Agmatidine, a modified cytidine in the anticodon of archaeal tRNA(Ile), base pairs with adenosine but not with guanosine

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Mandal, D. et al. “Agmatidine, a modified cytidine in the anticodon of archaeal tRNA(Ile), base pairs with adenosine but not with guanosine.” Proceedings of the National Academy of Sciences 107.7 (2010): 2872-2877. Copyright ©2011 by the National Academy of Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.0914869107">http://dx.doi.org/10.1073/pnas.0914869107</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences (U.S.)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sat Dec 08 13:13:36 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/61373">http://hdl.handle.net/1721.1/61373</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Agmatidine, a modified cytidine in the anticodon of archaeal tRNA\textsubscript{Ile}, base pairs with adenosine but not with guanosine

Debabrata Mandal\textsuperscript{1,4}, Caroline Köh rer\textsuperscript{1,4}, Dan Su\textsuperscript{4}, Susan P. Russell\textsuperscript{1}, Kady Krivos\textsuperscript{1}, Colette M. Castleberry\textsuperscript{1}, Paul Blum\textsuperscript{6}, Patrick A. Limbach\textsuperscript{1}, Dieter Söll\textsuperscript{7}, and Uttam L. RajBhandarya\textsuperscript{1,2

\textsuperscript{1}Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139; \textsuperscript{2}Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520; \textsuperscript{3}Department of Chemistry, University of Cincinnati, Cincinnati, OH 45221; \textsuperscript{4}George Beadle Center for Genetics, University of Nebraska, Lincoln, NE 68588

Contributed by Dieter Söll, December 24, 2009 (sent for review December 2, 2009)

Modification of the cytidine in the first anticodon position of the AUA decoding tRNA\textsubscript{Ile} (tRNA\textsubscript{Ile}\textsubscript{AU}) of bacteria and archaea is essential for this tRNA to read the isoleucine codon AUG and to differentiate between AUG and the methionine codon AUG. To identify the modified cytidine in archaea, we have purified this tRNA species from Haloarcula marismortui, established its codon reading properties, used liquid chromatography–mass spectrometry (LC-MS) to map RNase A and T1 digestion products onto the tRNA, and used LC-MS/MS to sequence the oligonucleotides in RNase A digests. These analyses revealed that the modification of cytidine in the anticodon of tRNA\textsubscript{Ile} adds 112 mass units to its molecular mass and makes the glycosidic bond unusually labile during mass spectral analyses. Accurate mass LC-MS and LC-MS/MS analysis of total nucleoside digests of the tRNA\textsubscript{Ile} demonstrated the absence in the modified cytidine of the C2-oxo group and its replacement by agmatine (decarboxy-arginine) through a secondary amine linkage. We propose the name agmatidine, abbreviation C, for this modified cytidine. Agmatidine is also present in Methanococcus maripaludis tRNA\textsubscript{Ile} and in Sulfolobus solfataricus total tRNA, indicating its probable occurrence in the AUA decoding tRNA\textsubscript{Ile} of euryarchaea and crenarchaea. The identification of agmatidine shows that bacteria and archaea have developed very similar strategies for reading the isoleucine codon AUG while discriminating against the methionine codon AUG.

\textsuperscript{4}D.M. and C.K. contributed equally to this work.
Freely available online through the PNAS open access option.

The genetic code table consists of sixteen four-codon boxes. In fourteen of the boxes, all four codons either specify the same amino acid or are split into two sets of two codons, with each set encoding a different amino acid. For example, the UUN box is split into UUU/UUG codons for phenylalanine and UUA/UGG coding for leucine. The wobble hypothesis of Crick proposes how a single phenylalanine tRNA with G in the first anticodon position can base pair with either U or C and a single leucine tRNA with a modified U (or 2-thioU) in the anticodon can base pair with either A or G (1-3). The two remaining boxes, UGN and AUN, are exceptions in that the UGN box is split into UGU/UUG coding for cysteine, UGG coding for tryptophan, and UGA being used as a stop codon, whereas the AUN box is split into AUU/AUA/CAA/CAU, indicating its possible occurrence in the AUA decoding tRNA\textsubscript{Ile} of euryarchaea and crenarchaea. The identification of agmatidine shows that bacteria and archaea have developed very similar strategies for reading the isoleucine codon AUG while discriminating against the methionine codon AUG.

Eukaryotes, on the other hand, contain a tRNA\textsubscript{Ile} with the anticodon IAU (I = inosine), which can read all three isoleucine codons using the wobble pairing rules of Crick. They also contain a tRNA\textsubscript{Ile} with the anticodon 1ΨAΨ, which is thought to read only the isoleucine codon AUG but not the methionine codon AUG (8). Given these two distinct mechanisms in bacteria and eukaryotes, a question of much interest is how the archaeal tRNA\textsubscript{Ile} accomplishes the task of reading the AUG codon.

In a recent paper, we showed that a tRNA with the anticodon sequence CAÜ, which was annotated as a methionine tRNA in the archaeon Haloarcula marismortui, was actually aminoaacetylated with inosine in vivo (9). We showed that the cytidine in the anticodon of this tRNA was modified, but not to lysidine as it is in bacteria, and that the same modification was likely present in other haloarchaea and in Methanocaldococcus jannaschii. We also showed that modification of the cytidine was necessary for aminoaacetylation of the tRNA with inosine in vitro. These findings raised the question of the nature of the modification in the archaeal AUA decoding tRNA\textsubscript{Ile} (tRNA\textsubscript{Ile}\textsubscript{AU}) and its relationship, if any, to lysidine in Escherichia coli tRNA\textsubscript{Ile}.

Here, we describe the purification of two tRNA\textsubscript{Ile} species from H. marismortui. We show that tRNA\textsubscript{Ile} binds to AUC but not to AUG or AU on the ribosome, whereas tRNA\textsubscript{Ile} binds to AUG but not to AUC, AUG, or AUA. Mass spectral analysis of nucleosides and oligonucleotides isolated from tRNA\textsubscript{Ile} show that the modified cytidine is located in the anticodon wobble position, lacks the C2-oxo group, and has instead agmatine (1-amino-4-guaindino butane) attached to it through a secondary amine linkage. We propose the name agmatidine and the abbreviation C+ for this modified nucleoside. Agmatidine is also present in tRNA\textsubscript{Ile} from Methanococcus maripaludis and in total tRNA isolated from Sulfolobus solfataricus, indicating its possible presence in tRNA\textsubscript{Ile} of euryarchaea and crenarchaea.

Agmatidine is in many ways similar to lysidine. In both cases the C2-oxo group of cytidine is replaced by either of two closely related basic amino acids, deoxyarginine or lysine. Thus, bacteria and archaea have developed very similar strategies for generating a tRNA\textsubscript{Ile} to read the isoleucine codon AUG without also reading the methionine codon AUG. The identification of agmatidine also suggests that agmatine, which is a known
neuromodulator (10, 11), and which is essential for polyamine biosynthesis in the archaeon *Thermococcus kodakaraensis* (12), may also be essential in euryarchaea and in crenarchaea (13) for modification of the cytidine required for decoding the AUA codon by tRNA\textsubscript{Ile}.

**Results**

**Purification and Characterization of Isoleucine tRNAs from *H. marismortui***. Starting from total tRNA, a two-step procedure, involving hybrid selection of tRNA with biotinylated DNA oligonucleotides bound to streptavidin sepharose followed by native polyacrylamide gel fractionation of the enriched tRNA, was used to isolate pure tRNA\textsubscript{Ile}s. The biotinylated oligonucleotides were complementary to nucleotides 54–73 of tRNA\textsubscript{Ile} or tRNA\textsubscript{Ile}\textsubscript{320} (Fig. L4) and contained biotin at the 3'-hydroxyl end. The final yields of purified tRNAs from approximately 1,000 A\textsubscript{260} units of total tRNA were 6–8 × 10\textsuperscript{4} A\textsubscript{260} units for tRNA\textsubscript{Ile} and 0.6–0.8 × 10\textsuperscript{4} A\textsubscript{260} unit for tRNA\textsubscript{Ile}\textsubscript{320}. The significantly lower yield of tRNA\textsubscript{Ile} compared to tRNA\textsubscript{Ile}\textsubscript{320} reflects the low abundance of this tRNA and the fact that AUA is a rarely used codon in haloarchaea as it is in *E. coli* (14).

Fig. 1B shows that the purified tRNAs are essentially homogeneous. tRNA\textsubscript{Ile} yielded only one band on a native polyacrylamide gel, with ethidium bromide, and this band hybridized to the corresponding complementary oligonucleotide in Northern blots. tRNA\textsubscript{Ile} yielded two bands, a strong band and a weaker one, with both bands hybridizing to the oligonucleotide complementary to this tRNA. Based on a comparison of intensities of the hybridization bands and the amount of total tRNA (0.5 × 10\textsuperscript{4} A\textsubscript{260} unit) and the purified tRNAs (0.005 × 10\textsuperscript{4} A\textsubscript{260} unit) loaded on the gel, the purified tRNA\textsubscript{Ile} is enriched approximately 57-fold compared to total tRNA, and the purified tRNA\textsubscript{Ile}\textsubscript{320} is enriched approximately 500-fold.

Additional evidence that the tRNA\textsubscript{Ile} is essentially homogeneous was derived from gel electrophoretic analysis of partial RNase T1 digests of 5'-\textsuperscript{32}P-labeled tRNA\textsubscript{Ile} (Fig. 1C). The 32P-labeled bands in the RNase T1 lane are due to cuts at G residues, and these are consistent with the pattern expected for tRNA\textsubscript{Ile}. There is no band due to cleavage by RNase T1 at G26 because G26 is dimethylated in the tRNA, and it is known that under conditions of partial RNase T1 digestion there is no cleavage of the phosphodiester bond on the 3'-side of the dimethyl-G residue (15).

The pattern obtained in partial alkali digests shows a pronounced shift in gel electrophoretic mobility of the 5'-\textsuperscript{32}P-labeled oligonucleotide going from nucleotide 33 to 34. We had shown previously that nucleotide 34 contained a modified cytidine (9). The presence of nine bands between G31 and G41 in the alkali digest shows that every phosphodiester bond between G31 and G41 is cleaved by alkali. This means that none of the nucleotides in between contain a substitution in the 2'-hydroxyl group. Therefore, the very pronounced shift in the gel electrophoretic mobility indicates strongly that the modified cytidine has at least one if not more positive charges in the ring, which would retard the electrophoretic migration of the oligonucleotide that contained the modified cytidine (16).

**Codon Reading Properties of tRNA\textsubscript{Ile} and tRNA\textsubscript{Ile}\textsubscript{320}**. Purified tRNA\textsubscript{Ile} and tRNA\textsubscript{Ile}\textsubscript{320} were aminoacylated in vitro with 32H-Ile and the 32H-Ile-tRNAs were used for measuring oligonucleotide-directed binding to *H. marismortui* ribosomes. The oligonucleotides used as template had the following sequences AUGAUC, AUGAU, AUGAUG and AUGAUU. Results of ribosome binding experiments show that tRNA\textsubscript{Ile} binds to AUC but not to AUG or AUA (Fig. 2A). In contrast, tRNA\textsubscript{Ile}\textsubscript{320} binds to AUA but not to AUG, AUC (Fig. 2B), or to AUU (Table S1). Thus, the modified cytidine in the anticodon wobble position of tRNA\textsubscript{Ile}\textsubscript{320} base pairs specifically with A but not with G, C, or U.

**Localization of the Modified Cytidine (C*) in the Anticodon of tRNA\textsubscript{Ile} and Determination of Its Mass**. Liquid chromatography–mass spectrometry (LC-MS)/MS RNase mapping and sequencing of nucleic acids has been used extensively for the localization of modified nucleosides in RNAs (17–21). Here, we used separate digests of tRNA\textsubscript{Ile} with RNase T1 and RNase A combined with bacterial alkaline phosphatase (BAP) followed by LC-MS/MS to confirm C34 as the site of the unknown modification and to identify the mass of the modified nucleoside. Knowledge of the mass of the modified nucleoside was then used as a search criterion for subsequent LC-MS and LC-MS/MS analysis of the nucleosides present in a total digest of tRNA\textsubscript{Ile}.

Digestion of tRNA\textsubscript{Ile} with RNase T1 is predicted to yield a decanucleotide 5'-CUC-AU[t\textsuperscript{6}A]ACCGp-3', which would

![Figure 1](https://example.com/figure1.png)

**Figure 1.** (A) Cloverleaf structure of tRNA\textsubscript{Ile} from *H. marismortui*. Location of modified nucleosides is based on Gupta (46) and LC-MS analysis presented in this work. G\textsuperscript{1} (arachosine); m\textsuperscript{1}G (N\textsubscript{6}, N\textsubscript{2}-dimethylguanosine); C\textsuperscript{*} (modified cytidine); t\textsuperscript{6}A (N\textsuperscript{2}-threonylcarbamoyladenosine); m\textsuperscript{5}C (5-methylcytidine); m\textsuperscript{1}ψ (1-methylpsuedouridin); ψ (pseudouridine); C\textsuperscript{m} (2'-O-methylcytidine); m\textsuperscript{1}I (1-methylinosine). (B) Purification of isoleucine tRNAs from *H. marismortui*. Native PAGE analysis of total tRNA, purified tRNA\textsubscript{Ile} (Ile2) and tRNA\textsubscript{Ile} (Ile1) is shown. tRNAs are visualized by ethidium bromide staining or Northern blot analysis using probes specific for tRNA\textsubscript{Ile} and tRNA\textsubscript{Ile}\textsubscript{320} as indicated. (C) Characterization of purified tRNA\textsubscript{Ile}. The homogeneity of tRNA\textsubscript{Ile} was confirmed by partial RNase T1 digest (lane T1) of 5'-\textsuperscript{32}P-labeled tRNA. 32P-labeled fragments were separated by denaturing PAGE and visualized by autoradiography; lane OH-, partial alkali digest.

Mandal et al.

PNAS  | February 16, 2010  | vol. 107  | no. 7  | 2873
include C34 (designated as C*), the first nucleotide of the anticodon. The predicted molecular mass of this digestion product without the C34 modification is 3237.4 u. Consistent with our previous results (9) showing that C34 is quantitatively modified in tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}, LC-MS analysis of the RNase T1 digestion mixture failed to identify any \(m/z\) values related to the above unmodified sequence. Instead, abundant ions were detected (Fig. S1) at \(m/z\) 1718.8 (doubly charged) and \(m/z\) 1145.5 (triply charged). These \(m/z\) values correspond to an oligonucleotide mass of 3339.6, an increase of 112 over the oligonucleotide with no modification at C34. Collision-induced dissociation tandem mass spectrometry (CID-MS/MS) was attempted on this oligonucleotide, but minimal sequence information could be obtained likely due to the relatively large size of this RNase T1 digestion product and the lability of the glycosidic bond involving the modified C34 (see below).

To obtain a shorter oligonucleotide more amenable to MS/MS, a combination of RNase A and BAP was used to digest tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}. The sequence of the RNase A and BAP digestion product of tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1} that covers the anticodon is the trinucleoside diphosphate 5'-C*pApU-3'. Based on elution time and \(m/z\) values, several oligonucleotides were detected in the total ion chromatograms (TIC) of the RNase A/BAP digest (Fig. 3A). Each of these was sequenced and all except one, \(m/z\) 989.3, could be mapped onto the tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1} sequence. The selected ion chromatogram (SIC) of the oligonucleotide at \(m/z\) 989.3 is shown in Fig. 3B, and the corresponding mass spectrum of the ions eluting at 4.5 min is shown in Fig. 3C. The mass spectrum shows that this oligonucleotide elutes with GpC (\(m/z\) 587.3) and ApG\textsubscript{p}U (\(m/z\) 958.3), which are derived from nucleotides 49–50 and 14–16, respectively, of tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}.

Because the difference in mass between the oligonucleotide at \(m/z\) 989.3 and the unmodified CpApU is 112 u and because there are no unmodified nucleotide base compositions that match 990.3 u (989.3 + H\textsuperscript{+}) nor any known modification mass of 112 u, this oligonucleotide is most likely C*pApU. This assignment is confirmed by MS/MS of \(m/z\) 989.3 (Fig. 3D). The major fragment ion is the loss of 223 (cytosine + 112), indicating once again the lability of the glycosidic bond and being consistent with the finding above (Fig. 1C) that C* contains at least one positive charge. The fragment ions at \(m/z\) 572.1, 652.1, and 744.9 correspond to cleavages of the C*pA phosphodiester bond, yielding \(y_2\) and \(w_2\) fragment ions, and the ApU phosphodiester bond, yielding a d\textsubscript{2}H\textsubscript{2}O fragment ion, respectively (Fig. S2) (22). Together, these data establish the sequence of the trinucleoside as [C + 112]pApU. Thus, both RNase T1 mapping and RNase A/BAP mapping and sequencing confirm that C34 is modified in tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1} with a net modification mass of 112 u. A modified nucleoside of mass 355 (cytidine + 112) was, therefore, used as a search criterion during subsequent LC-MS/MS analysis of total nucleosides derived from tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}.

**Fig. 2.** Ribosome binding of *H. marismortui* tRNA\textsubscript{\textsuperscript{H-Ile}}\textsuperscript{1} and tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{2}. Template-dependent binding of purified \(^{1}\text{H}-\text{tRNA}\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}\) (A) and \(^{1}\text{H}-\text{tRNA}\textsubscript{\textsuperscript{H-Ile}}\textsubscript{2}\) (B) to ribosomes isolated from *H. marismortui*. Oligonucleotides used were AUGAUG (Δ), AUGAUA (□) and AUGAUC (○).

**Fig. 3.** LC-MS/MS analysis of oligonucleotides present in RNase A/BAP digests of *H. marismortui* tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}. (A) TIC of RNase A/BAP digestion of *H. marismortui* tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1}. (B) SIC of the anticodon-derived trinucleoside diphosphate 5'-C*pApU-3' at \(m/z\) 989.3. (C) Mass spectrum of oligonucleotides eluting at 4.5 min. (D) MS/MS analysis of the trinucleoside diphosphate at \(m/z\) 989.3 (22).

**Modified Nucleosides Present in *H. marismortui* tRNA\textsubscript{\textsuperscript{H-Ile}}.** The purified tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1} was digested completely to nucleosides and analyzed by LC-UV-MS (Fig. S3). All of the expected modified nucleosides in tRNA\textsubscript{\textsuperscript{H-Ile}}\textsubscript{1} were detected. The presence of 2'-deoxyribo-nucleosides (dG, dA) indicates minor contamination with DNA oligonucleotide used for the purification of the tRNA. SICs for the molecular ion, \(m/z\) 356 (MH\textsuperscript{+}), and the base ion, \(m/z\) 224 (BH\textsuperscript{2+}), were obtained demonstrating that a nucleoside of the expected mass (from the RNase mapping and sequencing data) was present in this sample (Fig. 4A and 4B). The mass difference of 132 u between the molecular ion and the base ion is consistent with location of the unknown modification on the base and not on the ribose of the nucleoside. The mass spectrum from the SIC is shown in Fig. 4C. The unknown modified cytidine C* coelutes with the leading edge of archaeosine (\(m/z\) 325.2).
Identification of C* as Agmatidine. Accurate mass analysis of the protonated modified cytidine yielded an elemental composition of C_{14}H_{26}N_{7}O_{4} (measured 356.2038, expected 356.2040, error 0.6 ppm). Accurate mass analysis of the protonated modified base yielded an elemental composition of C_{9}H_{18}N_{7} (measured 224.1617, expected 224.1618, error 0.4 ppm). The difference in elemental composition of the nucleoside and base is consistent with the absence of the C2-oxo group in this modified cytidine. Accurate mass analysis of the CID-MS/MS fragments of the base ion (Fig. 5A) yielded elemental compositions consistent with the loss of NH_{3} (m/z 207.1352) and CH_{5}N_{3} (m/z 165.1135) from the modified cytidine.

The absence of the C2-oxo group, the relatively high nitrogen content, and the fragmentation pattern suggested that the modification was attachment of agmatine to the C2 of cytidine. To confirm that loss of ammonia and a guanidino group are consistent with the presence of agmatine (23), a solution of agmatine was electrosprayed and subjected to MS/MS (Fig. 5B). The major fragment ions were indeed loss of NH_{3} and CH_{5}N_{3}. Thus, taken together, these data support the structure of C* as shown in Fig. 6.

The mass spectral data by themselves do not rule out an alternate structure for agmatidine in which the C2 oxygen of cytidine is replaced by hydrogen and agmatine is attached to the C5 or C6. However, this is most unlikely for several reasons: (i) such a modified nucleoside would make tRNA_{ile}^{AUA} base pair with AUG instead of AUA (Fig. 2); (ii) the tRNA_{ile}^{AUG} would be aminoacylated with methionine instead of isoleucine (8, 9); (iii) it would not explain the block in reverse transcriptase reaction by agmatidine (9); and (iv) it would require, in addition to a reductase to reduce the C2 oxygen, the biochemically difficult linkage of C5 or C6 of cytidine to the amino group of agmatine.

Agmatidine Is also Present in tRNA_{ile}^{AUA} from M. maripaludis and in Total tRNA from S. solfataricus. The LC/MS analysis of nucleosides from digests of partially purified tRNA_{ile}^{AUA} from M. maripaludis and total tRNA from S. solfataricus (Fig. S4) shows that agmatidine is also present in tRNAs from these archaenal organisms. Identification of agmatidine was confirmed by the SICs and mass spectral data. Thus, agmatidine is present in the AUA decoding tRNA_{ile}^{AUA} of euryarchaea and crenarchaea.

Discussion

A clear conclusion of the work described in here is that archaenal and bacteria use very similar strategies to read the isoleucine codon AUA without also reading the methionine codon AUG. Agmatidine (abbreviation C*), identified here to be in the anticodon wobble position of the AUA decoding archaeal tRNA_{ile}^{AUA}, is in many respects similar to lysidine found in the corresponding tRNA of bacteria. In both cases, the C2-oxo group of cytidine is replaced by a long side chain derived from either lysine, an amino acid, or by agmatine, a decarboxylated amino acid. The only difference is that in lysidine, the linkage involves the side...
tRNAIle

As shown by us here and elsewhere for the H. marismortui tRNAs with the anticodon CAU, which are annotated as methionine, sixty-six of these archaeal genomes also contain at least three the isoleucine codons AUU and AUC but not AUA. Interestingly, GAU annotated as an isoleucine tRNA. This tRNA can read genomes sequenced have only one tRNA with the anticodon euryarchaea and crenarchaea. This conclusion is supported by the different from that used by eukaryotes is quite surprising.

In view of these substantial similarities between archaea and bacteria-like in lacking the 5′-terminal cap and 3′-terminal PolyA sequences and in using Shine–Dalgarno-like sequences and leaderless mRNAs for translation initiation. However, the archaeal translation initiation factors, of which six have been identified, are a subset of eukaryotic initiation factors (27, 28). Translation elongation and termination in archaea and eukaryotes are similar based on patterns of antibiotic sensitivity of the archaeal translation system (29) and on the biochemical and genomic analysis of translation elongation and termination factors, for example, the presence of diphthamide, a modified histidine, in the elongation factor a/εEF2 (30), the presence of hypusine, a critical modified lysine, in the elongation factor a/εIF5A (31), and the presence of a single Class I transcription termination factor a/εRF1, which reads all three stop codons instead of two termination factors (RF1 and RF2) in bacteria. In view of these substantial similarities between archaea and eukaryotes, our finding that archaea use a strategy for translation of the AUA codon that is similar to that used by bacteria but different from that used by eukaryotes is quite surprising.

As for lysidine, the identification of agmatidine in the AUA specific archaeal tRNAIle^65 raises the question of how agmatidine base pairs specifically with A on the ribosome. Agmatidine has the potential for forming several tautomeric structures (Fig. 6), and it is not known which tautomeric form is present in the tRNA and whether this form undergoes changes in the context of the ribosome. Detailed structural studies will be necessary to understand the mechanism of specific recognition of the AUA codon by the corresponding archaeal and bacterial tRNAIle^65s.

The identification of agmatidine in tRNAIle from H. marismortui and M. maripaludis and in total tRNA from S. solfataricus suggests that agmatidine is present in the AUA decoding tRNAIle^65 of euryarchaea and crenarchaea. This conclusion is supported by the finding that all sixty-six of the euryarchaeal and crenarchaeal genomes sequenced have only one tRNA with the anticodon GAU annotated as an isoleucine tRNA. This tRNA can read the isoleucine codons AUU and UAC but not UAA. Interestingly, all sixty-six of these archaeal genomes also contain at least three tRNAs with the anticodon CAU, which are annotated as methionine tRNA (KEGG database; http://www.genome.jp/kegg/).

As shown by us here and elsewhere for the H. marismortui tRNAIle^65 (9) and for the M. maripaludis tRNAIle^65, one of these putative methionine tRNAs is actually the AUA decoding isoleucine tRNA. Therefore, it is likely that in the rest of euryarchaea and crenarchaea also, the anticodon wobble nucleoside C of one of the putative methionine tRNAs is modified to agmatidine to generate a tRNA, which is aminoacylated with isoleucine and which reads the isoleucine codon AUA. It is interesting to note that in contrast to H. marismortui, in which AUA is a very rare codon (3.6/1000 codons), in S. solfataricus it is the most frequent codon (49.4/1000 codons) (http://www.kazusa.or.jp/codon/). Thus, the use of agmatidine in tRNAIle^65 does not depend on the frequency of utilization of the AUA codon in mRNA.

In contrast to euryarchaea and crenarchaea, in the only available genome sequences of a nanoarchaeum and a korarchaeum, two tRNA genes are annotated as coding for isoleucine tRNAs (32–34). One of them would have the anticodon GAU, whereas the other one would have the anticodon UAU, which is likely modified to VΨP as in eukaryotic tRNAIle^65 (8). This latter tRNA would have the potential for reading the AUA codon although how it would be prevented from also reading the methionine codon AUG is unknown. The genome sequences could suggest that nanoarchaea and korarchaea use a eukaryote-like strategy to read the isoleucine codon AUA. However, before reaching this conclusion, it is important to sequence additional nanoarchaeal and korarchaeal genomes to determine whether other isolates of nanoarchaea and korarchaea also have a tRNAIle with the anticodon UAU (possibly VΨP) or have C^65A (C^65 = agmatidine) as in euryarchaea and crenarchaea.

The very similar structures of agmatidine and lysidine suggest a simple pathway for biosynthesis of agmatidine, similar to that of lysidine, involving activation of the C2 of cytidine by adenyllylation followed by nucelophilic attack on the activated C2 by the primary amino group of agmatine (35, 36). Agmatine is derived by decarboxylation of arginine and occurs naturally as one of the intermediates in polyamine biosynthesis. In the thermophilic euryarchaeon T. kodakarenensis, agmatine is essential for polyamine biosynthesis (12). Agmatine is also an important neuronal modulator, which binds to various receptors such as the α2-adrenergic receptor, NMDA receptors, etc. (10, 11). It has also been suggested to be a neurotransmitter, although that remains to be proven (37). In view of the critical role of agmatidine in archaeal AUA decoding tRNAIle^65 function, the potential involvement of agmatine in agmatidine biosynthesis suggests that agmatine may also be essential in euryarchaea and crenarchaea for the reading of the genetic code.

Although agmatidine and lysidine are very similar modified nucleosides, both derived from cytidine, in spite of extensive bioinformatics searches, no homolog of TiIS, the enzyme responsible for biosynthesis of lysidine in bacteria, has been detected in archaea (38). This raises the interesting question of how similar the enzyme responsible for biosynthesis of agmatidine is to TiIS and whether the two enzymes are products of convergent evolution. Clearly, an analysis of the evolutionary history of these enzymes would be of much interest.

Materials and Methods

Purification of tRNAIle^65 and tRNAIle from H. marismortui. H. marismortui ATCC 43,049 was kindly provided by Peter Moore (Yale University). General manipulations of H. marismortui were performed according to standard procedures (39). Cells were grown at 37 ºC, harvested in late-log phase, and crude RNA was isolated by acid guanidinium thiocyanate-phenol-chloroform extraction (40). Ribosomal RNA was removed by precipitation with 1 M NaCl yielding total tRNA. The yield of total tRNA was approximately 1,000 A260 units from a 15 L-culture. Purification of tRNAs was carried out essentially as described by Suzuki and Suzuki (41). Details are given in SI Text.

Binding of ^3H-Ile-tRNAIle^65 to Ribosomes. H. marismortui ribosomes were isolated essentially as described (42) with slight modifications. Pelleted cells from a 2 L-culture were suspended in 10 mM Tris-HCl pH 7.5, 100 mM Mg(OAc)2, and 3.4 M KCl (buffer 1) and lysed by French Press (Constant Cell Disruption System), and the lysate was cleared of cell debris by centrifugation first at 27,000 x g for 15 min and then at 65,000 x g for
mRNA (Griffiths (43) with modifications. Details are given in Mandal et al. PNAS (to D.Söll), snake venom phosphodiesterase I (Worthington Biochemical Corporation) and antarctic phosphatase (New England Biolabs) as described by Crain (44).

Digestion of tRNA to Nucleosides. Purified H. marismortui tRNA\(^{\text{N}}\) (0.2 A\(_{260}\) unit), purified M. maripaludis tRNA\(^{\text{t}}\) (0.6 A\(_{260}\) unit), H. marismortui total tRNA (0.6 A\(_{260}\) unit) and S. solfataricus total tRNA (1.0 A\(_{260}\) unit) were digested to nucleosides using nuclease P1 (Sigma), snake venom phosphodiesterase I (Worthington Biochemical Corporation) and antarctic phosphatase (New England Biolabs) as described by Crain (44).


