Volume 325, number 3, 255–261 FEBS 12633 July 1993
© 1993 Federation of European Biochemical Societies 00145793/93/$6.00

The genes *aroA* and *trnQ* are located upstream of *psbO* in the chromosome of *Synechocystis* 6803

Steve R. Mayes*, Marta Dalla Chiesa, Zhihong Zhang** and James Barber

AFRC Photosynthesis Research Group, Wolfson Laboratories, Department of Biochemistry, Imperial College of Science, Technology and Medicine, London, SW7 2AY, UK

Received 5 May 1993

We have identified the existence of two genes, *trnQ* and *aroA*, located upstream of the *psbO* gene in *Synechocystis* sp. PCC 6803. The *trnQ* gene encodes a glutamine-specific transfer RNA ($tRNA^{Gln}$) and the sequence given is the first reported for any cyanobacterium. The gene seems to exist as a single copy since its deletion results in non-viable mutation. The *aroA* gene encodes for 5-enolpyruvylshikimate 3-phosphate synthase and its discovery in the genome of *Synechocystis* 6803 is the first genetic evidence for the existence of the shikimate biosynthetic pathway in cyanobacteria. Interestingly, the partial sequence shares close homologies with the sequences of *aroA* from Gram-positive bacteria.

5-Enolpyruvylshikimate 3-phosphate synthase; *aroA* gene; Shikimate pathway; Transfer ribonucleic acid; *trnQ* gene; Photosystem II; *psbO* gene; *Synechocystis* 6803

1. INTRODUCTION

The *psbO* gene encodes the so-called '33-kDa extrinsic protein' of photosystem II (PSII), also known as the 'manganese-stabilising protein' (MSP). The PsbO protein is an important component of the multi-subunit PSII complex, binding in close proximity to the Mn cluster on the lumenal side of the thylakoid membrane in both cyanobacteria and oxygenic photosynthetic eukaryotes [1].

The cyanobacterium *Synechocystis* spp. PCC 6803 has proved particularly useful in the study of PSII function, owing to the relative ease with which targeted mutations may be introduced into genes encoding subunits of PSII in this organism. Previously we determined that the *psbO* gene occurs as a single copy in *Synechocystis* 6803, and reported the in vivo inactivation of this gene to generate the insertion mutant IC1 and the deletion mutant IC2 [2]. These mutants were then characterised for their phenotypic properties [2,3]. Together with other researchers [4,5], we found that the PsbO protein optimises PSII-driven oxygen evolution, but is not absolutely required for PSII activity. Similar results concerning the function of the PsbO protein were reported for another cyanobacterium, *Synechococcus* spp. PCC 7942 [6].

In this paper, we firstly describe attempts to delete a larger region of the *Synechocystis* 6803 chromosome containing *psbO* than was successfully achieved in mutant strain IC2. This experiment led to the discovery of *trnQ* and *aroA* genes upstream of *psbO*. In particular the identification of an *aroA* gene in *Synechocystis* 6803 is significant as it is the first genetic evidence for the existence of the shikimate biosynthetic pathway in cyanobacteria. We also report a physical map of 24 kb of the *Synechocystis* 6803 chromosome containing *aroA*, *trnQ* and *psbO*.

Correspondence address: J. Barber, AFRC Photosynthesis Research Group, Wolfson Laboratories, Department of Biochemistry, Imperial College of Science, Technology and Medicine, London, SW7 2AY, UK.

**Present address:* Stratagene Ltd., Cambridge Innovation Centre, 140 Science Park, Cambridge, CB4 4GF, UK.

***Present address:* Dyson Perrins Laboratory, South Parks Road, Oxford, OX1 3QY, UK.

Abbreviations: ATP, adenosine 5′-triphosphate; EPSP, 5-enolpyruvylshikimate 3-phosphate synthase; kb, kilobase; Kmr, kanamycin resistance cassette; PSII, photosystem II; $tRNA^{Gln}$, glutamine-specific transfer RNA; $tRNA^{Glu}$, glutamate-specific transfer RNA.

2. MATERIALS AND METHODS

Routine DNA manipulations were carried out as outlined in [2]. The glucose-tolerant strain, *Synechocystis* 6803-G [50], was used throughout these studies and is referred to as *Synechocystis* 6803 or the wild-type. This strain of *Synechocystis* spp PCC 6803 was kindly provided by Dr. J.G.K. Williams (DuPont, Wilmington, DE). The routine maintenance of this strain on BG-11-based medium on agar plates and in liquid culture was as described in [2].

Transformation of *Synechocystis* 6803 with *psbO* deletion construct pSMO8 was performed as described by Mayes et al. [2]. After transformation, kanamycin-resistant cell lines were subjected to multiple rounds of re-streaking from single colonies onto BG-11 agar plates supplemented with 5 mM glucose, 100 $\mu g \cdot ml^{-1}$ kanamycin and 20 μM atrazine.

^{32}P-Labelling of DNA probes used in Southern analysis was accom-

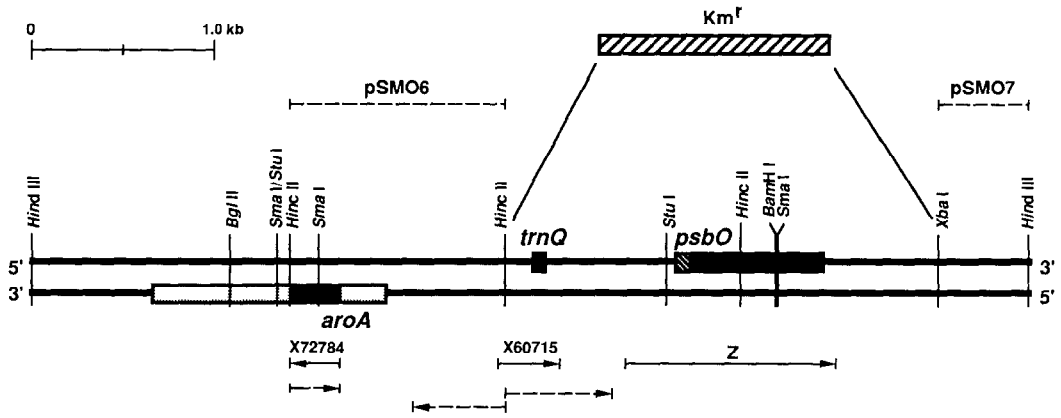


Fig 1. Restriction map of 5 5-kb of the *Synechocystis* 6803 chromosome containing *psbO*. Gene positions are indicated by boxed regions. The hatched part of *psbO* encodes the transit peptide. Also indicated is the attempted replacement of the 2.4-kb *Hinc*II–*Xba*I fragment by the Km^r cartridge that is described in the text. Dashed lines above the map show the extent of the inserts of plasmids pSMO6 and pSMO7. Arrowed horizontal lines beneath the map show the orientation and extent of sequencing data. Dashed lines indicate preliminary sequencing runs from one DNA strand. Named solid lines correspond to data sequenced from both strands and submitted to the EMBL database. The region sequenced by Philbrick and Zilinskas [33] is denoted 'Z'. The sequenced region of the putative *aroA* gene is shown by a solid box. The shaded box region shows the predicted extent of the *aroA* gene, assuming that it is the same length as the *aroA* gene of *B subtilis* [14].

plished by extension of random oligonucleotide primers, essentially as described by Feinberg and Vogelstein [7]. For the Southern blot shown in Fig. 1, samples of genomic DNA were isolated by the miniprep method for *Synechocystis* 6803 reported in [2]. Restricted DNA was fractionated on a 0.8% agarose gel and capillary-transferred to BA 85 nitrocellulose (Schleicher and Schuell, 0.45 μm pore size) Prehybridisation for 2 h, and then overnight hybridisation, were performed in 10 ml of 5 × SSC, 5 × Denhardt's, 0 5% SDS, 100 $\mu g \cdot ml^{-1}$ calf thymus DNA (solutions defined in [2]) at 65°C. The filter was then washed to a final stringency of 0.1 × SSC, 0.1% SDS at 65°C DNA to be sequenced was subcloned into M13 mp18 and mp19 vectors [8] Dideoxy DNA sequencing of single-stranded template was performed using the Sequenase cloning kit (United States Biochemical Corp.) using either the universal sequencing primer or custom-made primers synthesised on a Pharmacia LKB Gene Assembler Plus machine. DNA and deduced protein sequences were analysed using the PC/Gene microcomputer software (formerly Genofit, now Intelligenetics Inc.). Homology searches of the EMBL, Genbank and SwissProt nucleic acid/protein databases were performed using the University of Wisconsin Computing Group (UWGCG, Devereux et al., 1984) FASTA programme (which employs the algorithm of Lipmann and Pearson, 1985) running on the AFRC VAX Computer Service (Harpenden, UK). Multiple sequence alignments were assembled using the CLUSTAL programme [9], also installed on the AFRC VAX Computer.

The phage clones used to map DNA surrounding *psbO* were isolated from a genomic library constructed by Stratagene (USA) for Dr. J.G.K. Williams. The library (previously described in [2,10,20]) was constructed by ligating *Synechocystis* 6803 genomic DNA partially cut by *Sau*3AI into the *Bam*HI site of the λEMBL3 cloning vector [11].

3. RESULTS

We have previously reported a restriction map of plasmid pSMO1 which contains a 5.5 kb fragment of *Synechocystis* 6803 containing the entire *psbO* gene plus upstream and downstream flanking DNA. Early on in our studies on *psbO*, before the successful generation of *psbO* deletion mutant IC2 (see [2]), we attempted to replace the 2.4 kb *Hinc*II–*Xba*I fragment containing *psbO* with the kanamycin resistance (Km^r) cartridge from pUC4K (variant described in [12]). The plasmids pSMO6 and pSMO7 were used as the starting points for making the *psbO* deletion construct pSMO8. Plasmid pSMO6 consisted of the 1.2 kb *Hinc*II–*Hinc*II fragment upstream of *psbO* ligated into plasmid vector pUC18; plasmid pSMO7 consisted of the 0.5 kb *Xba*I–*Hind*III fragment downstream of *psbO* inserted into vector pBS$^+$ (Stratagene, USA). The extent of the pSMO6 and pSMO7 inserts are shown in Fig. 1. In a series of cloning

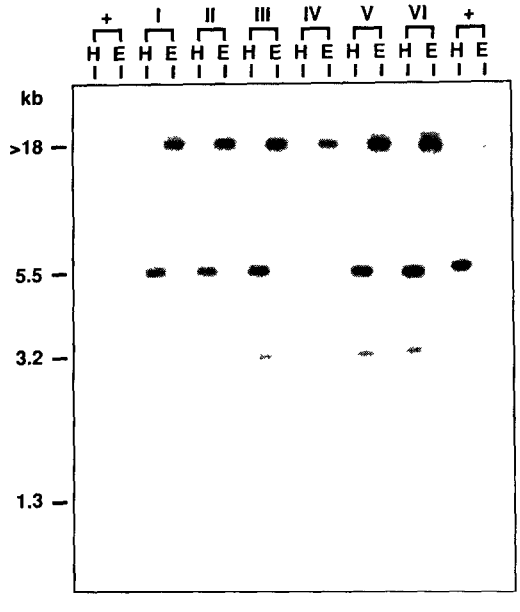


Fig. 2. Southern analysis of genomic DNA from wild-type *Synechocystis* 6803 and from six independent cell lines transformed with plasmid pSMO8 (I–IV) after seven segregation rounds. The filter was hybridised at high stringency with a radiolabelled DNA probe consisting of a 1.1:1 molar mixture of the isolated inserts from plasmids pSMO6, pSMO7 (see Fig. 1) and pUC4K (the Km^r cartridge). The figure shows the resulting autoradiogram after exposure for 14 h at room temperature.

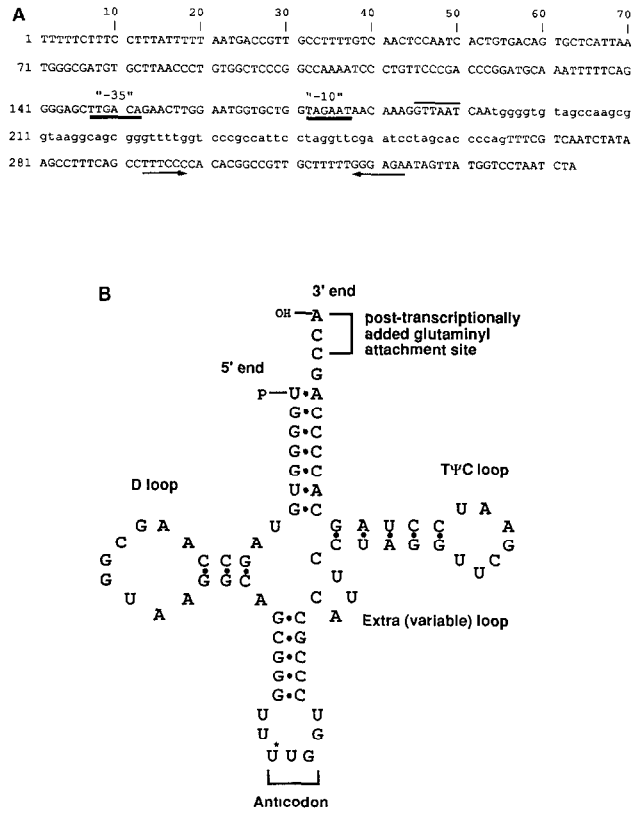


Fig. 3. (A) Nucleotide sequence X60715 which includes the *trnQ* gene of *Synechocystis* 6803. The position and orientation of X60715 are shown in Fig. 1. The first strand was sequenced away from the *Hin*cII site as also shown in Fig. 1. The second strand was then sequenced using the oligonucleotide 5′-CACCTCAGGACCCCCTTA-3′ to prime DNA synthesis from the single-stranded template of an appropriate M13 clone. The predicted extent of the $tRNA^{Gln}$ sequence is shown in lower case letters. Sequence elements similar to the '-35' and '-10' consensus regions of the *E. coli* σ^{70} promoter are underlined. A hexanucleotide region sharing some identity with the consensus sequence of stringent response elements from *B. subtilis* (GT(C/T)G(C/T)(T/Pu), see [27]) is overlined. In addition, the extent of a downstream inverted repeat is arrowed. (B) Secondary structure prediction for the $tRNA^{Gln}$ molecule of *Synechocystis* 6803. This prediction, showing a typical cloverleaf tRNA structure, was made by comparison to the predicted tobacco $tRNA^{Gln}$ structure [34] and application of the PC Gene TRNASRCH programme. The figure shows the primary sequence and no attempt has been made to indicate the probable positions of minor or rare bases arising from post-transcriptional modification. The 5′-CCA-3′ glutaminyl attachment site is not encoded by *trnQ* and must therefore be added post-transcriptionally. G–C and A–T base pairs are indicated by filled circles and the anticodon wobble position is marked with an asterisk.

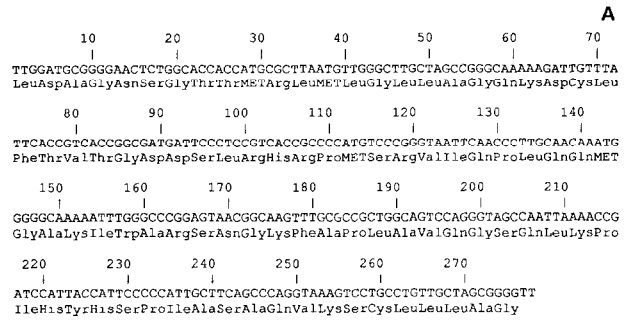
steps, plasmid pSMO8 was constructed so that the Km^r cartridge was placed between the *Synechocystis* 6803 inserts of pSMO6 and pSMO7, with the whole construct housed in pUC18. Plasmid pSMO8 was then transferred into *Synechocystis* 6803 with the intention of generating the *psbO* deletion strain, IC9.

Since *Synechocystis* 6803 contains about 12 copies of a single circular chromosome per cell [13] an attempt was made to segregate IC9 as a strain homozygous for the introduced deletion. Fig. 2 shows the results of Southern analysis of six independent kanamycin-resistant cell lines (denoted I–VI) following seven rounds of streaking from single isolated colonies on selective media. It can be seen that each transformed cell line contains copies of both the wild-type *psbO* locus, which is visible as the 5.5 kb *Hin*dIII–*Hin*dIII fragment, and the intended deletion product which is manifested as the 3.2 and 1.3 kb *Hin*dIII–*Hin*dIII fragments. Two signals arise from the deletion product in this experiment as the introduced Km^r cartridge contains an internal *Hin*dIII restriction site. Additional signals from the *Hin*dIII digests of IV and V are due to incomplete enzyme digestion. It can be concluded that transformed lines I–IV are all merodiploids, there being in each cell a mixed population of chromosome containing the wild-type *psbO* locus or the intended deletion. Streaking was continued through further rounds on selective media, but homozygous deletion mutant segregation was never achieved. In a final Southern analysis performed after more than 20 segregation rounds, it was observed that in lines I–IV, either the merodiploid state persisted or there had been a rearrangement of the *Synechocystis* 6803 chromosome in the vicinity of *psbO* (data not shown).

Since the 1.5 kb *Stu*I–*Xba*I fragment containing the entire *psbO* gene was subsequently successfully deleted in vivo to generate *psbO* deletion strain IC2 [2], it was concluded that the 0.9 kb *Hin*cII–*Stu*I fragment upstream of *psbO* contains DNA that is essential for the viability of *Synechocystis* 6803 and which therefore cannot be deleted. To investigate this possibility further, the DNA sequencing indicated in Fig. 1 was then performed. Initially sequence data were collected from only one of the two possible strands and then compared to entries in the EMBL and Genbank nucleic acid databases using the FASTA programme. Once regions of sequence identity with database entries had been determined, data from the complementary DNA strand were established using custom-made oligonucleotide sequencing primers. This approach led to the sequencing of an entire *trnQ* gene and a portion of an *aroA* gene the locations of which are documented in Fig. 1.

The *trnQ* gene encodes a glutamine-specific transfer RNA ($tRNA^{Gln}$) and is located upstream of *psbO* on the same DNA strand. The region sequenced in both orientations is shown in Fig. 3A. Based on data for $tRNA^{Gln}$ sequences from other species, the coding region of the *Synechocystis* 6803 *trnQ* gene is predicted to consist of 72 nucleotides (Fig. 3). The predicted secondary structure of the $tRNA^{Gln}$ molecule, which is the typical cloverleaf structure, is shown in Fig. 3B.

Sequence data from a region further upstream of *psbO* (see Fig. 1) that was determined from both DNA strands is shown in Fig. 4A and has been assigned EMBL Database Accession no. X72784. Database searches revealed that this DNA sequence shares 58% sequence identity along its entire length with the *aroA*



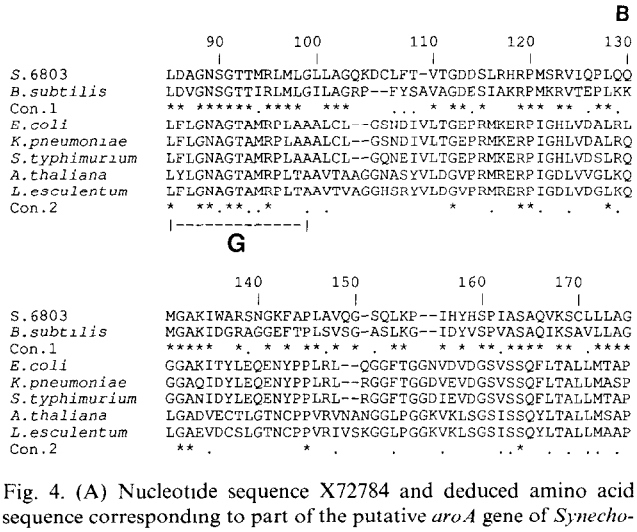


Fig. 4. (A) Nucleotide sequence X72784 and deduced amino acid sequence corresponding to part of the putative *aroA* gene of *Synechocystis* 6803. The map location and orientation of this sequence is shown in Fig. 1. (B) Alignment of the putative *aroA* gene product from sequence X72784 (S. 6803) with a number of other deduced EPSP synthase sequences. The alignment was generated with CLUSTAL (PAM250, Gap fixed = 10; Gap vary = 10). The sequences displayed are from: *Bacillus subtilis* [14]; *Escherichia coli* [35]; *Klebsiella pneumoniae* [19]; *Salmonella typhimurium* [18]; *Arabidopsis thaliana* [36]; *Lycopersicon esculentum* [17]. Numbering refers to the *B. subtilis* sequence. Con. 1 is the consensus match between the *Synechocystis* 6803 and *B. subtilis* sequences; Con. 2 is the consensus match across all the sequences. Asterisks denote identical amino acids; dots denote conservative amino acids (defined as scoring 8 or more in the Dayhoff matrix [37]). There is evidence that the domain labelled 'G' interacts with the herbicide, glycophosphate (see text).

gene of *Bacillus subtilis* [14], the gene encoding 5-enolpyruvylshikimate 3-phosphate synthase (EPSP synthase, also known as 3-phosphoshikimate 1-carboxyvinyltransferase; EC 2.5.1.19) which catalyses the condensation of phosphoenolpyruvate (PEP) and shikimate 3-phosphate in the shikimate biosynthetic pathway (see [15]). (Originally the EPSP synthase gene of *B. subtilis* was named *aroE* [14], but there is now consensus that this gene should be called *aroA*.)

Fig. 4B shows the deduced translation product aligned against the analogous region of other EPSP synthase protein sequences. Over the span shown in Fig. 4B the *Synechocystis* 6803 sequence shares amino acid identities of 56%, 23%, 23% and 22%, respectively, with the corresponding sequences from *B. subtilis*, *E. coli*, *Klebsiella pneumoniae* and from *Lycopersicon esculentum*. It can therefore be concluded that the sequence determined represents a portion of an *aroA* gene of *Synechocystis* 6803. The complete *aroA* sequences listed in Fig. 4B comprise 427 or 428 amino acids. By assuming that this length is conserved in *Synechocystis* 6803, the extent of the *aroA* gene has been predicted in Fig. 1.

Bacterial and plant EPSP synthases are inhibited by the herbicide glycophospate (reviewed in [16]), known commercially as 'Roundup' (Monsanto). As indicated in Fig. 4B, the region of EPSP synthase aligned includes a domain implicated in the glycophosphate interaction [17], and the mapping of point mutations conferring glycophosphate resistance [18,19].

Fig. 5 shows the location of the *psbO* and *trnQ* genes, as well as the predicted extent of the *aroA* gene, within a mapped 24 kb region of the *Synechocystis* chromosome. The restriction map was generated by digesting four phage clones isolated from a genomic library constructed in λEMBL3 [20] with combinations of restriction enzymes and then subjecting the digestion products to Southern analysis using the *psbO* gene from *Synechococcus* spp. PCC 7942 [21] as a radiolabelled probe.

4. DISCUSSION

4.1. *The* trnQ *gene of* Synechocystis *6803*

The sequencing of *trnQ* from *Synechocystis* 6803 represents the first *trnQ* sequence to be reported from a cyanobacterium. There are only two Gln codons in the universal genetic code, CAA and CAG, and the anticodon of $tRNA^{Gln}$ (the unmodified anticodon sequence is UUG) would be expected to interact with both of these through wobble in the first anticodon/third codon position. The amino acid attachment site 5′-CCA_{OH}-3′, invariant at the 3′ end of tRNA molecules and essential for the charging of the tRNA via the terminal adenosine residue, is not present in the *trnQ* sequence and must therefore be added post-transcriptionally by the enzyme nucleotidyl transferase. Thus the mature *Synechocystis* 6803 $tRNA^{Gln}$ is 75 nucleotides in length. Alignment of available, unmodified $tRNA^{Gln}$ sequences from other species demonstrates that this length is conserved across bacteria, mitochondria and chloroplasts (data not shown).

Normally *trn* genes are transcribed as larger precursor molecules. The RNA is then trimmed to its mature size by nucleases and a number of bases in the sequence become modified [22,23]. Presumably these events also occur during expression of the *Synechocystis* 6803 *trnQ* gene.

A DNA sequence element showing a high degree of similarity to the *E. coli* σ^{70} promoter consensus defined

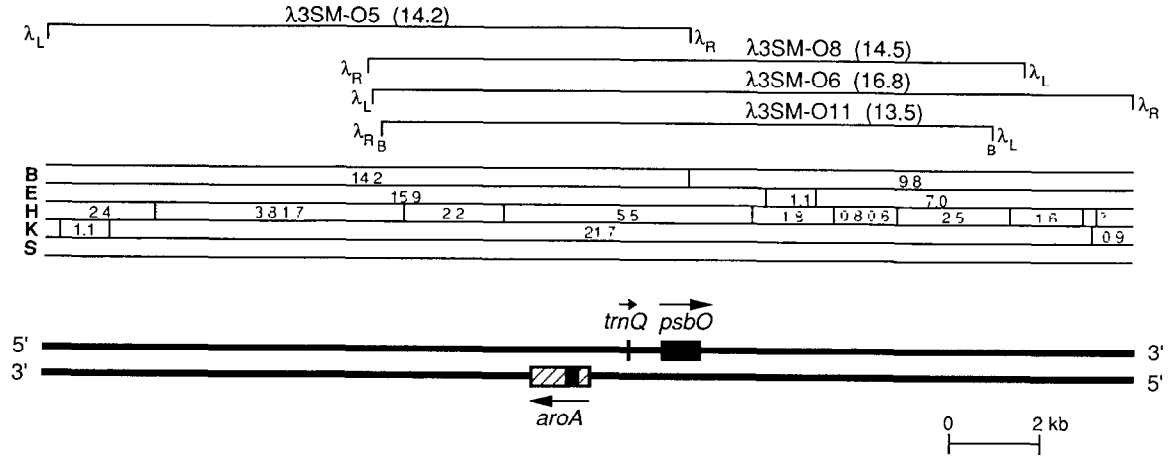


Fig. 5. Restriction map of 24-kb of the *Synechocystis* 6803 chromosome containing *aroA*, *trnQ* and *psbO*. All the restriction sites within this region for: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; S, *Sal*I; are shown. The filled boxes on the two chromosomal DNA strands indicate, respectively, the extent of coding regions of *trnQ* and *psbO* (including presequence), and the sequenced region of the putative *aroA* gene. The predicted extent of the *aroA* gene (see Fig. 1) is shown by the hatched box. Phage library clones 13SM-O5, 13SM-O6, 13SM-O8 and 13SM-O11, were analysed to generate the map. The λEMBL3 vector right and left arms are denoted, respectively, by λ_R and λ_L, and the numbers estimate restriction fragment sizes (kb). Both insert-vector junctions in λ3SM-O11 form reconstituted *Bam*HI sites and are accordingly labelled 'B'. The linear orders of pairs of restriction fragments within the two stippled regions of the *Hin*dIII map have not been established. '?' indicates an uncertain *Hin*dIII site.

by Harley and Reynolds [24] is located immediately upstream of *trnQ* (Fig. 3). A putative '-35' promoter element (TTGACA) is separated from a putative '-10' element (TAGAAT) by 19 bp. This spacing compares reasonably well with the gap of 16–18 bp most frequently observed between these elements in *E. coli*. The mature *trnQ* sequence commences 17 bp downstream of the '-10' element. Although *Synechocystis* 6803 promoters are not well-defined, there is reason to suppose that σ^{70}-type promoters do exist in this cyanobacterium [10,25]. Consequently this σ^{70}-type element appears to be the prime candidate for the *trnQ* promoter. Supporting this speculation, it has been noted that *trnE*, which is the only other *trn* gene so far sequenced from *Synechocystis* 6803, also possesses an upstream σ^{70}-like sequence motif [26].

Further comparison of the *Synechocystis* 6803 *trnQ* and *trnE* genes reveals that the regions immediately downstream of the putative '-10' elements of *trnQ* and *trnE* are also somewhat similar. Both regions contain a short A-rich domain, followed by a region with some similarity to the hexanucleotide consensus sequence of the stringent response element of *B. subtilis* [27]. Possibly these DNA elements mediate a stringent response in *Synechocystis* 6803 in a manner similar to that observed in *B. subtilis*, as well as *E. coli* [28]. An inverted repeat located downstream of *trnQ* (see Fig. 3) might also have functional significance in *trnQ* expression.

4.2. *An intact* trnQ *gene appears essential for the viability of* Synechocystis *6803*

As discussed above, the failure to generate a homozygous IC9 strain indicates that some part of the 0.9 kb *Hin*cII–*Stu*I region upstream of *psbO* is essential for the viability of *Synechocystis* 6803. All this region, except for approximately 0.1 kb, has now been sequenced on at least one DNA strand (see Fig. 1) and the only candidate coding sequence that has been identified is *trnQ*. It is therefore reasonable to conclude that this *trnQ* gene is absolutely required by *Synechocystis* 6803, almost certainly because its role in protein biosynthesis cannot be assumed by any other *trn* gene in the *Synechocystis* 6803 genome. This in turn means that *trnQ* is highly likely to be single-copy in *Synechocystis* 6803, which is the same situation as for the *trnE* gene in this same cyanobacterium [29].

Synechocystis 6803 does not appear to possess any glutaminyl-tRNA synthetase. Instead the organism employs an alternative pathway to charge $tRNA^{Gln}$ with glutamine and form $Gln\text{-}tRNA^{Gln}$ [30]. This alternative pathway is also apparently the sole pathway for forming $Gln\text{-}tRNA^{Gln}$ in chloroplasts, the mitochondria of plants and animals and Gram-positive bacteria, such as *B. subtilis* [30]. In these organelles and bacteria, $tRNA^{Gln}$ is firstly 'mischarged' with glutamate to form $Glu\text{-}tRNA^{Gln}$, which is then converted to $Gln\text{-}tRNA^{Gln}$ by a specific amidotransferase that requires ATP, Mg^{2+} and a suitable amide donor, such as glutamine or asparagine [31]. Two distinct $Glu\text{-}tRNA^{Gln}$ species ($tRNA^{Gln}_1$ and $tRNA^{Gln}_2$) have been resolved from *Synechocystis* 6803 by reverse-phase chromatography [26,29]. Therefore the conclusion that *trnQ* is single-copy in *Synechocystis* 6803 argues that in this organism both resolvable $tRNA^{Gln}$ species are likely to be derived from the same *trnQ* gene by post-transcriptional modification. Indeed, there are existing precedents for such a situation. Both the $tRNA^{Glu}_1$ and $tRNA^{Glu}_2$ isoacceptors of *Synechocystis* 6803 are derived from the same *trnE* gene [29]. Further-

Table I

Percentage nucleotide identities between the $tRNA^{Gln}$ sequences from *Synechocystis* 6803 and other species

E. coli $tRNA_1^{Gln}$	*E. coli* $tRNA_2^{Gln}$	*B. subtilis*	*M. capricolum*	*C. reinhardtii* MIT	wheat MIT $tRNA_1^{Gln}$	wheat MIT $tRNA_2^{Gln}$	*A. longa*	*E. gracilis* CP	*M. polymorpha* CP	tobacco CP	mustard CP	barley CP	rice CP
83.3	84.7	76.4	73.6	73.6	79.1	80.6	73.6	81.9	80.6	84.7	87.5	83.3	83.3

Identities between the primary, unmodified sequences, which all consist of 72 nucleotides, excluding the 5′-CAA glutaminyl attachment site, are listed. MIT, mitochondria; CP, chloroplasts. Gene sequences used in the comparison are from: *Escherichia coli*, where two $tRNA^{Gln}$ genes have been sequenced [38]; *Bacillus subtilis* [39]; cpDNA-like DNA of *Astasia longa* [40]; *Mycoplasma capricolum* [41]; mitochondrial DNA of the green alga, *Chlamydomonas reinhardtii* [42]; wheat (*Triticum aestivum*) mitochondrial DNA, where genes for two isoacceptors, $tRNA_1^{Gln}$ [43] and $tRNA_2^{Gln}$ [44] have been identified and sequenced; cpDNA of *Euglena gracilis* [45]; cpDNA of *Marchantia polymorpha* [46]; cpDNA of tobacco (Nicotiana tabacum, [47]; cpDNA of mustard (Sinapis alba; [48], cpDNA of rice (*Oryza sativa*; [49]). The unmodified sequences of the two distinguishable $tRNA^{Gln}$ isoacceptors of barley (*Hordeum vulgare*) chloroplasts have been determined directly and were found to be identical [30].

more, two barley chloroplast $tRNA^{Gln}$ isoacceptors are chromatographically distinguishable yet, since they possess the same primary sequence and differ only in their post-transcriptional modification, they are presumably encoded by the same *trnQ* gene [30].

Table I shows that the *trnQ* gene of *Synechocystis* 6803 does not share appreciably more sequence identity with *trnQ* genes of chloroplasts than it does with homologues encoded in wheat mitochondrial DNA, or in *E. coli*. The significance of this observation in terms of drawing phylogenetic inferences and assessing the constraints on evolution of *trnQ* genes is unclear; however, this pattern of homologies is markedly different from that observed for the *trnE* gene of *Synechocystis* 6803 the high sequence identity of which with chloroplast *trnE* sequences was interpreted as further evidence for the endosymbiotic origin of chloroplasts from an ancestral cyanobacterial species [26].

4.3. *The putative* aroA *gene of* Synechocystis *6803*

The mapping and partial sequencing of the *aroA* gene from *Synechocystis* 6803 provides the first genetic evidence for the existence of the shikimate biosynthetic pathway in cyanobacteria. In other prokaryotes and in plants this pathway produces chorismate, an essential intermediate in the synthesis of many aromatic compounds, including the aromatic amino acids and quinones [15,17]. As plastoquinone molecules are key electron acceptors within PSII it is an interesting observation that *aroA*, a gene involved in quinone biosynthesis, should map close to the PSII gene. It remains to be established whether this gene association is fortuitous or functional, e.g. co-regulation of *aroA* and *psbO* gene expression might permit some coordination of quinone biosynthesis and PSII assembly.

In light of speculation that plant chloroplasts evolved from the endosymbiosis of ancestral cyanobacteria into eukaryotic cells (e.g. [32]) it is curious that the sequenced region of the putative *Synechocystis* 6803 *aroA* gene shares more sequence identity with the Gram-positive bacterium *B. subtilis* than with available sequences from either plants or Gram-negative bacteria. Clearly, a detailed phylogenetic analysis will be of interest when the full *aroA* sequence from *Synechocystis* 6803 has been established.

Acknowledgements Particular thanks go to Dr. J.G.K. Williams (DuPont, Delaware, USA) for providing *Synechocystis* 6803-G and the λEMBL3 genomic library, to Dr. T. Kuwabara for sending the *psbO* gene of *Synechococcus* spp. PCC 7942, to Dr. C. Lichtenstein for synthesising oligonucleotides, and to C. O'Donnell (AFRC Computer Centre, Harpenden, UK) for his computing expertise. Thanks also to Drs. P.J. Nixon, K.M. Cook, J.M. Dubbs and other members of the AFRC Photosynthesis Research Group for help and advice. This work was supported by the Agricultural and Food Research Council and by CNPQ-Brazil (M.D.C.).

REFERENCES

[1] Barber, J. (1989) in: Oxford Surveys of Plant Molecular Biology, Vol. 6 (Miflin, B.J. ed.) pp. 115–162, Oxford University Press, Oxford, UK.
[2] Mayes, S.R., Cook, K.M., Self, S.J., Zhang, Z. and Barber, J. (1991) Biochim. Biophys. Acta 1060, 1–12.
[3] Biochenko, V.A., Klimov, V.V., Mayes, S.R. and Barber, J. (1993) Z. Naturforsch. (in press).
[4] Burnap, R.L. and Sherman, L.A. (1991) Biochemistry 30, 440–446.
[5] Philbrick, J.B., Diner, B.A. and Zilinskas, B.A. (1991) J. Biol. Chem. 266, 13370–13376.
[6] Bockholt, R., Masepohl, B. and Pistorius, E.K. (1991) FEBS Lett. 294, 59–63.
[7] Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266–267.
[8] Norrander, J.T., Kempe, T. and Messing, J. (1983) Gene 26, 101–106.
[9] Higgins, D.G. and Sharp, P.M. (1989) CABIOS 5, 151–153.
[10] Mayes, S.R., Dubbs, J.M., Vass, I., Hideg, E., Nagy, L. and Barber, J. (1993) Biochemistry 32, 1454–1465.
[11] Frischauf, A.-M., Lehrach, H., Poustka, A. and Murray, N. (1983) J. Mol. Biol. 170, 827–842.
[12] Taylor, L.A. and Rose, R.E. (1988) Nucleic Acids Res. 16, 358.
[13] Labarre, J., Chauvat, F. and Thuriaux, P. (1989) J. Bacteriol. 171, 3449–3457.
[14] Henner, D.J., Band, L.B., Flaggs, G. and Chen, E. (1986) Gene 49, 147–152.
[15] Hoch, J.A. and Nester, E.W. (1973) J. Bacteriol. 116, 59–66.
[16] Kishore, G.M. and Shah, D.M. (1988) Annu. Rev. Biochem. 57, 627–663.
[17] Gasser, C.S., Winter, J.A., Hironaka, C.M. and Shah, D.M. (1988) J. Biol. Chem. 263, 4280–4289.
[18] Stalker, D.M., Hiatt, W.R. and Comai, L. (1985) J. Biol. Chem. 260, 4724–4728.
[19] Sost, D. and Amrhein, N. (1990) Arch. Biochem. Biophys. 282, 483–486.
[20] Mayes, S.R., Cook, K.M. and Barber, J. (1990) in: Current Research in Photosynthesis, Vol. 3 (Baltscheffsky, M., ed.) pp. 617–620, Kluwer, The Netherlands.

[21] Kuwabara, T., Reddy, K.J. and Sherman, L.A. (1987) Proc. Natl. Acad. Sci. USA 84, 8230–8234.
[22] Björk, G.R., Ericson, J.U., Gustafsson, C.E.D., Hagervall, T.G., Jönsson, Y.H. and Wikström, P.M. (1987) Annu. Rev. Biochem. 56, 263–287.
[23] Lewin, B. (1990) Genes IV, Oxford University Press, Oxford, UK.
[24] Harley, C.B. and Reynolds, R.P. (1987) Nucleic Acids Res. 15, 2343–2359.
[25] Zhang, Z.H. (1992) PhD thesis, University of London.
[26] O'Neill, G.P., Peterson, D.V., Schön, A., Chen, M.-W. and Söll, D. (1988) J. Bacteriol. 170, 3810–3816.
[27] Vold, B.S. (1985) Microbiol. Rev. 49, 71–80.
[28] Fournier, M.J. and Ozeki, H. (1985) Microbiol. Rev. 49, 379–397.
[29] O'Neill, G.P. and Söll, D. (1990) J. Bacteriol. 172, 6363–6371.
[30] Schön, A., Kannangara, C.G., Gough, S. and Söll, D. (1988) Nature 331, 187–192.
[31] Jahn, D., Kim, Y.-C., Ishino, Y., Chen, M.-W. and Söll, D. (1990) J. Biol. Chem. 265, 8059–8064.
[32] Giovannoni, S.J., Turner, S., Olsen, G.J., Barns, S., Lane, D. and Pace, N.K. (1988) J. Bacteriol. 170, 3584–3592.
[33] Philbrick, J.B. and Zilinskas, B.A. (1988) Mol. Gen. Genet. 212, 418–425.
[34] Sugiura, M. (1987) Bot. Mag. Tokyo 100, 407–436.
[35] Duncan, K., Lewendon, A. and Coggins, J.R. (1984) FEBS Lett. 170, 59–63.
[36] Klee, H.J., Muskopf, Y.M. and Gasser, C.S. (1987) Mol. Gen. Genet. 210, 437–442.
[37] Dayhoff, M.O. (1978) in: Atlas of Protein Sequence and Structure, Vol. 5, Suppl. 3 (Dayhoff, M.O., ed.) pp. 345–348, National Biomedical Research Foundation, Washington, USA.
[38] Nakajima, N., Ozeki, H. and Shimura, Y. (1981) Cell 23, 239–249.
[39] Wawrousek, E.F., Narasimhan, N. and Hansen, J.N. (1984) J. Biol. Chem. 259, 3694–3702.
[40] Siemeister, G., Buchhoz, C. and Hatchtel, W. (1990) Mol. Gen. Genet. 220, 425–432.
[41] Muto, A., Andachi, Y., Yuzawa, H., Yamao, F. and Osawa, S. (1990) Nucleic Acids Res. 17, 5037–5043.
[42] Ma, D.P., Yang, Y.W. and Hasnain, S. (1989) Nucleic Acids Res. 17, 1256.
[43] Joyce, P.B. and Gray, M.W. (1988) Nucleic Acids Res. 16, 1210.
[44] Joyce, P.B. and Gray, M.W. (1989) Nucleic Acids Res. 17, 4885.
[45] Karabin, G.D. and Hallick, R.B. (1983) J. Biol. Chem. 258, 5512–5518.
[46] Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H. and Ozeki, H. (1986) Nature 322, 572–574.
[47] Deno, A. and Sugiura, M. (1983) Nucleic Acids Res. 11, 8407–8414.
[48] Neuhaus, H. (1989) Nucleic Acids Res. 17, 444.
[49] Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.R., Meng, B.Y., Li, Y.Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K. and Sugiura, M. (1989) Mol. Gen. Genet. 217, 185–194.
[50] Williams, J.G.K. (1988) Methods Enzymol. 167, 766–778.