

Self-splicing group I intron in cyanobacterial initiator methionine tRNA: evidence for lateral transfer of introns in bacteria

Detlev Biniszkiwicz, Egle Cesnaviciene¹ and David A. Shub²

Department of Biological Sciences and Center for Molecular Genetics, University at Albany, SUNY, Albany, NY 12222, USA

¹Present address: Fermentas Institute of Biotechnology, Graiciuno 8, 2028 Vilnius, Lithuania

²Corresponding author

Communicated by I.W. Mattaj

A group I self-splicing intron has been found in the anticodon loop of tRNA^{Met} genes in three cyanobacterial genera: *Dermocarpa*, *Scytonema* and *Synechocystis*; it is absent in nine others. The *Synechocystis* intron is also interrupted by an open reading frame (ORF) of 150 codons. Of these three bacteria, only *Scytonema* also contains the group I intron that has previously been reported in tRNA^{Leu} (UAA) genes in both cyanobacteria and chloroplasts. The presence of an ORF in the tRNA^{Met} intron, the sporadic distribution of the intron among cyanobacteria and the lack of correlation between relatedness of the intron sequences and the bacteria in which they reside, are all consistent with recent introduction of this intron by lateral transfer.

Key words: bacterial introns/cyanobacteria/group I intron/intron evolution/tRNA splicing

Introduction

The origin of introns, especially the self-splicing introns belonging to groups I and II, has been the subject of a lively debate (for a concise summary see Doolittle, 1991; Roger and Doolittle, 1993). One view holds that introns are ancient, perhaps predating the existence of DNA as genetic material, while the other considers them to be relatively recent invaders of genomes. The distribution of group I introns gives some support to both viewpoints. The sporadic occurrence of these introns in homologous genes of closely related species is consistent with their mobility. A mechanism for their lateral movement is provided by the presence in many of them of open reading frames (ORFs) that encode homing endonucleases (Dujon, 1989). On the other hand, closely related group I introns are inserted in the UAA anticodon of tRNA^{Leu} genes of chloroplasts and a wide variety of cyanobacteria, which is consistent with their antiquity and bacterial origin (Kuhse *et al.*, 1990; Xu *et al.*, 1990). The absence of ORFs in tRNA^{Leu} introns also argues against their recent spread. The recent demonstration of conserved intron position in the *cox1* genes of fungal and plant mitochondria gives further support for the antiquity of some group I introns (Ohta *et al.*, 1993). Although 'introns early' and 'introns late' have been argued as mutually exclusive

scenarios, others have proposed schemes in which ORFs invade otherwise immobile introns, rendering them competent for transposition and lateral transfer (for example see Perlman and Butow, 1989; Belfort, 1991).

Group I introns capable of self-splicing *in vitro* can be detected in crude RNA extracts by end-labeling with [³²P]GTP (Garriga and Lambowitz, 1984). The group I splicing reaction is dependent on GTP as a cofactor: the 3'-OH of GTP makes a nucleophilic attack on the 5' splice site and becomes covalently attached to the 5' end of the intron during excision. We used this GTP end-labeling assay in our demonstration of introns in two species of *Anabaena* (Xu *et al.*, 1990). In that study RNA extracts of *Synechocystis* PCC 6803 were also screened, and one major end-labeled band of ~650 nt was observed. However, our subsequent attempts to amplify the putative intron-containing gene using primers designed for tRNA^{Leu}, or to clone the gene in multicopy *Escherichia coli* plasmids were unsuccessful (unpublished). We report here the occurrence of a novel intron, with a highly toxic ORF, in the tRNA^{Met} gene of *Synechocystis* sp. PCC 6803, and the presence of homologous introns in two other, distantly related cyanobacteria.

Results

An intron in Synechocystis PCC 6803

Southern blots of *Synechocystis* DNA were probed with ³²P-labeled group I introns cloned from tRNA^{Leu} genes of various cyanobacteria (Kuhse *et al.*, 1990; Xu *et al.*, 1990). A clear signal at 3.3 kb was detected when *Bam*HI-digested DNA was probed with intron sequences from *Synechococcus* (*Anacystis nidulans* in Kuhse *et al.*, 1990) (Figure 1). The ~3.3 kb band was excised from the gel and the DNA was ligated with *Bam*HI-cleaved pBSM13+. Screening of transformed colonies with the *Synechococcus* probe identified a clone that contained the same ~3.3 kb fragment (pBB2.2). Southern blots (not shown) of the digested cloned fragment placed the hybridizing sequence at one end of the genomic insert (Figure 2).

Subclones of pBB2.2 were made, extending from the *Bam*HI site on the left of Figure 2 to the *Xba*I, proximal *Nhe*I and *Nco*I sites. The sequence (Figure 3) indicated that the *Bam*HI site lies within an ORF that is followed by the core region of a group I intron (Figure 4). The sequence immediately 3' of the putative terminal G of the intron resembles the 3' half of a tRNA, starting at the CAU anticodon (Figures 3 and 5). Three G–C base pairs flanking the anticodon stem were indicated, consistent with tRNA^{Met} (RajBhandary, 1994). It appeared that the 5' portions of the tRNA, the intron and the ORF were all missing.

Assuming that the 5' part of the putative tRNA had to match the 3' part of the sequence, and that it would

resemble *Synechococcus* tRNA^{Met} (Ecarot-Charrier and Cedergren, 1976), we designed primers for amplifying the entire tRNA gene together with its group I intron by the polymerase chain reaction (PCR) (Figure 6, lane 4). An oligonucleotide primer complementary to the sequence flanking the *Bgl*II site (Figure 3) allowed separate amplification and cloning of the 5' portion of the gene. Products from two independent PCRs were cloned into pBSM13+, and their sequences were determined on both strands. The nucleotide sequence of the putative intron can be folded into a secondary structure typical for group I introns (Michel and Westhof, 1990), containing all conserved base-paired regions (P1–P9), with splice site cleavage following U (paired with a G) at its 5'-end, and G at its 3'-end, respectively (Figure 4).

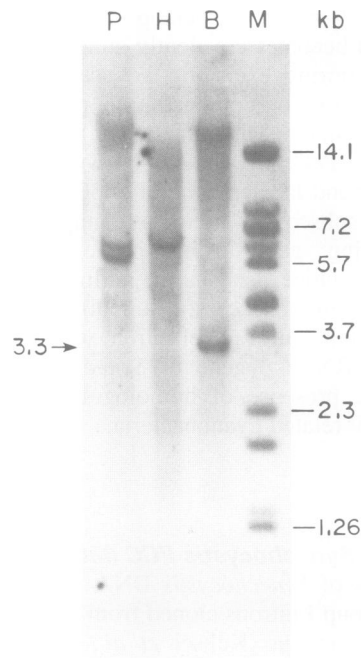


Fig. 1. Southern hybridization of genomic DNA of *Synechocystis* PCC 6803 to *Anacystis* R2 (*Synechococcus*) tRNA^{Leu} intron probe. Genomic DNA was digested with *Pst*I (P), *Hind*III (H) and *Bam*HI (B). Lane M contains a *Bst*EII digest of phage lambda DNA, with some fragments showing non-specific hybridization due to the high DNA concentration. The position of the cloned *Bam*HI fragment is indicated by an arrow.

An ORF within the intron

In contrast to other bacterial tRNA introns (Kuhnel *et al.*, 1990; Xu *et al.*, 1990; Reinhold-Hurek and Shub, 1992), whose structures conform to the minimal group I pattern, the *Synechocystis* intron has a 450 nucleotide ORF inserted in the P1 stem–loop, ending with a UGA stop codon on the 3' side of P2 (Figures 3 and 4). Database searches with BLASTP using the NCBI network service (Altschul *et al.*, 1990; Benson *et al.*, 1993) did not provide clues to the function of the ORF. The only significant match was to the 118 amino acid product of phage T7 gene 5.3 (Dunn and Studier, 1983), with 34% identity over the 74 amino acids that were aligned. Unfortunately, the function of gene 5.3 is not known. The four other weak matches that were obtained (using standard parameters) included the 63.4 kDa protein product of free-standing ORF516 of *Euglena gracilis* chloroplasts (Hallick *et al.*, 1993). The function of this protein is also unknown, but part of its sequence (which does not overlap with the short matches we obtained) has recently been shown to align with ORFs in group II introns (Mohr *et al.*, 1994). The absence of a recognizable strong ribosome binding sequence, and the possible involvement of the first AUG in a stem–loop structure ($\Delta G = -6.6$ kcal) (Figure 3), makes it unlikely that the ORF is expressed at a high level.

Expression of the ORF is apparently toxic for *E.coli*. Previous attempts at cloning the intron in multicopy plasmids were consistently unsuccessful, while the adventitious use (in this work) of a *Bam*HI site within the ORF was apparently responsible for our eventual success. The multiple cloning site of pBSM13+ is flanked on one side by a phage T7 promoter and on the other by tandem phage T3 and *E.coli lac* promoters. The PCR-amplified intron, presumably free of its promoter, was only obtained in the orientation of the T7 promoter (eight clones obtained from two independent PCRs). Even in the absence of a strong predicted ribosome binding sequence for *E.coli*, transcription from the *lac* promoter seems to be lethal. Introduction of an efficient ribosome binding site next to the T7 promoter, through cloning in plasmid pAII17 (Perler *et al.*, 1992), restores lethality: the only clone (of 15 isolates) in the correct orientation for expression carried an insertion that destroyed the reading frame (unpublished results).

As the 5' end of our sequence was derived from PCR

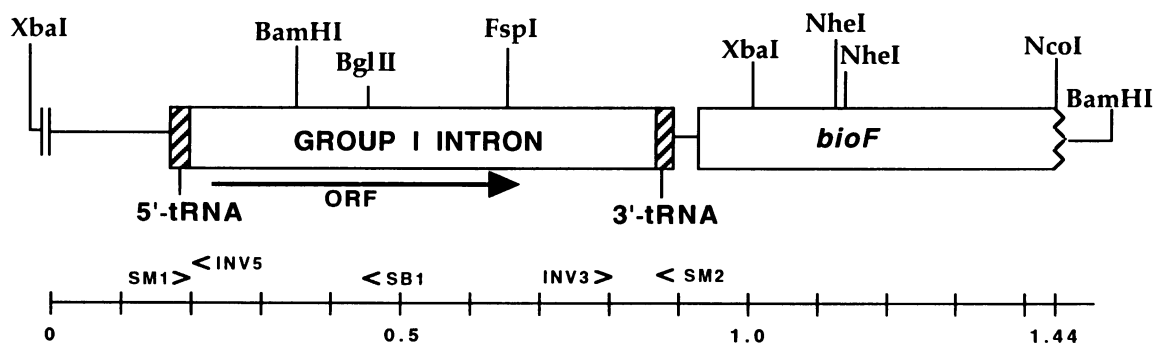


Fig. 2. Genomic location of tRNA^{Met} in *Synechocystis* PCC 6803. The tRNA gene, its group I intron and ORF, and the *bioF* fragment are presented in relationship to restriction enzyme cleavage sites mentioned in the text. Positions and directions of PCR primers are indicated by carets. The numbering (in kb) refers to Figure 3.

```

1  TTGATCGCGTCATGTCATCTAGTTGTGACCATTCAGTAAGTACCGCCCGTATACTTCG
61  AAAATCCCTAAAATTCTACTCTTCAGTGCCATAGACTATGGGATGAATCTGCCCTAAAA
121  TAAAGTTTGCCAAAATTCCCCCGATCAGTTATGATATTGGAAGCGACGCGGGATAGAGC
      5' - tRNAMet - -
-----↓-----
181  AGTCTGGTAGCTCGTCGGGCTCAGGTCGCÁAGATGTAAÁCCATGAGCACÁAGCTTAAGG
      M S T K L K G
-----
241  GTGACATTGCAGAGCAAGCCGCCATTCTTCTGTCATTGAÁGCTTGGTTGGGGGTTCTGÁ
      8 D I A E Q A A I L R A L K L G W G V L K
-----
301  AGCCATTGGGAGACCGTCTTTCTTACGATTGGTCTTTGÁTGTCGAAGGÁATCCTTCTGÁ
      28 P L G D R L S Y D L V F D V E G I L L K
      BamHI
361  AGGTGCAGGTTAAAAGCTCCTGGTTCAGCGAAAAGACTGGCAACTATGTÁGTTGACAATC
      48 V Q V K S S W F S E K T G N Y V V D N R
-----
421  GCAGAACAÁGGACAAACAGÁAGAAACATTGTTAGATCTCCTTATCGTGGÁATGACTTTG
      68 R T R T N R R N I V R S P Y R G N D F D
      BglII
481  ATTTTGTGTTGCCTACGTÁGAAACTTÁGCTGTTTTÁCGTCTTCCCGTTGATGTCÁ
      88 F A V A Y V E E L E L F Y V F P V D V F
541  TTATCAGTTÁCGGCAGTGAÁATTCACCTGÁTCGAAACTGÁTAAGCGTCAÁGAGAAAACCAÁ
108  I S Y G S E I H L V E T D K R Q R K P R
-----
601  GGTCATTCGÁTTACCGAGAGGCTTGGCATCTCATCTGCAAAAAGGGGCTGCGCAAAAAGG
      128 S F G Y R E A W H L I L Q K G A A Q K E
      FspI
661  AAÁCTTCTGÁTGATTATCTCTCAAATTCÁGGGAAGCCTTCAÁATGGTÁATCCCGAGCC
      148 T S A *
721  AAACCTAGGÁATGCTTGGTÁTTTCTGGGAÁGGTGTAGAGÁCTTAATGGGÁGACACCCTAA
781  CAGAAAAGCTGAGGGTGAAGÁGAAAGTCCÁGACCACAAAÁCTGACAGAGTÁAGGCAGTGAÁ
      ↓----- tRNAMet - 3' -----
841  AACTGTAGTTGGTAAGCATÁACCCGAAGGTCGGTGGTTCAÁATCCGCCTÁCCGCCATTAA
-----
901  AAAAAAAGÁAAAAATTCÁCCCGCCGGAÁATGTTGGGGGGTTGTTTÁATCAATCT
      1 M V G G F V L F N L
      bioF
961  AAAGTTTATGCTACTGGTÁCCACTGCCTÁCACTGGCTÁGACGATGCTCTAGAAACTAT
      11 K F M A T G S T A Y T W L D D A L E T I
      XbaI
1021  TCAGCGGGCCÁATTGGCATÁGCCATCCCCÁGATAATTACÁCAAGGGCCÁGGCCAGAAAT
      31 Q R A H W H R H P Q I I T Q G P G P E I
-----
1081  TAAGTTGGAÁGGGCAACGGCTAGTTAÁTTTGGCTAGCAATGATTATCTTGGCTTAGCTAG
      51 K L E G Q R L V N F A S N D Y L G L A S
      NbeI NbeI
1141  CCATCCCCÁCTAAAACAGCGGCCATTAÁGGCGATCGCCGAATGGGGAÁCCGGCAGTÁC
      71 H P H L K T A A I K A I A E W G T G S T
1201  AGGTTCCCGÁTTGCTGAGTGGCCACCGTCÁGCTCCATCAÁGACTTGGAAÁAGGCCATTGC
      91 G S R L L S G H R Q L H Q D L E Q A I A
1261  CCGCTGGAÁGGGACGGAÁGCGCTTGGTGTTCAGTTCTGGCTATTGGCTAÁTTGGG
      111 R W K G T E A A L V F S S G Y L A N L G
1321  CACCATCACÁGCGTTGGTGGGTAÁACGGGÁTTTGATTTTÁGCGGACGAÁTATAACCATTÁ
      131 T I T A L V G K R D L I L A D E Y N H S
-----
1381  CAGTCTGAÁGCGTGGAGCCÁAGTTGAGTGGÁGACCAAGGTÁATTAÁCTATÁCCATGG
      151 S L K R G A Q L S G A K V I N Y D H G
      NcoI

```

Fig. 3. Nucleotide sequence of tRNA^{Met} gene and surrounding regions (GenBank accession number U10482). Splice sites are indicated by arrows. Protein coding sequences are represented beneath the DNA sequence. Restriction enzyme recognition sites are underlined, while overlines indicate self-complementary regions discussed in the text.

primers, we could not be sure that the genomic sequence resembled a functional tRNA. To resolve this uncertainty, inverse PCR was used to obtain the complete split gene and flanking sequences. Southern blots of *Synechocystis* PCC 6803 DNA, using the *Bam*HI–*Xba*I fragment of pBB2.2 as probe, revealed an *Xba*I fragment of ~1.6 kb that was large enough to contain the entire gene (see Figure 2). After diluting and religating the genomic *Xba*I digest, primers that start within the intron were used to amplify the surrounding regions and the 1.1 kb amplifica-

tion product was cloned into pBSM13+. The sequence flanking the intron can be folded into a characteristic cloverleaf secondary structure (Figure 5) which conforms in every respect to conserved features of eubacterial tRNA^{Met} (RajBhandary, 1994). There are only seven changes between the DNA of the *Synechocystis* gene and the RNA sequence of tRNA^{Met} from *Synechococcus* (*Anacystis nidulans*; Ecarot-Charrier and Cedergren, 1976). As with tRNA^{Trp}, the only other tRNA whose gene has been sequenced in *Synechocystis* PCC 6803 (Schmidt

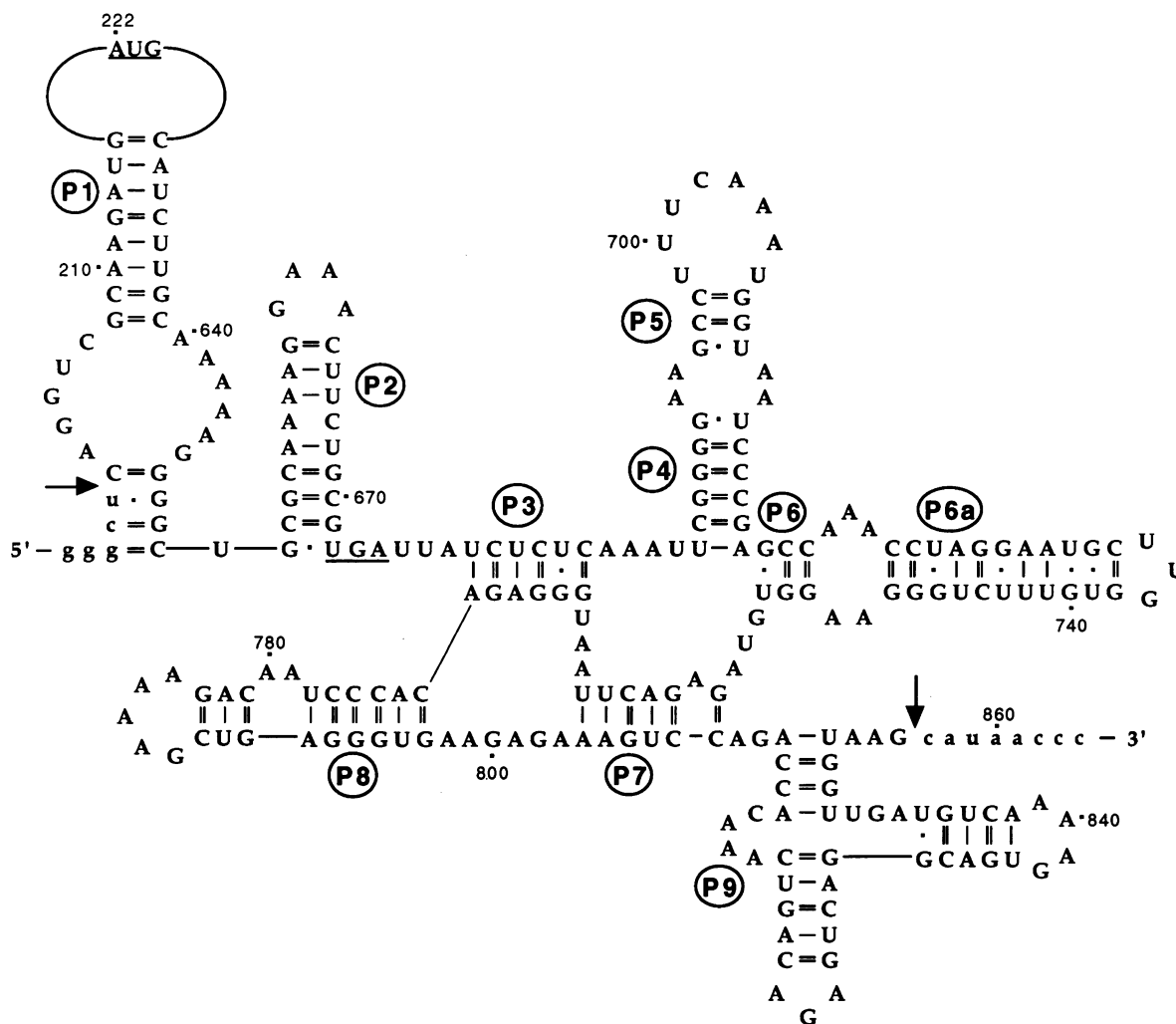


Fig. 4. Secondary structure of the *Synechocystis* PCC 6803 tRNA^{Met} intron. Exon sequence is in lowercase and intron sequence in uppercase, with arrows indicating the splice boundaries. Conserved structural elements P1–P9 are indicated. Start and stop codons for the intron ORF are underlined. Numbering corresponds to Figure 3.

and Subramanian, 1993), the 3'-terminal CCA sequence of tRNA^{Met} is not encoded in DNA.

Another ORF begins downstream of the tRNA gene, at position 932, and continues for 169 codons until the sequence ends at the *Nco*I site (Figure 3). It is not clear whether the AUG of the first or thirteenth codon is the translation start, as no obvious ribosome initiation sequence is apparent. Search of protein databases with BLASTP revealed a strong match (without gaps in the alignment) to the *bioF* genes of *Bacillus subtilis* (42% identity starting with codon 27) and *E.coli* (36% identity starting at codon 23). *bioF* encodes 8-amino-7-oxononanoate synthase, the first enzyme of biotin biosynthesis. The region between these genes is a long interrupted palindrome, a G+C-rich stem flanked 5' by a string of A and 3' by a string of U, with potential to serve as a factor-independent termination signal in both transcriptional directions. This may function to uncouple the biosynthesis of biotin operon mRNA from tRNA^{Met}.

Group I introns in tRNA^{Met} genes of other cyanobacteria

Although bacterial tRNAs^{Met} have relatively few invariant nucleotides, all have three consecutive G–C base pairs

closing the anticodon stem (RajBhandary, 1994). As our primers end at the base of the anticodon stem, we hoped that they would also amplify this tRNA gene from other bacteria. The expected size for PCR products of an uninterrupted gene is 71 bp, while larger PCR products would indicate a potential intron. Amplification of cyanobacterial DNA from *Xenococcus* sp. (ATCC 29373), *Synechococcus* R2, *Anabaena* sp. (PCC 7120), *Lyngbya kuetzingi* (UTEX 1547), *Chlorogloeopsis fritschii* (ATCC 27193), *Calothrix* sp. (PCC 7101), *Fischerella ambigua* (UTEX 1903) and *Chamaesiphon* sp. (PCC 7430) result in a single band of ~70 bp (not shown). The *Synechococcus* amplification product (Figure 6, lane 3) was cloned and sequenced, confirming an uninterrupted anticodon loop identical in sequence to that previously determined for its tRNA^{Met} (Ecarot-Charrier and Cedergren, 1976). However, PCRs with genomic DNAs of *Scytonema hofmanii* (UTEX 2349) and *Dermocarpa* sp. (ATCC 29371) resulted in amplification products of ~310 and 320 bp, respectively (Figure 6, lanes 5 and 6). As was the case with *Synechocystis*, no tRNA-sized PCR product was visible in these reactions, consistent with the absence of an uninterrupted gene in these organisms.

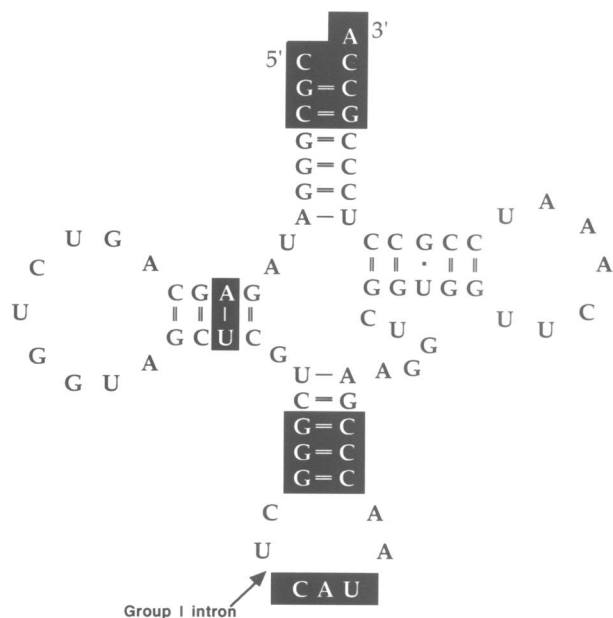


Fig. 5. tRNA^{fMet} structure. Complete *Synechocystis* tRNA^{fMet} sequence determined from inverse PCR. Features diagnostic of tRNA^{fMet} are shown as white against a black background. The intron insertion site is indicated by an arrow.

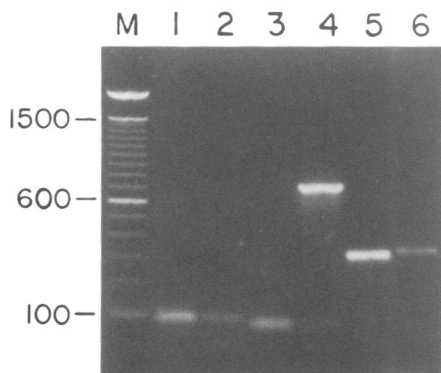


Fig. 6. PCR amplification products of cyanobacterial tRNA genes. *Synechocystis* (lane 1) and *Dermocarpa* (lane 2) DNA amplified with tRNA^{Leu} (UAA) primers. Lanes 3–6: tRNA^{fMet} primers using DNA from *Anacystis* R2 (*Synechococcus*) (lane 3), *Synechocystis* (lane 4), *Scytonema* (lane 5) and *Dermocarpa* (lane 6). M: 100 bp DNA marker (Gibco/BRL).

We sequenced the cloned amplification products from two or three independent PCRs of *Scytonema* and *Dermocarpa* DNA, respectively. Both have group I introns (Figure 7) whose insertion sites are consistent with the anticodon loop of tRNA^{fMet} genes, 5' of the anticodon. These new introns are 244 and 258 bp long, respectively, and are very similar to the *Synechocystis* intron in secondary structure as well as in sequence, displaying in addition the P10 interaction typical of most group I introns (not shown). Although the P1 stem-loop structures are more extended than is the case for the other bacterial tRNA introns, these introns do not have an ORF inserted into them. Alignment of the sequences from the beginning of P2 to the 3' end of these introns revealed that the *Scytonema* and *Dermocarpa* introns share >90% identity over 195 nucleotides. The main difference between these two introns is the variable length of P9. The *Synechocystis*

intron is more distantly related to the other two, with 76% and 74% identity, respectively. This result is not expected, as *Dermocarpa* is more closely related to *Synechocystis* than to the filamentous, heterocystous *Scytonema* (Rippka *et al.*, 1979; Giovannoni *et al.*, 1988).

As all other cyanobacteria so far investigated have introns in their tRNA^{Leu} (UAA) genes, we were curious to determine whether this is also true of the strains bearing introns in tRNA^{fMet}. Neither the *Synechocystis* nor the *Dermocarpa* tRNA^{Leu} (UAA) gene is disrupted by an intron. Rather, sequences of the cloned tRNA-sized amplification products (Figure 6, lanes 1 and 2) display the anticodon stem-loop and long extra arm characteristic of tRNA^{Leu} cloverleaf structures (Figure 8A and B). The *Scytonema* tRNA^{Leu} (UAA) gene was already reported to be interrupted by a group I intron (Kuhnel *et al.*, 1990), which we confirmed by the size of its amplification product (not shown). Thus, *Scytonema* is the only bacterium so far reported to have two genes divided by group I introns.

No group I introns detected in tRNA^{fMet} genes of other bacterial phyla

The primers used to amplify cyanobacterial tRNA^{fMet} genes were also used to amplify DNA from bacteria in other phyla, with no evidence so far for the presence of an intron (not shown). We detected a tRNA-sized product in all cases and, where additional more slowly migrating bands were also present, they turned out to be clusters of two or more tRNA^{fMet} genes (D.Biniszkiewicz and D.A.Shub, in preparation). While these results do not exclude finding split tRNA^{fMet} genes in diverse phyla, they are clearly not common.

Discussion

A group I intron in a novel bacterial gene

The few previously described examples of group I introns in bacteria are located within, or immediately 3' of the anticodons of tRNA^{Leu} (UAA) genes in various cyanobacteria (Kuhnel *et al.*, 1990; Xu *et al.*, 1990), and tRNA^{Arg} (CCU) and tRNA^{Ile} (CAU) in α - and β -proteobacteria, respectively (Reinhold-Hurek and Shub, 1992). The intron described in this work was originally detected on the basis of the GTP end-labeling reaction, characteristic of the group I splicing mechanism (Xu *et al.*, 1990), independent of any preconception of the gene it might inhabit. Yet, it is also located in the anticodon loop of a tRNA gene, in this case inserted 5' of the first anticodon nucleotide of tRNA^{fMet} in *Synechocystis* and two other cyanobacterial genera: *Dermocarpa* and *Scytonema*. Interestingly, like all other group I introns in tRNA genes, they are closely related members of the IC3 subfamily (Michel and Westhof, 1990). The tRNA^{fMet} introns described here are distinguished from the others by an extended P1 stem, which in *Synechocystis* is expanded into an ORF of 150 codons.

The *Synechocystis* intron ORF

Most ORFs that are found within group I introns are site-specific endonucleases that cleave double-stranded DNA near the intron insertion site. DNA that contains the intron is not cleaved, as the endonuclease recognition sequence is split. Subsequent repair of the cleaved gene, using its

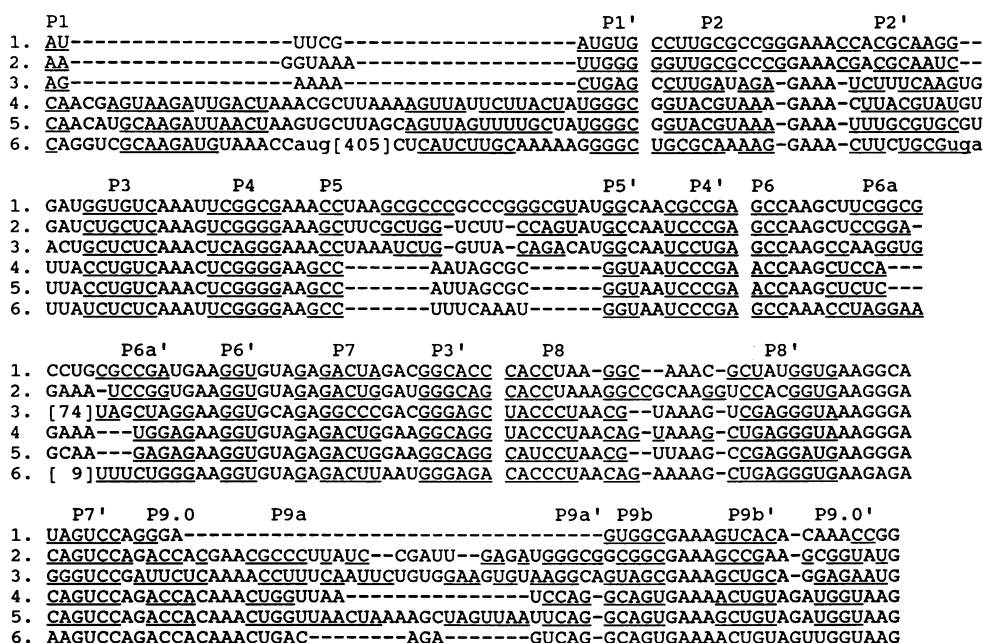


Fig. 7. Alignment of bacterial tRNA introns. (1) *Azoarcus* tRNA^{Ile}, (2) *Agrobacterium* tRNA^{Arg} (Reinhold-Hurek and Shub, 1992) and (3) *Scytonema* tRNA^{Leu} (Kuhse *et al.*, 1990) intron sequences are aligned with tRNA^{Met} introns from (4) *Scytonema* (GenBank accession number U10481), (5) *Dermocarpa* (U10480) and (6) *Synechocystis*. Underlining indicates base-paired portions of conserved secondary structural elements (see Figure 4). Numbers in brackets indicate residues not shown. Lower case letters indicate the start and stop codons of the *Synechocystis* intron ORF.

intact homolog as template, results in acquisition of the intron and co-conversion of nearby genetic markers, a phenomenon that has been called intron homing (Dujon, 1989; Lambowitz and Belfort, 1993). Although we are not able to assign a function to the *Synechocystis* intron ORF, its apparent toxicity is consistent with a presumed function as a homing endonuclease. At 150 amino acids it is only slightly smaller than the smallest of these enzymes.

The conserved nature of the anticodon stem and loop of bacterial tRNA^{Met} genes makes it a likely target for nuclease-containing introns, which tend to occupy genomic locations that do not tolerate sequence variation. Once inserted in such a highly conserved site the intron is difficult to dislodge: exact deletion merely provides a substrate for the homing endonuclease, whereas deletions that simultaneously change the homing site are deleterious (Shub and Goodrich-Blair, 1992). In contrast to the tRNA^{Leu} introns—which are maintained in widely divergent cyanobacteria and chloroplasts in spite of the absence of an ORF to provide the means for mobility—the introns in tRNA^{Met} are found only sporadically. Moreover, *Dermocarpa*'s intron is most similar to the one in *Scytonema*, in spite of the much closer evolutionary relatedness to *Synechocystis* (Giovannoni *et al.*, 1988). These observations are consistent with the hypothesis that the tRNA^{Met} introns are recent examples of lateral gene transfer in the cyanobacteria.

Although all bacterial tRNA introns are closely related (Reinhold-Hurek and Shub, 1992), the three tRNA^{Met} introns are more closely related to each other than they are to any of the others, including the tRNA^{Leu} introns in cyanobacteria (D.A.Shub, in preparation). This is true even in the case of *Scytonema*, which is the only example, so far, of a bacterium with two group I introns. One of

these is apparently ancient and is preserved due to an unknown selective value, while the other was probably obtained more recently through lateral transfer from a distant species.

Materials and methods

DNA preparation

Synechocystis and *Synechococcus* cultures were grown as previously described (Xu *et al.*, 1990). Cells obtained from 400 ml cultures were resuspended (10 ml/g wet cells) in STET buffer (0.1 M NaCl, 10 mM Tris-Cl, pH 8.0, 1 mM EDTA, 5% Triton X-100) containing 10 mg/ml lysozyme, and incubated for 1 h at 37°C. Pronase (1 mg/ml) and SDS (0.5%) were added and incubation continued for another hour, after which the SDS concentration was increased to 2% and incubation continued for an additional hour. The viscous, translucent suspension was subjected to standard phenol and phenol/chloroform extractions until the protein interface disappeared, and extracted once with chloroform. DNA was precipitated with 2 vols of ethanol, collected by spooling on a sterile glass rod, air dried and dissolved in 10 mM Tris-Cl (pH 8.0), 1 mM EDTA. Other cyanobacterial DNAs were the gift of S.Nierzwicki-Bauer.

DNA filter hybridization

Five micrograms of DNA from *Synechocystis* PCC 6803 were digested with restriction enzymes, and the resulting fragments were separated on a 1.0% agarose gel and vacuum blotted onto a positively charged nitrocellulose membrane (Hybond-N⁺, Amersham). Before blotting, the agarose gel was incubated for 20 min in 0.25 M HCl. The vacuum transfer solution contained 0.4 M NaOH and 0.6 M NaCl. After neutralization with 5 × SSC (5 min) and prehybridization for 1 h. at 50°C, the membrane was hybridized to the probe for 15 h at 50°C in 6 × SSC, 0.1% SDS, 5% dextran sulfate, 5 × Denhardt's solution, 10 mM EDTA, 100 µg/ml heparin and 10 µg/ml yeast RNA (type III, Sigma). The radiolabeled probe was generated using random hexamer primers (Feinberg and Vogelstein, 1983), and purified on Sephadex G50 columns. After incubation with the radiolabeled probes, all membranes were washed three times for 10 min in 6 × SSC and 0.1% SDS.

Polymerase chain reaction

PCRs were performed using 0.1–0.5 µg DNA, 20 pmol of each primer, 200 µM of each deoxynucleotide triphosphate, 2 U of *Taq* DNA

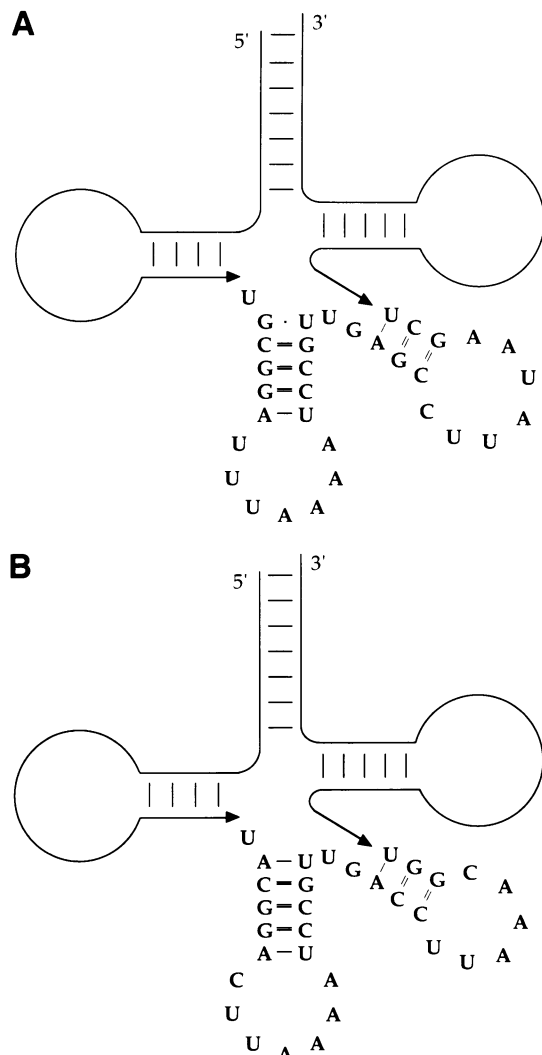


Fig. 8. Absence of introns in tRNA^{Leu} (UAA) genes of (A) *Synechocystis* and (B) *Dermocarpa*. Sequences of PCR amplification products using primers for tRNA^{Leu} (UAA) are shown. PCR primers are shown as solid lines.

polymerase and 1 × *Taq* buffer (Boehringer Mannheim) containing 1.5 mM MgCl₂. The reaction mixtures were heated to 92°C for 1 min, 48°C or 40°C (for tRNA^{Leu}) for 1 min, and 72°C for 2 min. After incubation for 30 cycles, one final extension step at 72°C for 10 min was added. PCR products were separated on a 1.8% agarose gel using a 0.5 mm comb.

Primers SM1 (5'-GCGGGGTAGAGCAGCCTGGTACCTCGTCGGG-3') and SM2 (5'-GCGGGATCCGGATTGAAACCACCGACCTTCGGG-3') were used to amplify tRNA^{Met}. The 5' part of the *Synechocystis* tRNA^{Met} gene was amplified using primers SM1 and SB1, which is complementary to residues 442–470 of the sequence, spanning a *Bgl*II site (Figure 3). Primers INV3 (5'-CAGAAAAGCTGCAGGTGAAGA-GAAAGTCC-3') and INV5 (5'-GCTTAGTGCTCATGGTCTAGATCT-TGCGAC-3') were used for inverse PCR. Positions of primers for PCR of tRNA^{Met} genes are shown in Figure 2. Primers TL25 [5'-GGGG(G/A)T(G/A)TGG(C/T)G(A/G)AAT(C/T)JGGTAGACGC-3'] and TL23 [5'-TGGGG(G/A)(C/T)(G/A)G(G/A)GGACTTGAACCC(C/A/T)CAGC-3'] were used to amplify tRNA^{Leu} (UAA) genes.

Cloning and sequencing

Plasmid pBSM13+ (Stratagene) was used as the vector for all cloning experiments. pBB2.2 was constructed by purifying the 3.3 kb fraction of a *Bam*HI digest of *Synechocystis* PCC 6803 DNA and ligating it into the *Bam*HI site. Subclones were obtained from the upstream *Bam*HI site to the *Xba*I, proximal *Nhe*I or *Nco*I sites and from the *Bgl*II or *Xba*I sites to the downstream *Bam*HI site (Figure 2). PCR amplification

products of the entire *Synechocystis* intron were cloned into the vector *Hinc*II site. Additionally, the amplification products were digested with *Bgl*II and *Fsp*I, and the fragments from *Bgl*II to *Fsp*I, and *Fsp*I to the 3' end of the tRNA (Figure 2) were cloned. The 5' part of the intron was cloned separately after amplification using primers SM1 and SB1.

All sequences were obtained by the dideoxy chain termination method, on both DNA strands. Intron sequences were from clones obtained from two independent PCRs. In the case of *Dermocarpa* tRNA^{Met}, a third independent clone was used to resolve a discrepancy. Overlap between sequences of PCR clones of the *Synechocystis* intron was confirmed on one strand of genomic subclones. The 3' exon and *bioF* sequences were obtained from genomic subclones, while the 1.13 kb inverse PCR product was used to determine the 5' exon and upstream flanking sequence.

Inverse PCR

Synechocystis PCC 6803 DNA was digested with restriction endonucleases and Southern blots were probed with the *Bam*HI–*Xba*I fragment of pBB2.2. A hybridizing band of 1.6 kb was observed for the *Xba*I digest, which is large enough to contain the entire gene plus flanking regions (Figure 1). *Xba*I-digested and religated DNA was amplified with primers INV5 and INV3 (Figure 2). PCR conditions were: 92°C for 1 min, 42°C for 1 min and 72°C for 2 min, and after 30 cycles, a final extension step at 72°C for 10 min. Products were separated on a 1.2% agarose gel, and the band of the expected size (1.1 kb) was extracted for cloning.

Acknowledgements

We are very grateful to Sandra Nierzwicki-Bauer for cyanobacterial DNAs, and to Carole Keith for help with the figures. This work was supported by NIH grant GM37746.

References

- Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) *J. Mol. Biol.*, **215**, 403–410.
- Belfort,M. (1991) *Cell*, **64**, 9–11.
- Benson,D., Lipman,D.J. and Ostell,J. (1993) *Nucleic Acids Res.*, **21**, 2963–2965.
- Doolittle,W.F. (1991) *Curr. Biol.*, **1**, 145–146.
- Dujon,B. (1989) *Gene*, **82**, 91–114.
- Dunn,J.J. and Studier,F.W. (1983) *J. Mol. Biol.*, **166**, 477–535.
- Ecarot-Charrier,B. and Cedergren,R.J. (1976) *FEBS Lett.*, **63**, 287–290.
- Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
- Garriga,G. and Lambowitz,A.M. (1984) *Cell*, **38**, 631–641.
- Giovannoni,S.J., Turner,S., Olsen,G.J., Barns,S., Lane,D.J. and Pace,N.R. (1988) *J. Bacteriol.*, **170**, 3584–3592.
- Hallick,R.B., Hong,L., Drager,R.G., Favreau,M.R., Monfort,A., Orsat,B., Spielmann,A. and Stutz,E. (1993) *Nucleic Acids Res.*, **21**, 3537–3544.
- Kuhnel,M.G., Strickland,R. and Palmer,J.D. (1990) *Science*, **250**, 1570–1573.
- Lambowitz,A.M. and Belfort,M. (1993) *Annu. Rev. Biochem.*, **62**, 587–622.
- Michel,F. and Westhof,E. (1990) *J. Mol. Biol.*, **216**, 585–610.
- Mohr,G., Perlman,P.S. and Lambowitz,A.M. (1993) *Nucleic Acids Res.*, **21**, 4991–4997.
- Ohta,E., Oda,K., Yamato,K., Nakamura,Y., Takemura,M., Nozato,N., Akasky,K., Ohya,K. and Michel,F. (1993) *Nucleic Acids Res.*, **21**, 1297–1305.
- Perler,F.B. et al. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 5577–5581.
- Perlman,P.S. and Butow,R.A. (1989) *Science*, **246**, 1106–1109.
- RajBhandary,T.L. (1994) *J. Bacteriol.*, **176**, 547–552.
- Reinhold-Hurek,B. and Shub,D.A. (1992) *Nature*, **357**, 173–176.
- Rippka,R., Deruelles,J., Waterbury,J.B., Herdman,M. and Stanier,R.Y. (1979) *J. Gen. Microbiol.*, **111**, 1–61.
- Roger,A.J. and Doolittle,W.F. (1993) *Nature*, **364**, 289–290.
- Schmidt,J. and Subramanian,A.R. (1993) *Nucleic Acids Res.*, **21**, 2519.
- Shub,D.A. and Goodrich-Blair,H. (1992) *Cell*, **71**, 183–186.
- Xu,M.-Q., Kathe,S.D., Goodrich-Blair,H., Nierzwicki-Bauer,S.A. and Shub,D.A. (1990) *Science*, **250**, 1566–1570.

Received on May 10, 1994; revised on July 2, 1994