The petFI gene encoding ferredoxin I is located close to the str operon on the cyanelle genome of *Cyanophora paradoxa*

C. Neumann-Spallart, M. Brandtner, M. Kraus, J. Jakowitsch, M.G. Bayer, T.L. Maier, H.E.A. Schenk and W. Löffelhardt

Institut für Allgemeine Biochemie und Ludwig-Boltzmann-Forschungsstelle für Biochemie, A-1090 Wien, Währinger Str. 38, Austria and *Institut für Biologie I der Universität Tübingen, D-7400 Tübingen, FRG*

Received 17 May 1990; revised version received 31 May 1990

The petFI gene encoding ferredoxin I was localized in the large single copy region of cyanelle DNA by heterologous hybridization. Sequence analysis revealed an ORF of 99 amino acids (including the N-terminal processed methionine) at a position 477 bp from the 3′ end of tufA but on the opposite strand. The 25 amino-terminal residues well corresponded to partial sequences obtained with purified cyanelle ferredoxin. The assignment of yet another gene that is not found on the genomes of chlorophyll *b*-type plastids to cyanelle DNA again corroborates the special position of cyanelles serving as a model for plastid evolution from endocytobiotic cyanobacteria.

Cyanophora paradoxa; Cyanelle; PetFI; Ferredoxin I; Sequence analysis

1. INTRODUCTION

Ferredoxin fulfills important functions in plastids, not only in the photosynthetic electron transport chain but also as electron donor to nitrite reductase, sulphite reductase, glutamate synthase and thioredoxin reductase [1]. In higher plants and green algae the major plastid ferredoxin (ferredoxin I) is encoded by a single copy nuclear gene, petFI [2–4]. A N-terminal transit sequence directs the cytoplasmically synthesized ferredoxin I precursor polypeptide towards the plastid where it is cleaved off by a stroma protease after ATP-dependent import of preferredoxin I [5]. Cyanobacterial ferredoxin genes [6–8] also seem to occur only once per genome but lack transit sequences. Numerous amino acid sequences determined for plastidic and cyanobacterial ferredoxin I are also available allowing the construction of a phylogenetic tree [9]. A second molecular species, ferredoxin II, whose function still has to be elucidated, has been reported for many plastids and cyanobacteria, whereas nitrogen-fixing species contain in addition a heterocyst ferredoxin [10].

Cyanelles, the photosynthetic organelles of *Cyanophora paradoxa* are distinct from plastids by a surrounding murein sacculus [11] and by the presence in their genome of certain genes that are nuclear in higher plants and green algae. The corresponding proteins range from SSU of Rubisco [12] to at least 8 ribosomal proteins [13–15]. Recently, the synthesis and translation of ferredoxin mRNA has been shown to occur within the cyanelles [16]. Here we report the mapping as well as sequence and transcript analysis of cyanelle petFI.

Correspondence address: W. Löffelhardt, Institut für Allgemeine Biochemie der Universität Wien, A-1090 Wien, Währinger Str. 38, Austria

The nucleotide sequences reported will appear in the EMBL, GenBank and DDBJ databases under the accession number X52143.

2. MATERIALS AND METHODS

Cyanophora paradoxa was grown and DNA extracted from purified cyanelles as described [17]. Southern hybridization was performed under low stringency conditions [18] using a 450 bp probe comprising petFI from *Silene pratensis* [2] labelled either by means of the digoxigenin kit from Boehringer Mannheim or the primed synthesis kit from Amersham. Cyanelle RNA isolation and Northern hybridization followed published methods [19]. Clone *Bgl*II-11 was contained in a cyanelle gene bank of *Bgl*II fragments cloned into the *Bam*HI site of pEMBL8 [20]. The supercoil sequencing method [21] was adopted on exoIII deletion subclones [22] obtained with the 'erase-a-base' system from Promega. Computer analysis of the sequence was done using the Dave Mount program 5.07 from the University of Arizona.

3. RESULTS AND DISCUSSION

Low stringency hybridization using a petFI probe from *Silene pratensis* yielded signals for *Bgl*II fragment 11 (4.9 kb) and its 1.25 kb *Pst*I/*Bgl*II subfragment, respectively. This placed the ferredoxin gene approximately in the middle of the large single copy region of cyanelle DNA [18], surrounded by tRNA genes [23] and genes for components of the translation apparatus [24,25] as shown in Fig. 1. Sequence analysis revealed an open reading frame of 300 bp (Fig. 2) whose

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

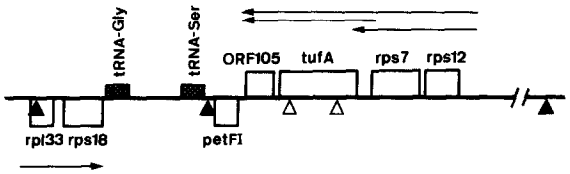

Fig. 1. Gene arrangement in the central part of the large single copy region on the cyanelle genome comprising *Bgl*II fragments no. 17 and 11. The distribution of the respective genes over both DNA strands is also given. Transcript sizes [16,28] are indicated by arrows. Restriction enzyme cleavage sites are symbolized by ▲ (*Bgl*II) and by △ (*Pst*I), respectively.

translated N-terminal amino acids corresponded to those determined for purified cyanelle ferredoxin [26]. Upstream petFI is flanked by trnS, situated on the opposite strand and separated from it by 440 bp of noncoding sequences, very rich in A/T. Our sequence is contiguous with that reported for *Bgl*II-17 [23]. Downstream, also on the opposite strand an ORF of 105 codons was found with sequence homology to *E.coli* rps10 (C. Neumann-Spallart et al., in preparation), another gene unreported to date for any plastid DNA. The 72 bp intergenic region (Fig. 2) contained a perfect inverted repeat forming a 22 bp stem and a 4 b loop that might act as a transcription termination signal for petFI as well as for the opposite str operon [25].

A putative ribosome binding site, complementary to the 3′ end of cyanelle 16 S rRNA [27] was localized 12 bp upstream of the petFI initiation codon. Sequence motifs resembling '−10' and '−35' promoter boxes were found on the 5′ side of the Shine-Dalgarno sequence but more deviation from the prokaryotic consensus was observed than with cyanelle psbA and str promoters [19,25]. The observed size of the transcript, approximately 450 b (Fig. 3), correlates with a 3′ terminus between the end of the inverted repeat and the stop codon of ORF 105 and with a 5′ untranslated region of about 90 b. A comparison of ferredoxin amino acid sequences is given in Fig. 4. All data are taken from [9] with the exception of those for *Chlamydomonas reinhardtii* [28] and *Marchantia polymorpha* [29]. Clearly the cyanelle gene encodes the type I protein and as usual the N-terminal methionine is processed posttranslationally yielding a molecular mass of 10.5 kDa. For the cyanelle sequence N-terminal amino acids also obtained by protein sequencing are indicated by bold letters. While the amino-terminal region shows some variation [26], 4 successive blocks of highly conserved amino acid sequences underline the strong structure-function interrelationship for this protein. The 4 cystein residues involved in the formation of the Fe/S-cluster (pos. 41, 46, 49 and 79) are located within two of the conserved regions. The highest identity score relative to *C. paradoxa* is obtained for *Synechococcus sp.* PCC 7418 (79.6%) that was also found closest with respect to the N-terminal amino

```
   BglII
    ↓
1   GATCTTATCTAAGATATGTAAATAAATAAAAATATATATCTATATTTATA

51  GTATATATTAATTTTTTTTAAAAATCGATACTAAATTTAAATTTTCCTTT

101 TTTTTCTTTATAAAAATTTAATTTTAAATAGAAAAAATTAAGTTTTTCGA

151 AAAAAGCAATTAAAACATATTAAAAAAAAATTAATAAACATGGTAAACTT

201 TAAATATAAATTTATAATTAACTGAAAAAATAATAAAAATAAATTTATAT
                  .....      .....

251 ATATATATATTTTAGATTAAAATAATTTAAATTAAATTATTAAAAGTTCT
        .....      .....
                                  ⟶petFI
301 ACCTTGTAACTATAATTATTTAGGAGATAGTATTTTATGGCAGTATATAA
                  ═══          M  A  V  Y  K

351 AGTTCGTCTTATTTGTGAAGAACAAGGTTTAGATACCACTATTGAATGTC
      V  R  L  I  C  E  E  Q  G  L  D  T  T  I  E  C

401 CAGATGATGAGTACATTCTTGATGCAGCAGAAGAACAAGGTATTGATTTA
      P  D  D  E  Y  I  L  D  A  A  E  E  Q  G  I  D  L

451 CCATACTCCTGTCGTGCAGGTGCATGTTCTACTTGTGCAGGTAAAGTGGT
      P  Y  S  C  R  A  G  A  C  S  T  C  A  G  K  V  V

501 AGAAGGAACTGTAGATCAATCTGATCAATCTTTCTTAGATGACGCTCAAT
      E  G  T  V  D  Q  S  D  Q  S  F  L  D  D  A  Q

551 TAGCAGCTGGTTATGTATTAACTTGTGTAGCATACCCATCTTCTGACTGT
      L  A  A  G  Y  V  L  T  C  V  A  Y  P  S  S  D  C

601 ACAGTTAAAACTCACCAAGAAGAATCTCTTTACTAAAAAATAAAAAATCT
      T  V  K  T  H  Q  E  E  S  L  Y              ═══

651 AAATAATAAAATAGAAATCTCTATTTTATTATTTAGATTTTCTTAATTCA
    ═══════════          ═══════════════

701 AAAAAAAACTAAAGTTTAACTTCCACATCAACACCTGCTGGTAAATCTAA
              ⟵ORF105
```

Fig. 2. Cyanelle DNA sequence starting at the left border of *Bgl*II-11 (Fig. 1) containing petFI and the 3′-terminus of ORF105. Putative promoter sequences are marked by dotted lines whereas the ribosome binding site is indicated by a double line. Start and stop codons of petFI and the stop codon of ORF105 are shown in bold letters. The inverted repeat in the intergenic space is underlined.

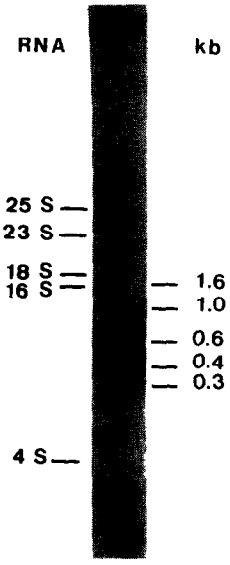

Fig. 3. Northern hybridization with a ^{32}P-labelled 25mer oligonucleotide probe complementary to positions 456–479 (Fig. 2). 20 μg of *C. paradoxa* poly(A)$^-$ RNA per lane were separated on 1% agarose gels. The endogenous rRNAs visualized by ethidium bromide staining of an adjacent lane and RNA standards obtained from Boehringer Mannheim served as size markers.

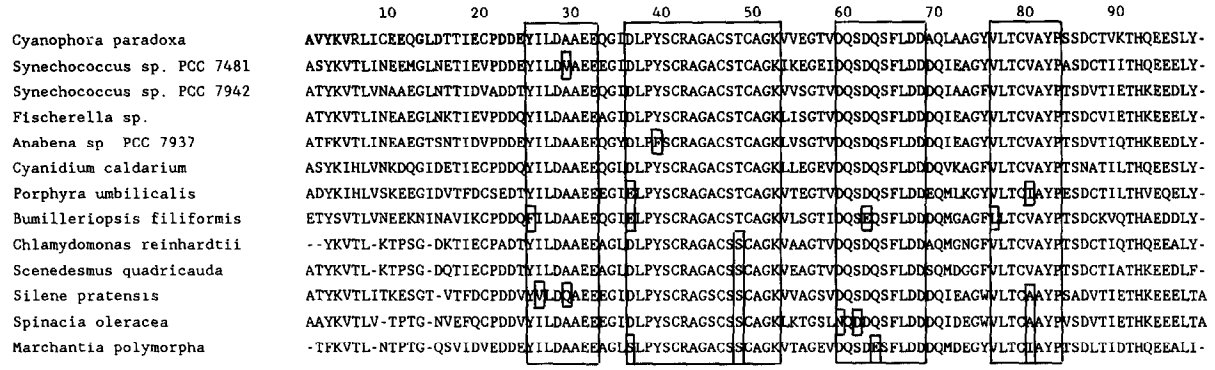

Fig. 4. Amino acid sequence conservation among ferredoxins I from *C. paradoxa*, Cyanobacterin and plastids. The boxes indicate arrays of almost invariant amino acids.

acids as determined by protein sequencing [26]. In general the identity scores decrease in the order: cyanobacteria > red algae > green algae > plants. The designation of the cyanelle gene as petFI is justified by the fact that the identity scores for comparison to ferredoxin II genes are always lower than those with petFI from the same organism, i.g. 67% versus 72.4% for *Nostoc MAC* [30]. In addition, ferredoxin II sequences show deviations within the highly conserved blocks outlined in Fig. 4 [6]. A comparably large body of data covering members from many groups of organisms performing oxygenic photosynthesis is available only with respect to 16 S rRNA sequences [31]. Upon pairwise comparison cyanelle 16 S rRNA is closest related to a number of cyanobacteria including *Fischerella*. However, some cyanobacteria are superseded in that respect by *Marchantia* or tobacco. Accordingly the identity scores for comparison to cyanelle ferredoxin I range from 65.3% to 79.6% for cyanobacteria and from 62.2% to 70.4% for higher plants, respectively. Interestingly the ferredoxin from *Cyanidium caldarium* shows the highest degree of identity (76.9%) to the cyanelle protein among all eukaryotes investigated. This is in agreement with the observation that monospecific antibodies directed against cyanelle ferredoxin react with the counterpart from *C. caldarium* but not from spinach (M.G. Bayer, unpublished). In conclusion, the features of *C. paradoxa* ferredoxin support the view that the cyanelles have been derived from cyanobacteria-like descendants ensuring that this organism will continue to serve as an excellent model for the study of plastid evolution [32].

Acknowledgements: We thank Hans J. Bohnert and Christine Michalowski who provided the cloned *Bgl*II-11 fragment and helped with sequence comparisons. This work was supported by grants from the Fonds zur Förderung der wissenschaftlichen Forschung (P 7349-CHE) to W.L. and from the Deutsche Forschungsgemeinschaft (SPP Intrazelluläre Symbiose, Sche 98/105) to H.E.A.S. Thanks are also due to Prof. P. Weisbeek (Univ. of Utrecht) for the ferredoxin gene probe and to Prof. F. Götz (Univ. of Tübingen) for synthesizing the oligonucleotide.

REFERENCES

[1] Arnon, D.I. (1988) Trends Biochem. Sci. 13, 30–33.
[2] Smeekens, S., van Binsbergen, J. and Weisbeek, P. (1985) Nucleic Acids Res. 13, 3179–3194.
[3] Dobres, M.S., Elliott, R.C., Watson, J.C. and Thompson, W.F. (1987) Plant Mol. Biol. 8, 53–59.
[4] Wedel, N., Bartling, D. and Herrmann, R.G. (1988) Bot. Acta 101, 295–300.
[5] Smeekens, S., Bauerle, C., Hagemann, J., Keegstra, K. and Weisbeek, P. (1986) Cell 46, 365–375.
[6] van der Plas, J., de Groot, R., Woortman, M., Cremers, F., Borrias, M., van Arkel, G. and Weisbeek, P. (1988) Photosynth. Res. 18, 179–204.
[7] Alam, J., Whitaker, R.A., Krogmann, D.W. and Curtis, S.E. (1986) J. Bacteriol. 168, 1265–1271.
[8] Reith, M.E., Laudenbach, D.E. and Straus, N.A. (1986) J. Bacteriol. 168, 1319–1324.
[9] Matsubara, H. and Hase, T. (1983) in: Proteins and Nucleic Acids in Plant Systematics (Jensen, U. and Fairbrothers, D.E. eds) pp. 169–181, Springer-Verlag, Berlin.
[10] Böhme, H. and Schrautemeier, B. (1987) Biochim. Biophys. Acta 891, 1–7.
[11] Schenk, H.E.A. (1970) Z. Naturforschung 25b, 640–656.
[12] Heinhorst, S. and Shively, J.M. (1983) Nature 304, 373–374.
[13] Evrard, J.-L., Johnson, C., Janssen, I., Löffelhardt, W., Weil, J.-H. and Kuntz, M. (1990) Nucleic Acids Res. 18, 1115–1119.
[14] Bryant, D.B. and Stirewalt, V.M. (1990) FEBS Lett. 259, 273–280.
[15] Michalowski, C.B., Pfanzagl, B., Löffelhardt, W. and Bohnert, H.J. (1990) Mol. Gen. Genet. submitted.
[16] Bayer, M.G. and Schenk, H.E.A. (1989) Curr. Genet. 16, 311–313.
[17] Bohnert, H.J., Crouse, E., Pouyet, F., Mucke, H. and Löffelhardt, W. (1982) Eur. J. Biochem. 126, 381–388.
[18] Bohnert, H.J., Michalowski, C., Bevacqua, S., Mucke, H. and Löffelhardt, W. (1985) Mol. Gen. Genet. 201, 565–574.
[19] Janssen, I., Jakowitsch, J., Michalowski, C.B., Bohnert, H.J. and Löffelhardt, W. (1989) Curr. Genet. 15, 335–340.
[20] Dente, L., Cesareni, G. and Cortese, R. (1983) Nucleic Acids Res. 11, 1645–1655.
[21] Chen, E.Y. and Seeburg, P.H. (1985) DNA 4, 165–170.
[22] Henikoff, S. (1984) Gene 28, 351–359.
[23] Kuntz, M., Evrard, J.L. and Weil, J.H. (1988) Nucleic Acids Res. 16, 8733.
[24] Evrard, J.L., Kuntz, M. and Weil, J.H. (1990) J. Mol. Evol. 30, 16–25.
[25 Kraus, M., Götz, M. and Löffelhardt, W. (1990) Plant Mol. Biol. (in press).

[26] Stevanovic, S., Bayer, M.G., Tröger, W. and Schenk, H.E.A. (1990) Endocyt. Cell Res. 6, 219–226.
[27] Janssen, I., Mucke, H., Löffelhardt, W. and Bohnert, H.J. (1987) Plant Mol. Biol. 9, 479–484.
[28] Schmitter, J.M., Jacquot, J.P., de la Motte-Guery, F., Beauvallet, C., Dutka, S., Gadal, P. and Decottignies, P. (1988) Eur. J. Biochem. 172, 405–412.
[29] Minami, Y., Wakabayashi, S., Imoto, S., Ohta, Y. and Matsubara, H. (1985) J. Biochem. 98, 649–655.
[30] Hase, T., Matsubara, H., Hutber, G.N. and Rogers, L.J. (1982) J. Biochem. 92, 1347–1355.
[31] Giovannoni, S.J., Turner, S., Olsen, G.J., Barns, S., Lane, D.J. and Pace, N.R. (1988) J. Bacteriol. 170, 3584–3592.
[32] Schenk, H.E.A. (1990) in: Endocytobiology IV (Nardon, P., Gianinazzi-Pearson, V., Grenier, A.M., Margulis, L. and Smith, D.C. eds.) pp. 199–209, INRA, Paris.