Synthesis and investigation of the 5-formylcytidine modified, anticodon stem and loop of the human mitochondrial tRNAMet

Hrvoje Lusic¹, Estella M. Gustilo², Franck A. P. Vendeix², Rob Kaiser³, Michael O. Delaney³, William D. Graham², Virginia A. Moye², William A. Cantara², Paul F. Agris²,* and Alexander Deiters¹

¹Department of Chemistry, ²Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695 and ³Dharmacon, 2650 Crescent Drive #100, Lafayette, CO 80026, USA

Received July 29, 2008; Revised September 28, 2008; Accepted September 29, 2008

ABSTRACT

Human mitochondrial methionine transfer RNA (hmtRNAMetCAU) has a unique post-transcriptional modification, 5-formylcytidine, at the wobble position-34 (f5C34). The role of this modification in hmtRNACAU for the decoding of AUA, as well as AUG, in both the peptidyl- and aminoacyl-sites of the ribosome in either chain initiation or chain elongation is still unknown. We report the first synthesis and analyses of the tRNA's anticodon stem and loop domain containing the 5-formylcytidine modification. The modification contributes to the tRNA's anticodon domain structure, thermodynamic properties and its ability to bind codons AUA and AUG in translational initiation and elongation.

INTRODUCTION

Mitochondria generate over 90% of the energy used by mammalian cells through oxidative phosphorylation. Thirteen proteins, components of the electron transfer chain and the ATP synthase, are the products of mitochondrial DNA. The synthesis of these proteins is carried out by a specific protein synthesizing machinery within this organelle. During the almost-three decades since the sequencing of the human mitochondrial genome (1), the mitochondrial genetic code has been found to differ significantly from the universal code. The human mitochondrial gene for the one methionine specific tRNA (hmtRNAMetCAU, where CAU is the anticodon) plays a unique role since it must provide the tRNA used for both the initiation of protein synthesis and the elongation of the protein chain by responding to the codon AUA, normally an isoleucine codon in the cytoplasm, as well as the universal methionine code, AUG. This is highly unusual since all cytoplasmic protein biosynthetic systems employ two different tRNA Met species, one for initiation and one for elongation, and both respond to the single methionine codon, AUG. Maternally inherited mutations in the gene of this tRNA, including an A37 to G37 mutation adjacent to the anticodon nucleosides that read the two codons (Figure 1), are responsible for some devastating diseases (2–5). Moreover, the hmtRNACAU50 has a unique modification, 5-formylcytidine (Figure 1), at the wobble position-34 (f5C34) seen only in one other tRNA, a bovine liver, cytoplasmic tRNA Leu with a f 5C34 further modified with a 2'-O-methyl (6). Nothing, however, is known about the decoding characteristics of tRNA CAU 5.

Since its discovery in bovine and nematode mitochondrial tRNA Met in 1994 (7,8), f5C34 also has been found in the mitochondrial tRNA Met of squids, frogs, chickens, rats and fruit flies (9–11). The contribution of the f5C34 modification to the structure of the hmtRNAMetCAU, its role in the decoding of AUG and AUA and its possible participation in either chain initiation or chain elongation by this unique tRNA Met is still unknown.

We speculate that mitochondria have a unique mechanism to partition this single hmtRNACAU species between initiation and elongation. While a tRNA CAU unmodified at the wobble position-34 can read AUG, we hypothesize that the 5-formyl modification allows one tRNA Met to expand codon reading to include recognition of the AUA codon in mitochondrial mRNAs. Toward proving the hypothesis that this wobble modification affords the single tRNA Met the ability to decode AUG and AUA, we are reporting the first synthesis of a 5-formylcytidine-modified RNA and the initial structural and biological

*To whom correspondence should be addressed. Tel: +1 919 515 6188; Fax: +1 919 515 2047; Email: Paul_Agris@ncsu.edu
Correspondence may also be addressed to Alexander Deiters. Tel: +1 919 513 2958; Fax: +1 919 515 5079; Email: Alex_Deiters@ncsu.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

© 2008 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
investigations. Previously, bovine mitochondrial tRNA\textsubscript{CAU} had been shown to translate AUG and AUA in an \textit{Escherichia coli} translational system \textit{in vitro}, where AUG coded for methionine and AUA for isoleucine (12). However, site-specific binding was not investigated. Here we compare the codon binding affinities of the fC\textsubscript{54}-modified anticodon stem and loop of human mitochondrial tRNA\textsubscript{CAU} (hm\textsubscript{ASLMet-f5C34}) with that of the unmodified ASL. Our results for both AUA and AUG codons at both the A and P-site of \textit{E. coli} ribosomes increase our understanding of the modification's contributions to decoding and are consistent with previous results from the translation of poly(AUA) by bovine mitochondria (12).

**MATERIALS AND METHODS**

Experimental procedures and analytical data for the synthesis of the fC\textsubscript{54} phosphoramidite (9)

All reagents used in the following experiments are of the highest purity and dryness possible. Before use, glassware was thoroughly cleaned and dried (oven at 110°C for 30 min). NMR analysis of intermediates was conducted in the appropriate deuterated solvent (referenced accordingly for CDCl\textsubscript{3}: \textit{H} 7.24 p.p.m., \textit{C} 77.23 p.p.m.; and CD\textsubscript{3}CN: \textit{H} 1.94 p.p.m.) using a Bruker Avance UltraShield 300 MHZ spectrometer. Phosphorus, \textit{P}, NMR experiments were referenced according to an external H\textsubscript{3}PO\textsubscript{4} standard (0.00 p.p.m.). Mass-spectra analysis of the samples was performed on a Micromass LCT ESI-TOF or an Agilent LC-TOF. Analytes were dissolved in acetonitrile and flown against a Leucin Enkcephalin lock mass standard. Chemical and physical properties of the intermediates were those of the expected compounds (Supplementary Data).

5-(Hydroxymethyl)-2',3'-O-isopropylideneacetate (3).

To a solution of 2 (Figure 2, 1 g, 3.5 mmol) in 15 ml of 0.5 M KOH was added paraformaldehyde (1.05 g, 35 mmol). The reaction was stirred at 55°C for 4 h, after which it was cooled to room temperature and neutralized with 6 M HCl. The solution was filtered and evaporated. The oily residue was dissolved in MeOH:DCM (40:60), filtered and evaporated again. Purification by silica gel chromatography using MeOH:DCM (gradient 3:97, 8:92, 12:88) with 2% TEA, afforded 416 mg of 5-(hydroxymethyl)-2',3'-O-isopropylideneacetate (3) as a white foam (38% yield, 60% yield based on 357 mg of recovered starting material 2',3'-O-isopropylideneacetate (2).

5-Formyl-2',3'-O-isopropylideneacetate (4).

To a solution of 4 (Figure 2, 100 mg, 0.32 mmol) in 3 ml dioxane was added 500 mg ruthenium dioxide hydrate (five weight equivalents). The reaction mixture was refluxed for 12 h and filtered. Purification by silica gel chromatography using MeOH:DCM (5:95) with 2% TEA, afforded 81 mg of the 5-formyl-2',3'-O-isopropylideneacetate (4) as a white solid (82% yield).

5-Formylcytidine (5). The acetonide, 5-formyl-2',3'-O-isopropylideneacetate (4, Figure 2, 500 mg, 1.61 mmol) was suspended in 1 M HCl (15 ml) at room temperature. The reaction progress was monitored by TLC. Upon disappearance of the starting material, the solution was neutralized with TEA. Water was subsequently evaporated. Recrystallization from MeOH afforded 414 mg of fC\textsubscript{54} (5) as a white solid (95% yield).

\textit{N}^4-[(\textit{Diisobutylamino})methylidene]-3',5'-O-(1,1,3,3-tetraisopropyl-1,3-disiloxanediyl)-5-formylcytidine (6). 5-Formylcytidine (5, Figure 2, 0.47 g, 1.72 mmol) was dissolved in a mixture of 20 ml of pyridine and 2 ml of DMF. The solution was cooled to 0°C and TIPDSCl\textsubscript{2} (0.59 g, 1.89 mmol) in 2 ml of pyridine was added dropwise over a period of 1 h. The reaction was allowed to gradually warm to room temperature overnight. The following morning the reaction was quenched with 5 ml of MeOH and evaporated to dryness. The resulting paste was coevaporated twice with 20 ml of toluene and the crude material was purified by flash chromatography on 30 ml of silica gel using a gradient of MeOH in DCM (3-4%). Product fractions were pooled and evaporated to afford the TIPDS protected intermediate (0.88 g, 100%) as light yellow oil that is contaminated with residual pyridinium salts. The above compound was used as is without further purification to remove the residual pyridinium salts. TIPDS protected fC\textsubscript{54} (0.88 g, 1.72 mmol) was dissolved in an E300 MHZ spectrometer. Phosphorus, \textit{31}P, NMR analysis of intermediates was conducted in the appropriate deuterated solvent (referenced accordingly for CDCl\textsubscript{3}: \textit{H} 7.24 p.p.m., \textit{C} 77.23 p.p.m.; and CD\textsubscript{3}CN: \textit{H} 1.94 p.p.m.) using a Bruker Avance UltraShield 300 MHZ spectrometer. Phosphorus, \textit{P}, NMR experiments were referenced according to an external H\textsubscript{3}PO\textsubscript{4} standard (0.00 p.p.m.). Mass-spectra analysis of the samples was performed on a Micromass LCT ESI-TOF or an Agilent LC-TOF. Analytes were dissolved in acetonitrile and flown against a Leucin Enkcephalin lock mass standard. Chemical and physical properties of the intermediates were those of the expected compounds (Supplementary Data).
in 20 ml DMF and N,N-diisobutylformamidine dimethyl acetal (14) (0.70 g, 3.44 mmol) was added. The reaction was stirred for 16 h and evaporated under high vacuum. The resulting loose oil was coevaporated twice with 20 ml of toluene and the crude material was purified by flash chromatography on 30 ml of silica gel using a gradient of MeOH in DCM (1–2%). Product fractions were pooled and evaporated to afford 0.80 g of 6 as light yellow oil in 71% overall yield from \( f \) (Figure 2, 0.5°C).

- **Figure 2.** Synthesis of the 5-formylcytidine phosphoramidite. The starting compound cytidine is numbered compound 1, intermediates are numbered 2–8 and the protected 5C phosphoramidite is compound 9. The synthetic transformations are: (a) acetone, dimethoxypropane, cat HClO\(_4\) (83%); (b) paraformaldehyde, 0.5 M KOH, 55°C (38%, 60% brsm); (c) RuO\(_2\) xH\(_2\)O, dioxane, reflux (82%); (d) 1 M HCl (95%); (e) TIPDSCl, Pyr, DMF; (f) DBF-CH(OMe)\(_2\), DMF (71% over 2 steps); (g) ACE-orthoester, PPTS, TBDMS-pentanedione, DCM; (h) HF-TEMED, CH\(_3\)CN (53% over 2 steps); (i) BZH-Cl, DIA, DCM (84%); (j) P(OMe)(DIA)\(_2\), DIA, S-Et Tetrazole, DCM (93%).

hexanes with 0.1% TEMED. This material was concentrated to near dryness and taken directly onto the desilylation reaction.

A freshly made solution of TEMED (0.71 g, 6.15 mmol) in 10 ml of acetonitrile at 0°C was added 48% HF (0.15 ml, 4.30 mmol). This solution was allowed to stir for 5 min and added to the foregoing material from above at room temperature. The reaction was stirred for 2 h and concentrated to dryness. The crude material was purified by flash chromatography on 50 ml silica gel using a gradient of 20% hexanes in ethyl acetate with 0.1% TEMED to 1% methanol in ethyl acetate with 0.1% TEMED. Product fractions were pooled and evaporated to leave the 2’-O-protected compound 7 as a light yellow oil (0.41 g) in 53% yield from the nucleobase protected compound 6.

- **Figure 2.** Synthesis of the 5-formylcytidine phosphoramidite. The starting compound cytidine is numbered compound 1, intermediates are numbered 2–8 and the protected 5C phosphoramidite is compound 9. The synthetic transformations are: (a) acetone, dimethoxypropane, cat HClO\(_4\) (83%); (b) paraformaldehyde, 0.5 M KOH, 55°C (38%, 60% brsm); (c) RuO\(_2\) xH\(_2\)O, dioxane, reflux (82%); (d) 1 M HCl (95%); (e) TIPDSCl, Pyr, DMF; (f) DBF-CH(OMe)\(_2\), DMF (71% over 2 steps); (g) ACE-orthoester, PPTS, TBDMS-pentanedione, DCM; (h) HF-TEMED, CH\(_3\)CN (53% over 2 steps); (i) BZH-Cl, DIA, DCM (84%); (j) P(OMe)(DIA)\(_2\), DIA, S-Et Tetrazole, DCM (93%).

2’-O-[Bis(2-acetoxyethoxy)methyl]-N\(^4\)-[(diisobutylamino)methylidene]-5-formylcytidine (8). Diisopropylamine (0.07 g, 0.65 mmol) was added to a solution of the 2’-O- and N\(^4\)-protected nucleoside, 7 (Figure 2, 0.41 g, 0.65 mmol) in 7 ml of DCM and the solution was cooled
to 0°C. In a separate flask BZHCl (0.34 g, 0.81 mmol) was diluted in 5 ml of DCM. Diisopropylamine (0.10 g, 0.98 mmol) was added to the silylating solution and the solution was allowed to stir for 2 min before being added dropwise to the nucleoside solution. The addition was completed within 30 min and the reaction was allowed to stir for 3 h and the reaction was quenched with 1 ml of MeOH and evaporated to dryness. The crude material was purified by flash chromatography on 30 ml silica gel using a gradient of 10% acetone in hexanes containing 0.1% (v/v) TEA to 20% acetone in hexanes containing 0.1% (v/v) TEA. Product fractions were pooled and evaporated to afford the 5′-O-protected compound 8 as a colorless oil. The yield was 0.56 g (84%).

5′-O-[(Benzyldihydroxy-bis(trimethylsilyloxy)silyl)-2′-O-(bis(2-acetoxyethoxy) methyl)-N4-[(diisobutyl-amino)methylidene]-5-formylcytidine-3′-(methyl-N,N-diisopropyl) phosphoramidite (9). Bis(diisopropyl-amino)methoxy phosphate (0.21 g, 0.82 mmol) was dissolved in 3 ml of DCM and a 0.5-M solution of 5-ethylthio-1-H-tetrazole in anhydrous acetonitrile (0.08 ml, 0.55 mmol) was added. Diisopropylamine (0.06 g, 0.55 mmol) was then added and the phosphine solution was allowed to stir for 5 min at ambient temperature. In a separate flask, the 2′′′-O-, 5′-O- and N4-protected 5′C, compound 8 (Figure 2, 0.56 g, 0.55 mmol) and diisopropylamine (0.06 g, 0.55 mmol) were dissolved in 5 ml of DCM. The activated phosphine solution was added into the nucleoside solution and the reaction was stirred at room temperature. After 16 h the reaction was quenched with 2 ml of absolute ethanol and concentrated to dryness. The resulting white paste was purified by flash chromatography on 30 ml of silica gel using a mixture of DCM in hexanes [5: 95 (v/v)] containing 2% (v/v) TEA followed by acetone in hexanes [1: 9 (v/v)] to 2: 8 (v/v) containing 0.5% (v/v) TEA. Product fractions were pooled and evaporated to afford the protected 5′C phosphoramidite 9 (Figure 2), as a colorless oil.

**Polyribonucleotide synthesis of 5′-ΨCGGGCC-f5C-AUA CCCCGA-3′**

The above sequence was synthesized on a 1-μmol scale using an ABI 394 DNA synthesizer using previously published procedures (15,16). The 5′C phosphoramidite (9, 0.067 M in anhydrous acetonitrile) was coupled to the growing polyribonucleotide chain for 3.5 min using 5-ethylthio-1H-tetrazole (0.5 M in anhydrous acetonitrile) as the activator. Once the synthesis of the polyribonucleotide chain was completed, the phosphate protecting groups were removed from the immobilized polyribonucleotide by treatment with disodium 2-carbamoyl-2-cyanooxyethylene-1,1-dithiolate trihydrate in DMF for 10 minutes. The support was washed excessively with water for 5 min and then flushed with Argon gas for 5 min to dry the support. The support was then transferred to a 2-ml Eppendorf tube and the polyribonucleotide was cleaved from the support and the exocyclic amine protecting groups were removed with 1:3 (v/v) tert-butylamine:water for 6 h at 60°C. The sample was cooled to room temperature, filtered and lyophilized to obtain the crude polyribonucleotide. The hmASLMetCAU-Ψ27-f5C34 was deprotected with acetic/TEMED according to standard Dharmaco protocols, purified by ion exchange HPLC (17), and dialyzed extensively against H2O. The hmASLMetCAU-Ψ27 was synthesized and deprotected under standard conditions (18).

**Confirmation of nucleoside composition by nucleoside HPLC and NMR of the hmASLMetCAU constructs**

Incorporation of f5C modification within the hmASLMet was confirmed by NMR, including the two-dimensional NOESY (Figures 3 and 4). The nucleoside composition of the hmASLMet products was confirmed by enzymatic hydrolysis of the RNA to its constituent nucleosides (17) and then subjected to HPLC monitored by diode array UV spectrometry, the peaks identified, integrated and quantified (19) (Figure 4).

**Analysis of thermodynamic stability, circular dichroism and molecular dynamics simulations**

The ASL samples were dissolved to obtain a concentration of ~4μM in 20 mM Na–K phosphate buffer (pH 6.8). UV-monitored, thermal denaturations and renaturations were replicated five times and monitored by measuring UV absorbance (260 nm) using a Cary 3 spectrophotometer as published (20,21). The data points were averaged over 20 s and collected four times a minute with a temperature change of 0.5°C/min from 4 to 90°C. The data were analyzed (22), and the thermodynamic parameters were determined (Origin software, Microcal, Inc.) (Figure 5A). CD spectral ellipticity data were collected using a Jasco 600 spectropolarimeter and an interfaced computer (Jasco, Inc.). hmASLMetCAU-Ψ or the hmASLCAU-Ψ27 f5C34 (0.2 A260/ml, 20 mM Na–K phosphate buffer, pH 6.8) was placed in a temperature-regulated, 1-cm path-length quartz cell. Each sample was scanned 10 times at 25°C. The final data are an average of the 10 scans (Figure 5B). The molecular dynamics simulation (MDS) were performed by following standard published protocol (23) with the exception of using a truncated octahedral TIP3P water box (24).

**Ribosomal binding assay**

The 27-mer mRNA oligos used in codon binding assays were designed from that of T4 gp32 mRNA (25) and purchased (Dharmacon RNA Technologies). They were chemically deprotected and HPLC-purified in our lab. Each mRNA sequence was entered into the program RNA Structure 4.2 (26) and was found to have a low probability of folding into any stable conformation. The mRNA sequences are as follows (mitochondrial methionine codons AUA and AUG are in bold):

(i) 5′-GGCAAGGAGGUAUUUUAUGUAGCACGU-3′;
(ii) 5′-GGCAAGGAGGUAUUUUAUGUAGCAGCU-3′;
(iii) 5′-GGCAAGGAGGUAUUUUAUGUAGCAGCU-3′;
Figure 3. NMR spectra of the hmASL Met CAU. (A) One-dimensional 1H-NMR spectrum (in H2O) of hmASL Met CAU-ψ C9 f 5C (top) is compared to that of the unmodified hmASL Met CAU-ψ C9 (bottom). The formyl proton’s chemical shift in the RNA is almost identical to that of the mononucleoside fC. *Denotes impurities. (B) Superimposed 1H-13C HMBC spectra of cytidine (blue) and 5-formylcytidine (red).

Figure 4. HPLC nucleoside composition and NMR analyses of hmASL Met CAU and f5C. (A) The upper section of the figure depicts the HPLC separation of standard ribonucleosides where C, G and A (solid line) had been injected together, and U, G and f5C (dotted line) had been injected in a separate control experiment. The lower section depicts the chromatography of nucleosides composing the hmASL Met CAU to include f5C, as well as f5C. (B) The NOESY connectivities between imino protons at 500 MHz in a 2D 1H NOESY NMR spectrum (mix = 250 ms) of hmASL Met CAU (90% H2O + 10% 2H2O; 20 mM PO43-, 50 mM NaCl, pH 6.2; 28°C) with water suppression using the WATERGATE sequence. (C) 1D 1H NMR spectra of (a) cytidine (1) and (b) 5-formylcytidine (5) (500 MHz; 100% 2H2O; 20 mM PO43-, 50 mM Na+, 50 mM K+; pH = 6.2; 25°C) using the presaturation NMR sequence to suppress the water peak. The spectral regions and peaks corresponding to the various proton types in the molecules are labeled.
 Circular dichroism spectra. Spectra of the hmASL Met/C9 from three denaturations and two renaturations for the hmASL stability of the ASLs. UV-monitored, thermal data were averaged over the wavelength range of 200 to 300 nm.

**Figure 5.** Thermal denaturations and circular dichroism spectra of the hmASLMet/C9 (thick black line) at 4°C for at least 20 min and each well of the filtration apparatus was washed with 100 μl of cold binding buffer. Cold binding buffer (100 μl) was added to each sample, and the entire 120-μl volume was quickly filtered. Each well was then washed twice with 100 μl of cold binding buffer. The nitrocellulose was dried out on ‘kim’ wipes, and the radioactivity was measured using a phosphorimager (Molecular Dynamics, GE Healthcare). Data were measured for radioactivity intensity using ImageQuant (Amersham).

Nonspecific binding was determined by the binding of ASLs to ribosomes without mRNA and subtracted from the experimental data. The final data are a result of at least three separate experiments, each done with samples in triplicate, i.e. minimally nine determinations for each binding (Figure 6).

**RESULTS AND DISCUSSION**

A 5-formylcytidine (fC) has previously been synthesized from 5-(hydroxymethyl)cytosine (13) and from 5-methyluridine (30), but not incorporated into an RNA sequence. First, we developed a short (four steps) and facile synthesis from 5-(hydroxymethyl)cytosine (13) and from 5-methylcytidine (30), starting by protecting cytidine as the 5′-end (13) with an 83% yield. The acetonide protecting group was subsequently removed to deliver fC (KCl- HCl buffer for pH values between 1.1 and 2.0 and citrate-phosphate buffer for pH values between 2.2 and 5.0). The line fitting and data analysis was conducted with Prism v3.00 (Graphpad Software, Inc.) (Figure 7).

**Analysis of the fC pKa**

UV spectra were compiled (220–320 nm) using a Varian Cary3 Spectrophotometer at different pH values for cytidine and 5-formylcytidine. The spectra were normalized to 0.2 OD at 260 nm. Entire spectra were collected to ensure that they all intersected at 260 nm at an OD of ~0.2. However, the absorbance maximum at 280 nm was plotted against the pH, a previously published method of assessing the pKa of nucleosides (29). A pH range of 2.2–7.0 (citrate-phosphate buffer) was used for cytidine, and a pH range of 1.1–5.0 was used for fC (KCl–HCl buffer for pH values between 1.1 and 2.0 and citrate–phosphate buffer for pH values between 2.2 and 5.0). The line fitting and data analysis was conducted with Prism v3.00 (Graphpad Software, Inc.) (Figure 7).

**RESULTS AND DISCUSSION**

A 5-formylcytidine (fC) has previously been synthesized from 5-(hydroxymethyl)cytosine (13) and from 5-methyluridine (30), but not incorporated into an RNA sequence. First, we developed a short (four steps) and facile synthesis from 5′-end (13) with an 83% yield. The acetonide protecting group was subsequently removed to deliver fC (KCl- HCl buffer for pH values between 1.1 and 2.0 and citrate-phosphate buffer for pH values between 2.2 and 5.0). The line fitting and data analysis was conducted with Prism v3.00 (Graphpad Software, Inc.) (Figure 7).

**RESULTS AND DISCUSSION**

A 5-formylcytidine (fC) has previously been synthesized from 5′-end (13) with an 83% yield. The acetonide protecting group was subsequently removed to deliver fC (KCl- HCl buffer for pH values between 1.1 and 2.0 and citrate-phosphate buffer for pH values between 2.2 and 5.0). The line fitting and data analysis was conducted with Prism v3.00 (Graphpad Software, Inc.) (Figure 7).

**RESULTS AND DISCUSSION**

A 5-formylcytidine (fC) has previously been synthesized from 5′-end (13) with an 83% yield. The acetonide protecting group was subsequently removed to deliver fC (KCl- HCl buffer for pH values between 1.1 and 2.0 and citrate-phosphate buffer for pH values between 2.2 and 5.0). The line fitting and data analysis was conducted with Prism v3.00 (Graphpad Software, Inc.) (Figure 7).
The equilibrium binding of the two ASLs to the cognate and noncognate codons, AUG and AUA respectively, was assessed using programmed E. coli ribosomes. (The ASL unmodified and the wobble position is designated as ‘ψ27’; ASL modified at the wobble position is designated as ‘ψ27;f5C34’.) The ASLs were bound to the A-site with E. coli ASL Val1 bound to its cognate codon in the P-site. The ASLs were bound to the P-site with the ASL Val1 bound to its cognate codon in the A-site.

Figure 6. Codon binding by hmASL Val1, ψ27 and the hmASL Val1, ψ27;f5C34. The equilibrium binding of the two ASLs to the cognate and noncognate codons, AUG and AUA respectively, was assessed using programmed E. coli ribosomes. (The ASL unmodified and the wobble position is designated ‘ψ27’; ASL modified at the wobble position is designated ‘ψ27;f5C34’.) The ASLs were bound to the A-site with E. coli ASL Val1 bound to its cognate codon in the P-site. The ASLs were bound to the P-site with the ASL Val1 bound to its cognate codon in the A-site.

Figure 7. Analysis of the pKa of (filled squares) cytidine in comparison to that of (filled triangles) 5-formylcytidine. UV spectra were collected for the two nucleosides over a range of pH values, pH 2.2–7.0 for cytidine, and a pH 1.1–5.0 for 5-formylcytidine. The absorbance at 280 nm was normalized and plotted against the pH.

F2 = 9.40 p.p.m.) corresponding to the CH group of the f5C modification. A complete and unambiguous assignment of the non-exchangeable protons was achieved by using the two dimensional (2D) 1H-13C Heteronuclear Multiple Quantum Correlation (HMOC) method (31,32). As expected, the NMR peaks observed between (F1 = 50.00–85.00 p.p.m.; F2 = 4.50–6.00 p.p.m.) corresponded to the protons (H1’-H5’/H5’’/H5’’’) of the ribose moiety (Figure 3B) (33). Conversion of f5C into the 2’,5’-protected f5C phosphoramidite (9, Figure 2) commenced with the protection of the 3’and 5’ hydroxy groups as a disiloxane followed by protection of the 4-NH2 group as the formamidine 6 (71% yield). Installation of a 2’-ACE orthoester [2’-O-bis(acetoxyethoxy)methyl]- followed by fluoride treatment delivered the diol 7 in 53% yield over two steps (15). The synthesis of the f5C phosphoramidite 9 was completed through 5’-BZH (5’,O-benzhydroxy-bis(trimethylsiloxy)silyl-) protection (8, 84% yield) and phosphor-amidite formation (93% yield). A major concern for the incorporation of f5C phosphoramidite into synthetic RNA oligomers was the formation of imine adducts under cleavage and deprotection conditions. However, we decided not to protect the formyl group since the final deprotection of the 2’-ACE groups under mildly acidic conditions would potentially hydrolyze any imine formation that results during base-deprotection back to the formyl group.

In order to investigate the contribution of f5C34 to the structure of the anticodon loop and to the decoding of both the AUG and AUA codons at both the A- and P-sites, f5C was incorporated into the anticodon stem and loop domain of hmRNA CAU (hmASL Met CAU) at the wobble position 34, along with pseudouridine, ψ27. The oligonucleotide was synthesized with ψ27 and with and without f5C34 (hmASL Met CAU, ψ27;f5C34, and hmASL Met CAU, ψ27) using care not to oxidize the formyl group. Incorporation of f5C into RNA was accomplished by activating with S-ethyl tetrazole and coupling of the activated species for 3.5 min to the growing polynucleotide on the solid-support. Cleavage from the support and deprotection of the exocyclic amines was tested using NH2OH at room temperature for 24 h, methylamine at room temperature for 6 h, and t-butyl amine in water (1:3, v/v) at 60°C for 6 h. Only the t-butyl amine conditions resulted in the correct mass upon MALDI-TOF analysis of the crude products. There was no indication of any t-butyl-amine adducts present from the MALDI-TOF results. Successful incorporation of f5C was confirmed by NMR measurements (Figures 3 and 4) and HPLC of constituent nucleosides (Figure 4). The proton resonance of the formyl group is observed in the low field region of the 1D spectrum of hmRNA Met CAU, ψ27;f5C34 and absent from that of the hmASL Met CAU, ψ34 (Figure 3). The formyl proton in hmASL Met CAU, ψ27;f5C34 resonates at the same chemical shift as that of the mononucleoside f5C34, as observed in the superimposed 1H-13C HMOC spectra of cytidine and 5-formylcytidine (Figure 3B). The HPLC
nucleoside composition analysis confirms the presence of the fC (Figure 4A).

The structure of the resulting hmASL_Met-CAU-Ψ27;f5C34 was characterized by 1D 1H and 2D 1H NOESY NMR experiments conducted in H2O at 2°C (34), and by determining the thermodynamic contributions of fC to the RNA. The formyl proton resonance was found at 9.45 p.p.m. corresponding almost exactly to that of the nucleoside alone (Figure 3A). The imino protons of the stem of hmASL_Met-CAU-Ψ27;f5C34 were found to resonate between 12 and 13.5 p.p.m. on the 1H 1D NMR spectrum (Figure 3A). The NMR spin systems that involve the exchangeable imino protons of hmASL_Met-CAU-Ψ27;f5C34 were identified by conducting NMR experiments in H2O at 2°C (Figures 3A and Figure 4B). The identification and assignment of the exchangeable protons were indicative of the overall stability the hmASL_Met-CAU-Ψ27;f5C34 in solution, and the comparison with hmASL_Met-CAU-Ψ27 (Figure 3A) demonstrated the successful incorporation of f5C34 into the sequence of hmASL_Met-CAU.

The modified RNA synthesis has allowed us to begin examining the role of f5C34 in thermal stability and decoding activity of hmRNA_Met. Thermodynamic parameters were extracted from the repeated denaturations and renaturations of both hmASL_Met-CAU-Ψ27;f5C34 and hmASL_Met-CAU-Ψ27 (Table 1 and Figure 5A). Introduction of f5C34 lowered the melting temperature and standard free energy (∆G°) considerably, but did not alter the ASL’s hyperchromicity. The circular dichroism spectrum of the hmASL_Met-CAU-Ψ27 exhibited a greater ellipticity at 270 nm than that of the hmASL_Met-CAU-Ψ27;f5C34. The lower degree of ellipticity of hmASL_Met-CAU-Ψ27;f5C34 is indicative of a decrease in base stacking. The increased base stacking must be attributed to the anticodon loop nucleosides because of the location of the modification. These differences in thermodynamics and circular dichroism ellipticity between hmASL_Met-CAU-Ψ27;f5C34 and hmASL_Met-CAU-Ψ27 indicated that f5C34 may enhance the motional dynamics of the loop. This difference in motional dynamics was observed by a molecular dynamics simulation (MDS) performed on the hmASL_Met-CAU-Ψ27 and the hmASL_Met-CAU-Ψ27;f5C34 using AMBER 9 (35). The hmASL_Met-CAU-Ψ27 displayed an average root mean square deviation from the starting structure of 2.18 ± 0.23 as opposed to hmASL_Met-CAU-Ψ27;f5C34 for which higher fluctuations of 2.60 ± 0.60 were detected (Supplementary Data). The enhanced motional dynamics may be important for the decoding of AUG, as well as AUG.

The tRNA_Met anticodon CAU is a cognate pair for the Met codon AUG. According to Crick’s Wobble Hypothesis (36), the binding of anticodon CAU to codon AUA would be unlikely due to the C-A mismatch at the wobble position (wobble pair nucleosides in bold). However, the mitochondrial ribosome decodes both AUG and AUA using one tRNA with the anticodon CAU. This one tRNA consists of the modification f5C34. In contrast, two tRNAs decode the one Met codon AUG in the cytoplasm (37). One of the tRNAs is an initiator tRNA that decodes AUG in the ribosome’s peptidyl- or P-site at the initiation of translation, where AUG is the first codon to be translated on the mRNA. This initiator tRNA_Met consists of an unmodified CAU anticodon. The second cytoplasmic tRNA_Met is responsible for elongation and recognizes AUG located within the mRNA, and thus responds only to the aminoacyl- or A-site codon. In E. coli, this elongator tRNA_Met is modified with N4-acetylcytidine at the wobble position (acC34) (38). Thus, at the anticodon, one of the main distinguishing factors between the cytoplasmic initiator and elongator tRNA_Met is the modification at the wobble position. We used a codon-binding assay to observe the affinity of the hmASL_Met-CAU-Ψ27;f5C34 and hmASL_Met-CAU-Ψ27 for the codons AUA and AUG at either A-site or P-site of E. coli 70S ribosomes. To ensure binding of the two ASL_Met to the A- or the P-site, the ribosomal site not in observation (P- or A-site, respectively) was saturated with the unmodified E. coli ASL Val3 in response to its cognate codon GUA. The unmodified ASL Val3 binds its cognate codon with high affinity and specificity (39).

The hmASL_Met-CAU-Ψ27 bound AUG in the A-site and the P-site with an affinity comparable to what have been observed previously for certain ASLs with unmodified wobble positions responding to cognate codons (Figure 6) (21,39). In contrast, the hmASL_Met-CAU-Ψ27 bound poorly to AUG in both the A- and P-sites. Surprisingly, introduction of f5C34 enhanced binding to AUG by 2-fold (Figure 6). Our results indicated that of the two codons at either of the two ribosomal sites, the f5C34 modification appears to be most important for reading AUG.

Both the hmASL_Met-CAU-Ψ27 and the fully modified hmASL_Met-CAU-Ψ27;f5C34 exhibited considerable affinity for AUG, and at both the A-site and the P-site. However, only the hmASL_Met-CAU-Ψ27;f5C34 exhibited significant affinity for AUA. There was a doubling in the affinity of ASL_Met for the AUA codon when f5C34 was present. This increase in affinity of the f5C34-modified ASL in comparison to that of the hmASL_Met-CAU-Ψ27, unmodified at the wobble position, was not observed on AUG and may therefore be the sole contributor to the efficient translation of AUA codons. A 2-fold increase in affinity of tRNA toward a codon has been shown to be significant in translation (40). Although some ASL modifications cause small increases in codon-binding affinity, others can dramatically increase affinity to codons (39).
Of particular interest is the chemical and conformational mechanisms by which a stable, but noncanonical base pair occurs between $f^5C_{34}$ and the third base of the AUA codon, an adenosine, on the ribosome. C-A base pairs are extremely unusual. Although the C-A pairing has been found in the folded structure of some RNAs such as ribosomal RNAs (rRNAs), it is rarely found in anticodon:codon pairs. An anticodon:codon C-A mismatch has been detected when $C_{34}$ of tRNA$^{\text{Leu}}_{\text{CAU}}$ modified with lysidine (k$^2$C34) at the wobble position, is paired to the cytoplasmic isoleucine codon AUA. The lysine moiety of C34 on the anticodon provides an amino group which hydrogen bonds to A of the codon, thus allowing the wobble position C-A mismatch to occur. One could imagine that the 5-formyl modification raises the $pK_a$ of cytidine’s N3 to the physiological range where an additional hydrogen bond could be formed to AUA. However, the $pK_a$ of $f^5C_{34}$ determined by UV spectral analysis was lower than that of C (2.3 and 4.2, respectively; Figure 7), corresponding well with previous determinations (41) including those for $df^5C$ (42) and for $f^5U$ (43). Thus, $f^5C_{34}$ must contribute to the decoding of the mitochondrial genome through a different mechanism. Another C-A anticodon:codon mismatch may occur at the wobble position when tRNA$^{\text{Leu}}_{\text{CmAA}}$ pairs with the leucine codon UUA. Similar to hmtRNA Met, this tRNA Leu isoacceptor has a 5-formylated, 2’-O-methylated C at the wobble position, $f^5C_{34}$. The wobble modifications are thought to be a general characteristic of mammalian cytoplasmic tRNA$^{\text{Leu}}_{\text{CmAA}}$ that may aid in the decoding of leucine codons UUG and UUA and prevent the miscoding of the similar codons of phenylalanine, UUU and UUC. However, there is a lack of information on the decoding properties of tRNA$^{\text{Leu}}_{\text{CmAA}}$ and therefore, there is the possibility that an isoacceptor other than tRNA$^{\text{Leu}}_{\text{CmAA}}$ is responsible for specifically reading UUA (6,38).

Eighty percent of the methionine codons internal to mitochondrial mRNA are the AUA codon. Thus, the enhanced affinity of hmtRNA$^{\text{Met}}_{\text{CAU}}$-$f^5C_{34}$ for AUA in the A-site of the E. coli ribosome has important implications for the affinity and kinetics of decoding AUA during elongation. The enhanced A-site binding of AUA by the $f^5C$-modified, hmtRNA$^{\text{Met}}_{\text{CAU}}$ may be even more evident on the mitochondrial ribosome, a concept not studied here. Also, the disease-related A37-G37 (A4435G) mutation, associated with an increased penetrance and expression of the primary Leber hereditary optic neuropathy mutation (G11778A), LHON (44), may critically alter the anticodon architecture such that either or both decoding events do not occur. This has yet to be examined. The synthesis of the wild-type modified and unmodified anticodon stem and loops of the hmtRNA$^{\text{Met}}_{\text{CAU}}$ and their physical, chemical characterizations will be important in understanding the contributions of the modification to biological function and in characterization of the human disease-relevant mutant tRNA.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are very grateful to Dr. Glenn Björk (University of Umeå, Sweden) for the HPLC analysis of the synthesized RNA. AD is a Beckman Young Investigator and a Cottrell Scholar. We thank Mr. Antonio M. Munoz for his contribution to the molecular dynamics simulations.

FUNDING

The North Carolina State University RNA Biology Group (to A.D.); a grant from United Mitochondrial Disease Foundation, jointly with Dr. Linda Spremulli of the University of North Carolina—Chapel Hill (grant number 05-20 to P.F.A.); NSF (grant number MCB0548602 to P.F.A.); an NSF Graduate Fellowship (to E.M.G.); and Dharmacon RNA Technologies, Inc. (ThermoFisher, Inc.). Funding for open access charge: National Science Foundation.

Conflict of interest statement. The perception of conflict may arise due to Drs. Michael O. Delaney and Rob Kaiser being employees of Dharmacon RNA Technologies (ThermoFisher). The company provided the polymer synthesis in which the f5C nucleotide was incorporated into RNA for the first time.

REFERENCES


