Sequences of halobacterial tRNAs and the paucity of U in the first position of their anticodons

Xiang-Rong Gu*, K.Nicoghosian*, R.J.Cedergren* and J.Tze-Fei Wong†

*Département de Biochimie, Université de Montréal, Montréal, Québec H3C 3J4, and †Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Received 27 May 1983; Revised and Accepted 15 July 1983

ABSTRACT

The sequence of three tRNAs from Halobacterium cutirubrum have been determined. The sequences of tRNA<sub>Val</sub> and tRNA<sub>Val</sub> differ by only one nucleotide which is in the 5' terminal anticodon position. These tRNAs as well as that of tRNA<sub>Ala</sub> are compared to other known halobacterial tRNAs. An observed paucity (or absence) of U in the first anticodon position is unique to archaebacterial tRNAs and may be indicative of unusual decoding properties of these organisms.

INTRODUCTION

The identification of archaebacteria as a grouping of life forms, distinct from eukaryotes and eubacteria has been one of the major contributions to evolutionary biology in recent years (1,2). In order to bring new data to bear on this hypothesis of Woese and co-workers, we initiated structural studies on the tRNAs of halobacteria, a major phylum of archaebacteria. As previously discussed (3), tRNA sequences can be useful indicators of evolutionary relationships between organisms. Moreover, the intimate role played by tRNA during the translation of mRNA often permits the exploitation of tRNA structure data in the inference of aspects of the translation mechanism. This was illustrated recently by the discovery that some mitochondrial tRNAs decode mRNA by a novel mechanism (4-6), which is not predicted by the Wobble hypothesis (7). In this study, we present three sequences from Halobacterium cutirubrum: tRNA<sub>Ala</sub>, tRNA<sub>Val</sub> and tRNA<sub>Val</sub>. Analysis of these and other known tRNA sequences in halophiles leads to the conclusion that message decoding in these organisms may differ substantially from that in eukaryotes and eubacteria.

MATERIALS AND METHODS

RNA was isolated from Halobacterium cutirubrum by phenol extraction (8). The low molecular weight RNA fraction was obtained by DEAE cellulose column
chromatography. Unique RNA species were isolated after two or three electrophoretic dimensions of varying temperatures (5°C or 20°C) and/or urea concentrations (4M or 7M) (9). Figure 1 shows the band pattern of the crude tRNA sample obtained in the first electrophoretic dimension. tRNA sequencing was carried out by chemical or enzymatic hydrolysis of 3' or 5' labelled material (9,10). In addition, major portions of each tRNA sequence were verified by the limited hydrolytic technique (11). All positions of the tRNAs reported here were determined by at least two sequencing techniques.

RESULTS

The sequences of the three H. cutirubrum tRNAs: tRNA$^{\text{Ala}}_{\text{GCC}}$, tRNA$^{\text{Val}}_{\text{GAC}}$ and tRNA$^{\text{Val}}_{\text{CAC}}$ are shown in Figure 2. Modified nucleotides were isolated as the 5'-terminally labelled nucleotides in the hydrolytic degradation technique (11) and identified by thin layer chromatography (12). The modified nucleotide in position 54 (numbering by the Cold Spring Harbor Convention) was particularly difficult to identify, since it behaves like ribothymidylate (T) in many solvent systems. In the two-dimensional system of Nishimura (12), however, this nucleotide was identified as the newly isolated 1-methyl-pseudouridine (13). The identification of the modified nucleotide, 1-methyl-inosine in position 57 is tenuous and based on the presence of this nucleotide at position 57 in closely related tRNAs. The dinucleotide pCmpx has an rf value of 0.38 (with pU as 1.0) in isopropanol:HCl:water (70:15:15).

The H. cutirubrum tRNA$^{\text{Ala}}_{\text{GCC}}$ (Fig. 2a) shares with all other alanine tRNAs the feature of contiguous G-C, G-C and G-U base pairs in the terminal positions of the acceptor stem. Significantly, no tRNA of any other amino acid family contains this feature (14). The two valine tRNAs, shown in Fig. 2b, differ by only one nucleotide, that of the first anticodon position. This result was quite unexpected, since the tRNAs were well separated on gels (see Fig. 1) and both were major components of their gel band. Subsequently, in order to isolate and identify other potential tRNA$^{\text{Val}}$ isoacceptors, crude tRNA was aminoacylated with $[^{14}\text{C}]$valine using a heterologous yeast synthetase preparation, since these synthetases have been shown to aminoacylate H. cutirubrum tRNA even better than some eukaryotic tRNA (15). The RPC-5 pattern of the resulting mixture is shown in Figure 3. The two radioactive fractions peak I and II were identified as tRNA$^{\text{Val}}_{\text{GAC}}$ and tRNA$^{\text{Val}}_{\text{CAC}}$ respectively, since their electrophoretic mobility corresponded to the gel bands from which the tRNAs were initially isolated (see Fig. 1, tracks c and d). These data are only strongly suggestive that no other valine isoacceptor tRNA exists.
Figure 1. Gel separations of total and valine tRNAs. Deacylated tRNA samples (0.1 O.D. 260 unit) were labelled with $[^{32}P]Cp$ (2000 Ci/m mole and T₄ RNA ligase (9). These samples were run on a 10% polyacrylamide gel at 800 volts and 5° until the xylene cyanol marker moved 40 cm on the gel. X-ray film exposure was for 15 min. Track a is a labelled 16S RNA control. Track b shows the pattern obtained from crude H. cutirubrum tRNA. The two bands 5 and 7 were extracted and run on a second 20% polyacrylamide 7M urea gel at room temperature. Band 7 gave rise to two major bands: tRNA$^{Ala}$, the faster moving band, and tRNA$^{Val}_{CAC}$, both reported in this paper. Band 5 in the second dimension gave rise to one major band, tRNA$^{Val}_{CAC}$ and 4 minor bands. The two lanes (c and d) are the fractions II and I respectively, isolated from the RPC-5 chromatography shown in Figure 3. These two tRNA fractions had been precipitated using 16S RNA as carrier before labelling.
Figure 2. The cloverleaf structure of tRNAs. a) the tRNA$^{\text{Ala}}_{\text{CGC}}$ is shown in the cloverleaf conformation. b) the tRNA$^{\text{Val}}_{\text{GAC}}$ sequence. The nucleotides linked by arrow to the cloverleaf are those of tRNA$^{\text{Val}}_{\text{GAC}}$ which differ from the GAC isoacceptor tRNA. The identity of the modification at G$_9$ is tentative. The m$^4$C at position 5 is the unique occurrence of this modification in that position.
Figure 3. RPC-5 chromatography of *H. cutirubrum* tRNA. A sample of 5 O.D. units of total *H. cutirubrum* tRNA was aminoacylated with [14C]-valine (50 mCi/mmole) in 0.5M Hepes (pH 7.5), 0.05M MgCl₂ and 0.0125M ATP using 413 units of baker's yeast aminoacyl-tRNA synthetases (Sigma) for 20 min at 37°. The reaction was stopped with 0.1 volume of 2M NaOAc (pH 4.5), precipitated with ethanol and dried under reduced pressure. Total charging was 160 pmoles of valine per O.D₂₆₀ unit. The tRNA was dissolved in 3 ml of 0.01M NaOAc (pH 4.5) containing 0.01M MgCl₂ and 0.1M NaCl and applied to a 0.7 x 23 cm RPC-5 column previously equilibrated in the acetate buffer. Elution was performed by 150 ml of a 0.3M to 1.2M NaCl gradient in the acetate buffer. Fractions of 1 ml were taken and the radioactivity in 25 μl was counted in a scintillation cocktail. Peaks I and II were precipitated using 16S RNA as a carrier and deacylated in 100 μl of 0.2M Tris-acetate buffer (pH 9.0) for 30 min at 37°. These samples were then treated as in Figure 1. The first large peak elutes in the void volume of the 0.1 M salt loading buffer.

since it could be argued that a particular isoacceptor is poorly recognized by the heterologous enzyme preparation.

The finding of two isoaccepting tRNAs differing only in the anticodon is quite rare. Among sequenced tRNA only mammalian proline (16) and, perhaps significantly, mammalian valine (17,18) tRNA isoacceptors exhibit this feature. Although it is unlikely that this observation reveals a close phylogenetic relationship between mammals and halobacteria, a similarity in genome structure, particularly repeated DNA arrangements (19,20), could be instrumental in directing gene conversion between two originally distinct tRNA genes. Indeed, evidence is now accumulating which demonstrates that two
tRNA gene sequences can be homogenized by inter- or intra-chromosomal recombination (W. McClain, personal communication and Ref. 21).

There are at present 21 archaeabacterial tRNA sequences, which we analyze here by pairwise comparison (22): beside the three sequences from this report, the tRNA list includes 13 from Halobacter volcanii (Gupta and Woese, personal communication and Ref. 14), the initiator tRNAs from Sulfolobus acidocaldarius, Halococcus morrhuae and Thermoplasma acidophilum (23), the initiator tRNA from H. cutirubrum (S. Bayley and U. L. RajBhandary, personal communication) and tRNA$_{m}$ from T. acidophilum (24). The tRNA sequences of H. volcanii and H. cutirubrum show a high degree of similarity. Their initiator initiator tRNAs and tRNA$_{Val}$ differ only at three and five positions respectively. A slightly more distant relationship holds between this pair and H. morrhuae. Considerably more distant, although still significant, are the relationships with Thermoplasma and Sulfolobus; the latter being the most distant from the three halobacteria and only slightly closer to Thermoplasma. In tree form, this topology is entirely consistent with that derived by Stackebrandt and Woese (25) based on 16S RNA catalogue data.

DISCUSSION

The finding of only two tRNA$_{Val}$ isoacceptors with GAC and CAC anticodons bear further consideration, since it raises the question as to the generality and the reason for the lack or paucity of A (as inosine) and U in the first position of the anticodon. Examination of the 21 sequences shows none with an identified A or modified A, a U or modified U at the first anticodon position. Furthermore, that position is occupied by a C or a modified C in 13 of the 21 sequences. Excluding the six tRNAs for methionine which are not expected to have A or U at that position, there remains 15 tRNA sequences uniformly excluding A or U. While not a large sample, this nonetheless provides important support to the indication from the H. cutirubrum tRNA$_{Val}$ sequences of an absence or paucity of A and U. Comparable values for eukaryotes are 17 A's and 17 U's from 83 sequences; for eubacteria: 3 A's and 32 U's for 83 sequences and from fungal mitochondria: 2 A's and 32 U's from 53 sequences (26). Thus in both eubacteria and eukaryotes, all four nucleotides can be used in the first anticodon position, even though their distribution is not random: for example, A is definitely underused and among tRNA families for amino acids having 3' terminal purine codons only Ala, Ile and Thr exclude C, while U is excluded in Met and Trp tRNAs (the single codon amino acids). Halobacterial tRNAs, therefore, distinguish themselves by
underusing first position U in the anticodon. Figure 4 summarizes the
distribution of nucleosides at each position of the cloverleaf for the 21
archaeabacterial sequences (cf. Ref. 26). Other positions show unusual base
distributions, but these will not be discussed in this paper.

In support of the above discussion, the frequency of codon use in the
bacteriorhodopsin gene in _H. halobium_ mirrors anticodon structure in that the
valine codon distribution is 3 GUU, 8 GUC, 3 GUA and 9 GUG (27). Therefore
there are close to three times more G + C than A + U at the third valine
codon position. Since the G + C content for this gene is 61.5%, that for
total _H. halobium_ DNA is 67% (28), the G + C enrichment is not enough to
account for the codon usage data. Similar skewed distributions at the third
codon position is found for other amino acids using 4-fold degenerate codons:

On the other hand, the more than trace amounts of the A + U type codons
clearly require some sort of translation. At least in the case of valine in
H. cutirubrum, a strict two-out-of-three reading mechanism (29) in translation can probably be excluded due to the existence of isoacceptor tRNAs (30). The apparent paucity or absence of A and U at the first anticodon position virtually requires consideration of the decoding of A + U type codons in halobacteria by G or C in the anticodon. Whereas G is known to decode C and U terminating codons (7), the decoding of A in the codon by C or G in the anticodon must be extremely rare. A recent article has dealt with the possibility of C-A pairing in codon-anticodon interaction (31) however, and in this light, C reading A of the message might obtain for isoleucine tRNA of chloroplasts (32) and the T₄ bacteriophage (33). Likewise evidence for C-A pairing emanates from in vitro studies of anticodon-anticodon interactions (34).

With the exception of organelle and T₅ phage, almost no sequenced tRNA contains an unmodified A or U in the first anticodon position. This has led Nishimura (12) to suggest that unmodified A or U at that position might be lethal to cells by causing message misreading. To minimize and avoid this hazard, eubacteria and eukaryotes could have developed the variety of modifications of A and U seen in sequence data. The halobacteria, by limiting the synthesis of tRNA sequences carrying A or U at that position, could achieve protection against unmodifieds A or U by simple avoidance of the parent nucleotides. The nature of the postulated toxicity of unmodified A or U remains to be elucidated for any organism however. Another attractive, alternate hypothesis involves the energetic needs of the codon-anticodon interaction during translation. Given the generally hostile environment which these organisms frequent, the higher G and C stacking energies (35) may well help stabilize the tRNA-message interaction.

A range of characteristics already delineate the archaeabacteria from eubacteria and eukaryotes (1,2,19,23,36,37). A unique strategy of anticodon usage would confirm at a particularly fundamental level their classification as a third form of life. The evolutionary rationale and consequence of their strategy, when more fully understood, will hopefully throw light on the intricacy of the modern and ancient genetic apparatus.

ACKNOWLEDGEMENTS

The authors express their gratitude to the National Science and Engineering Research Council and the Medical Research Council of Canada for financial support, to N. Beauchemin and P. M. Bronskill for technical aid, and to Dr. H. Grosjean for assistance in compiling Figure 4. We are
particularly indebted to Drs. C. R. Woese and U. L. RajBhandary for communicating sequences done in their laboratories prior to publication.

REFERENCES
27. Dunn, R., McCoy, J., Simsek, M., Majumdar, A., Chang, S. H.,

NOTE ADDED: Preliminary results suggest that \( tRNA_{Val}^{\text{CAC}} \) responds well to GUG triplets, but poorly to GUA triplets. It remains to be seen whether low binding of the GUA triplet to mixed RNA results from the poor binding to \( tRNA_{Val}^{\text{CAC}} \) or to a different minor tRNA.