Post-transcriptional modifications in the small subunit ribosomal RNA from *Thermotoga maritima*, including presence of a novel modified cytidine

REBECCA GUYMON,1,3 STEVEN C. POMERANTZ,1,4 J. NICHOLAS ISON,1 PAMELA F. CRAIN,1 and JAMES A. MCCLOSKEY1,2

1Department of Medicinal Chemistry, University of Utah, Salt Lake City, Utah 84112, USA
2Department of Biochemistry, University of Utah, Salt Lake City, Utah 84112, USA

ABSTRACT

Post-transcriptional modifications of RNA are nearly ubiquitous in the principal RNAs involved in translation. However, in the case of rRNA the functional roles of modification are far less established than for tRNA, and are subject to less knowledge in terms of specific nucleoside identities and their sequence locations. Post-transcriptional modifications have been studied in the SSU rRNA from *Thermotoga maritima* (optimal growth 80°C), one of the most deeply branched organisms in the Eubacterial phylogenetic tree. A total of 10 different modified nucleosides were found, the greatest number reported for bacterial SSU rRNA, occupying a net of 1414 sequence sites, compared with a similar number of sites recently reported for *Thermus thermophilus* and 11 for *Escherichia coli*. The relatively large number of modifications in *Thermotoga* offers modest support for the notion that thermophile rRNAs are more extensively modified than those from mesophiles. Seven of the *Thermotoga* modified sites are identical (location and identity) to those in *E. coli*. An unusual derivative of cytidine was found, designated N-330 (Mr 330.117), and was sequenced to position 1404 in the decoding region of the rRNA. It was unexpectedly found to be identical to an earlier reported nucleoside of unknown structure at the same location in the SSU RNA of the archaeal mesophile *Haloferax volcanii*.

Keywords: post-transcriptional modification; *Thermotoga maritima; Haloferax volcanii*; small ribosomal subunit RNA; liquid chromatography–mass spectrometry; electrospray ionization

INTRODUCTION

Relatively little is known concerning the specific roles of post-transcriptional modifications in rRNA (summarized by Decatur and Fournier 2002), in contrast to tRNA, in which a number of modification functions are slowly being unraveled (Agris 2004). In the case of rRNA, the influence of nucleotide modification on ribosome assembly and maturation (Bachellerie and Cavaille 1998; Maden 1998; Mason 1998) is perhaps the best documented (Vaughan et al. 1967; Cunningham et al. 1990, 1991; Rydén-Aulin et al. 1993; Sirum-Connolly and Mason 1993; Sirum-Connolly et al. 1995; Green and Noller 1996), but mechanistic details at the molecular level have been scant. There is strong circumstantial evidence in the archaeal thermophiles that ribose methylation at O2’ provides a significant mechanism for thermal stabilization of rRNA (Noon et al. 1998) and of archaeal tRNA (Kowalak et al. 1994; Noon et al. 2003), primarily through reduced flexibility resulting from strong thermodynamic enforcement of the C3’-endo sugar conformation (Kawai et al. 1992). Also, from the standpoint of function, selective methylation of rRNA nucleotides provides, in some instances, an effective mechanism for
conferring resistance to antibiotics that target the bacterial ribosome (Douthwaite et al. 2005).

Many of the structures and sequence locations of rRNA modifications are relatively well conserved (Maden 1990; Rozenski and McCloskey 2005) and exhibit distinct differences among the three principal evolutionary domains, Archaea, Bacteria, and Eukarya. The complete modification map for SSU rRNA of the bacterial thermophile Thermus thermophilus was recently completed (Guymon et al. 2006) and found to exhibit a notably lower level of modification than that of the archaeal thermophile Sulfolobus solfataricus (Noon et al. 1998) growing in the same temperature range (70°C–75°C), and a pattern of modification (nucleoside structures and sequence locations) unexpectedly similar to that in the mesophile Escherichia coli. For a tabulation of E. coli rRNA modifications and leading references see http://medlib.med.utah.edu/RNAmods/.

We have examined the SSU rRNA modifications in a second bacterial thermophile, Thermotoga maritima (optimal growth at 80°C) (Huber et al. 1986), an organism placed by small subunit ribosomal RNA (SSU rRNA) phylogeny as one of the most deeply rooted organisms in the Euarchaeal tree. Analysis of the complete Thermotoga genome sequence (Nelson et al. 1999) as well as earlier data (Aravind et al. 1998) led to conclusions that extensive lateral gene transfer between T. maritima and Archaea have occurred, reflected in the unusually high percentage (24%) of genes in T. maritima that are similar to those in Archaea. Examination of T. maritima SSU rRNA in the present study was intended in part to establish whether its modifications are to any extent archaeal in nature, and thus reflecting possible horizontal transfer of RNA modification enzymes. Further, this study serves to extend the knowledge base of bacterial SSU rRNA modification patterns (Rozenski and McCloskey 2005; Guymon et al. 2006) beyond E. coli and T. thermophilus, the only bacteria for which complete SSU modification maps are available, and beyond organisms for which RNase T1 catalog data were reported (Rozenski and McCloskey 2005) in which modifications were sometimes reported to occur, often without chemical identity of the modified nucleoside or firm knowledge of its sequence location in the 16S RNA molecule.

RESULTS

Modified nucleoside content of T. maritima 16S rRNA

Identities and approximate numbers of modified nucleosides, determined by LC/ESI-MS analysis of total nucleoside digests of 16S rRNA (Supplemental Fig. S1, at http://library.med.utah.edu/mccloskey) are: Ψ (3.8 residues), m3U (1.0), m4Cm (1.2), m2C (0.73), Cm (0.82), unknown nucleoside N-330 (UV molar absorptivity not known, so stoichiometry not estimated), m5G (2.3), m7G (1.1), Am (trace level to 1.3; see text), and m2A (2.1).

Identification and sequence location of modified residues in T. maritima 16S rRNA

RNA modifications represented by changes in mass from the parent nucleotide were mapped. The locations of pseudouridines, which are mass silent, were not measured beyond the estimation of 3.8 net residues in the molecule. Chromatographic separation of RNase T1 digestion products of T. maritima SSU rRNA is shown in Figure 1, with the corresponding mass data for modified oligonucleotides given in Table 1. A section of the RNase U2 chromatogram in which modified oligonucleotides were eluted is shown in Supplemental Figure S2.

To a certain extent the procedures used for identification and sequence placement of modified residues are analogous to those recently used and described in a study of T. thermophilus SSU rRNA (Guymon et al. 2006). Therefore, descriptions of results for seven modified residues in the present work can be found in the Supplemental Data section. These are: m7G-527 (Supplemental Fig. S3), m4G-966 and m4C-967 (Supplemental Fig. S4), m5G-1051 (Supplemental Table S2), m2Cm-1402 (Supplemental Fig. S5), Cm-1409 (Supplemental Figs. S6,S7), and 1518-m2Am2A-1519 placement. Results for N-330–1404, m7U-1498, and Am-1499 follow below.

N-330–1404

The presence of a novel structurally unknown nucleoside of M, at position 1404 is indicated from the following four lines of evidence. (1) The sequencing spectrum of T1 oligonucleotide M, 1393 (Supplemental Fig. S5) can be assigned only if a unique nucleotide residue mass of 392 (N-330p minus H2O) is used in the third nucleotide position. (2) The base- (m/z 197) and N > p- (m/z 391) monomer fragment ion signals (Supplemental Table S1) were assigned from the data in Supplemental Figure S5 and time aligned as shown in Supplemental Figure S8. Ions
### Table 1. Modified nucleosides in RNases T1 and U2 hydrolysis products from *Thermotoga maritima* 16S rRNA

<table>
<thead>
<tr>
<th>Modification</th>
<th>Sequence&lt;sup&gt;a&lt;/sup&gt;</th>
<th>RNase T1</th>
<th>RNase U2</th>
<th>Monomer ions, m&lt;zsup&gt;e&lt;/zsup&gt;</th>
<th>Monomer ions, m&lt;zsup&gt;e&lt;/zsup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>m&lt;sup&gt;7&lt;/sup&gt;G-527</td>
<td>525-CCm&lt;sup&gt;7&lt;/sup&gt;GCCGp&lt;sup&gt;6&lt;/sup&gt;-529</td>
<td>1637.29 (0.05)</td>
<td>164</td>
<td>164</td>
<td>1964.31 (0.03)</td>
</tr>
<tr>
<td>m&lt;sup&gt;7&lt;/sup&gt;G-966</td>
<td>964-AUm&lt;sup&gt;7&lt;/sup&gt;Gp-966</td>
<td>1012.16 (0.01)</td>
<td>164, 358, 376</td>
<td>965-Um&lt;sup&gt;7&lt;/sup&gt;Gm&lt;sup&gt;6&lt;/sup&gt;CUAp-969</td>
<td>1619.26 (0.04)</td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;C-967</td>
<td>967-m&lt;sup&gt;2&lt;/sup&gt;CUAAGp-971</td>
<td>1646.25 (0.01)</td>
<td>124, 336</td>
<td>164, 358, 376, 124, 336</td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1051-m&lt;sup&gt;2&lt;/sup&gt;Gp-1051</td>
<td>359.03 (0.03)</td>
<td>164</td>
<td>164, 358</td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;5&lt;/sup&gt;Cm-1402, N-330–1404</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>1393.28 (0.01)</td>
<td>124, 350, 197</td>
<td>1049-UCm&lt;sup&gt;5&lt;/sup&gt;Gm&lt;sup&gt;6&lt;/sup&gt;GP&lt;sup&gt;6&lt;/sup&gt;-1053</td>
<td>1661.25 (0.04)</td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>1619.26 (0.04)</td>
<td>164, 358</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>3270.53 (0.06)</td>
<td>124, 350, 197</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>1399-CCm&lt;sup&gt;5&lt;/sup&gt;CmCN-330GCUAp-1408</td>
<td>391 (N-330 base), 125, 319, 337, 342</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>1605.26 (0.06)</td>
<td>1620.16 (0.17) avg.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>1620.16 (0.17) avg.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>125, 319, 337, 342</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>Mass not found through tracking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>Mass not found through tracking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>Mass not found through tracking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m&lt;sup&gt;2&lt;/sup&gt;G-1051</td>
<td>1402-m&lt;sup&gt;5&lt;/sup&gt;CmCN-330Gp&lt;sup&gt;6&lt;/sup&gt;-1405</td>
<td>Mass not found through tracking</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup><sup>E. coli</sup> sequence numbering used throughout.

<sup>b</sup>Mass was experimentally measured; error is found minus calculated mass for the modified oligonucleotide shown.

<sup>c</sup>See also Supplemental Table S1.

<sup>d</sup>Sequenced by mass spectrometry.

---

16S rRNA modifications in *Thermotoga*.
corresponding to the modified residue m^4Cm, whose presence is indicated from the total nucleoside analysis data in Supplemental Figure S1, were also time aligned. The exact coelution pattern of ions from N-330 and m^4Cm is shown in Supplemental Figure S8, and supports the presence of both of those nucleosides in the T1 oligonucleotide M_r 1393. The characteristic bacterial nucleoside m^3Cm serves as an internal marker for position 1402 (Rozenski and McCloskey 2005) and correlates with the full oligonucleotide sequence derived by mass spectrometry in Supplemental Figure S5. (3) The presence of a base-modified unknown nucleoside of M_r 330 (MH^+, m/z 331; BH^+, m/z 199; λ_max 270) eluting at 3.85 min in the total nucleoside digest of Thermotoga 16S rRNA, Supplemental Figure S1. (4) Modification of the universal C-1404 was also reported in several other bacterial SSU rRNAs (although not in E. coli) (Rozenski and McCloskey 2005), but not the identity of the modification.

m^3U-1498 and Am-1499

Assignment of the unusual tandem methylation sites 1498 and 1499 was made primarily through the T1 product M_r 2318 (Table 1). The experimental mass 2318.43, using the corresponding gene sequence to calculate allowable RNA sequences, solely designates the unique sequence segment corresponding gene sequence to calculate allowable RNA sequences, solely designates the unique sequence segment corresponding to nucleotides 1498–1504 with two methyl groups. The sequence mass spectrum displayed in Figure 2, including the absence of ion m/z 225. The sequence of the Am-containing region is therefore assigned as shown in Table 1. The occurrence of Am is further supported by its presence in the total nucleoside digest (Fig. S1), with MH^+ and BH^+ ions appearing at the expected retention time (Pomerantz and McCloskey 1990). Interestingly, a second culture of T. maritima showed an otherwise very similar overall modification data set, but with the 1495–1499 fragment occurring primarily as M_r 1605.2 (see Table 1), i.e., with one fewer methyl group. We interpret this finding, taken in conjunction with variable amounts of Am found in nucleoside total digests from two different T. maritima cultures, as an indication that the level of Am-1499 is both substoichiometric and variable in the rRNA. To our knowledge, this is the first report of Am in RNA from bacteria; however, using LC/ESI-MS we have observed Am in unfractionated tRNA from E. coli MRE 600 (F. Qiu and J. McCloskey, unpubl., 1998).

3’-Terminal RNase T1 fragment from T. maritima 16S RNA

The terminal fragment is a 15-mer, expected to contain the anti-Shine–Dalgarno mRNA-binding sequence. Because of the recent unusual finding of adjacent pseudouridines at 1540 and 1541 in the 3’-tail of T. thermophilus (Guymon et al. 2006), the Thermotoga 3’-terminal fragment was subjected to a Ψ detection protocol (Pomerantz and McCloskey 2005) using LC/ESI-MS/MS. No Ψ was detected. The terminal fragment was mass measured as 4591.7 (4591.60 calc.), and MS-sequenced as AUCACCUCCUUUCUAOH (Supplemental Fig. S9), thus revealing the molecular terminus as AOH-1545 (E. coli numbering) in the processed 16S molecule.

Presence of nucleoside N-330 in Haloferax volcanii 16S RNA

The discovery of a new modified nucleoside in the Thermotoga SSU at the same position previously described as containing a modified cytidine in H. volcanii 16S rRNA (Gupta et al. 1983) prompted a revisit of our previously
published study of the *Haloferax* SSU RNA modifications (Kowalak et al. 2000). When the ion chromatograms for the diagnostic MH$^+$ and BH$_2^+$ ions for N-330 nucleoside (m/z 331 and 199, respectively) were extracted from our *Haloferax* total nucleoside digest data, their presence was confirmed by the appearance of these two ions in a peak exactly coeluting at 3.6 min (Fig. 3). If N-330 is, in fact, the “C***” nucleotide previously reported, the calculated mass of the predicted *Haloferax* N-330-containing T1 oligonucleotide (CCC*Gp) would be M, 1365.7. Reconstructed ion chromatograms were therefore generated using the M$^+$ and M$^{2+}$ ions (m/z 1364.7 and m/z 681.8, respectively), and for the B$^-$ ion for the base of N-330 (m/z 197). As shown in Supplemental Figure S10, each of the three ions forms a peak eluting at 16.0 min, indicating N-330 to be the modified C$^*$ nucleotide found by Woese and colleagues (Gupta et al. 1983).

**DISCUSSION**

**Modification identities and levels in *T. maritima* SSU RNA**

The finding of 10 different modified nucleosides at a net occupancy level of ~14 sequence sites marks *T. maritima*, and similarly *T. thermophilus* (Guymon et al. 2006), as the most extensively modified bacterial SSU RNAs presently known. Examination of a number of catalogued RNase T1 modification maps, although by their nature less complete than the full modification maps from *E. coli* (leading citations in ref. Bakin et al. 1994; see also http://medlib.med.utah.edu/RNAmods), and *T. thermophilus* (Guymon et al. 2006) SSU RNAs, suggests that the levels and identities of modifications are perhaps narrower and more conserved in Bacteria than in the Archaea and Eukarya. For example, modification levels in Archaeal 16S RNA range from five residues in *H. volcanii* (Gupta et al. 1983) to ~38 in *S. solfataricus* (Noon et al. 1998). Modification at ~8–11 SSU RNA sequence sites appears to be most common in bacteria (*E. coli* has 11) while the elevation by about 30% of modification levels in the bacterial thermophiles (versus mesophiles) supports the conclusion that post-transcriptional modification in general serves to support structural stabilization of RNA (Sampson and Uhlenbeck 1988; Derrick and Horowitz 1993; Kowalak et al. 1994). The presence of 11 modifications was recently reported for *T. thermophilus* 23S rRNA (Mengel-Jorgensen et al. 2006), compared with ~23 in *E. coli* LSU RNA. This relative concentration is significantly lower than expected, based on our finding of 14 sites in the *Thermus* 16S RNA. However, the methodology used (Mengel-Jorgensen et al. 2006) was not designed to provide a complete census of modifications that required placement. As in the case of *T. thermophilus* (Guymon et al. 2006), rRNA modification levels in *Thermotoga* are characteristically much lower than in the Archaeal thermophiles growing at about the same temperature (Noon et al. 1998), with considerably less reliance on ribose O-2′ methylation as a major means of structural stabilization. The finding of a net 3.8 residues of Ψ in Thermotoga SSU RNA is notable in that Ψ has been shown to play a strong, although often overlooked role in RNA stabilization (Davis 1995, 1998), and the level in *Thermotoga* appears to be the highest concentration of Ψ reported in bacterial SSU RNAs (Ofengand and Rudd 2000).

The availability of three bacterial SSU RNA modification maps, *E. coli* (Brosius et al. 1978), *Thermus* (Guymon et al. 2006), and now *Thermotoga*, when coupled with information from T1 catalogs (Rozenski and McCloskey 2005) permits an estimation of the most highly conserved modification sites that are unique to bacteria, even though most of the T1 catalog data identify sites but not chemical structures of modifications. In bacteria these sites are C*-967 in h31 and U*-1498 in the center of the decoding region in h44. An exception to the bacterial uniqueness of a modified C-967 appears to occur in the archaeon *Thermoproteus tenax* (Woese et al. 1984). The position 967 modification is reported to be m3C in all four cases in which the nucleoside structure has been established (*E. coli*, *Thermus*, *Thermotoga*, and *Proteus*). The adjacent modification m3G-966, implicated in P-site tRNA binding (von Ahsen and Noller 1995; Korostelev et al. 2006), represents an essentially universal SSU modification site in all phylogenetic domains (Rozenski and McCloskey 2005) and is represented by an interesting diversity of modified nucleoside structures in Archaea and Eukarya (Kowalak et al. 2000). The identity of U*-1498 has been established specifically as m3U in the same four organisms. The unusual dimethylcytidine m4Cm-1402 is unique to bacterial rRNA (Limbach et al. 1994), but this site appears to be...
modified in fewer than half of the reported cases (Rozenski and McCloskey 2005). In *T. thermophilus* the m^4^Cm residue serves to stabilize the third tRNA nucleotide by H binding of its phosphate to the 4-methylamino moiety (Korostelev et al. 2006).

**Modifications at the functional center of the ribosome**

As deduced in earlier literature, modification sites in bacterial SSU RNAs tend to occur, in three-dimensional space, near the decoding center of the RNA (Brimacombe et al. 1993; Decatur and Fournier 2002). Four of the 16S modifications (m^6^G-966, m^6^C-1400, m^6^Cm-1402, and m^3^U-1498) were determined by X-ray crystallography to support interaction between 16S RNA and the P-site codon and anticodon stem–loop (Korostelev et al. 2006). These observations reflect the net importance of modification to efficient ribosomal function, as has been stated (Decatur and Fournier 2002). The high level of modification in the upper portion of h44 in SSU RNA occurs at the interface with the LSU RNA, forming a cavity in which translation occurs. These modifications in h44 include six methyl groups each in *Thermotoga* and *Thermus*, opposite an analogous concentration of modifications on the 23S side of the *Thermus* LSU (Mengel-Jorgensen et al. 2006). Mengel-Jorgensen et al. (2006) have concluded that the occurrence of modifications in the 23S RNA of *Thermus* principally at the RNA–RNA interface suggests that they play a role in modulating the RNA–RNA interface contact. Their conclusions are supported by the locations of modifications in 16S RNA from *Thermotoga* (present work), *Thermus* (Guyon et al. 2006), and a series of studies on *E. coli* (Brimacombe et al. 1993; Decatur and Fournier 2002), showing that the ribosomal subunit interface is intimately associated with post-transcriptional modifications.

**Modified nucleoside N-330–1404 in the decoding region of the RNA**

Unknown N-330 is remarkable in two ways: first, in terms of its structural properties as inferred thus far; second, its unexpected occurrence in another phylogenetic domain, at the same location in the SSU RNA of the archaeal mesophile *H. volcanii*. Interestingly, C-1404 has been reported as modified but with unknown structure in the RNase T1 SSU maps of five bacteria (Rozenski and McCloskey 2005), and two archaea, *H. volcanii* (Gupta et al. 1983) and *S. solfataricus* (Woese et al. 1984). C-1404 pairs with essentially universal G-1497 (Cannone et al. 2002) near the top of h44. N-330 is implied to be a derivative of cytidine by the corresponding gene sequence (Gupta et al. 1983). The molecular mass of 330 Da is unique among all known modified nucleosides in RNA (see http://medlib.med.utah.edu/RNAmods) and suggests its structure to be more complex than any of the nine known modified cytidines in rRNA (Rozenski and McCloskey 2005). The accurate molecular mass of N-330 was measured as 330.117 ± 0.002 using a Micromass Q-Tof mass spectrometer. This value is consistent with an elemental composition of C_{12}H_{18}N_{4}O_{7} (calc. 330.1175), thus requiring a side chain having one N atom, with three Ns accounted for by the cytosine heterocycle. Unfortunately, the exact chemical structure of N-330-1404 cannot be further pursued due to laboratory closure and retirement of the corresponding author.

It is interesting to consider the possibility that the structures and distributions of post-transcriptionally modified nucleosides—the specific products of a sizable array of RNA modification enzymes—could, in some instances, be indicators of past horizontal transfer of DNA coding for modification enzymes. This notion is supported by the sharp lines of phylogenetic demarcation of some modified nucleoside families (Limbach et al. 1994; Motorin and Grosjean 1998). The possibility of lateral transfer was earlier raised (McCloskey et al. 2001) in light of the strong clustering of the (otherwise bacterial) anticodon nucleoside mm^5^s^3^U in tRNAs from the *Methanococci*, a lineage of methanogenic marine archaea. Interestingly, nucleoside N-330 is the only rRNA nucleoside occurring in *Thermotoga* that is shared only by bacteria and archaea, all others (with the exception of m^4^Cm-1402, which is uniquely bacterial) being found in all three phylogenetic domains. Nucleoside N-330 might therefore be, in part, the product of RNA modification enzyme genes transferred horizontally between *Halofexax* and *Thermotoga*. This speculation is based on the report of extensive lateral transfer between *Thermotoga* and archaea (Aravind et al. 1998; Logsdon and Faguy 1999), such that almost one-quarter of the *Thermotoga* genome (24%) was found to be archaeal in nature (Nelson et al. 1999). Work in the area of RNA modification enzymes (Leung et al. 1998), however, while extensive in some notable cases, such as the yeast tRNA modification enzymes (Johansson and Byström 2005) and the pseudouridine synthases (Ofengand and Fournier 1998), by its nature generally lags behind knowledge of modified nucleoside structures (now numbering about 106; see http://medlib.med.utah.edu/RNAmods) and their phylogenetic distributions. Pursuit of this possibility that N-330 is a product of lateral gene transfer will require knowledge of the chemical structure of N-330, and also of the sequences of specific enzymes responsible for its formation.

**MATERIALS AND METHODS**

*Thermotoga maritima* MSB8 cells were grown at 80°C at The University of Georgia Bioexpression and Fermentation Facility, under the supervision of T.E. Davies, and RNA from this source was acquired in two ways. First, purified 16S rRNA was obtained as a gift from V. Ramakrishnan (Department of Biochemistry, University of Utah; present address MRC Laboratory of Molecular
Biology, Cambridge, U.K.). Second, T. maritima 30S ribosomal subunits were obtained from the same source and 16S RNA was extracted from the 30S ribosomes with TRI Reagent (Chomczynski and Sacchi 1987) following the manufacturer’s protocol. 16S rRNAs (typically, 100 μg; 200 pmol) were digested totally to nucleotides with nuclease P1, phosphodiesterase I, and BAP (Crain 1990). Aliquots of RNA (5–100 pmol) were digested with 1000 units of RNase T1 (Ambion) in 10 mM Tris+ 1 mM EDTA, pH 7 for 30 min at 37°C or 45 min at 55°C. Alternatively, the RNA was digested with RNase U2 (Industrial Research Ltd.) in 20 mM diammnonium citrate (DAC), pH 5, 1 mM EDTA, or in 20 mM DAC, pH 5, 1 mM EDTA, 8 M urea, with 5–10 units RNase U2 for 15 min at 60°C, then an additional 5–10 units of enzyme was added and digestion continued for 15 min. Details of the mass spectrometry methods used in this study can be found in the Supplemental Data section.

SUPPLEMENTAL DATA

The following supplemental material can be found at http://library.med.utah.edu/mccloskey: LC/ESI-MS analysis of nucleosides in Thermotoga 16S rRNA; partial chromatogram from LC/ESI-MS analysis of RNA U2 digest of Thermotoga 16S rRNA; assignments for monomer ions used in Table 1; data and discussion for placements of mG, m2G (two each), m5C, m6A (two each), and Cm; additional comments on nucleoside N=300; sequence mass spectrum of T1 oligonucleotide M, 4591.7; mass spectrometry Materials and Methods.

ACKNOWLEDGMENTS

This work was supported by NIH Grant R01 GM29812 (J.A.M.). We are indebted to V. Ramakrishnan for T. maritima ribosomes and rRNA, and to M. Tarry for preparative work in the V. Ramakrishnan laboratory. We thank H. F. Noller for a copy of his 70S ribosome-tRNA manuscript prior to publication, and F. Kirpekar for a preprint of his paper on T. thermophilus 23S rRNA modifications.

Received October 27, 2006; accepted December 13, 2006.

REFERENCES

Post-transcriptional modifications in the small subunit ribosomal RNA from *Thermotoga maritima*, including presence of a novel modified cytidine

Rebecca Guymon, Steven C. Pomerantz, J. Nicholas Ison, et al.

*RNA* published online January 25, 2007

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to *RNA* go to: http://rnajournal.cshlp.org/subscriptions

Copyright © 2007 RNA Society