Halobacterium cutirubrum tRNA sequences

K. Nicoghosian, Xiang-Rong Gu+ and Robert Cedergren*

Département de Biochimie, Université de Montréal, Montréal, Québec H3C 3J7, Canada

Received 30 September 1985

The primary structures of nine *Halobacterium cutirubrum* tRNAs are presented. These tRNAs are compared with other archaebacterial tRNAs and in particular with their *H. volcanii* counterparts. Striking similarities are observed among isoacceptor tRNA species; observed differences between a given tRNA pair of the two halophiles having the same anticodon vary from 0 to 14 nucleotides. Particular features including modified nucleotides of each of the nine tRNAs are indicated.

tRNA sequence *Evolution* *Halophile* *Methanogen*

1. INTRODUCTION

A major effort in our laboratory over the past few years has involved the structural determination of tRNAs and the interpretation of these structures within an evolutionary framework. This strategy has been applied recently to the study of the tRNA structures from archaebacteria by ourselves and others [1–5]. Archaebacteria have a growing following among scientists, ever since Woese and co-workers [6] defined this kingdom as distinct from other cell types, i.e. eukaryotes and other bacteria (eubacteria). Among members of this kingdom which have been subjected to tRNA structural scrutiny are *Halobacterium cutirubrum, H. volcanii, Halococcus morrhuae, Methanobacterium thermoautotrophicum, Methanococcus vanielli, Sulfolobus acidocaldarius*, and *Thermoplasma acidophilum* [1–5].

Here we present 8 new sequences from *H. cutirubrum*: $tRNA^{Arg}_{GCG}$, $tRNA^{Asn}_{GUU}$, $tRNA^{Gly}_{GCG}$, $tRNA^{Gln}_{CUG}$, $tRNA^{Thr}_{GGU}$, $tRNA^{Ser}_{ac^4CGA}$, $tRNA^{His}_{GUG}$, $tRNA^{Met}_{CAU}$, along with a correction of an earlier published sequence ($tRNA^{Val}_{CAC}$). These sequences are subsequently compared to other archaebacterial tRNAs, and in particular to tRNAs from *H. volcanii* [3].

2. MATERIALS AND METHODS

[γ-^{32}P]ATP was purchased from Amersham; [^{32}P]pCp was synthesized in our laboratory from [γ-^{32}P]ATP with polynucleotide kinase. Yeast hexokinase and polynucleotide kinase were from Boehringer Mannheim, Canada, while T_4 RNA ligase (RNase-free) was from PL-Biochemicals. Sankyo brand ribonucleases T_1 and T_2 were purchased from Calbiochem. Ribonucleases U_2, *B. cereus*, Phy-M and nuclease P_1 were from PL-Biochemicals. PEI-cellulose thin-layer chromatography plates were purchased from Brinkman and cellulose thin-layer plates from Eastman.

2.1. *Purification of tRNAs*

Bulk tRNA was extracted from *H. cutirubrum* cells as described [2] and subjected to electrophoresis for the isolation of pure tRNA species. Individual bands were eluted and re-electrophoresed sequentially in 10 and 15% polyacrylamide-7 M urea gels at 4°C and room temperature, respectively. In most cases these steps were sufficient to obtain pure individual tRNA species. In cases where tRNAs were not completely

* To whom correspondence should be addressed
+ On leave from the Biochemistry Institute of Shanghai, Academia Sinica, Shanghai, People's Republic of China

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

purified, a third 20%-7 M urea polyacrylamide gel electrophoresis was performed. The electrophoresis buffer was 50 mM Tris-boric acid, pH 8.3, 20 mM EDTA (TBE buffer) in all separations.

2.2. *Preparation and sequencing of ^{32}P-labelled tRNAs*

End labelling of the tRNAs was carried out as described by Peattie (3′-end) [7] and Donis-Keller et al. (5′-end) [8]. The end-labelled tRNAs were electrophoresed for 15 h at 1500 V in thin (0.8 mm) 15% polyacrylamide-7 M urea gels in 50 mM TBE buffer (50 mM Tris-boric acid, 1 mM EDTA at pH 8.3). Under these conditions one main band for each tRNA was observed on the gel. tRNA sequencing was carried out by enzymatic hydrolysis of the labelled tRNA with RNases T_1, U_2, *B. cereus* and Phy-M. The chemical degradation method of 3′-labelled tRNAs was essentially as described by Peattie [7] with the following changes in experimental conditions: the G reaction was for 1 min and the A reaction for 5 min, both at 90°C, and the C and U reactions for 5 min on ice; the aniline cleavage reaction was in the dark at 60°C, for 5 min. A variation of the technique described by Gupta and Randerath [9] was also used to identify modified nucleoside diphosphates: 1 μg tRNA was partially hydrolyzed in 5 μl deionized formamide at 100°C for 3–4 min. The resulting fragments were labelled with 100–200 μCi of [γ-^{32}P]ATP (3000 Ci/mmol) and polynucleotide kinase and separated on 15 and 20% polyacrylamide-7 M urea gels at 1500 V and room temperature, in the above TBE buffer. Individual bands were excised from the gel, extracted with 300 μl of 300 mM NaCl, 0.1% SDS and precipitated with ethanol in the presence of 5 μg carrier tRNA. Digestion of each fragment was carried out overnight with 0.5 U RNase T_2, 1.0 U RNase T_1 and 0.5 μg RNase A in 100 mM Na acetate, pH 4.5, at 37°C. 5′-labelled nucleoside 3′,5′-diphosphates were identified by thin-layer chromatography on PEI-cellulose sheets in 0.55 M ammonium sulfate at 4°C.

2.3. *Partial hydrolysis of tRNA*

To aid in the identification of nucleotides in strongly compacted regions of sequencing gels certain tRNAs were subjected to partial cleavage by the mung bean nuclease, by first renaturing 1 μg unlabelled tRNA in 9 μl of 0.3 M Na acetate, 0.05 M NaCl, 1 mM $ZnSO_4$, 5 mM $MgCl_2$ and 5% glycerol for 3 min at 60°C. The solution was then equilibrated at 37°C for a few minutes and 0.8 U mung bean nuclease was added. Labelling with [^{32}P]pCp followed, and the resulting products were electrophoresed on a 15% polyacrylamide-7 M urea gel overnight at 800 V and room temperature.

2.4. *Modified nucleotides*

5′-Nucleoside monophosphates were identified by nuclease P_1 digestion of the fragments resulting from formamide hydrolysis, and 2-dimensional thin-layer chromatography on cellulose plates. The first dimension was in isobutyric acid/conc. NH_4OH/H_2O (66:1:33, by vol.) and the second, 0.1 M sodium phosphate, pH 6.8/ammonium sulfate/*n*-propanol (100:60:2, v/w/v).

Modified bases were also detected by 5′-labelling of a complete T_2 digest of the tRNA with [γ-^{32}P]ATP then treating the labelled pNp bases with nuclease P_1 and performing 2-dimensional thin-layer chromatography on cellulose plates in the 2 systems described.

3. RESULTS AND DISCUSSION

Due to the ambiguities in sequencing gels (mainly compression of bands), sequencing methods were applied to both full-length molecules and mung bean nuclease generated fragments. Standard sequencing techniques were used except for the modification given in section 2. The sequences determined here are shown in fig.1; they include $tRNA^{Arg}_{GCG}$, $tRNA^{Asn}_{GUU}$, $tRNA^{Gly}_{GCC}$, $tRNA^{Gln}_{CUG}$, $tRNA^{Thr}_{GGU}$, $tRNA^{Ser}_{ac^4CGA}$, $tRNA^{His}_{GUG}$, $tRNA^{Met}_{CAU}$ and $tRNA^{Val}_{CAC}$. These sequences share the characteristics of other known halobacterial sequences: (i) the lack of T, m^7G and D; (ii) the presence of the sequence N_1UCmN_2A in the common loop, where N_1 is a derivative of U, and N_2 a derivative of I (we adopt the nomenclature of Gupta [3] in that the former T-loop is called the common loop and the D-loop is changed to the GG-loop); (iii) the high G-C content ranging from 55 to 64%; (iv) the occasional modification of G_{15}.

3.1. *$tRNA^{Arg}_{GCG}$*

This tRNA is 76 nucleotides long and differs from the cognate tRNA in *H. volcanii* at 6 posi-

tions. The distribution and content of the modified bases are identical in both cases, with the exception of the m^5C in the common arm. Its position is difficult to determine because of compacting in the extra loop region, therefore it is doubtful that the observed difference is significant.

3.2. $tRNA^{Asn}_{GUU}$

This tRNA (of 76 nucleotides) has 12 differences with its cognate tRNA in *H. volcanii*. The mispair, G_{10}-U_{69}, present in *H. volcanii* is a normal pair G_{10}-C_{69} in *H. cutirubrum*. Most significant, however, are those differences present in the common and anticodon loops. Also, we note the inversion of C_{30}-G_{40} in *H. volcanii* to G_{30}-C_{40} in *H. cutirubrum*.

3.3. $tRNA^{Gly}_{GCC}$

This tRNA is 74 nucleotides long and is identical to the isocoding tRNA sequence (1A) from *H. volcanii* [3], and therefore exhibits the same double base pair inversion when compared with $tRNA^{Gly}_{1B}$ in *H. volcanii* (C_4-G_{69} and U_5-A_{68} inverted to U_4-A_{69} and C_5-G_{68}). Comparison with the tRNA from *M. thermoautotrophicum* (a methanogen) however, shows 23 differences [1], a value that could be related to the divergence of these organisms.

3.4. $tRNA^{Gln}_{CUG}$

The glutamine tRNA is composed of 76 nucleotides and differs from that of *H. volcanii* in 12 positions. Interestingly, 8 of these differences are due to inverted base pairs, i.e., A_{27}, A_{28}, U_{42} and U_{43} in *H. cutirubrum* are U_{27}, U_{28}, A_{42} and A_{43} in *H. volcanii*. In the same way, G_{49}, G_{50}, C_{64} and C_{65} in *H. cutirubrum* are C_{49}, C_{50}, G_{64} and G_{65} in *H. volcanii*. It might be relevant that both inversions take place at the end of stems which have other symmetry characteristics. This tRNA contains an unusual nucleotide modification at position 9 (N) which is cut by Phy-M and formamide, but is resistant to T_1, U_2 and *B. cereus* ribonucleases.

3.5. $tRNA^{His}_{GUG}$

This tRNA of 77 nucleotides differs from that of *H. volcanii* by 3 nucleotides. As in all other histidine tRNAs an extra nucleotide is present at the 5′-position which is complementary to the fourth nucleotide from the 3′-terminus. The U_7-A_{66} present in *H. volcanii* is C_7-G_{66} in *H. cutirubrum*, these positions are commonly A_7-U_{66} in $tRNA^{His}$. This tRNA also contains the noted modification at position 9 (N).

3.6. $tRNA^{Met}_{CAU}$

The sequence of 78 nucleotides determined here differs in 2 positions from the commonly known, unpublished sequence of RajBhandary and Bayley. This sequence is identical with one exception (U_{17} is C_{17} in *H. volcanii*) to that from *H. volcanii*. Both tRNAs contain an unknown modified G_{15} (G*). This modification is resistant to both formamide and T_1 hydrolysis.

3.7. $tRNA^{Thr}_{GGU}$

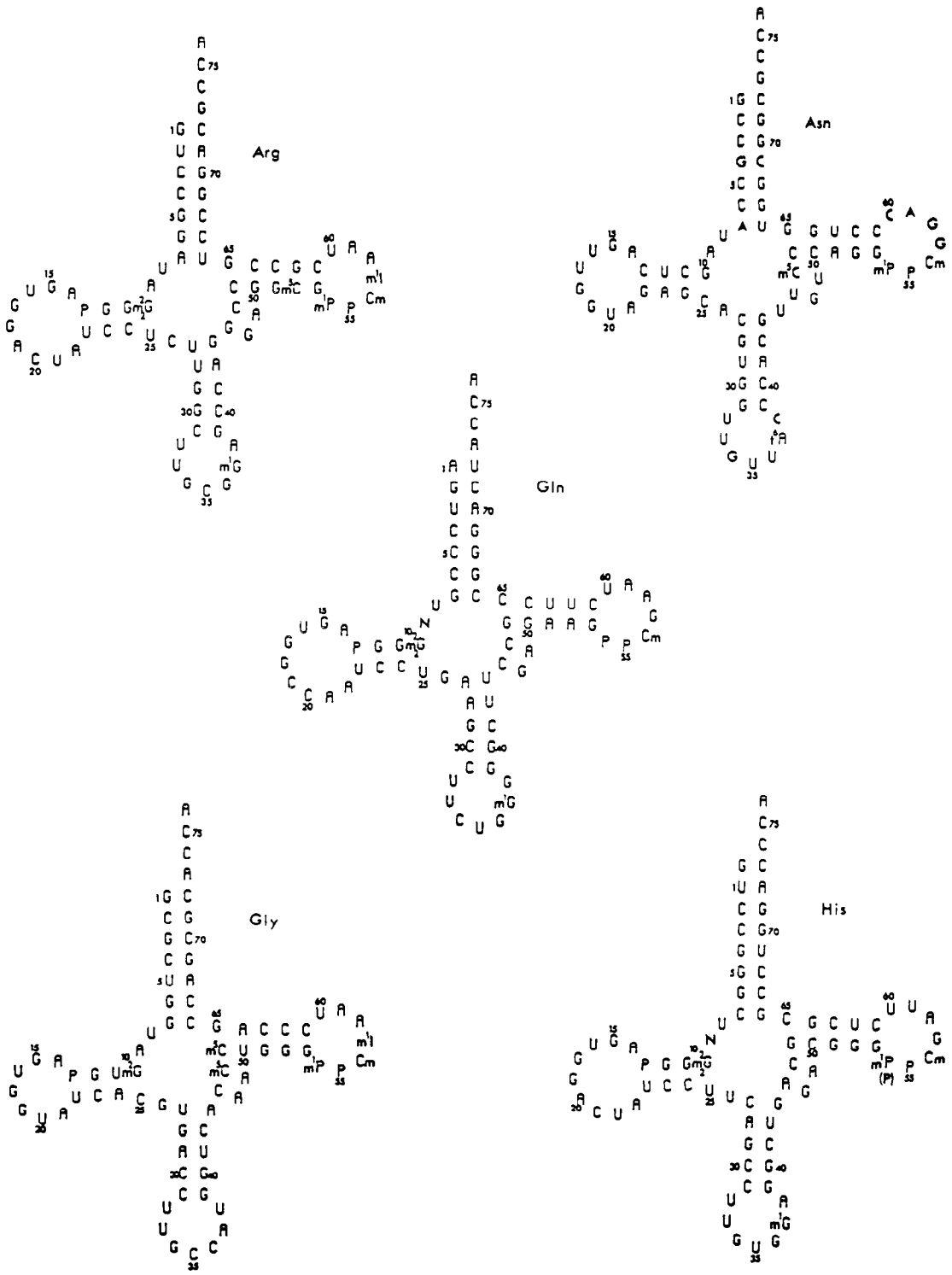
This tRNA of 75 nucleotides has one surprising characteristic that of an A_{53}-U_{61} base pair instead of the canonical G-C at these positions. This region is very clear on the sequence gel, so that it is doubtful that the given sequence is in error. This feature is shared by some mitochondrial and one bacteriophage sequence [10]. In effect these positions represent the only differences with the sequence from *H. volcanii*. Again the positions of the m^5C are in variance with the identical regions in *H. volcanii*.

3.8. $tRNA^{Ser}_{ac^4CGA)}$

This tRNA has 88 nucleotides and therefore belongs to the class of tRNAs having a long extra arm. It differs in 21 positions (including the extra arm) with a Ser tRNA from *H. volcanii* having an NGA anticodon. This large number of differences may be explained by the fact that the 2 tRNAs do not read the same codon. Also the extra arms, which have identical lengths, contain a large number of the differences between the 2 tRNAs. This tRNA contains the modified G_{15} (G*).

3.9. $tRNA^{Val}_{CAC}$

This sequence is a corrected version of one previously published [2]. The 2 isoacceptors of *H. cutirubrum* differ in 9 positions. Similarly this sequence differs in 6 positions with the cognate sequence in *H. volcanii*. It lacks the modification at G_{15} (G*) that is present in *H. volcanii*. After considerable effort we have not been able to isolate or even show the existence of a third tRNA with the


Arg
Asn
Gln
Gly
His

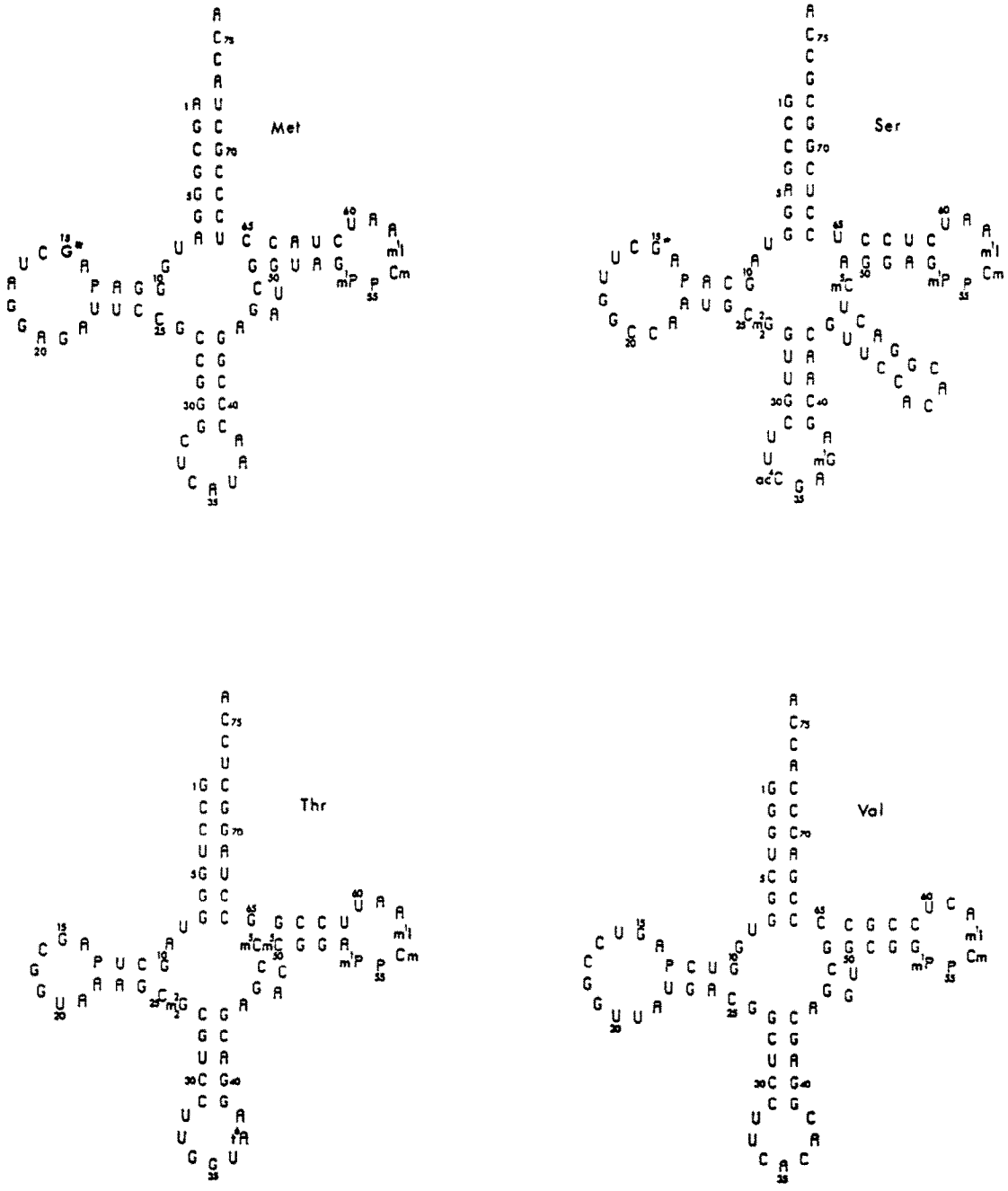


Fig.1. tRNA sequences reported in this paper, in cloverleaf-folded patterns. m_2^2G, N^2,N^2-dimethyl-G; G*, modification of G that is resistant to both formamide and T_1 hydrolysis; N, an unknown modification that is cut by the ribonuclease from Phy-M and formamide but is resistant to T_1, U_2 and *B. cereus* ribonucleases; m^2G, N^2-methyl-G; m^1G, 1-methyl-G; t^6A, *N*-[(9-β-D-ribofuranosylpurin-6-yl)carbamoyl]threonine; m^5C, 5-methyl-C; m^1P ($m^1\Psi$), 1-methylpseudouridine; C_m, 2′-*O*-methylcytidine; m^1I, 1-methylinosine; ac^4C, N^4-acetyl-C.

UAC anticodon, which should be necessary to read GUA codons (see also [3]).

In [2] we suggested that A and U were avoided in the so-called wobble position of the anticodon of halophilic tRNAs. That hypothesis was supported by a compilation of halobacterial sequences that were known at that time. Since then in a major work, Gupta [3] has published sequences which have U in this position. In spite of this fact, we and Gupta [3] have been thwarted in efforts to find a valine tRNA containing a UAC codon; therefore even if it does exist, it is definitely underrepresented as are some other U_{34}-containing tRNAs in halophiles [3]. These observations are reinforced by the sequence data from 2 halophile genes [11]. The use of A or U terminating codons is distinctly low for most amino acids. In the case of valine, we must conclude that until a tRNA is found which can decode the GUA codon, another tRNA decodes this codon by 'mispairing' [12]. This codon-anticodon recognition may be implicated in a translational regulation mechanism as has been suggested in other cases of minor species of tRNAs [13].

Comparison of our data with other tRNA data from archaebacteria demonstrates the large evolutionary distance between methanogens and halobacteria, demonstrated particularly by the glycine tRNA family. On the other hand, the 2 halophiles which have been extensively studied are seemingly very much closer. However, several comparisons of isocoding tRNAs from *H. cutirubrum* and *H. volcanii* are more distant (i.e. $tRNA^{Gln}$ and $tRNA^{Asn}$). At present it cannot be determined whether these differences are due to the comparison of products of different genes duplicated long before the divergence of the 2 orgnisms from a common ancestor or whether these results reflect a more rapid evolution of these tRNA genes. Finally, comparison of some halobacterial sequences has shown a number of base-paired positions which have been inverted, i.e. $tRNA^{His}$ and $tRNA^{Gln}$. In addition the data of Gupta point to inversions at position 4,69 and 5,68 of $tRNA_{1A}^{Gly}$ and $tRNA_{1B}^{Gly}$. Whether this observation is in fact related to some biological process is unknown at the present.

ACKNOWLEDGEMENTS

This work was supported by the Natural Science and Engineering Research Council of Canada. We thank J.T.F. Wong for crude tRNA preparations.

REFERENCES

[1] Gu, Xiang-Rong, Nicoghosian, K. and Cedergren, R.J. (1984) FEBS Lett. 176, 462–466.
[2] Gu, Xiang-Rong, Nicoghosian, K., Cedergren, R.J. and Wong, J.T.F. (1983) Nucleic Acids Res. 11, 5433–5442.
[3] Gupta, R. (1984) J. Biol. Chem. 259, 9461–9471.
[4] Kuchino, Y., Ihara, M., Yahusaki, Y. and Nishimura, S. (1982) Nature, 684–685.
[5] Wich, G., Jarsch, M. and Böck, A. (1984) Mol. Gen. Genet. 196, 146–151.
[6] Woese, C.R. and Fox, G.E. (1977) Proc. Natl. Acad. Sci. USA 74, 5088–5090.
[7] Peattie, D.E. (1979) Proc. Natl. Acad. Sci. USA 76, 1760–1764.
[8] Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527–2535.
[9] Gupta, R. and Randerath, K. (1979) Nucleic Acids Res. 6, 3443–3458.
[10] Gauss, D.H. and Sprinzl, M. (1984) Nucleic Acids Res. 12, r1–r57.
[11] Betlach, M., Friedman, J., Boyer, W.J. and Pfeifer, F. (1984) Nucleic Acids Res. 12, 7949–7959.
[12] Mikelsaar, R. (1983) J. Theor. Biol. 105, 221–232.
[13] Grosjean, H., Nicoghosian, K., Haumont, E., Söll, D. and Cedergren, R.J. (1985) Nucleic Acids Res. 13, 5697–5706.