Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 2306–2309, April 1985
Biochemistry

UGA is read as tryptophan in *Mycoplasma capricolum*

(ribosomal proteins/opal tRNA/*Mycoplasma* genetic code/*Mycoplasma* tRNA)

FUMIAKI YAMAO, AKIRA MUTO, YASUSHI KAWAUCHI*, MASAFUMI IWAMI†, SHOJI IWAGAMI, YOSHITAKA AZUMI, AND SYOZO OSAWA

Laboratory of Molecular Genetics, Department of Biology, Faculty of Science, Nagoya University, Chikusa-ku, Nagoya 464, Japan

Communicated by Motoo Kimura, December 11, 1984

ABSTRACT **UGA is a nonsense or termination (opal) codon throughout prokaryotes and eukaryotes. However, mitochondria use not only UGG but also UGA as a tryptophan codon. Here, we show that UGA also codes for tryptophan in *Mycoplasma capricolum*, a wall-less bacterium having a genome only 20–25% the size of the *Escherichia coli* genome. This conclusion is based on the following evidence. First, the nucleotide sequence of the S3 and L16 ribosomal protein genes from *M. capricolum* includes UGA codons in the reading frames; they appear at positions corresponding to tryptophan in *E. coli* S3 and L16. Second, a $tRNA^{Trp}$ gene and its product tRNA found in *M. capricolum* have the anticodon sequence 5′ U-C-A 3′, which can form a complementary base-pairing interaction with UGA.**

We recently have sequenced a part of the *Mycoplasma capricolum* ribosomal-protein gene cluster that codes for polypeptides highly homologous to the *Escherichia coli* ribosomal proteins S3 and L16. The sequence contains four UGA codons in the reading frames; three appear at the sites corresponding to tryptophan, and one, at a site corresponding to arginine in the *E. coli* proteins. No "universal" UGG codon for tryptophan has so far been found. We have also isolated a clone containing a pair of *M. capricolum* tRNA genes, the sequence of both of which resembles that of $tRNA^{Trp}$ of *E. coli*. The anticodon sequence of one of these tRNA genes is 5′-T-C-A-3′, which can base-pair with both opal codon UGA and universal tryptophan codon UGG. That of the other is 5′-C-C-A-3′, which may base-pair exclusively with UGG. These two tRNA genes are expressed in the cell. All these findings suggest strongly that, in *M. capricolum*, UGA codes for tryptophan using the opal $tRNA_{UCA}$ but not $tRNA_{CCA}$.

RESULTS AND DISCUSSION

UGA Codons in *M. capricolum* S3 and L16 Genes. As reported in a previous paper (1), we isolated the recombinant plasmid pMCB1088 containing a 9-kilobase-pair fragment of *M. capricolum* DNA. The fragment contains the genes for at least nine ribosomal proteins—S3, S5, S8, S14, S17, L5, L6, L16, and L18—as deduced from its encoded protein sequences being highly homologous with the corresponding *E. coli* ribosomal protein sequences (refs. 1 and 2; unpublished results). Fig. 1 shows the complete nucleotide sequence of a 629-base-pair (bp) *Hin*dIII fragment which is a part of the insert of pMCB1088 (see refs. 1 and 2). The DNA corresponds to the 3′ half of the S3 gene and about 90% of the L16 gene from the 5′ terminus. When the *M. capricolum* sequences are aligned with the *E. coli* protein sequences (3, 4) (Fig. 1), four UGA (opal) codons are found within the reading frames. The possibility that these UGA codons are termination signals can be excluded by their occurrence in the regions having extensive sequence homologies with the *E. coli* proteins. More importantly, three out of the four UGA codons appear at the positions corresponding to tryptophan in the *E. coli* proteins. This suggests that UGA is a sense codon, probably for tryptophan, in *M. capricolum*. No UGG codon for tryptophan has so far been found. One UGA appears in S3 at a site corresponding to arginine in the *E. coli* L16.

Genes for Tryptophan tRNAs. Since UGA, which is a stop codon in the universal code, seems to be read as tryptophan in *M. capricolum*, one would expect the occurrence of an opal tRNA that can decode the UGA codon. Plasmid pMCH964, having a 2.0-kilobase-pair *Hin*dIII fragment in pBR322, was isolated as one of the clones that hybridize with unfractionated *M. capricolum* tRNAs. By restriction mapping of this fragment followed by hybridization with ^{32}P-labeled total *M. capricolum* tRNAs, the tRNA genes were localized within a 600-bp *Alu* I subfragment that had been derived from the middle part of the *Hin*dIII fragment (data not shown). The DNA sequence of this region (Fig. 2) revealed the presence of a pair of tRNA genes with a 40-bp spacer between them. The tRNA encoded by the first gene has an anticodon sequence 5′-U-C-A-3′ that can decode both opal codon UGA and universal tryptophan codon UGG, whereas the second one has an anticodon sequence 5′-C-C-A-3′ for the universal UGG codon for tryptophan (Fig. 3). The structural gene region for these two tRNAs is preceded by the expected promoter structures: a Pribnow-box-like sequence (underlined in Fig. 2) ≈20 bp upstream and a −35 sequence (also underlined in Fig. 2) 45 bp upstream from the coding sequences. The tRNA genes are followed by a probable termination signal: a dyad symmetrical structure and a stretch of thymidine residues (indicated by two arrows and by a broken line, respectively, in Fig. 2) 24 bp downstream from the coding sequence for $tRNA_{CCA}$. The above structure suggests that the two tRNA genes are arranged in a single operon. The $tRNA_{UCA}$ gene could have emerged by duplication of the $tRNA_{CCA}$ gene, since the two tRNA genes are closely related to each other not only in their tandem linkage on the chromosome but also in their high sequence homology (78% identity) and both $tRNA_{UCA}$ and $tRNA_{CCA}$ can be charged with tryptophan *in vitro* (see below).

Expression of $tRNA^{Trp}$ Genes. To determine whether the two tRNA genes described above are expressed *in vivo*, we purified tRNAs that hybridize with the DNA fragment containing these two tRNA genes. The purification procedure consists of hybridization of crude tRNAs with the DNA fragment, followed by identification of the hybridized tRNAs by sequencing. Since $tRNA_{UCA}$ is one base longer than $tRNA_{CCA}$, as deduced from their DNA sequences, they may

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: bp, base pair(s).

*On leave from the Department of Biochemistry and Biophysics, Research Institute for Nuclear Medicine and Biology, Hiroshima University, Hiroshima 734, Japan.
†Present address: Meiji Institute of Health Science, 540 Naruda, Odawara 250, Japan.

```
     AA GCT TTA AAA GCT GGA GCT AAA GGA ATT AAA ACT GCT GTA AGT GGA AGA TTA GGT GGA GTT  062
M.c.    Ala Leu Lys Ala Gly Ala Lys Gly Ile Lys Thr Ala Val Ser Gly Arg Leu Gly Gly Val
E.c.    Ala Met Arg Leu Gly Ala Lys Gly Ile Lys Val Glu Val Ser Gly Arg Leu Gly Gly Ala

        GAA ATG GCA CGT ACT GAA GGA TAT TTA GAA GGT TCA GTA CCA CTA TCA ACT TTA AGA AAT  122
M.c.    Glu Met Ala Arg Thr Glu Gly Tyr Leu Glu Gly Ser Val Pro Leu Ser Thr Leu Arg Asn
E.c.    Glu Ile Ala Arg Thr Glu Trp Tyr Arg Gln Gly Arg Val Pro Leu His Thr Leu Arg Ala

        AAT ATT GAT TAT GCT TTA TAT GAA GCT CCA ACA ACA TAT GGT CAA ATT GGA GTT AAA GTA  182
M.c.    Asn Ile Asp Tyr Ala Leu Tyr Glu Ala Pro Thr Thr Tyr Gly Gln Ile Gly Val Lys Val
E.c.    Asp Ile Asp Tyr Asn Thr Ser Gln Ala His Thr Thr Tyr Gly Val Ile Gly Val Lys Val

         *
        TGA ATT AAT CAT GGT GAA GTA TTT --- --- --- --- AAA AAA GAA AGA ATG AAT AAT TCA  230
M.c.    TRP Ile Asn His Gly Glu Val Phe --- --- --- --- Lys Lys Glu Arg Met Asn Asn Ser
E.c.    Trp Ile Phe Lys Gly Glu Ile Leu Gly Gly Met Ala Ala Val Glu Gln Pro Glu Lys Pro
                                                        S3  End     L16 Start
        CAA ATA ATG GCA AAA CCA AGA ACT AAT AAA GGA GGT AAA AGA TAA TT- ATG TTA CAA CCA  289
M.c.    Gln Ile Met Ala Lys Pro Arg Thr Asn Lys Gly Gly Lys Arg ---     Met Leu Gln Pro
E.c.    Ala Ala Gln Pro Lys Lys Gln Gln Arg Lys Gly Arg Lys             Met Leu Gln Pro

        AAA AGA ACA AAA TAT CGT AAA CCT CAT AGA GTT AGT TAT GAA GGA AAA GCT AAA GGA GCT  349
M.c.    Lys Arg Asn Lys Tyr Arg Lys Pro His Arg Val Ser Tyr Glu Gly Lys Ala Lys Gly Ala
E.c.    Lys Arg Thr Lys Phe Arg Lys Met His Lys Gly Arg Asn Arg Gly Leu Ala Gln Gly Thr

                                                                    *
        AAA GAA ATT AAC TTT GGT GAA TTT GGT TTA ATG GCT TTA GAT GGT GCT TGA ATT GAT AAT  409
M.c.    Lys Glu Ile Asn Phe Gly Glu Phe Gly Leu Met Ala Leu Asp Gly Ala TRP Ile Asp Asn
E.c.    Asp --- Val Ser Phe Gly Ser Phe Gly Leu Lys Ala Val Gly Arg Gly Arg Leu Thr Ala

        CAT CAA ATA GAA GCT GCG CGT ATT GCT ATG ACA CGT TAT ATG AAG CGT GAT GGA AAA ATT  469
M.c.    His Gln Ile Glu Ala Ala Arg Ile Ala Met Thr Arg Tyr Met Lys Arg Asp Gly Lys Ile
E.c.    Arg Gln Ile Glu Ala Ala Arg Arg Ala Met Thr Arg Ala Val Lys Arg Gln Gly Lys Ile
         *
        TGA ATG AGA ATT TTC CCA CAT ATG GCA ATG ACT AAA AAA CCT GCT GAA GTT CGT ATG GGT  529
M.c.    TRP Met Arg Ile Phe Pro His Met Ala Met Thr Lys Lys Pro Ala Glu Val Arg Met Gly
E.c.    Trp Ile Arg Val Phe Pro Asp Lys Pro Ile Thr Glu Lys Pro Leu Ala Val Arg Met Gly

                                         *
        TCA GGA AAA GGA AAT CCT GAA AAA TGA GTA GCA GTA GTT AAA AAA GGA ACA ATT ATG TTT  589
M.c.    Ser Gly Lys Gly Asn Pro Glu Lys TRP Val Ala Val Val Lys Lys Gly Thr Ile Met Phe
E.c.    Lys Gly Lys Gly Asn Val Glu Tyr Trp Val Ala Leu Ile Gln Pro Gly Lys Val Leu Tyr

        GAA GTT GCT CAA GTA AAT GAG CAA GTA GCT AGA GAA GCT T                            629
M.c.    Glu Val Ala Gln Val Asn Glu Gln Val Ala Arg Glu Ala
E.c.    Glu Met Asp Gly Val Pro Glu Glu Leu Ala Arg Glu Ala
```

FIG. 1. DNA sequence of a part of *M. capricolum* S3 and L16 ribosomal protein genes. The DNA sequence of the mRNA-like strand, together with the predicted amino acid sequence, of a *Hin*dIII fragment from a plasmid pMCB1088 (see ref. 1) that contains a part of *M. capricolum* [American Type Culture Collection 27343 (Kid)] ribosomal protein S3 and L16 genes (M.c.) was aligned with the corresponding *E. coli* protein sequences (E.c.). The sites for identical amino acids and those of conservative amino acid substitution are boxed with solid and dotted lines, respectively. TGA triplets are marked with asterisks. DNA sequencing was performed by the chain-termination methods (5, 6).

```
TTTAACTATT CAAGAAATAT ATTAATAATT TAGAATCTGA AATTAACAAT ATTAATGTAA TGAGTGTTTA TAATACCATT GATTTATTAT TAAAAGAACA  100
TATCGTTTTT GCTAATACTT TTAATGGAAA AGATATTTCT TATGAAATAG CAGCTGATAA ATCTGTTCAT TTAAAGTGTG ATGAATGTTT AAAAGTAATT  200
CACTTAGATG ATAAGAACAT AAAAAATTAT CACTTTTTAG AATTATTAGA TTTATGTGAA AAATATAATA TTAAATTAAC TCATTTCAAA TCGAAGGTCA  300
TGGGTATTGT TTAAATGTTC AAATAAAGAA ATAAATAAGT AGGTAAGTTA GCTAATGATA CGCATTTGCC TTGAAAAACT CATTAATTGA CTTTATAATT  400
ATAGGTGAGT TTTATAGGGG CATAGTTCAG TAGGTAGAAC ATCGGTCTTC AAAACCGAGT GTCACGAGTT CGAGTCTTGT TGCCCCTGCC ATTTTGAAAG  500
CAAATCACAC TTTGTGTGAT TTTTTTATAG GAGAGTAGTT CAATGGTAGA ACGTCGGTCT CCAAAACCGA GCGTTGAGGG TTCGATTCCT TTCTCTCCTG  600
CCATAAGAAA TAAAAAAAAA CTGGAAATTC CAGTTTTTTT ATTCTTCAAT TGCAACAAAA CCTATAGTTT CAATTCCTGC ATGAATAGTG TAAATATTTG  700
GCACATATCC ATGAATAAAT TTTACTTTTT CATCACTAAG AATTTGCTTA ACAATTTCAA CAGTTTTGCT TGATGTTAGT GGAGTTGATA AAAAATATAA  800
TTTATATTTA TTTTTTTTAA ACTTGTTAGA AAGATTTTTT AATCAATTTT TCAATAAGAC TATTGTAAGT TCTTCCAATG GCTTCTTTTT TAGGTTCTTT  900
TGCTCAAACG ATTAGTAATT TAGTTTTTAA AAGATTTAAA ACAGTTGTAA TAACTCCTTT AGCTCTACCA CCACTTGATA ATTTTTTTAG ATCT       1000
```

FIG. 2. DNA sequence of two tryptophan tRNA genes and their flanking region. Coding sequences for $tRNA_{UCA}$ and $tRNA_{CCA}$ are boxed. A Pribnow box-like structure and −35 sequences are underlined. Probable transcription-termination (dyad symmetrical) structure is shown by two arrows. A stretch of thymidine residues is shown by a broken line.

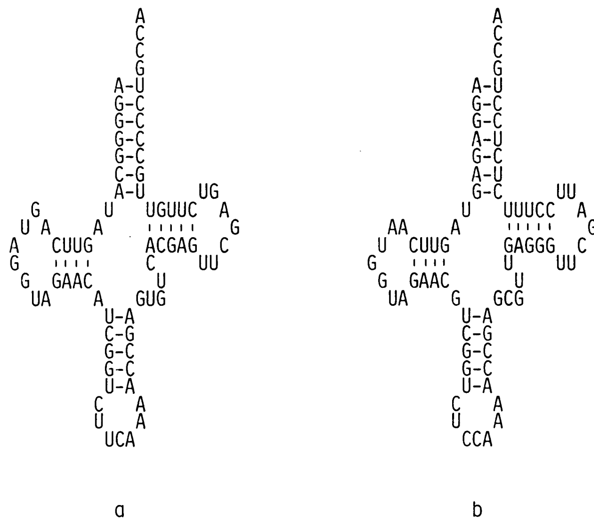

FIG. 3. Cloverleaf structures of $tRNA_{UCA}$ (*a*) and $tRNA_{CCA}$ (*b*) deduced from DNA sequence.

be easily distinguished by polyacrylamide gel electrophoresis. Fig. 4 shows an electrophoretic separation of tRNAs, the 3′ ends of which were labeled with $[^{32}P]pCp$ and T4 RNA ligase after hybridization with the DNA fragment (see lane a). Only two tRNA species, differing slightly in length, predominated. Partial sequence analysis of these tRNAs by the chemical degradation method revealed that the fast migrating one agreed with the DNA sequence for the $tRNA_{CCA}$ gene, and the other, with that for the $tRNA_{UCA}$ gene (data not shown). Thus, in the cells, both the genes for $tRNA_{UCA}$ and $tRNA_{CCA}$ are transcribed and processed. However, when tRNAs were purified without first removing amino acids, $tRNA_{CCA}$ was predominantly labeled with $[^{32}P]pCp$ (Fig. 4, lane b). This suggests that the bulk of the $tRNA_{UCA}$ molecules but not $tRNA_{CCA}$ have been already charged with amino acids *in vivo*, because amino acid bound to the -C-C-A terminus would prevent the ligation of $[^{32}P]pCp$ to the 3′ end. To verify this and identify the amino acid bound to $tRNA_{UCA}$, the following experiment was performed. The deacylated tRNAs were incubated for a short period with amino acid in the presence of the *M. capricolum* S-100 fraction to reacylate the tRNAs. Charging the tRNAs with tryptophan inhibited the incorporation of $[^{32}P]pCp$ into both $tRNA_{CCA}$ and $tRNA_{UCA}$ (Fig. 4, lane d), whereas other amino acids, leucine for example, did not affect the labeling efficiency (Fig. 4, lane c). These results indicate that both $tRNA_{CCA}$ and $tRNA_{UCA}$ accept tryptophan *in vitro*.

The presence of $tRNA^{Trp}_{UCA}$ in *M. capricolum* strongly supports the idea that UGA is translated as tryptophan by using this tRNA. Tryptophan is "universally" coded for by a single codon, UGG, which is decoded by $tRNA_{CCA}$ throughout prokaryotes (9) and eukaryotes (10). In mitochondria, not only UGG but also UGA are used as tryptophan codons (11–13), both of which are translated by a single tRNA with the anticodon UCA (14–17). Thus, the discovery of two tRNA species having anticodon sequences of, respectively, CCA and UCA in one genetic system contrasts with previous observations in other systems. Although the UCA anticodon can decode both UGA and UGG according to the wobble theory (18), we have not so far found UGG codons in the reading frames not only in the S3 and L16 genes but in other ribosomal protein genes (see also ref. 2). This suggests that UGA is predominantly, if not exclusively, used as a tryptophan codon in *M. capricolum*. It is thus interesting to see whether and how $tRNA_{CCA}$ participates in translation. The failure to find $tRNA_{CCA}$ appreciably charged *in vivo* with tryptophan might mean that its role is subsidiary.

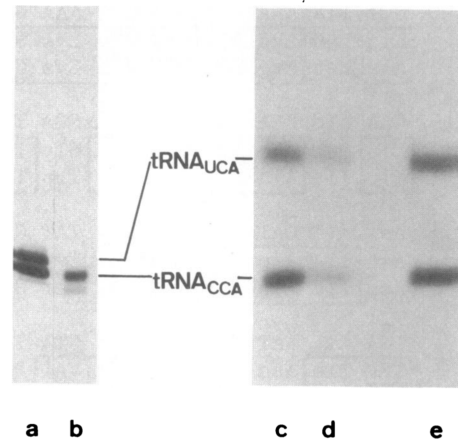

FIG. 4. Hybridization of tRNAs with pMCH964 DNA. Total tRNAs were prepared from *M. capricolum* by direct phenol extraction, followed by hybridization with pMCH964 DNA. For hybridization, pMCH964 DNA (1 mg) was fragmented by sonication and bound to a Sephacryl S-500 column according to the method of Bünemann and Westhoff (7). Hybridization was carried out at 42°C for 24 hr in a hybridization buffer containing 20 mM Pipes [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)] (pH 6.5), 0.6 M NaCl, 1 mM EDTA, 0.2% $NaDodSO_4$, and 50% (vol/vol) formamide. After unbound material was removed by washing with the hybridization buffer at 42°C, the hybridized tRNAs were eluted with the buffer at 90°C, precipitated, and washed with 70% (vol/vol) ethanol. The tRNAs so obtained were incubated at 37°C in 100 mM Tris·HCl (pH 10.0) for 30 min to "strip" amino acid (deacylate), if necessary. The 3′ ends of the tRNAs were labeled with $[^{32}P]pCp$ and T4 RNA ligase as described by Peattie (8), followed by electrophoresis in 12% polyacrylamide gel and autoradiography. Electrophoresis was at 50 V/cm for 5 hr (lanes a and b) or at 10 V/cm for 30 hr (lanes c–e). Lanes a and b: tRNAs hybridized to pMCH964 DNA and labeled with $[^{32}P]pCp$ after deacylation (a) or without deacylation (b). Lanes c and d: tRNAs were aminoacylated with leucine (c) or tryptophan (d) before $[^{32}P]pCp$ labeling. Lane e: tRNAs were incubated with the S-100 fraction from *M. capricolum* without amino acid. Aminoacylation was carried out at 37°C for 5 min in a reaction mixture consisting of 50 μl of 0.2 M sodium cacodylate buffer (pH 7.0) containing 10 mM magnesium acetate and 8 mM ATP, 10 μl of amino acid solution (1 mM), 20 μl of the *M. capricolum* S-100 fraction, and 10 μl of H_2O in a total volume of 100 μl.

Evolutionary Aspects. The A+T content of the *M. capricolum* genome is about 75%, one of the highest among all organisms. Reflecting this, an obvious preference for adenosine and thymidine in the *M. capricolum* genome has been seen in various regions: e.g., rRNA genes (19), their spacers (20), and codons (2). In a previous paper (2), we have demonstrated that the codons used for the *M. capricolum* ribosomal proteins S8 and L6 are strongly biased to those rich in A and U and that more than 90% of the codons have A or U at the third position. The same tendency of codon usage can be seen in the S3 and L16 genes, as shown in Fig. 1. The total A+T content of the coding regions is about 67%, and only 12 out of the total 207 codons have G or C at the third position [9 of these 12 are AUG (methionine) codons]. This suggests that the constraint for the preferential use of A and T in the protein genes is operating at the DNA level as a selection force. Thus, the use of UGA rather than UGG as a tryptophan codon in *M. capricolum* may be the consequence of this evolutionary pressure.

It has been suggested that mitochondria have evolved from certain prokaryotes by endosymbiosis (21, 22). The use of UGA as a tryptophan codon in *M. capricolum* as in mitochondria raises an interesting possibility in the phylogenetic relationship between mycoplasmas and mitochondria. Mycoplasmas are parasitic in eukaryotes, and the A+T-richness of their genomic DNAs resembles that of the mitochondrial

DNAs of lower eukaryotes. The codon usage of the yeast mitochondrial protein genes is strongly biased to the A- and U-rich codons (13, 23, 24), as is the case in *M. capricolum* (2). Furthermore, the *M. capricolum* $tRNA^{Trp}_{UCA}$ sequence is more similar to yeast mitochondrial $tRNA^{Trp}_{UCA}$ (14) (66% identity) than to cytoplasmic $tRNA^{Trp}_{CCA}$ (25) (55% identity). Thus, the mycoplasma-like organisms might have played some role in the evolution of mitochondria.

The deviation from the universal codons that occurs in mitochondria, such as AUA for methionine instead of isoleucine and AGA for nonsense instead of arginine (11), may not be the case in *M. capricolum*, because AUA can be seen in the reading frame of the *M. capricolum* genes at the positions corresponding to isoleucine in the *E. coli* protein sequences (e.g., the 47th codon of the L16 gene in Fig. 1) and AGA is the most abundantly used codon for arginine in this organism (2).

We thank Dr. Hiroshi Hori of this laboratory for many stimulating discussions. This work was supported by the Ministry of Education, Science, and Culture, Japan (Grants 59400007, 59107004, 59107010, and 59214010) and the Naito Science Foundation (Research Grant 81-106).

1. Kawauchi, Y., Muto, A., Yamao, F. & Osawa, S. (1984) *Mol. Gen. Genet.* **196**, 521–525.
2. Muto, A., Kawauchi, Y., Yamao, F. & Osawa, S. (1984) *Nucleic Acids Res.* **12**, 8209–8217.
3. Brauner, D. & Röming, R. (1979) *FEBS Lett.* **106**, 352–357.
4. Brosius, J. & Chen, R. (1976) *FEBS Lett.* **68**, 105–109.
5. Messing, J. & Vieira, J. (1982) *Gene* **19**, 269–276.
6. Sanger, F., Wicken, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
7. Bünemann, H. & Westhoff, P. (1983) *Methods Enzymol.* **100**, 400–407.
8. Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1760–1764.
9. Hirsh, D. (1971) *J. Mol. Biol.* **58**, 439–458.
10. Harada, F., Sawyer, R. C. & Dahlberg, J. E. (1975) *J. Biol. Chem.* **250**, 3487–3497.
11. Barrell, B. G., Bankier, A. T. & Drouin, J. (1979) *Nature (London)* **282**, 189–194.
12. Macino, G., Coruzzi, G., Nobrega, F. G., Li, M. & Tzagoloff, A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3784–3785.
13. Fox, T. D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6534–6538.
14. Martin, N. C., Pham, H. D., Underbrink-Lyon, K., Miller, O. L. & Donelson, J. E. (1980) *Nature (London)* **285**, 579–581.
15. Heckman, J. E., Sarnoff, J., Alzner-DeWeerd, B., Yin, S. & RajBhandary, U. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3159–3163.
16. Barrell, B. G., Anderson, S., Bankier, A. T., DeBruijn, M. H. L., Chen, E., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3164–3166.
17. Köchel, H. G., Lazarus, C. M., Basak, N. & Küntzel, H. (1981) *Cell* **23**, 625–633.
18. Crick, F. H. C. (1966) *J. Mol. Biol.* **19**, 548–555.
19. Iwami, M., Muto, A., Yamao, F. & Osawa, S. (1984) *Mol. Gen. Genet.* **196**, 317–322.
20. Sawada, M., Muto, A., Iwami, M., Yamao, F. & Osawa, S. (1984) *Mol. Gen. Genet.* **196**, 311–316.
21. Margulis, L. (1981) *Symbiosis in Cell Evolution* (Freeman, San Francisco).
22. Küntzel, H. & Köchel, H. G. (1981) *Nature (London)* **293**, 751–755.
23. Dujon, B. (1980) *Cell* **20**, 185–197.
24. Macreadie, I. G., Novitski, C. E., Maxwell, R. J., John, U., Ooi, B., McMullen, G. L., Lukins, H. B., Linnane, A. W. & Nagley, P. (1983) *Nucleic Acids Res.* **11**, 4435–4451.
25. Keith, G., Roy, A., Ebel, J. P. & Dirheimer, G. (1972) *Biochimie* **54**, 1405–1426.