The chromosomal location of genes for elongation factor Tu and ribosomal protein S10 in the cyanobacterium *Spirulina platensis* provides clues to the ancestral organization of the *str* and S10 operons in prokaryotes

ANNA M. SANANGELANTONI* and ORSOLA TIBONI

Dipartimento di Genetica e Microbiologia 'A. Buzzati-Traverso', Università di Pavia, via Abbiategrasso 207, 27100 Pavia, Italy

(Received 2 March 1993; revised 9 June 1993; accepted 14 June 1993)

The structural gene (*rps10*) encoding ribosomal protein S10 of the cyanobacterium *Spirulina platensis* has been localized both on chromosomal DNA and the previously characterized recombinant plasmid pSp7 harbouring the 3'-terminal portion of the gene for elongation factor G (*fus*) and the gene for elongation factor Tu (*tuf*). Alignment of the predicted S10 sequence of *S. platensis* with the homologous sequences from cyanelles, bacteria, archaea and eukarya showed that the cyanobacterial S10 shares a high degree of sequence homology (74% amino acid identity) with the cyanellar protein. Unlike the situation in *Escherichia coli*, the *rps10* gene of *S. platensis* is unlinked to the S10 operon genes, being adjacent to the *str* operon genes. Since a similar organization could be observed in cyanelles of *Cyanophora paradoxa* and in all archaea so far analysed, this probably represents the ancestral state.

Introduction

The organization and regulation of genes encoding components of the translational apparatus have been extensively investigated in *Escherichia coli* (for a review, see Lindahl & Zengel, 1986). As more data become available in other prokaryotes, it appears that the chromosomal organization of these genes is substantially conserved amongst bacteria and archaea, although differences do exist in detail.

In the cyanobacteria *Spirulina platensis* and *Anacystis nidulans* the genes encoding ribosomal proteins S12, S7 and elongation factors EF-G (*fus*) and EF-Tu (*tuf*) are arranged in a transcriptional unit resembling the *str* operon of *Escherichia coli* (Buttarelli *et al.*, 1989; Meng *et al.*, 1989). An *E. coli*-like organization of these genes is also maintained in the cyanelles of *Cyanophora paradoxa*, except that the *fus* gene has been recruited by the nuclear genome (Kraus *et al.*, 1990).

According to the endosymbiotic theory for the origin of cell organelles, cyanelles represent an evolutionary intermediate between the free-living cyanobacteria, the putative plastidial ancestor, and modern chloroplasts. Although the cyanellar genome resembles that of chloroplasts both in structure and size, cyanelles share with cyanobacteria certain structural similarities such as the presence of unstacked thylakoids, phycobilisomes, and a lysozyme-sensitive cell wall. In addition, similar to cyanobacteria, the genes encoding the two Rubisco subunits are present on the cyanellar genome and are arranged in an operon (Wasmann *et al.*, 1987).

It has been reported that the gene encoding ribosomal protein S10 (*rps10*) of cyanelles is linked to the gene for EF-Tu (*tufA*) and is cotranscribed with the *str* operon genes (Neumann-Spallart *et al.*, 1991; Bryant *et al.*, 1991). This linkage has been found also in all of the archaea so far investigated (Creti *et al.*, 1991; Tesch & Klink, 1992). This is unlike the situation in *E. coli*, where the *rps10* gene is located in a different operon, termed the S10 operon, that lies 15 kb downstream from the *str* operon (Zurawski & Zurawski, 1985). The situation in the extreme thermophile *Thermotoga maritima*, representing one of the deepest offshoots of the bacterial tree (Wheelis *et al.*, 1992), is similar to that found in cyanelles. In this micro-organism the *tufA* and *rps10* gene belong to two different operons, but are separated by an extremely

* Author for correspondence. Tel. +39 382 391578; fax +39 382 528496.

Abbreviation: EF, elongation factor.

The nucleotide sequence data reported in this paper have been submitted to the EMBL Data Library and assigned the accession number Z21676.

short distance (only 50 bp) compared to that (15 kb) between the *tufA* and *rps10* genes in *E. coli* (O. Tiboni and others, unpublished data). Given that the linkage between the *tuf* and *rps10* genes is maintained in cyanelles, in all of the archaea, in *Tg. maritima* and, as shown in this paper, in the cyanobacterium *Spirulina platensis*, the possibility exists that linkage of the *rps10* gene with those of the *str* operons represents the ancestral situation.

Here we report the cloning, sequencing and localization of the *rps10* gene relative to the *tuf* gene in *S. platensis*.

Methods

Bacterial strains and plasmids. *Spirulina platensis* cells were grown as reported by Riccardi *et al.* (1981). A plasmid containing a 1250 bp *EcoRI*/*PstI* fragment of *Cyanophora paradoxa* cyanellar DNA (kindly supplied by Dr W. Löffelhardt) in the vector pEMBL8 was used to prepare the heterologous probe. Plasmid pSp7 (Tiboni & Di Pasquale, 1987), containing a 5.2 kb fragment of *S. platensis* genomic DNA, was used for subcloning. Plasmids pBR322 (Bolivar *et al.*, 1977) and pUC19 (Vieira & Messing, 1982) were used as vectors and *E. coli* HB101 as the host. Plasmid-containing strains were grown in LB medium supplemented with 50 µg ampicillin ml⁻¹.

Recombinant DNA techniques. Chromosomal DNA from *S. platensis* was prepared as described previously (Tiboni *et al.*, 1984). Plasmid DNA was isolated and purified using Qiagen tips (Diagen) following the manufacturer's instructions. Recovery of DNA fragments from low-melting-point agarose gel was performed as reported by Higuchi *et al.* (1981). Transformation experiments were carried out as described by Dagert & Ehrlich (1979). Hybridization experiments (Southern, 1975) were performed at 37 °C in the presence of 35% (v/v) formamide (for heterologous DNA) or at 42 °C in the presence of 50% formamide (homologous DNA). Probes were labelled by the Multiprime DNA Labelling System using [α -³²P]dCTP [Amersham; sp. act. 3000 Ci mmol⁻¹ (111 TBq)] and the Multiprime DNA Labelling System kit supplied by Amersham. Restriction endonuclease enzymes and T4 DNA ligase were obtained from Boehringer and used according to the manufacturer's instructions.

Sequencing. The inserts of plasmids pSp14 and pSp7 were digested with several restriction enzymes and fragments of 800–1000 bp were subcloned in the polylinker cloning sites of pUC19 and sequenced by the dideoxy chain-termination method of Sanger *et al.* (1977). In the case of longer inserts, synthetic oligonucleotides were used as primer. The DNA was radiolabelled with [α -³⁵S]dATP [> 1000 Ci mmol⁻¹ (37 TBq)] and primed as specified by the manufacturer using a T7 Sequencing kit (Pharmacia).

Analysis of the DNA sequence was performed using a Macintosh IIfx computer and the DNA Strider 1.1 program (Marck, 1988).

Phylogenetic analysis. Amino acid sequences were aligned by the hierarchical clustering method of Corpet (1988) as implemented in the program MULTALIN with a gap penalty of 8. The amino acid sequence alignment was converted into a colinear alignment of first and second codon positions and phylogenetic trees were constructed by distance matrix and maximum parsimony methods from clearly homologous gap-free regions of the DNA alignments using the programs CONSENSE, FITCH, DNADIST, DNAPARS and SEQBOOT as implemented in the PHYLIP 3.5c package. Evolutionary distances were calculated from the pairwise nucleotide sequence identities using the Jukes–Cantor model of

sequence evolution as implemented in the DNADIST program (Felsenstein, 1990). The reliability of the internal branches was determined by bootstrap analysis of at least 100 pseudoreplications of the original sample.

Results and Discussion

Cloning of the *S. platensis rps10* gene

To identify DNA fragments carrying the *rps10* gene of *S. platensis*, genomic DNA was digested with several restriction endonucleases and probed with the 1.25 kb *EcoRI*–*PstI* fragment of *C. paradoxa* cyanelles DNA containing the entire coding region of the *rps10* gene as well as the entire *petFI* gene and the 3'-terminal region of the *tuf* gene (Neumann-Spallart *et al.*, 1991; Bryant *et al.*, 1991). Among all digestions exhibiting different positive signals, the *HaeIII* digestion was selected because the size of the two fragments (1800 and 600 bp) hybridizing to the probe allowed discrimination between the signal due to the previously cloned *tuf* gene (600 bp fragment, Buttarelli *et al.*, 1989) and that due to *rps10* and/or *petFI* genes. Thus, a partial genomic library of *S. platensis* DNA (1.6–2.0 kb *HaeIII* fragments) was constructed in *E. coli* HB101 in the *EcoRV* site of pBR322.

Since preliminary experiments demonstrated that the signal produced by the *S. platensis rps10* gene was as weak as that produced by the gene of the host, it was impossible to carry out colony hybridizations. Thus, the recombinant plasmids carrying sequences homologous to the probe were selected by analysis of the inserts. Plasmid DNA was prepared from 30 pools of 18 recombinant clones each, and digested with *HaeIII*. The fragments were separated by gel electrophoresis and hybridized with the cyanellar probe. The individual transformants from the positive pool were analysed for plasmid content and a recombinant plasmid, containing a 1.8 kb *HaeIII* fragment (designated pSp14) was isolated. A preliminary sequence analysis indicated that the 5' end of the insert contained part of the *S. platensis rps10* gene.

Because in *C. paradoxa* cyanelles the *rps10* gene is located 87 bp downstream from the stop codon of the *tuf* gene (Neumann-Spallart *et al.*, 1991), experiments were performed to ascertain if the same organization was present in the phylogenetically related cyanobacteria. To this end, the 5.2 kb fragment of the plasmid pSp7 (Tiboni & Di Pasquale, 1987) harbouring part of the *S. platensis str* operon (namely, part of the *fus* gene and the entire *tuf* gene) and a downstream region of 2200 bp, was probed with the insert of pSp14 demonstrating that the entire *HaeIII*/*HaeIII* fragment cloned in plasmid pSp14 was contained within this region. Thus, starting from the 3' end of the *S. platensis tuf* gene, the nucleotide sequence of 1215 bp was determined. In addition, and because in

EF-Tu
ATC GCT ATT GAG CAA GGG ATG CGT TTC GCT ATC CGT GAG GGT GGT CGT ACT
ile ala ile glu gln gly met arg phe ala ile arg glu gly gly arg thr
GTT GGT GCG GGT GTA GTT GCT AAG ATT CTG GCT TAGAAGTAGAATCCGATCGCTAG
val gly ala gly val val ala lys ile leu ala TER
ATTCGAACTGATCTGATCAATTAAGTGGCAGGAAGTCTCCGAGTTGAGGACACTTTTCTGCTCT
CTCTGACCTTGCTTATTATTAATCAACTGAACAGCAGAACTAAACAATTGAGACCATG GCT ACC
MET ala thr
ATT CAG CAG CAA AAA ATT CGC ATC CGT CTT AAA GGC TTC GAT CGC CGC TTA
ile gln gln gln lys ile arg ile arg leu lys ala phe asp arg arg leu
CTT GAT ACC TCC TGC GAG AAG ATC GTC GAT ACG GCT AAC CGC ACC GGA GCT
leu asp thr ser cys glu lys ile val asp thr ala asn arg thr gly ala
ACG GCC CTC GGT CCG ATT CCT TTA CCA ACT AAA AGA CGC ATT TAC TGT GTC
thr ala leu gly pro ile pro leu pro thr lys arg arg ile tyr cys val
CTA CGT TCT CCC CAC GTT GAT AAG GAT TCA CGA GAA CAT TTT GAA ACC CGG
leu arg ser pro his val asp lys asp ser arg glu his phe glu thr arg
ACT CAC CGC CGC ATT ATT GAT ATT TAT CAA CCT TCC TCA AAA ACA ATT GAT
thr his arg arg ile ile asp ile tyr gln pro ser ser lys thr ile asp
GCT CTC ATG AAA CTC GAT TTA CCC GCT GGG GTG GAT ATT GAG GTG AAA TTG
ala leu met lys leu asp leu pro ala gly val asp ile glu val lys leu
TAGGGTCACTAGGGGCGATCGCTCTTCAATATATTCGGTGATCGAAGCAGGGGTCTCCGAAATCGG
TER
CTAAGTATAAAGAAAGCGCACACTGTACTAAGGGACA ATG GCA TCA TCT TCC TCT
MET ala ser ser ser ser
CTT GCT GTT CGA GAA CTT CCC CTG TTT CCT CTA CCG GAA GTG GTT CTG TTT
leu ala val arg glu leu pro leu phe pro leu pro glu val val leu phe

Fig. 1. Nucleotide sequence of the *rps10* gene of *S. platensis* flanked by the 3'-terminal region of *tuf* gene and 5' portion of the unknown ORF. The deduced amino acid sequence is shown below the DNA sequence. The underlined sequence of 15 nucleotides was used to prepare the synthetic primer.

bacteria (*E. coli* and *Mycoplasma capricolum*) the *rps10* gene is the promoter distal gene of the S10 operon, the sequence was extended to cover the 3' end of the pSp7 insert using synthetic 15-mer oligonucleotides (data not shown).

Nucleotide sequence of the *S. platensis* S10 structural gene

The sequence of the 1215 bp region, whose relevant part is shown in Fig. 1, revealed the presence of an ORF of 318 bp located 143 bp downstream of the *tuf* gene. The

derived amino acid sequence shared a high degree of similarity with the cyanellar ribosomal protein S10. The length of the intergenic region between the *tuf* and *rps10* genes was comparable to that separating the genes in the *S. platensis* *str* operon. In addition, no potential ribosome binding site sequence, and no obvious *p*-independent termination structures, could be identified within the intergenic region, suggesting the possibility that *rps10* and *tuf* genes are cotranscribed as in the case of the cyanelles of *C. paradoxa*.

Analysis of the sequence, downstream of the *rps10* gene, did not reveal any ORFs whose derived polypeptide sequence exhibited significant homology with any of the other ribosomal proteins constituting the bacterial S10 operon (Zurawski & Zurawski, 1985).

Amino acid sequence comparison of *S. platensis* ribosomal protein S10 with those of cyanelles, archaea and bacteria

The deduced amino acid sequence of the *S. platensis* ribosomal protein S10 was compared with available homologues from bacteria (*E. coli*, *Tg. maritima*, *Myco. capricolum*), archaea (*Sulfolobus acidocaldarius*, *Methanococcus vannielii*, *Pyrococcus woesei*, *Thermoplasma acidophilum*), eukarya (rat) and cyanelles of *C. paradoxa*. No other plastidial S10 sequences were available for comparison. The alignment presented in Fig. 2 clearly shows that all of the S10-equivalent sequences contain the conserved heptapeptide GPIPLPT that could be related to the functioning of S10 in the ribosome. S10 is, in fact, located on the external surface of the small subunit of the *E. coli* ribosome together with proteins S3, S5, S14 and S19, all of which appear to be involved in tRNA binding (Kahan *et al.*, 1981). As the conserved heptapeptide occurs in all of the prokaryotic S10-equivalent proteins examined until now, it is reasonable

Fig. 2. Alignment of the *S. platensis* S10 amino acid sequence with the sequence of ribosomal protein S10 from: cyanelles of *Cyanophora paradoxa* (Neumann-Spallart *et al.*, 1991); *E. coli* (Zurawski & Zurawski, 1985); *Myco. capricolum* (Ohkubo *et al.*, 1987); *Tg. maritima* (O. Tiboni and others, unpublished data); *Meth. vannielii* (Lechner *et al.*, 1989); *Tpl. acidophilum* (Tesch & Klink, 1992); *Sulf. acidocaldarius* (Auer *et al.*, 1991); *P. woesei* (Creti *et al.*, 1991) and rat (Chan & Wool, 1990). The conserved heptapeptide is indicated by asterisks. The underlined region was used for phylogenetic analysis.

Table 1. Identity matrices from nucleotide and amino acid sequence alignments of ribosomal protein S10

Lower left quadrant: percentage of identity calculated from ungapped region underlined in Fig. 2 between amino acid sequences. Upper right quadrant: identity matrix calculated from nucleotide sequence alignment (first and second position) colinear with the amino acid sequence alignment of ribosomal protein S10 from cyanelles (Cya), bacteria (*S. platensis*, Spl; *E. coli*, Eco; *Myc. capricolum*, Mca; *Tg. maritima*, Tma), archaea (*Meth. vannielii*, Mva; *Tpl. acidophilum*, Tac; *Sulf. acidocaldarius*, Sac; *P. woesei*, Pwo) and eukarya (Rat).

	Spl	Cya	Eco	Mca	Tma	Mva	Tac	Sac	Pwo	Rat
Spl		79.5	73.5	65.7	70.6	47.8	58.5	51.5	49.5	50.5
Cya	74.0		65.7	63.7	69.6	50.0	55.0	54.0	50.0	52.4
Eco	59.4	52.1		66.8	67.5	47.7	59.3	52.1	50.5	50.0
Mca	50.0	47.9	51.0		68.8	52.3	60.3	56.2	56.2	50.5
Tma	64.6	61.5	55.2	56.3		52.8	58.2	55.6	55.7	51.0
Mva	34.1	31.8	32.9	35.3	35.3		70.3	65.9	76.9	59.3
Tac	38.9	34.7	38.9	35.8	41.1	54.7		69.3	72.5	59.9
Sac	35.8	34.7	32.6	30.5	35.8	44.2	45.8		72.0	53.5
Pwo	35.8	32.6	36.8	35.8	41.1	67.4	58.4	47.9		59.0
Rat	24.2	31.6	25.3	24.2	28.4	39.5	38.5	41.7	28.1	

to argue that it is involved in this step of protein synthesis.

The identity matrices, reported in Table 1, were calculated from gap-free sites of the amino acid sequence alignment in Fig. 2 (lower left quadrant) and from a colinear alignment of the corresponding DNA sequence (first and second codon position; upper right quadrant). As expected the degree of similarity between the *S. platensis* ribosomal protein S10 and the cyanellar S10 is higher (74% amino acid identity) than the similarity between the *Spirulina* protein and that of other bacteria. A much lower degree of similarity (about 30% amino acid identity) exists between the cyanobacterial and the archaeal S10 protein. The mutual relationships between the S10 sequences are concisely illustrated by the distance tree in Fig. 3(a), constructed from DNA sequence identities using the Jukes-Cantor correction for determination of evolutionary distances, and by the tree in Fig. 3(b) constructed from DNA sequence alignment using the maximum parsimony method. In agreement with results obtained using other phylogenetic probes, such as elongation factors (Auer *et al.*, 1991; Cammarano *et al.*, 1992) and DNA polymerase core subunits (Pühler *et al.*, 1989), both DNA sequence distance and parsimony trees show a clear segregation of S10 sequences from the three primary domains, the archaeal sequences being definitely closer to the eukaryal than to the bacterial ones. Also, in both trees cyanobacterial and cyanelles form sister branches at high (95–100%) bootstrap confidence levels, and the overall tree topologies from the two methods are highly congruent, both being insensitive to removal of eukaryal and archaeal

Fig. 3. Consensus phylogenetic trees inferred from the nucleotide sequence alignment (first and second codon positions) colinear with the amino acid sequence alignment in Fig. 2. (a) Unrooted distance matrix tree calculated by the least-square method of Fitch & Margoliash (FITCH) from a set of Jukes-Cantor distance matrices (DNADIST) generated from 100 bootstrap resamplings of the DNA alignment (SEQBOOT) and (b) unrooted maximum parsimony tree constructed using DNAPARS from 100 bootstrap resamplings of the DNA sequence alignment (first and second codon positions). In both trees the figures superimposed on internal branches represent the number of bootstrap replications supporting the grouping to the right of that branch. The topologies of the consensus trees were obtained using CONSENSE.

representatives. The deep branching position of *Mycoplasma* (*Mycoplasma* not *Thermotoga* being the deepest offshoot in the bacterial branch) most probably reflects a disparity of evolutionary rates due to the reported fast evolution of the former species (Woese, 1987).

Organization of the *tuf* and *rps10* gene

Unlike those of eukaryotes, ribosomal protein genes of archaea and bacteria are arranged in transcriptional units. The organization of those operons has been

Fig. 4. Organization of the gene encoding EF-Tu and ribosomal protein S10 in cyanelles, bacteria (*S. platensis*, *E. coli*, *Myco. capricolum*, *Tg. maritima*) and archaea (*Meth. vannielii*, *Sulf. acidocaldarius*, *P. woesei*). P and T indicate promoter and terminator sites, respectively.

completely elucidated in *E. coli* (Lindahl & Zengel, 1986) and less so in the archaea *Meth. vannielii* (Auer *et al.*, 1989) and *Halobacterium marismortui* (Arndt *et al.*, 1990), but little information exists with regard to the situation in other prokaryotes. Attention has been focused mainly on *str* and S10 operons (Fig. 4). In *E. coli* the *tuf* gene is the most distal gene of the *str* operon (comprising genes for ribosomal proteins S12 and S7, the *fus* gene and the *tuf* gene, in that order), whilst the *rps10* gene is the first unit of another operon (S10 operon) located far downstream (15 kb) from the *tuf* gene. Although no information is available concerning the *tuf* gene in *Myco. capricolum*, the *rps10* gene has been cloned and sequenced together with the genes constituting the S10-*spc* operon (Ohkubo *et al.*, 1987). In this micro-organism the *rps10* gene and the *tuf* gene are spaced at least 1200 bp apart as the sequence upstream from the *rps10* gene does not exhibit any significant homology with the EF-Tu-encoding gene. In the extreme thermophile *Tg. maritima*, the *str* (Tiboni *et al.*, 1991) and S10-equivalent operons are separated by an extremely short sequence (50 bp) compared to *E. coli* (O. Tiboni

and others, unpublished data). A peculiar gene organization has been found in all of the archaea so far analysed and in the cyanelles of *C. paradoxa* where the *rps10* gene is located at the 3' end of *tuf* gene rather than at the 5' end of the S10 operon (Creti *et al.*, 1991; Neumann-Spallart *et al.*, 1991; Bryant *et al.*, 1991). Therefore, the close linkage between *rps10* and *tuf* genes in *S. platensis* resembles the situation in cyanelles and archaea and is unlike the situation in bacteria where the two genes belong to two distinct operons (Fig. 4). As cyanobacteria are evolutionarily related to cyanelles, the close association of these two genes is not unsurprising. It appears of interest that none of the other genes, constituting the S10 operon of *E. coli*, are closely linked to the *rps10* gene in *S. platensis*, as already observed in cyanelles and archaea. In fact, in cyanelles the S10 operon genes are located 2.5 kb downstream of the *rps10* gene (Neumann-Spallart *et al.*, 1990); in *Meth. vannielii* these genes are located 30 kb apart (Auer *et al.*, 1989); and in *H. marismortui* a gene cluster equivalent to the S10 operon has been cloned and sequenced but the S10-encoding gene appears to be located elsewhere in the genome (Arndt *et al.*, 1990). In *P. woesei* (Creti *et al.*, 1991) and *Sulf. acidocaldarius* (Auer *et al.*, 1991) the S10-encoding gene is linked to the *tuf* gene and immediately followed by a tRNA^{ser} gene.

This is not the first example of deviation of the cyanobacterial ribosomal protein gene organization from the *E. coli* pattern. The genes for ribosomal proteins L10 and L12, which in *E. coli* cluster in the β operon with the *rpoB* and *rpoC* genes, appear to be unlinked from the RNA polymerase subunits genes in *Synechocystis* 6803 (Sibold & Subramanian, 1990). In archaea the same genes are also widely separated (Matheson *et al.*, 1990). Similarity in gene organization between cyanobacteria and archaea has also been detected in the case of the genes encoding the larger subunit of RNA polymerase (Bergsland & Haselkorn, 1991).

Some conclusions can be drawn from the results of this study. Firstly, the strong similarity between the cyanellar and cyanobacterial S10 sequence provides further support for the endosymbiotic origin of cyanelles from a free-living cyanobacterium. Secondly, the close linkage between *rps10* and *tuf* genes suggests that this organization has been retained in the cyanelles. Thirdly, this new case (after *rpo* genes) of colinearity in gene organization between such distantly related organisms as cyanobacteria and archaea suggests the possibility that this gene arrangement represents the ancestral situation.

This work was supported by grants of the MURST and from a CNR grant Progetto Finalizzato Biotecnologie e Biostrumentazione, sottoprogetto Bioingegneria Molecolare e Cellulare. We are greatly indebted to Dr Piero Cammarano for his most generous help in deriving phylogenetic trees.

References

- ARNDT, E., KRÖMER, W. & HATAKEYAMA, T. (1990). Organization and nucleotide sequence of a gene cluster coding for eight ribosomal proteins in the archaeobacterium *Halobacterium marismortui*. *Journal of Biological Chemistry* **265**, 3034–3039.
- AUER, J., LECHNER, K. & BÖCK, A. (1989). Gene organization and structure of two transcriptional units from *Methanococcus* coding for ribosomal proteins and elongation factors. *Canadian Journal of Microbiology* **35**, 200–204.
- AUER, J., SPICKER, G., MAYERHOFER, L., PÜLHER, G. & BÖCK, A. (1991). Organization and nucleotide sequence of a gene cluster comprising the translation elongation factor EF-1 α from *Sulfolobus acidocaldarius*. *Systematic and Applied Microbiology* **14**, 14–22.
- BERGSLAND, K. J. & HASELKORN, R. (1991). Evolutionary relationships among eubacteria, cyanobacteria and chloroplasts: evidence from the *rpoC1* gene of *Anabaena* sp. strain PCC 7120. *Journal of Bacteriology* **173**, 3446–3455.
- BOLIVAR, F., RODRIGUEZ, R. L., GREENE, P. J., BETLACH, M. C., HEYNEKER, H. L., BOYER, H. W., CROSA, J. H. & FALKOW, S. (1977). Construction and characterization of new cloning vehicles II. A multipurpose cloning system. *Gene* **2**, 95–113.
- BRYANT, D. A., SCHLUCHTER, W. M. & STIREWALT, V. L. (1991). Ferredoxin and ribosomal protein S10 are encoded on the cyanelle genome of *Cyanophora paradoxa*. *Gene* **98**, 169–175.
- BUTTARELLI, F. R., CALOGERO, R. A., TIBONI, O., GUALERZI, C. O. & PON, C. L. (1989). Characterization of the *str* operon genes from *Spirulina platensis* and their evolutionary relationship to those of other prokaryotes. *Molecular and General Genetics* **217**, 97–104.
- CAMMARANO, P., PALM, P., CRET, R., CECCARELLI, E., SANANGELANTONI, A. M. & TIBONI, O. (1992). Early evolutionary relationships among known life forms inferred from elongation factor EF-2/EF-G sequences: phylogenetic coherence and structure of the archaeal domain. *Journal of Molecular Evolution* **34**, 396–405.
- CHAN, Y.-L. & WOOL, I. G. (1990). The primary structure of rat ribosomal protein S20. *Biochimica et Biophysica Acta* **1049**, 93–95.
- CORPET, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* **16**, 10881–10890.
- CRET, R., CITARELLA, F., TIBONI, O., SANANGELANTONI, A., PALM, P. & CAMMARANO, P. (1991). Nucleotide sequence of a DNA region comprising the gene for elongation factor 1 α (EF-1 α) from the ultrathermophilic Archaeote *Pyrococcus woesei*: Phylogenetic implications. *Journal of Molecular Evolution* **33**, 332–342.
- DAGERT, M. & EHRLICH, S. D. (1979). Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* **6**, 23–28.
- FELSENSTEIN, J. (1990). *Phylogenetic inference program (PHYLIP) manual version 3.3*. Seattle: University of Washington.
- HIGUCHI, R., STANG, H. D., BROWNE, J. K., MARTIN, M. O., HUOT, M., LIPELES, J. & SALSER, W. (1981). Human ribosomal RNA gene spacer sequences are found interspersed elsewhere in the genome. *Gene* **15**, 177–186.
- KAHAN, L., WINKELMANN, D. A. & LAKE, J. A. (1981). Ribosomal proteins S3, S6, S8 and S10 of *Escherichia coli* localized on the external surface of the small subunit by immune electron microscopy. *Journal of Molecular Biology* **145**, 193–214.
- KRAUS, M., GÖTZ, M. & LÖFFELHARDT, W. (1990). The cyanelle *str* operon from *Cyanophora paradoxa*: sequence analysis and phylogenetic implications. *Plant Molecular Biology* **15**, 561–573.
- LECHNER, K., HELLER, G. & BÖCK, A. (1989). Organization and nucleotide sequence of a transcriptional unit of *Methanococcus vannielii* comprising genes for protein synthesis elongation factors and ribosomal proteins. *Journal of Molecular Evolution* **29**, 20–27.
- LINDAHL, L. & ZENGEL, J. M. (1986). Ribosomal genes in *Escherichia coli*. *Annual Review of Genetics* **20**, 297–326.
- MARCK, C. (1988). 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Research* **16**, 1829–1836.
- MATHESON, A. T., AUER, J., RAMIREZ, C. & BÖCK, A. (1990). Structure and evolution of archaeobacterial ribosomal proteins. In *The Ribosome. Structure, Function & Evolution*, pp. 617–633. Edited by W. E. Hill, A. Dahlberg, R. A. Garrett, P. B. Moore, D. Schlessinger & J. R. Warner. Washington, D.C.: American Society for Microbiology.
- MENG, B. Y., SHINOZAKI, K. & SUGIURA, M. (1989). Genes for the ribosomal proteins S12 and S7 and elongation factors EF-G and EF-Tu of the cyanobacterium, *Anacystis nidulans*: structural homology between 16S rRNA and S7 mRNA. *Molecular and General Genetics* **216**, 25–30.
- NEUMANN-SPALLART, C., JAKOWITSCH, J., KRAUS, M., BRANDTNER, M., BOHNERT, H. J. & LÖFFELHARDT, W. (1991). *rps10*, unreported for plastid DNAs, is located on the cyanelle genome of *Cyanophora paradoxa* and is cotranscribed with the *str* operon genes. *Current Genetics* **19**, 313–315.
- OHKUBO, S., MUTO, A., KAWAUCHI, Y., YAMAO, F. & OSAWA, S. (1987). The ribosomal protein gene cluster of *Mycoplasma capricolum*. *Molecular and General Genetics* **210**, 314–322.
- PÜHLER, G., LEFFERS, H., GROPP, F., PALM, P., KLENK, H., LOTTSPREICH, H. P., GARRETT, R. & ZILLIG, W. (1989). Archaeobacterial DNA-dependent RNA polymerases testify to the evolution of the eukaryotic nuclear genome. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 4569–4573.
- RICCARDI, G., SORA, S. & CIFERRI, O. (1981). Production of amino acids by analog-resistant mutants of the cyanobacterium *Spirulina platensis*. *Journal of Bacteriology* **147**, 1002–1007.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SIBOLD, C. & SUBRAMANIAN, A. R. (1990). Cloning and characterization of the genes for ribosomal proteins L10 and L12 from *Synechocystis* sp. PCC 6803: comparison of gene clustering pattern and protein sequence homology between cyanobacteria and chloroplasts. *Biochimica et Biophysica Acta* **1050**, 61–68.
- SOUTHERN, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* **98**, 503–517.
- TESCH, A. & KLING, F. (1992). Nucleotide sequence of the gene coding for ribosomal protein S10 from the archaeum *Thermoplasma acidophilum*. *Nucleic Acids Research* **20**, 4090.
- TIBONI, O. & DI PASQUALE, G. (1987). Organization of genes for ribosomal proteins S7 and S12, elongation factors EF-Tu and EF-G in the cyanobacterium *Spirulina platensis*. *Biochimica et Biophysica Acta* **908**, 113–122.
- TIBONI, O., DI PASQUALE, G. & CIFERRI, O. (1984). Two *tuf* genes in the cyanobacterium *Spirulina platensis*. *Journal of Bacteriology* **159**, 407–409.
- TIBONI, O., CANTONI, R., CRET, R., CAMMARANO, P. & SANANGELANTONI, A. M. (1991). Phylogenetic depth of *Thermotoga maritima* inferred from analysis of the *fus* gene: amino acid sequence of elongation factor G and organization of the *Thermotoga str* operon. *Journal of Molecular Evolution* **33**, 142–151.
- VIEIRA, J. & MESSING, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.
- WASMANN, C. C., LÖFFELHARDT, W. & BOHNERT, H. J. (1987). Cyanelles: organization and molecular biology. In *The Cyanobacteria*, pp. 303–324. Edited by P. Fay & C. Van Baalen. Amsterdam: Elsevier.
- WHEELIS, M. L., KANDLER, O. & WOESE, C. R. (1992). On the nature of global classification. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 2930–2934.
- WOESE, C. R. (1987). Bacterial evolution. *Microbiological Reviews* **51**, 221–271.
- ZURAWSKI, G. & ZURAWSKI, S. M. (1985). Structure of the *Escherichia coli* S10 ribosomal protein operon. *Nucleic Acids Research* **13**, 4521–4526.