of Utah and the National Institutes of Health awards 1S10OD25241-01.

REFERENCES

- (1) Roundtree, I. A.; Evans, M. E.; Pan, T.; He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* **2017**, *169*, 1187–1200.
- (2) Cantara, W. A.; Crain, P. F.; Rozenski, J.; McCloskey, J. A.; Harris, K. A.; Zhang, X.; Vendeix, F. A. P.; Fabris, D.; Agris, P. F. The RNA Modification Database, RNAMDB: 2011 Update. *Nucleic Acids Res.* **2011**, *39*, D195–D201.
- (3) Shepherd, J.; Ibba, M. Bacterial Transfer RNAs. *FEMS Microbiol. Rev.* **2015**, *39*, 280–300.
- (4) Thiaville, J. J.; Kellner, S. M.; Yuan, Y.; Hutinet, G.; Thiaville, P. C.; Jumpathong, W.; Mohapatra, S.; Brochier-Armanet, C.; Letarov, A. V.; Hillebrand, R.; Malik, C. K.; Rizzo, C. J.; Dedon, P. C.; de Crécy-Lagard, V. Novel Genomic Island Modifies DNA with 7-Deazaguanine Derivatives. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, E1452–E1459.
- (5) Hutinet, G.; Kot, W.; Cui, L.; Hillebrand, R.; Balamkundu, S.; Gnanakalai, S.; Neelakandan, R.; Carstens, A. B.; Fa Lui, C.; Tremblay, D.; Jacobs-Sera, D.; Sassanfar, M.; Lee, Y.-J.; Weigele, P.; Moineau, S.; Hatfull, G. F.; Dedon, P. C.; Hansen, L. H.; de Crécy-Lagard, V. 7-Deazaguanine Modifications Protect Phage DNA from Host Restriction Systems. *Nat. Commun.* **2019**, *10*, 1–12.
- (6) Schaefer, M.; Pollex, T.; Hanna, K.; Tuorto, F.; Meusburger, M.; Helm, M.; Lyko, F. RNA Methylation by Dnmt2 Protects Transfer RNAs against Stress-Induced Cleavage. *Genes Dev.* **2010**, *24*, 1590–1595.
- (7) Helm, M.; Motorin, Y. Detecting RNA Modifications in the Epitranscriptome: Predict and Validate. *Nat. Rev. Genet.* **2017**, *18*, 275–291.
- (8) Khoddami, V.; Yerra, A.; Mosbrugger, T. L.; Fleming, A. M.; Burrows, C. J.; Cairns, B. R. Transcriptome-Wide Profiling of Multiple RNA Modifications Simultaneously at Single-Base Resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 6784–6789.
- (9) Squires, J. E.; Patel, H. R.; Nusch, M.; Sibbritt, T.; Humphreys, D. T.; Parker, B. J.; Suter, C. M.; Preiss, T. Widespread Occurrence of 5-Methylcytosine in Human Coding and Non-Coding RNA. *Nucleic Acids Res.* **2012**, *40*, S023–S033.
- (10) Schwartz, S.; Bernstein, D. A.; Mumbach, M. R.; Jovanovic, M.; Herbst, R. H.; León-Ricardo, B. X.; Engreitz, J. M.; Guttman, M.; Satija, R.; Lander, E. S.; Fink, G.; Regev, A. Transcriptome-Wide Mapping Reveals Widespread Dynamic-Regulated Pseudouridylation of ncRNA and mRNA. *Cell* **2014**, *159*, 148–162.
- (11) Kimura, S.; Dedon, P. C.; Waldor, M. K. Comparative TRNA Sequencing and RNA Mass Spectrometry for Surveying TRNA Modifications. *Nat. Chem. Biol.* **2020**, *16*, 964–972.
- (12) Schwartz, M. H.; Wang, H.; Pan, J. N.; Clark, W. C.; Cui, S.; Eckwahl, M. J.; Pan, D. W.; Parisien, M.; Owens, S. M.; Cheng, B. L.; Martinez, K.; Xu, J.; Chang, E. B.; Pan, T.; Eren, A. M. Microbiome Characterization by High-Throughput Transfer RNA Sequencing and Modification Analysis. *Nat. Commun.* **2018**, *9*, 1–13.
- (13) Ross, R.; Cao, X.; Yu, N.; Limbach, P. A. Sequence Mapping of Transfer RNA Chemical Modifications by Liquid Chromatography Tandem Mass Spectrometry. *Methods* **2016**, *107*, 73–78.
- (14) Nakayama, H.; Akiyama, M.; Taoka, M.; Yamauchi, Y.; Nobe, Y.; Ishikawa, H.; Takahashi, N.; Isobe, T. Ariadne: A Database Search Engine for Identification and Chemical Analysis of RNA Using Tandem Mass Spectrometry Data. *Nucleic Acids Res.* **2009**, *37*, No. e47.
- (15) Sample, P. J.; Gaston, K. W.; Alfonzo, J. D.; Limbach, P. A. RoboOligo: Software for Mass Spectrometry Data to Support Manual and de Novo Sequencing of Post-Transcriptionally Modified Ribonucleic Acids. *Nucleic Acids Res.* **2015**, *43*, No. e64.
- (16) Yu, N.; Lobue, P. A.; Limbach, P. A. RNAModMapper: RNA Modification Mapping Software for Analysis of Liquid Chromatography Tandem Mass Spectrometry Data. *Anal. Chem.* **2017**, *89*, 10744–10752.
- (17) Wein, S.; Andrews, B.; Sachsenberg, T.; Santos-Rosa, H.; Kohlbacher, O.; Kouzarides, T.; Garcia, B. A.; Weisser, H. A Computational Platform for High-Throughput Analysis of RNA Sequences and Modifications by Mass Spectrometry. *Nat. Commun.* **2020**, *11*, 1–12.
- (18) Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M.; Jungkamp, A.-C.; Munschauer, M.; Ulrich, A.; Wardle, G. S.; Dewell, S.; Zavolan, M.; Tuschl, T. Transcriptome-Wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell* **2010**, *141*, 129–141.
- (19) Coleman, R. S.; Siedlecki, J. M. Synthesis of a 4-thio-2'-deoxyuridine containing oligonucleotide. Development of the thio-carbonyl group as a linker element. *J. Am. Chem. Soc.* **1992**, *114*, 9229–9230.
- (20) Secrist, J. A.; Barrio, J. R.; Leonard, N. J. Attachment of a Fluorescent Label to 4-Thiouracil and 4-Thiouridine. *Biochem. Biophys. Res. Commun.* **1971**, *45*, 1262–1270.
- (21) Scheit, K. H. The reaction of 4-thiopyrimidine nucleosides and 4-thiouridine 5'-phosphate with ethyleneimine. *Biochim. Biophys. Acta, Nucleic Acids Protein Synth.* **1969**, *195*, 294–298.
- (22) Besada, P.; Shin, D. H.; Costanzi, S.; Ko, H.; Mathé, C.; Gagneron, J.; Gosselin, G.; Maddileti, S.; Harden, T. K.; Jacobson, K. A. Structure–Activity Relationships of Uridine 5'-Diphosphate Analogues at the Human P2Y6 Receptor. *J. Med. Chem.* **2006**, *49*, 5532–5543.
- (23) Herzog, V. A.; Reichholf, B.; Neumann, T.; Rescheneder, P.; Bhat, P.; Burkard, T. R.; Wlotzka, W.; Von Haeseler, A.; Zuber, J.; Ameres, S. L. Thiol-Linked Alkylation of RNA to Assess Expression Dynamics. *Nat. Methods* **2017**, *14*, 1198–1204.
- (24) Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.

- Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L. BioMagResBank. *Nucleic Acids Res.* **2008**, *36*, D402.
- (25) Young, A. P.; Bandarian, V. TYW1: A Radical SAM Enzyme Involved in the Biosynthesis of Wybutosine Bases. *Methods in Enzymology*; Academic Press Inc., 2018; Vol. 606, pp 119–153.
- (26) Lauhon, C. T.; Erwin, W. M.; Ton, G. N. Substrate Specificity for 4-Thiouridine Modification in Escherichia Coli. *J. Biol. Chem.* **2004**, *279*, 23022–23029.
- (27) Loiseau, L.; Ollagnier-de-Choudens, S.; Nachin, L.; Fontecave, M.; Barras, F. Biogenesis of Fe-S Cluster by the Bacterial Suf System. *J. Biol. Chem.* **2003**, *278*, 38352–38359.
- (28) Kellenberger, C. A.; Hammond, M. C. In Vitro Analysis of Riboswitch-Spinach Aptamer Fusions as Metabolite-Sensing Fluorescent Biosensors. *Methods in Enzymology*; Elsevier Inc., 2015; Vol. 550, pp 147–172.
- (29) Bolger, A. M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120.
- (30) Dobin, A.; Davis, C. A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29*, 15–21.
- (31) Trapnell, C.; Roberts, A.; Goff, L.; Pertea, G.; Kim, D.; Kelley, D. R.; Pimentel, H.; Salzberg, S. L.; Rinn, J. L.; Pachter, L. Differential Gene and Transcript Expression Analysis of RNA-Seq Experiments with TopHat and Cufflinks. *Nat. Protoc.* **2012**, *7*, 562–578.
- (32) Li, H. A Statistical Framework for SNP Calling, Mutation Discovery, Association Mapping and Population Genetical Parameter Estimation from Sequencing Data. *Bioinformatics* **2011**, *27*, 2987–2993.
- (33) Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P. Integrative Genomics Viewer. *Nature Biotechnology*; Nature Publishing Group, 2011; pp 24–26.
- (34) Rould, M. A.; Perona, J. J.; Söll, D.; Steitz, T. A. Structure of E. coli Glutamyl-tRNA Synthetase Complexed with tRNA Gln and ATP at 2.8 Å Resolution. *Science* **1989**, *246*, 1135–1142.
- (35) Singer, C. E.; Smith, G. R. Histidine Regulation in Salmonella Typhimurium. 13. Nucleotide Sequence of Histidine Transfer Ribonucleic Acid. *J. Biol. Chem.* **1972**, *247*, 2989–3000.
- (36) Röst, H. L.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weisser, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.-C.; Gutenbrunner, P.; Kenar, E.; Liang, X.; Nahnsen, S.; Nilse, L.; Pfeuffer, J.; Rosenberger, G.; Rurik, M.; Schmitt, U.; Veit, J.; Walzer, M.; Wojnar, D.; Wolski, W. E.; Schilling, O.; Choudhary, J. S.; Malmström, L.; Aebersold, R.; Reinert, K.; Kohlbacher, O. OpenMS: A Flexible Open-Source Software Platform for Mass Spectrometry Data Analysis. *Nat. Methods* **2016**, *13*, 741–748.
- (37) Juhling, F.; Morl, M.; Hartmann, R. K.; Sprinzl, M.; Stadler, P. F.; Putz, J. TRNAdb 2009: Compilation of TRNA Sequences and TRNA Genes. *Nucleic Acids Res.* **2009**, *37*, D159–D162.
- (38) Boccaletto, P.; MacHnicka, M. A.; Purta, E.; Piątkowski, P.; Bagiński, B.; Wirecki, T. K.; de Crécy-Lagard, V.; Ross, R.; Limbach, P. A.; Kotter, A.; Helm, M.; Bujnicki, J. M. MODOMICS: A Database of RNA Modification Pathways. 2017 Update. *Nucleic Acids Res.* **2018**, *46*, D303–D307.
- (39) Fox, J. J.; Van Praag, D.; Wempen, I.; Doerr, I. L.; Cheong, L.; Knoll, J. E.; Eidinoff, M. L.; Bendich, A.; Brown, G. B. Thiation of Nucleosides. II. Synthesis of 5-Methyl-2'-deoxycytidine and Related Pyrimidine Nucleosides I. *J. Am. Chem. Soc.* **1959**, *81*, 178–187.
- (40) Lipsett, M. N. The Isolation of 4-Thiouridylic Acid from the Soluble Ribonucleic Acid of Escherichia Coli. *J. Biol. Chem.* **1965**, *240*, 3975–3978.
- (41) Carbon, J.; David, H.; Studier, M. H. Thiobases in Escherichia Coli Transfer RNA: 2-Thiocytosine and 5-Methylaminomethyl-2-Thiouracil. *Science* **1968**, *161*, 1146–1147.
- (42) Favre, A.; Michelson, A. M.; Yaniv, M. Photochemistry of 4-thiouridine in Escherichia coli transfer RNA1Val. *J. Mol. Biol.* **1971**, *58*, 367–379.
- (43) Bergstrom, D. E.; Leonard, N. J. Photoreaction of 4-Thiouracil with Cytosine. Relation to Photoreactions in Escherichia Coli Transfer Ribonucleic Acids. *Biochemistry* **1972**, *11*, 1–9.
- (44) Favre, A.; Roques, B.; Fourrey, J.-L. Chemical Structures of the TU-C and TU-Cred Products Derived from E. Coli TRNA. *FEBS Lett.* **1972**, *24*, 209–214.
- (45) Thomas, G.; Favre, A. 4-Thiouridine as the Target for near-Ultraviolet Light Induced Growth Delay in Escherichia Coli. *Biochem. Biophys. Res. Commun.* **1975**, *66*, 1454–1461.
- (46) Ramabhadran, T. V.; Fossum, T.; Jagger, J. In Vivo Induction of 4-thiouridine-cytidine Adducts in TRNA of E. Coli B/r by Near-ultraviolet Radiation. *Photochem. Photobiol.* **1976**, *23*, 315–321.
- (47) Thomas, G.; Favre, A. 4-Thiouridine Triggers Both Growth Delay Induced by Near-Ultraviolet Light and Photoprotection. *Eur. J. Biochem.* **1980**, *113*, 67–74.
- (48) Kramer, G. F.; Baker, J. C.; Ames, B. N. Near-UV Stress in Salmonella Typhimurium: 4-Thiouridine in TRNA, PpGpp, and ApppGpp as Components of an Adaptive Response. *J. Bacteriol.* **1988**, *170*, 2344–2351.
- (49) Tsai, S.-C.; Jagger, J. The Roles of the Rel+gene and of 4-Thiouridine in Killing and Photoprotection of Escherichia Coli by Near-ultraviolet Radiation. *Photochem. Photobiol.* **1981**, *33*, 825–834.
- (50) Ramabhadran, T. V.; Jagger, J. Mechanism of Growth Delay Induced in Escherichia Coli by near Ultraviolet Radiation. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 59–63.
- (51) Blanchetot, A.; Hajnsdorf, E.; Favre, A. Metabolism of tRNA in near-ultraviolet-illuminated Escherichia coli The tRNA repair hypothesis. *Eur. J. Biochem.* **1984**, *139*, 547–552.
- (52) Hajnsdorf, E.; Favre, A. Metabolism of TRNAs in Growing Cells of Escherichia Coli Illuminated with Near-ultraviolet Light. *Photochem. Photobiol.* **1986**, *43*, 157–164.
- (53) Favre, A.; Moreno, G.; Salet, C.; Vinzens, F. 4-Thiouridine Incorporation into the Rna of Monkey Kidney Cells (Cv-1) Triggers Near-uv Light Long-term Inhibition of Dna, Rna and Protein Synthesis. *Photochem. Photobiol.* **1993**, *58*, 689–694.
- (54) Melvin, W. T.; Milne, H. B.; Slater, A. A.; Allen, H. J.; Keir, H. M. Incorporation of 6-Thioguanosine and 4-Thiouridine into RNA. Application to Isolation of Newly Synthesised RNA by Affinity Chromatography. *Eur. J. Biochem.* **1978**, *92*, 373–379.
- (55) Shibata, H.; Ochiai, H.; Uchida, I. Incorporation of 4-Thiouridine into RNA in Germinating Radish Seeds. *FEBS Lett.* **1980**, *119*, 85–89.
- (56) Bezerra, R.; Favre, A. In Vivo Incorporation of the Intrinsic Photolabel 4-Thiouridine into Escherichia Coli RNAs. *Biochem. Biophys. Res. Commun.* **1990**, *166*, 29–37.
- (57) Miles, Z. D.; McCarty, R. M.; Molnar, G.; Bandarian, V. Discovery of Epoxyqueuosine (OQ) Reductase Reveals Parallels between Halorespiration and TRNA Modification. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 7368–7372.
- (58) McCarty, R. M.; Somogyi, Á.; Lin, G.; Jacobsen, N. E.; Bandarian, V. The Deazapurine Biosynthetic Pathway Revealed: In Vitro Enzymatic Synthesis of PreQ0 from Guanosine 5'-Triphosphate in Four Steps. *Biochemistry* **2009**, *48*, 3847–3852.
- (59) Okada, N.; Nishimura, S. Isolation and Characterization of a Guanine Insertion Enzyme, a Specific TRNA Transglycosylase, from Escherichia Coli. *J. Biol. Chem.* **1979**, *254*, 3061–3066.
- (60) Slany, R. K.; Boesl, M.; Crain, P. F.; Kersten, H. A New Function of S-Adenosylmethionine: The Ribosyl Moiety of AdoMet Is the Precursor of the Cyclopentenediol Moiety of the TRNA Wobble Base Queuine. *Biochemistry* **1993**, *32*, 7811–7817.
- (61) Young, A. P.; Bandarian, V. Radical Mediated Ring Formation in the Biosynthesis of the Hypermodified TRNA Base Wybutosine. *Curr. Opin. Chem. Biol.* **2013**, *17*, 613–618.
- (62) Noma, A.; Kirino, Y.; Ikeuchi, Y.; Suzuki, T. Biosynthesis of Wybutosine, a Hyper-Modified Nucleoside in Eukaryotic Phenylalanine TRNA. *EMBO J.* **2006**, *25*, 2142–2154.
- (63) Lauhon, C. T. Requirement for IscS in Biosynthesis of All Thionucleosides in Escherichia Coli. *J. Bacteriol.* **2002**, *184*, 6820–6829.

(64) Lauhon, C. T.; Kambampati, R. The *iscS* Gene in *Escherichia coli* Is Required for the Biosynthesis of 4-Thiouridine, Thiamin, and NAD. *J. Biol. Chem.* **2000**, *275*, 20096.

(65) Kingston, R. E.; Chomczynski, P.; Sacchi, N. Guanidine Methods for Total RNA Preparation. *Curr. Protoc. Mol. Biol.* **1996**, *36*, 4.2.1–4.2.9.

(66) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24*, 2534–2536.

(67) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30*, 918–920.

(68) Andrews, S. *FastQC A Quality Control Tool for High Throughput Sequence Data*; Babraham Bioinformatics, 2018; pp 3–5.

(69) Depristo, M. A.; Banks, E.; Poplin, R.; Garimella, K. V.; Maguire, J. R.; Hartl, C.; Philippakis, A. A.; Del Angel, G.; Rivas, M. A.; Hanna, M.; McKenna, A.; Fennell, T. J.; Kernysky, A. M.; Sivachenko, A. Y.; Cibulskis, K.; Gabriel, S. B.; Altshuler, D.; Daly, M. J. A Framework for Variation Discovery and Genotyping Using Next-Generation DNA Sequencing Data. *Nat. Genet.* **2011**, *43*, 491–498.

(70) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, Ş.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. The PRIDE Database and Related Tools and Resources in 2019: Improving Support for Quantification Data. *Nucleic Acids Res.* **2019**, *47*, D442–D450.