Pseudouridine in the Anticodon of *Escherichia coli* tRNA$_{\text{Tyr}}$(QΨA) Is Catalyzed by the Dual Specificity Enzyme RluF

Received for publication, July 11, 2016, and in revised form, August 15, 2016 Published, JBC Papers in Press, August 22, 2016, DOI 10.1074/jbc.M116.747865

© Balasubramanyam Addepalli and Patrick A. Limbach

From the Department of Chemistry, Rieveschl Laboratories for Mass Spectrometry, University of Cincinnati, Cincinnati, Ohio 45221

Pseudouridine is found in almost all cellular ribonucleic acids (RNAs). Of the multiple characteristics attributed to pseudouridine, making messenger RNAs (mRNAs) highly translatable and non-immunogenic is one such feature that directly implies this modification in protein synthesis. We report the existence of pseudouridine in the anticodon of *Escherichia coli* tyrosine transfer RNAs (tRNAs) at position 35. Pseudouridine was verified by multiple detection methods, which include pseudouridine-specific chemical derivatization and gas phase dissociation of RNA during liquid chromatography tandem mass spectrometry (LC-MS/MS). Analysis of total tRNA isolated from *E. coli* pseudouridine synthase knock-out mutants identified RluF as the enzyme responsible for this modification. Furthermore, the absence of this modification compromises the translational ability of a luciferase reporter gene coding sequence when it is preceded by multiple tyrosine codons. This effect has implications for the translation of mRNAs that are rich in tyrosine codons in bacterial expression systems.

Of 150-plus chemical modifications (1, 2) found in cellular RNA, 5-ribosyluridine or pseudouridine (Ψ) (3) is the most abundant but still poorly understood (4) modification. Pseudouridine, which has been found in the functionally important regions of RNAs (5), is known to modulate codon-anticodon interactions between mRNA and tRNA (6) and assist assembly of the ribosome (7) and spliceosome (8). The presence of pseudouridine in mRNA makes the message non-immunogenic and highly translatable, thus offering therapeutic opportunities in medical applications (9).

Although the exact mechanism of isomerization is still not clear (4), the incorporation of C5 of the uracil base into the glycosidic bond is catalyzed by either standalone proteins (referred to as Ψ synthases) (10, 11) or RNA–protein complexes (referred to as H/ACA box ribonucleoproteins) (12). Following site-specific recognition, the nitrogen–carbon glycosidic bond is cleaved, and uracil is rotated and then reattached to the sugar, forming a carbon–carbon glycosidic bond on the polyribonucleotide (Fig. 1) by Michael addition-like or glycal mechanisms (13).

The Ψ synthases have been initially classified based on the *Escherichia coli* enzymes RluA, RsuA, TruA, TruB, and TruD (10, 11). A sixth family with members from archaea and eukaryotes but not from bacteria was added after discovery of Pus10 (14). The enzyme active site carries a catalytically essential aspartate, which is the only absolutely conserved residue in all Ψ synthases (15, 16). Substrate recognition by the Ψ synthase is usually in the context of the sequence or structure of the target site in RNA. Site specificity also varies depending on the Ψ synthase. For example, RsuA isomerizes US16 of 16S rRNA (17) exclusively, whereas the universally conserved Ψ55 of the ΨCS loop in all elongator tRNAs is catalyzed by TruB (18). RluA catalyzes modification of two distinct RNAs, 23S rRNA (position 746) and some tRNAs (position 32) (19). Members of the RluA and RsuA families share three conserved sequence motifs (Motifs I, II, and III) apart from domains similar to ribosomal protein S4 (20).

Enzymatic activity of pseudouridine synthase is measured by a $^3$H release assay involving *in vitro* transcripts (21). Pseudouridine in cellular RNA is assayed quantitatively by enzymatic hydrolysis and thin-layer or liquid chromatography where differences in migration distance or elution time, respectively, distinguish pseudouridine from uridine. Additional specificity in detection is available by coupling liquid chromatography (LC) with electrospray ionization mass spectrometry (MS) (22).

A major drawback of nucleoside hydrolysates is the loss of sequence information. A widely used site-specific detection method involves alkali-stable, Ψ-specific irreversible derivatization using the reagent N-cyclohexyl-N′-β-(4-methylphospholinium)ethyl carbodiimide p-tosylate (CMCT) followed by extension of a radiolabeled primer with reverse transcriptase, separation, and detection on a denaturing polyacrylamide gel (23). Recently, CMCT-based derivatization has been adapted to high throughput next generation sequencing to detect dynamic pseudouridylations in mRNAs (24, 25). Mass spectrometry-based RNA modification mapping in combination with chem-
Pseudouridine in E. coli tRNA\textsuperscript{Tyr} Anticodon

FIGURE 1. Pseudouridine in RNA. Uridine is isomerized into pseudouridine by pseudouridine synthases through base cleavage from ribose sugar and reattachment to generate C–C glycosidic bond.

A direct tandem mass spectrometry (MS/MS) approach based on the unique fragmentation pathways of the stable C–C glycosidic bond to generate and detect Ψ-specific fragment/product ions has been developed to confirm the presence of pseudouridine in oligoribonucleotides (32). This selected reaction monitoring (SRM) approach, compatible with MS-based RNA modification mapping, has been used to quantify the levels of pseudouridine in RNA (33). Other quantification methods involve radiolabeling of the phosphate backbone, RNase H-mediated cleavage, and thin layer chromatography of hydrolyzed nucleosides for phosphorimaging (34), a ligation-based method that facilitates the recognition of the modification site and quantification on denaturing PAGE (35), relative quantification from base hydrolysis of oligonucleotides (36), and stable isotope labeling of RNA (37).

The original E. coli tRNA\textsuperscript{Tyr} sequences described in the literature are reported to contain two pseudouridines at position 39 and 55 (38). Prior reports discuss the possibility of an additional pseudouridine (31) or off-target derivatization (28), but until now, those possibilities remained uninvestigated. Here we have identified the additional pseudouridine derivatization site and quantification from base hydrolysis of oligonucleotides (36), and stable isotope labeling of RNA (37).

The original E. coli tRNA\textsuperscript{Tyr} sequences described in the literature are reported to contain two pseudouridines at position 39 and 55 (38). Prior reports discuss the possibility of an additional pseudouridine (31) or off-target derivatization (28), but until now, those possibilities remained uninvestigated. Here we have identified the additional pseudouridine derivatization site and quantification from base hydrolysis of oligonucleotides (36), and stable isotope labeling of RNA (37).

Results

LC-MS/MS Analysis of RNase T1 Digest of E. coli tRNA\textsuperscript{Tyr}—The Tyr I and Tyr II tRNAs from E. coli share an identical sequence except at two nucleotide positions in the variable region (UC versus CA) (Table 1). LC-MS analysis of tRNA\textsuperscript{Tyr} following treatment with RNase T1 and bacterial alkaline phosphatase (BAP) revealed a single response (40) for the triply charged oligonucleotide anion ACU[Q][UA][m5\textsuperscript{2}A]A[Ψ]CUG (m/z 1338.9, where U is the putative site of pseudouridine), which is a signature ion for E. coli tRNA\textsuperscript{Tyr} (41) (supplemental Fig. S1). Collision-induced dissociation (CID) MS/MS of this mass selected precursor ion yielded all the expected sequence-informative product ion series, viz. \(c_n\) (sharing the 3′-end) and \(y_n\) (sharing the 3′-end) (supplemental Table S1), from which the original sequence could easily be reconstructed. This digestion product contains the anticodon, [Q]UA, responsible for decoding tyrosine codons, UAY (Y = U or C), in the mRNA. This underivatized RNA digest did not exhibit any interfering MS signal for m/z values corresponding to additional carbodiimide or cyanoethylation units, indicating the potential for interference-free pseudouridine detection of a derivatized digestion product.

CMCT-treated RNA—LC-MS analysis of CMCT-derivatized T1 and BAP digest of E. coli tRNA\textsuperscript{Tyr} I or II yielded very little signal for the underivatized oligonucleotide anion but produced a high abundance signal for m/z values that correspond to the digestion product with one and two carbodiimide units (addition of 251 and 502 Da), respectively (Fig. 2, A and B), consistent with previous observations (28). CID MS/MS of these m/z values (Figs. 3 and 4) revealed differences with regard to the location of the second carbodiimide. When m/z values of product ions resulting from the dissociated precursor ion that have the common 3′-end (\(y_n\) ion series where \(n\) refers to nucleotide position) were scored, an additional mass corresponding to the carbodiimide was detected in the MS/MS data from \(y_n\) (m/z 1199.8 → m/z 1450.8) through \(y_1\), which is consistent with the presence of the known pseudouridine modification at position 39 (38). The \(c_n\) series yielded no product ions with carbodiimide for uridine at the third position, but carbodiimide-containing product ions were observed from \(c_n\) (m/z 1717.1 → m/z 1967.5) through \(c_1\) (Fig. 3 and supplemental Table S1). This result led to speculation that there could be pseudouridine at U35 in this tRNA. Furthermore, this observation also indicated that this precursor ion population is a mixture of oligonucleotides with carbodiimide at position 35 or 39 of the tRNA sequence (Fig. 3).

Sequence of ACU[Q][UA][m5\textsuperscript{2}A]A[Ψ]CUG with two carbodiimide units also revealed that the carbodiimide-tagged product ions started with \(y_n\) (m/z 1450.8) and \(c_5\) (m/z 1967.5) (Fig. 4). However, the mass of detected product ions shifted to two carbodiimide units from \(c_5\) (m/z 1687.7 → m/z 1813.3) onward and \(y_n\) (m/z 1429.0 → m/z 1554.6) onward (supplemental Table S1). These data are in contrast with a previous finding (28) where RNase A and snake venom phosphodiesterase-digested tRNA\textsuperscript{Tyr} was analyzed by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) after CMCT derivatization. Those earlier data also showed that two carbodiimide units were added to the anticodon region of tRNA\textsuperscript{Tyr}. However, the exonuclease digestion of the RNase A

TABLE 1

Sequences of E. coli tRNA\textsuperscript{Tyr} I and II that differ at two nucleotide positions (shown in italics)

<table>
<thead>
<tr>
<th>Sequence</th>
<th>E. coli tRNA\textsuperscript{Tyr} I</th>
<th>E. coli tRNA\textsuperscript{Tyr} II</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGUGGGGG[Ψ][UA][CUG][m5\textsuperscript{2}A]A[Ψ]CUG</td>
<td>GGGCACAGCUQ[UA][CUG][m5\textsuperscript{2}A]A[Ψ]CUG</td>
<td></td>
</tr>
<tr>
<td>CCGUCA[CA]GACUUGAAG</td>
<td>CCGUCA[CA]GACUUGAAG</td>
<td></td>
</tr>
<tr>
<td>AAUCCU[UU]CCCAACCA</td>
<td>AAUCCU[UU]CCCAACCA</td>
<td></td>
</tr>
</tbody>
</table>

Nucleotides in bold are part of RNase T1 signature digestion products (see text).

Uridine is isomerized into pseudouridine by pseudouridine synthases through base cleavage from ribose sugar and reattachment to generate C–C glycosidic bond.
product (A[ms2i6A]Ap) suggested carbodiimide tagging of ms2i6A at position 37. A key difference between the current study and that in Ref. 28 is the use of CID MS/MS in the present study to precisely define the location of carbodiimide tagging. Although both the present findings and the prior publication identified two sites of carbodiimide tagging in the anticodon loop of tRNATyr, only in this present work is the exact location of an additional pseudouridine revealed.

Given the disparity between the MALDI-MS data from the earlier report and the LC-MS/MS data obtained here, tRNATyr was digested with RNase U2 to yield the digestion product CU[Q]UA[p (m/z 857.7, where >p represents a 2’3’-cyclic phosphate) (supplemental Fig. S2) that is specific to the anticodon region (42). CMCT-based derivatization and subsequent examination of product ions obtained by CID of m/z 983.1 resulted in a 251-Da mass shift of c4 and y2, which is consistent with RNase T1 derivatization data that the second site for carbodiimide tagging is pseudouridine at position 35.

Acrylonitrile-treated RNA—To confirm the CMCT derivatization results, an alternative derivatization strategy was explored based on the specificity of acrylonitrile for pseudouridine (30, 31). LC-MS analysis of the acrylonitrile-treated T1 digest revealed the presence of triply charged oligonucleotide anions at m/z 1356.8 and 1374.3 that correspond to one and two cyanoethylations (addition of 53 and 106 Da), respectively (supplemental Fig. S3 and supplemental Table S1). CID MS/MS of m/z 1356.8 showed the presence of product ions that correspond to ACU[Q]UA[ms2i6A]A[CUG with a mass shift of 53 Da from y4 to y11 and c5 to c11. The profile of these product ions in the tandem mass spectrum followed a pattern...
Pseudouridine in E. coli tRNA<sub>Tyr</sub> Anticodon

![Graph](image_url)

**Figure 4.** CID-based sequencing of the double carbodiimide-tagged RNase T1 digestion product from the anticodon of E. coli tRNA<sub>Tyr</sub>UAA. Shown is the CID MS/MS spectrum of ACU[Q]UA[ms<sup>2</sup>PA][Ψ]CUG with two carbodiimide units (m/z 1506.6). The appearance of unique product ions with two carbodiimide adducts, c<sub>4</sub> and y<sub>6</sub>, defines the location of two pseudouridines in the oligonucleotide. Each addition of carbodiimide (+251) is denoted with * for the y product ion series and † for the c product ion series.

identical to that observed during carbodiimide derivatization. However, the MS/MS spectrum was more complex due to fragment ions corresponding to side chain fragmentation of quaeosine and nucleobase loss for cytosine, adenine, uracil, and ms<sup>2</sup>PA. Similarly, the MS/MS spectrum of m/z 1374.3 exhibited a mass shift of ~53 Da from y<sub>4</sub> to y<sub>6</sub> and c<sub>5</sub> to c<sub>6</sub>, which thereafter shifted by another 53 Da through y<sub>11</sub> and c<sub>11</sub> (supplemental Table S1). Thus, two independent derivatization strategies yielded data that point to the presence of pseudouridine at position 35 of tRNA<sub>Tyr</sub>UAA.

**SRM-based Pseudouridine Detection—**Pseudouridine at position 35 of tRNA<sub>Tyr</sub>UAA was further validated by subjecting the RNase U2 digest to LC-MS/MS-based SRM analysis. This experiment revealed the presence of the pseudouridine-specific SRM transition (m/z 207 → m/z 164) that corresponded with the extracted ion chromatogram of intact CU[Q]UA->p (supplemental Fig. S4) indicating pseudouridine in this oligomer.

**Nucleoside Analysis—**The CU[Q]UA->p digestion product was purified after post-column split and digested to nucleosides. LC-MS/MS analysis of the total nucleoside digest confirmed the presence of pseudouridine as it could be distinguished from uridine by its retention time (supplemental Fig. S5). The modified quaeosine expected in this RNase U2 digestion product was also detected, providing additional confidence that pseudouridine is associated with the anticodon of tRNA<sub>Tyr</sub>UAA.

**Analysis of Pseudouridine Synthase Disruption Mutants—**After confirming the presence of pseudouridine at position 35 of E. coli tRNA<sub>Tyr</sub>UAA, we embarked on identifying the pseudouridine synthase(s) responsible for this modification by analyzing gene knock-out strains for each E. coli pseudouridine synthase. Total RNA was isolated, digested with RNase T1, and analyzed for the doubly cyanoethylated signature digestion product of tRNA<sub>Tyr</sub>UAA by LC-MS/MS. A comparison of the computed peak areas for the singly and doubly cyanoethylated ions of ACU[Q]UA[ms<sup>2</sup>PA][Ψ]CUG from all of the pseudouridine synthase mutants should indicate the putative enzyme responsible for this modification.

Fig. 5A depicts the peak areas of m/z 1356.8 and 1374.3 from the acrylonitrile-treated total tRNA digest. A signal at m/z 1356.8 is present in all treated samples, although the ion abundance varied among different mutants. In contrast, the peak area corresponding to m/z 1374.3 showed a dramatic decrease for ΔtruA and ΔrluF. To ensure the derivatization conditions were optimal across all samples, the cyanoethylation status of a tRNA<sup>Tyr</sup> gene knock-out strain for each pseudouridine synthase mutants should indicate the putative enzyme responsible for this modification.

To distinguish whether TruA or RluF could complement the loss of modification at position 35, the coding regions of TruA and RluF<sup>4</sup> were amplified from an E. coli K12 strain, and the confirmed sequence was cloned into the pET22b vector for subsequent transformation of the ΔrluF or ΔtruA strains. Fig. 5C depicts the cyanoethylation status of the signature digestion product of tRNA<sub>Tyr</sub>UAA in complemented lines. The data clearly show that the ΔrluF strain can only be complemented with the RluF coding sequence as the peak area of m/z 1374.3 for this complementation is higher by many orders of magnitude compared with the complementation with TruA. Similarly, the peak area for m/z 1374.3 increased only after complementation of the ΔtruA strain with TruA.

**In Vitro Modification of tDNA<sup>Tyr</sup>UAA—**The in vitro transcribed tDNA<sup>Tyr</sup>UAA gene was incubated with purified RluF or TruA, digested with RNase T1 and BAP, and then derivatized by cyanoethylation. Due to the lack of quaeosine at position 34 of the in vitro transcript, RNase T1 digestion will generate the oligonucleotide UAAAUCUG (m/z 1245.1) where the 5′-uridine corresponds to position 35 of the native tRNA. LC-MS analysis of both samples yielded the underivatized and cyanoethylated digestion products that correspond to the anticodon region of the tDNA transcript. MS/MS analysis (Fig. 6) revealed differences with respect to the position of cyanoethylation on the nucleotide sequence. RluF-treated RNA exhibited cyanoethylation starting with c<sub>1</sub> and continuing through c<sub>7</sub> with no mass shift for any y-type product ions (Fig. 6A). In contrast, the

---

*Citation:* Journal of Biological Chemistry, DOI: 10.1074/jbc.M116.762508.

*Authors:* J. Smith, A. Johnson, K. Williams.

*Institution:* University of California, San Francisco.

*Keywords:* Pseudouridine, tRNA, E. coli, RNA, MS/MS, SRM.
TruA-treated RNA exhibited cyanoethylation from c₅ through c₇ only and from y₄ to y₇ (Fig. 6B).

A comparison of the peak areas corresponding to m/z 1245.2 and 1272.2 indicated that of 50 pmol of RNA incubated with 5 pmol of protein about 35% was cyanoethylated upon treatment with RluF and acrylonitrile. In contrast, about 48% of 1500 pmol of RNA (incubated with 0.2 pmol of protein) was cyanoethylated upon treatment with TruA and acrylonitrile. The observed levels of in vitro modification match with the previously reported tritium release assays by other groups (44, 45).

When TruA or RluF was omitted but the transcript was treated with acrylonitrile, less than 5% of the RNA was cyanoethylated at random positions (data not shown). The RluF-treated RNA transcript was also analyzed using the pseudouridine-specific SRM assay. SRM signals from the RluF-treated RNA align with the peak corresponding to UAAAUCUG (supplemental Fig. S6).

FIGURE 5. Identification of synthase gene responsible for pseudouridine modification at position 35 of E. coli tRNA^{Tyr} (QUA) transcript. Four micrograms of tRNA isolated from each of the pseudouridine synthase knock-out mutant cell lines (three to five biological replicates) were digested with RNase T1 and BAP before treating with acrylonitrile. The derivatized RNA was subjected to LC-MS. A, peak areas corresponding to one and two cyanoethylation groups on AC[Q]UA[ms²FA]AΨCCUG of tRNA^{Tyr} were plotted against each synthase gene. The truA and rluF gene knock-out tRNA samples reveal a significant decrease in cyanoethylation. B, peak areas corresponding to one cyanoethylation group for AA[ms²FA]AΨCCCG of tRNA^{Phe} were plotted against each synthase gene. Only truA exhibited a dramatic decrease in peak area, whereas the mutation in the other genes including rluF strain had no adverse impact on the peak area. C, complementation effects of TruA or RluF gene under truA or rluF genetic background on pseudouridine modification at position 35 of tRNA^{Tyr}. Note that the complementation ability of the wild type genes is possible only under the respective knock-out mutation background, i.e., rluF-RluF and truA-TruA, but not in the cases of rluF-TruA and truA-RluF. Error bars represent S.D. of replicate measurements.

TruA-treated RNA exhibited cyanoethylation from c₅ through c₇ only and from y₄ through y₇ (Fig. 6B).

A comparison of the peak areas corresponding to m/z 1245.2 and 1272.2 indicated that of 50 pmol of RNA incubated with 5 pmol of protein about 35% was cyanoethylated upon treatment with RluF and acrylonitrile. In contrast, about 48% of 1500 pmol of RNA (incubated with 0.2 pmol of protein) was cyanoethylated upon treatment with TruA and acrylonitrile. The observed levels of in vitro modification match with the previously reported tritium release assays by other groups (44, 45).

When TruA or RluF was omitted but the transcript was treated with acrylonitrile, less than 5% of the RNA was cyanoethylated at random positions (data not shown). The RluF-treated RNA transcript was also analyzed using the pseudouridine-specific SRM assay. SRM signals from the RluF-treated RNA align with the peak corresponding to UAAAUCUG (supplemental Fig. S6).

RluF Modifies 23S rRNA—RluF is known to isomerize uridine to pseudouridine at position 2604 of 23S rRNA. To confirm this behavior remains when detecting pseudouridine at position 35 of tRNA^{Tyr}, the pseudouridylation status of 23S rRNA was...
FIGURE 7. LC-MS analysis of acrylonitrile-derivatized RNase T1/BAP digests of E. coli 23S rRNA. A, XIC for m/z 1305.2 corresponding to [ΨΨΨ]CG with two cyanoethylation. B, mass spectrum corresponding to XIC peak at 23.4 min showing the presence of singly charged oligonucleotide [ΨΨΨ]CG anion with two cyanoethylation. C, MS/MS spectrum of m/z 1305.2. D, XIC for m/z 1252.6 corresponding to [U]ΨCG with one cyanoethylation. E, mass spectrum corresponding to XIC peak at 23.1 min showing the presence of singly charged oligonucleotide [U]ΨCG anion with one cyanoethylation. F, MS/MS spectrum of m/z 1252.6. Each addition of cyanoethylation (+53) is denoted with * for the w and y product ion series and † for the c product ion series. Rel Abund, relative abundance.

FIGURE 8. Luciferase assay measuring decreased gene expression under rluF gene knock-out background when the coding sequence is preceded by multiple tyrosine codons. A recombinant firefly luciferase gene was constructed with a set of 10 tyrosine codons (TAY)₁₀FF or 10 asparagine codons ([AAAY]₁₀FF) preceding the firefly luciferase open reading frame. Expression of the luciferase gene in combination with either Renilla luciferase or RluF gene in the rluF gene knock-out background was conducted using a pETDuet vector. The type of recombinant construct used in the luciferase assay is denoted on the x axis. At least four to five biological replicates were tested for each construct. Error bars represent S.D. of replicate measurements.

Pseudouridine in E. coli tRNA<sup>Tyr</sup> Anticodon

compared between K12 and ΔrluF strains by digesting with RNase T1 and BAP followed by acrylonitrile treatment. Fig. 7, A–C, depicts the presence of the RNase T1 digestion product ΨΨΨCG with two cyanoethylation groups (m/z 1199.2 + 106 = m/z 1305.2) from K12 23S rRNA. However, 23S rRNA purified from the ΔrluF strain exhibited only one cyanoethylation (m/z 1199.2 + 53 = m/z 1252.1), and MS/MS analysis revealed the mass shift from c₂ onward, and only y₁ exhibited the 53-Da mass shift (Fig. 7, D–F). These data indicate that 23S rRNA was not modified at position 2604 in the ΔrluF strain.

Luciferase Reporter Assay—The above experimental data make a strong case for pseudouridine modification in the anti-

codon at position 35 of tRNA<sup>Tyr</sup>. To understand the role of this modification in decoding Tyr codons, a recombinant reporter gene was constructed with multiple Tyr codons (TAY)₁₀ (to magnify any subtle effects it may have) preceding the open reading frame of the firefly luciferase coding sequence. Luciferase expression was tested under the ΔrluF background in combination with Renilla (Ren) luciferase as an internal control (construct pETDuet “Ren Luc and (TAY)₁₀FF Luc”). A complementation construct, “RluF and (TAY)₁₀FF Luc,” tested the ability of the RluF protein to rescue luciferase activity under the ΔrluF background. A third construct containing “Ren Luc and (AAAY)₁₀FF Luc” was also tested to probe the influence of an identical number of multiple codons preceding the reporter gene. As depicted in Fig. 8, transformation of ΔrluF cells with pETDuet Renilla reniformis luciferase (Ren Luc) and (TAY)₁₀ firefly (Photinus pyralis) luciferase (FF Luc) and subsequent induction by isopropyl β-D-1-thiogalactopyranoside (IPTG) at mid-log phase showed very little expression of luciferase as determined by bioluminescence. However, the luminescence from Renilla luciferase was unaffected in these cells. Conversely, when RluF was co-expressed with tyrosine codon-rich FF luciferase ((TAY)₁₀FF Luc), luciferase could easily be detected, indicating that the induced RluF protein complements the ΔrluF knock-out. The wild type firefly luciferase (without the preceding multiple tyrosine codons) expression under a ΔrluF background is more or less similar to the RluF complement, indicating that ΔrluF is inefficient in translating only those mRNAs that are rich in Tyr codons. Furthermore, luminescence levels were unaffected when recombinant (TAY)₁₀FF Luc was expressed in BL21 cells (data not shown), indicating that addition of multiple tyrosine codons has no impact on general expression and activity of FF Luc in E. coli under wild type conditions. Thus, these experimental data suggest that the pseudouridine modification at position 35 of
Pseudouridine in E. coli tRNA\textsuperscript{Tyr} Anticodon
tRNA\textsuperscript{Tyr} has a very subtle impact on mRNA translation under normal conditions, but the impact can be significant if the mRNA is rich in Tyr codons.

**Discussion**

The original *E. coli* tRNA\textsuperscript{Tyr} sequences described in the literature are reported to contain two pseudouridines at positions 39 and 55 (38). Prior reports discussed the possibility of an additional pseudouridine (31) or off-target derivatization (28), but until now, those possibilities remained uninvestigated. To clarify this open question, we conducted an exhaustive examination of *E. coli* tRNA\textsuperscript{Tyr} using multiple pseudouridine detection strategies. Two different derivatization strategies along with SRM-based direct detection confirmed that the originally annotated uridine at position 35 of tRNA\textsuperscript{Tyr} is actually modified to pseudouridine. We found no difference in \(\Psi\)35 modification levels between the two isoacceptors of tRNA\textsuperscript{Tyr} (Tyr I and Tyr II), indicating that both tRNAs serve as a substrate for the tRNA:Ψ35 synthase.

*E. coli* is known to contain 11 pseudouridine synthase genes, and any one of them could be responsible for pseudouridine at position 35 of tRNA\textsuperscript{Tyr}. Four *E. coli* genes are known to be tRNA-specific, six are rRNA-specific, and only one (rluA) is known to exhibit dual specificity, catalyzing pseudouridine formation for 23S rRNA (position 746) and tRNA\textsuperscript{Phe} (position 32) (46). We reasoned that screening each of the pseudouridine synthase knock-out strains for the double derivatization status of the tRNA\textsuperscript{Tyr} signature digestion product (41), ACU[Q]UA[ms\textsuperscript{T}A][Ψ]-CUG, could identify the putative modifying enzyme. Indeed, this strategy identified significantly reduced levels of double cyanoethylated signature digestion products from the Δ\textit{trua} and Δ\textit{rluF} strains.

TruA is known to exhibit broad specificity, catalyzing pseudouridine formation in the anticodon loop and stem of multiple tRNAs (43). Absence of this enzyme would affect the pseudouridylation status of tRNAs beyond tRNA\textsuperscript{Tyr}. To differentiate TruA effects, the pseudouridylation status of the *E. coli* tRNA\textsuperscript{Phe} signature digestion product, AA[ms\textsuperscript{T}A][Ψ]-CCCCCG, was also examined for each pseudouridine synthase knock-out. Thus, by using mass spectrometry to examine two tRNA-specific signature digestion products, it was found that RluF is the enzyme responsible for pseudouridine modification at position 35 of tRNA\textsuperscript{Tyr}. Complementation of Δ\textit{trua} and Δ\textit{rluF} strains with their respective wild type sequences confirmed the studies of the knock-outs. Moreover, the site-specific pseudouridine modification of an \textit{in vitro} tRNA\textsuperscript{Tyr} transcript by RluF was consistent with all other pseudouridine synthase data. These studies, therefore, have determined that RluF is a second dual specificity pseudouridine synthase responsible for catalyzing the uridine isomerization in both 23S rRNA (position 2604) and tRNA\textsuperscript{Tyr} (position 35).

As RluF is now found to be a dual specificity synthase, it is interesting to examine the mechanistic differences for pseudouridine synthases. Although RluB modifies U2605 of 23S rRNA through stabilization of the A2602 bulge (47), RluF modifies U2604 through a frameshift in RNA base pairing. This protein binds to an RNA stem-loop to rearrange the base pair in such a way that it moves the A2602 bulge into the stem, thereby translating the target U2604 by flipping it into the active site. Thus, if a mutation facilitates A2602 refolding into the stem, enzymatic activity increases substantially (44).

The pseudouridines at positions 2604 and 2605 of *E. coli* 23S rRNA are both located in domain V, which is part of the peptidyltransferase center involved in peptide bond formation. It is possible that these modifications could aid peptidyltransferase activity through enhanced base pairing available with pseudouridine rather than uridine, although the precise function of these pseudouridines is still undetermined.

TruA, a functional dimer, modifies uridines at position 38, 39, and/or 40 of tRNAs with highly divergent sequences and structures (43) where the protein utilizes the intrinsic flexibility of the anticodon stem-loop for substrate promiscuity. In contrast, TruB modifies a single uridine at position 55 of the T-loop. This enzyme directly docks to the preformed T-loop through shape complementarity and extensive interactions with nucleotides near the target site. Furthermore, the substrate is accessed by flipping out the uridine and disrupting the tertiary structure involving interaction between the T- and D-loops of the tRNA (48). Although the mechanism of action of RluF on the tDNA\textsuperscript{Tyr} transcript is not clear in the current study, it can be speculated that it may be recognizing the anticodon loop similarly to TruB based on its site-specific isomerization behavior.

The cytoplasmic tRNA\textsuperscript{Tyr} of several eukaryotes is known to contain pseudouridine at position 35 (49), which is introduced in an intron-dependent manner (for a review, see Grosjean et al. (50)). Studies with *Arabidopsis* pre-tRNA\textsuperscript{Tyr} (49) indicated that the intron sequence helps in orienting and exposing the 5′-exonic consensus sequence U[^\(23\text{N}^\text{A}\text{U}^\text{A}\text{C}\text{G}^\text{A}^\text{C}][\Psi]-\text{CCCCG}^\text{C}^\text{C}^\text{C}^\text{C}^\text{C}^\text{G}^\text{C}^\text{T}^\text{G}^\text{U}^\text{A}^\text{G}^\text{G}^\text{U}]$, where N stands for any one of four nucleosides and Pu stands for purine) to the tRNA:Ψ35 synthase. The *E. coli* tRNA\textsuperscript{Tyr} anticodon sequence (U[^\(23\text{N}^\text{A}\text{U}^\text{A}\text{C}\text{G}^\text{A}^\text{C}][\Psi]-\text{CCCCG}^\text{C}^\text{C}^\text{C}^\text{C}^\text{C}^\text{G}^\text{C}^\text{T}^\text{G}^\text{U}^\text{A}^\text{G}^\text{G}^\text{U}]$) matches this exonic consensus sequence. In fact, no other tRNA, except tRNA\textsuperscript{Tyr}, has this consensus sequence in *E. coli*, indicating that pseudouridine at position 35 could be unique to tRNA\textsuperscript{Tyr}. However, RluF catalyzes pseudouridine formation in an intron-independent manner as indicated by the \textit{in vitro} assay involving the tDNA\textsuperscript{Tyr} transcript. It is possible that RluF binds to 23S rRNA and tRNA\textsuperscript{Tyr} by recognizing the stem-loop to flip the base into the active site for modification. Further studies will reveal whether the anticodon loop is sufficient for pseudouridine introduction or RluF requires the whole anticodon stem-loop or other parts of the tRNA\textsuperscript{Tyr} sequence for recognition and docking.

It has been shown that pseudouridine can stabilize RNA duplexes by forming Ψ-A, Ψ-G, Ψ-U, and Ψ-C pairs by hydrogen bonding through the N3 and N1 imino protons depending on the position of Ψ within the RNA duplex and the type of Watson-Crick pair (51). Such interactions seem to play a crucial role in aminoacyl fidelity and efficiency in yeast (52) where Ψ interactions with adenosine strengthen the anticodon-codon/stop codon interactions (51, 53, 54). These findings might explain why incorporation of pseudouridine into mRNA makes it highly stable and translatable, offering a number of opportunities for medical applications (9).

Based on the enhanced base-pairing interactions mediated by pseudouridine (as described above), it is tempting to specu-
Pseudouridine in E. coli tRNA<sup>Tyr</sup> Anticodon

late that such interactions might be the reason for the observed luciferase expression differences in our assays. Here, the expression of luciferase is affected only when its coding sequence is preceded by multiple tyrosine codons; otherwise expression is unaffected in the Δrluf<sup>+</sup> strain. Complementation of Δrluf<sup>+</sup> with the wild type RluF gene restores the modification, thereby facilitating the luciferase gene expression even if the protein coding sequence is preceded by multiple tyrosine codons. A cursory examination of the E. coli K12 MG1655 proteome indicates that there are 47 genes with three consecutive tyrosine residues. As the Δrluf<sup>+</sup> strain does not exhibit any growth phenotype, the impact of three consecutive tyrosine residues on gene expression in this strain might be minimal because the codon-anticodon interactions remain sufficient over such a small region of translation. It would be interesting to study the variations of reporter gene expression if the coding sequence is preceded by four to nine tyrosine codons.

A phylogenetic classification of pseudouridine synthases based on their conserved domains reveals four families: TruA-like (cd00497), TruB-like (cd00506), Rsu_Rlu-like (cd02550), and TruD (cd02552) (supplemental Fig. S7A). RluF falls under the subfamily of Rsu_Rlu-like proteins. However, the subfamily hierarchy of the RluF branch (cd02554) contains 11 more homologous proteins that are of purely bacterial origin (supplemental Fig. S7B) where the homology extends beyond the shared common fold and active site cleft core structure (supplemental Fig. S8). One other interesting feature is that although the 23S rRNA from these 11 bacteria contains uridine at position 2604 pseudouridine is reported only in E. coli, indicating the unknown role of these pseudouridine synthases. In such a situation, it is tempting to speculate that these homologous proteins may be responsible for catalyzing pseudouridine modification at position 35 of tRNA<sup>Tyr</sup>. Further experiments are needed to clarify the exact role of these proteins in their respective bacteria.

Moreover, these efforts emphasize the advantages of detection methods of increasing sensitivity and accuracy. As noted by the discovery of V35 in tRNA<sup>Tyr</sup> from the well studied model organism E. coli, modern approaches, which provide multiple means of detection/validation, may reveal additional information relating to RNA modifications that have until now remained undiscovered.

Experimental Procedures

RNase T1 and U2 Digestion—RNase T1 (Roche Applied Science) was purified on C<sub>18</sub> Sep-Pak cartridges (Waters) as described (55). Two micrograms of E. coli tRNA<sup>Tyr</sup> or II (R0383 or R0258, Sigma-Aldrich) or 4 μg of E. coli total tRNA were digested with purified RNase T1 (25 units/μg of RNA) and BAP (0.01 unit/μg of RNA) (Worthington Biochemical Corp.) in 120 mM ammonium acetate, pH 6.5, at 37 °C for 2 h (56, 57). The digest was dried in a SpeedVac (Thermo Fisher Scientific, Inc.) and further purified for CMCT derivatization using C<sub>18</sub> ZipTips (Millipore Corp.) according to the manufacturer’s instructions used or as is for acrylonitrile treatment. RNase U2 was purified as described (42). Two micrograms of E. coli tRNA<sup>Tyr</sup> were digested with 2 μg of purified ribonuclease U2 in 120 mM ammonium acetate, pH 5.1, at 55 °C for 30 min. The dried RNA digest was purified by C<sub>18</sub> ZipTips for subsequent CMCT derivatization.

Chemical Derivatization—For CMCT derivatization, the RNA digest was derivatized with 10 μl of CMCT (10 mg/ml) solution made with 50 mM Tris-HCl, pH 8.3, 4 mM EDTA, and 7 M urea (28). This concentration was found to be optimal based on the independent LC-MS/MS analysis of Ψ-containing (ΨGG, ΨFCAG, and UAACΨAΨGACG) and Ψ-lacking (UGG, UACAG, and UAACUUGACG) oligonucleotides (Dharmacon, GE Healthcare) following derivatization and alkali treatment (10 μl of 50 mM ammonium bicarbonate, pH 10.4, at 75 °C for 90 min). Incubation of Ψ-lacking oligomers at 37 °C for 16 h at this optimal concentration, ZipTip purification, and alkali treatment resulted in no derivatization at uridine or guanosine. However, alkali-stable Ψ-specific derivatization was observed for >99% of Ψ-containing oligomers (addition of 251.4 Da per pseudouridine; data not shown). The acrylonitrile-based derivatization was similar to the previous procedure (31) except that 1 μl of acrylonitrile (Sigma)/2 μg of RNA digest was used. The procedure involved incubation of the RNA digest (2 μg) in 30 μl of 41% ethanol and 1.1 m triethylamine ammonium acetate (Sigma) and 1 μl of acrylonitrile at 70 °C for 2 h.

LC-MS/MS of RNase Digests—All undervatized and derivatized RNase digestion products were reconstituted in mobile phase A (400 mM hexafluoroisopropanol (Sigma), 16.3 mM triethylamine (Sigma), pH 7.0) before analysis. HPLC was conducted using a 1 × 150-mm XBridge C<sub>18</sub> column (Waters) using mobile phase A and mobile phase B (200 mM hexafluoroisopropanol, 8.15 mM triethylamine in 50% methanol (Burdick and Jackson), pH 7.0) at a flow rate of 30 μl/min. The chromatographic gradient and mass spectrometric conditions are identical to the conditions described before (56).

Mongo Oligo Mass Calculator was used to predict the m/z values of RNase T1 and U2 digestion products and associated MS/MS fragment ions that include the addition of carbodiimide (+251.4 Da) or cyanoethylation (+53 Da) when appropriate (supplemental Table S1). The Quan browser feature of Xcalibur was utilized to integrate and quantify the extracted ion chromatogram (XIC) peak areas corresponding to m/z values of interest.

SRM Detection of Pseudouridine—Undervatized RNA was analyzed for pseudouridine essentially as described previously (33). Alignment of the XIC of the m/z corresponding to any particular RNase digestion product (from the full scan) with the Ψ-specific SRM peak (m/z 207–164 or m/z 225–165) denotes the presence of pseudouridine in that RNase digestion product (32).

Nucleoside Analysis by LC-MS—The E. coli tRNA<sup>Tyr</sup> RNase U2 digestion product, CUQUA>G<p> (m/z 857.7), was collected during reversed phase ion pair HPLC after a postcolumn split, which allowed for LC-MS monitoring to identify the appropriate digestion product fraction. The fractionated oligomer was digested with snake venom phosphodiesterase and Antarctic phosphatase before subjecting it to subsequent LC-MS (56, 58) with detection in positive polarity.

E. coli Strains, Cell Culture, and RNA Isolation—E. coli strains harboring an insertion mutation in a specific pseudou-
Pseudouridine in E. coli tRNA<sub>Tyr</sub> Anticodon

The pseudouridine synthase gene as well as the gene associated with pseudouridine degradation (YeiN; supplemental Table S2) were procured from the Yale Coli Genetic Stock Center. The tRNA was isolated from mutant cells grown to mid-log phase (A<sub>600</sub> ~0.5–0.6) using TRI Reagent as described (33, 59). The 23S rRNA was gel-purified from total RNA by electrophoresis on a 1% low melting agarose gel followed by phenol-chloroform extraction and ethanol precipitation.

Genotyping of Insertion Mutants—E. coli strains exhibiting a significant decrease in pseudouridine (based on cyanoethylation) levels at position 35 of tRNA<sub>Tyr</sub> were genotyped and validated for kanamycin insertion. This was done by amplifying and sequencing the DNA at the junction of insertion using dated for kanamycin insertion. This was done by amplifying and testing for luciferase activity in (EMD Biosciences) vector. Four different constructs were made possessing suitable restriction sites (supplemental Table S3). 

Gene Cloning and Protein Purification—The coding regions of pseudouridine synthase genes TruA (gi|556503834) and RluF (gi|556503834) were amplified using primers (supplemental Table S3) that have compatible restriction sites (BamHI and HindIII) to enable subcloning into the IPTG-inducible gene expression cassette in the pET22b vector (EMD-Millipore) as described (61). Recombinant proteins with C-terminal His<sub>6</sub> tag fusion were expressed and purified as described elsewhere (62).

Pseudouridine Synthase Activity Assay—The tDNA<sub>Tyr</sub> template required for in vitro transcription was initially amplified and cloned from E. coli K12. The transcription template produced by PCR using the primers Tyr-5′-T7 and Tyr-3′; (supplemental Table S3) was in vitro transcribed, and the RNA was purified by an Ampliscribe kit (Epicon Biotechnologies) and a MEGAClear kit (Ambion), respectively, as described (62).

The pseudouridine synthase activity assay was performed on the in vitro transcript with recombinant pseudouridinyl synthase as described (44, 45). Briefly, the pseudouridylation assay mixture consisting of purified protein (50 nM for RluF and 2 nM for TruA) and in vitro transcript (0.5 μM for RluF and 1.5 μM for TruA) in the assay buffer (20 mM Tris-HCl, pH 8.0, 0.1 M NH₄Cl, 2 mM DTT) was incubated at room temperature for 2 h. Following digestion with RNase T1 (150 units) and BAP (0.02 unit) at 37°C for 2 h, the resulting oligonucleotides were extracted with phenol-chloroform, dried in a SpeedVac, and dissolved in sterile water for derivatization with acrylonitrile.

Genetic Manipulations for Luciferase Reporter Gene Assays—The coding regions of FF Luc and Ren Luc were amplified from pGL3 and pRL vectors (Promega), respectively, with primers possessing suitable restriction sites (supplemental Table S3) so that the amplified DNA could be cloned into the pETDuet-1 (EMD Biosciences) vector. Four different constructs were made and tested for luciferase activity in E. coli with the RluF enzyme knock-out background. Construct 1 was pETDuet with FF Luc and/or Ren Luc. For Construct 2, pETDuet Ren Luc and (AAY)<sub>10</sub> FF Luc, Ren Luc was cloned into the MCS1 (SacI and HindIII), whereas the annealed and digested (S′ Ndel and BamH1 3′) double-stranded oligomeric (AAY)<sub>10</sub> along with FF Luc (BglII and XhoI) were cloned into the Ndel-XhoI sites of MCS2. For Construct 3, pETDuet Ren Luc and (TAY)<sub>10</sub> FF Luc, cloning was similar to the previous construct except that the oligomer (TAY)<sub>10</sub> was used instead of (AAY)<sub>10</sub>. For Construct 4, pETDuet RluF and (TAY)<sub>10</sub> FF Luc, RluF was cloned into the BamH1 and HindIII sites of MCS1, and the MCS2 contained (TAY)<sub>10</sub> FF Luc.

Luciferase assays were carried out using the Dual-Glo® luciferase assay system (Promega) and a Turner TD-20/20 luminometer. The assay involves transforming the ΔrulF strain of E. coli with each of the four constructs and culturing 10 ml of cells to mid-log phase (0.5–0.6 OD) to induce the expression of cloned genes with IPTG for 2 h at 37°C. Subsequently, the cells were harvested by centrifugation (6,000 rpm for 3 min) and resuspended in lysis buffer (150 μl of 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mg/ml lysozyme). The cells were lysed by exposure to 4°C for 10 min, −80°C for 10 min, and a 25°C water bath for 5 min (63). The resulting lysate was centrifuged (13,000 × g for 12 min) to remove cell debris. Ten microliters of lysate were added to 50 μl of Luciferase Assay Reagent II and mixed thoroughly by repeated pipetting, and the resulting luminescence for FF Luc was recorded. Subsequently, 50 μl of Stop & Glo reagent was mixed to take the second luminescence measurement for Renilla luciferase.

Author Contributions—B. A. designed, executed, and analyzed the experiments. B. A. and P. A. L. formulated the study and wrote the paper. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments—We acknowledge the guidance received from Sitakanta Pattanaik (Kentucky Tobacco Research AND Development Center) in conducting luciferase assays and Christian Hong (University of Cincinnati Medical Center) for allowing the use of the luminometer.

References
Pseudouridine in E. coli tRNA\textsuperscript{Tyr} Anticodon


274, 2232–2233


