Complete Sequence of One Copy of the *psb*A Gene from the Thermophilic Cyanobacterium *Synechococcus elongatus**

Ralf Kloos, Edward Stevens, and Walter Oettmeier

Lehrstuhl Biochemie der Pflanzen, Ruhr-Universität, D-44780 Bochum,
Bundesrepublik Deutschland

Z. Naturforsch. **48c,** 799–802 (1993); received June 1/July 2, 1993

D-1 Protein, Herbicide Binding Protein, Q_B-Site, *psb*A Gene, *Synechococcus elongatus*

One copy of the *psb*A gene which codes for the photosystem II reaction center D-1 protein from the thermophilic cyanobacterium *Synechococcus elongatus* has been sequenced. It is feasible that a disulfide bridge between D-1 Cys_{212} and D-2 Cys_{212} is reponsible for the thermostability of the photosystem II reaction center from *Synechococcus elongatus*.

One of the three copies of the *psb*A gene in *Synechococcus elongatus* was sequenced completely and compared with the three copies of the *psb*A gene in the cyanobacterium *Anacystis nidulans* R 2 (PCC 7942), formerly called *Synechococcus* R 2 [1]. Like in *Anacystis nidulans* R 2 there are also most probably three copies of the *psb*A gene in *Synechococcus elongatus* (unpublished results). An EcoRI-ApaI restriction endonuclease fragment of 2.6 kb contains one complete copy of one *psb*A gene and in addition 330 base pairs of a second *psb*A gene. The third *psb*A gene is not located on this fragment. The complete *psb*A gene contains an open reading frame of 360 amino acids which is homologous to *psb*A genes of other organisms and encodes the photosystem II reaction center D-1 protein (Fig. 1). The *psb*A gene from *Synechococcus elongatus* has an *E. coli* like consensus promotor sequence at −10: TATATT but differs at the −35 sequence: TTTCAA.

```
-81  TTAAGTTAATGCTTAAGCCCTTGCAGTTGATGACTGTTTTCAAGGGTTTGTAAATCAAA
                                          TATATTTGGACTTATCACGACT

  1  ATG ACC ACA ACT CTC CAA CGT CGC GAA AGC GCG AAT TTG TGG GAG
  1   M   T   T   T   L   Q   R   R   E   S   A   N   L   W   E
      M   T   T   A   L   Q   R   R   E   S   A   S   L   W   Q
      *   *   *       *   *   *   *   *   *   *       *   *

 46 CGG TTT TGT AAC TGG GTG ACG AGC ACC GAT AAC CGC CTT TAT GTG
 16  R   F   C   N   W   V   T   S   T   D   N   R   L   Y   V
     Q   F   C   E   W   V   T   S   T   D   N   R   L   Y   V
         *   *       *   *   *   *   *   *   *   *   *   *   *

 91 GGC TGG TTT GGG GTG ATC ATG ATC CCC ACC CTA TTA GCC GCA ACC
 31  G   W   F   G   V   I   M   I   P   T   L   L   A   A   T
     G   W   F   G   V   L   M   I   P   T   L   L   T   A   T
     *   *   *   *   *       *   *   *   *   *   *       *   *

136 ATC TGC TTT GTG ATT GCC TTC ATC GCT GCT CCC CCT GTG GAC ATC
 46  I   C   F   V   I   A   F   I   A   A   P   P   V   D   I
     I   C   F   I   V   A   F   I   A   A   P   P   V   D   I
     *   *   *           *   *   *   *   *   *   *   *   *   *

181 GAT GGC ATC CGT GAG CCT GTT TCT GGC TCT TTG CTC TAT GGC AAC
 61  D   G   I   R   E   P   V   S   G   S   L   L   Y   G   N
     D   G   I   R   E   P   V   A   G   S   L   M   Y   G   N
     *   *   *   *   *   *   *       *   *   *       *   *   *
```

---

* EMBL accession no. Z 16081.

Reprint requests to Prof. Dr. Walter Oettmeier.

Verlag der Zeitschrift für Naturforschung,
D-72072 Tübingen
0939–5075/93/0900–0799 $ 01.30/0

```
226 AAC ATC ATC ACG GGT GCA GTT GTC CCC TCT AGC AAC GCC ATT GGC
 76  N   I   I   T   G   A   V   V   P   S   S   N   A   I   G
     N   I   I   S   G   A   V   V   P   S   S   N   A   I   G
     *   *   *       *   *   *   *   *   *   *   *   *   *   *

271 TTG CAC TTC TAC CCC ATT TGG GAA GCT GCT TCC CTC GAT GAG TGG
 91  L   H   F   Y   P   I   W   E   A   A   S   L   D   E   W
     L   H   F   Y   P   I   W   E   A   A   S   L   D   E   W
     *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

316 CTC TAC AAC GGT GGG CCC TAC CAA CTG ATC ATC TTC CAC TTC CTG
106  L   Y   N   G   G   P   Y   Q   L   I   I   F   H   F   L
     L   Y   N   G   G   P   Y   Q   L   V   V   F   H   F   L
     *   *   *   *   *   *   *   *   *           *   *   *   *

361 TTG GGT GCC TCC TGC TAC ATG GGT CGC CAG TGG GAA CTC ACG TAC
121  L   G   A   S   C   Y   M   G   R   Q   W   E   L   T   Y
     I   G   V   F   C   Y   M   G   R   E   W   E   L   S   Y
         *           *   *   *   *   *       *   *   *       *

406 CGC CTC GGT ATG CGG CCT TGG ATC TGC GTG GCC TAC TCT GCC CCC
136  R   L   G   M   R   P   W   I   C   V   A   Y   S   A   P
     R   L   G   M   R   P   W   I   C   V   A   Y   S   A   P
     *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

451 CTG GCT TCT GCC TTT GCA GTC TTC TTG ATC TAC CCC ATT GGT CAA
151  L   A   S   A   F   A   V   F   L   I   Y   P   I   G   Q
     V   A   A   A   T   A   V   F   L   I   Y   P   I   G   Q
         *       *       *   *   *   *   *   *   *   *   *   *

496 GGC AGC TTC TCT GAC GGG ATG CCC CTC GGT ATC TCT GGT ACC TTC
166  G   S   F   S   D   G   M   P   L   G   I   S   G   T   F
     G   S   F   S   D   G   M   P   L   G   I   S   G   T   F
     *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

541 AAC TTT ATG ATT GTG TTC CAA GCG GAG CAC AAC ATT CTC ATG CAC
181  N   F   M   I   V   F   Q   A   E   H   N   I   L   M   H
     N   F   M   F   V   F   Q   A   E   H   N   I   L   M   H
     *   *   *       *   *   *   *   *   *   *   *   *   *   *

586 CCC TTC CAC CAA CTG GGT GTA GCC GGT GTC TTT GGT GGG GCG CTG
196  P   F   H   Q   L   G   V   A   G   V   F   G   G   A   L
     P   F   H   M   L   G   V   A   G   V   F   G   G   S   L
     *   *   *       *   *   *   *   *   *   *   *   *       *

631 TTC TGC GCC ATG CAC GGT TCT CTG GTG ACC TCC AGC TTG ATC CGT
211  F   C   A   M   H   G   S   L   V   T   S   S   L   I   R
     F   S   A   M   H   G   S   L   V   T   S   S   L   V   R
     *       *   *   *   *   *   *   *   *   *   *   *       *

676 GAA ACC ACC GAA ACC GAA TCC GCC AAC TAC GGT TAC AAA TTT GGT
226  E   T   T   E   T   E   S   A   N   Y   G   Y   K   F   G
     E   T   T   E   T   E   S   Q   N   Y   G   Y   K   F   G
     *   *   *   *   *   *   *       *   *   *   *   *   *   *

721 CAA GAG GAA GAA ACC TAC AAC ATC GTG GCT GCC CAC GGT TAC TTT
241  Q   E   E   E   T   Y   N   I   V   A   A   H   G   Y   F
     Q   E   E   E   T   Y   N   I   V   A   A   H   G   Y   F
     *   *   *   *   *   *   *   *   *   *   *   *   *   *   *

766 GGC CGG TTG ATC TTC CAA TAC GCC AGC TTC AAC AAC AGC CGC TCC
256  G   R   L   I   F   Q   Y   A   S   F   N   N   S   R   S
     G   R   L   I   F   Q   Y   A   S   F   N   N   S   R   S
     *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
```

```
 811 CTG CAC TTC TTC TTG GCC GCT TGG CCG GTG GTG GGT GTG TGG TTT
 271  L   H   F   F   L   A   A   W   P   V   V   G   V   W   F
      L   H   F   F   L   A   A   W   P   V   V   G   I   W   F
      *   *   *   *   *   *   *   *   *   *   *   *       *   *

 856 ACC GCC TTG GGG ATC AGC ACC ATG GCC TTT AAC CTG AAT GGC TTC
 286  T   A   L   G   I   S   T   M   A   F   N   L   N   G   F
      T   S   L   G   I   S   T   M   A   F   N   L   N   G   F
      *       *   *   *   *   *   *   *   *   *   *   *   *   *

 901 AAC TTC AAC CAC TCG GTC ATT GAT GCC AAG GGC AAC GTG ATC AAC
 301  N   F   N   H   S   V   I   D   A   K   G   N   V   I   N
      N   F   N   Q   S   V   L   D   S   Q   G   R   V   I   N
      *   *   *       *   *       *           *       *   *   *

 946 ACC TGG GCG GAC ATC ATC AAC CGT GCC AAC TTG GGT ATG GAA GTG
 316  T   W   A   D   I   I   N   R   A   N   L   G   M   E   V
      T   W   A   D   V   L   N   R   A   N   L   G   M   E   V
      *   *   *   *           *   *   *   *   *   *   *   *   *

 991 ATG CAC GAG CGC AAT GCT CAC AAC TTC CCC CTC GAC TTG GCC AGC
 331  M   H   E   R   N   A   H   N   F   P   L   D   L   A   S
      M   H   E   R   N   A   H   N   F   P   L   D   L   A   A
      *   *   *   *   *   *   *   *   *   *   *   *   *   *

1036 GCT GAG TCT GCT CCT GTG GCC ATG ATT GCT CCC AGC ATC AAC GGC
 346  A   E   S   A   P   V   A   M   I   A   P   S   I   N   G
      G   E   A   T   P   V   A   L   T   A   P   A   I   N   G
          *           *   *   *           *   *       *   *   *

1081 TAAGCAACGTTATCTACTGAACTTTTACACAGCCGCTCCTAAACCTAGGGGCGGTTTTT
1040 TATTGCGTTA CGGCCCACGG GGAGTTTAAG GG
```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of one of the *psb*A genes from *Synechococcus elongatus*. For comparison, the sequence of the gene product from ANPSBA II from *Anacystis nidulans* R2 is given in the third row. Identical amino acids are marked by an asterisk.

The predicted amino acid sequence of the coded D-1 protein was compared to the two different polypeptides coded by the three *psb*A genes from *Anacystis nidulans* R2 (ANPSBA I and ANPSBA II/III). The amino acid sequence differs in 47 residues from the ANPSBA I D-1 protein and in 44 residues from the ANPSBA II/III D-1 protein. The homologies are 86.7% and 87.8%, respectively. The first 16 amino acids of *Synechococcus elongatus* differ in 12 amino acids from either the gene product of ANPSBA I or ANPSBA II/III.

The purpose of this study was to determine whether there are some specific changes in the amino acid sequence of the D-1 protein that are responsible for the fact that *Synechococcus elongatus* has optimal photosynthetic rates at a temperature of 57 °C [2]. At this temperature the photosystem II reaction center proteins D-1 and D-2 of other organisms are rapidly denatured. At 46 °C, the rate of electron transport in *Synechococcus elongatus* is reduced by 50% [3]. Most of the changes in the *Synechococcus elongatus* D-1 protein are conservative as compared to other D-1 proteins. There is however, a prominent exchange of Ser_{212} to Cys in transmembrane helix IV. Based on the model of the photosystem II reaction center [4] it is feasible that D-1 Cys_{212} is connected *via* a disulfide bridge to D-2 Cys_{212}, located in transmembrane helix IV (distance of the α-C atoms of the two Cys is 7.2 Å). This might be one reason for the thermostability of the photosystem II reaction center in *Synechococcus elongatus*.

Acknowledgements

This work was supported by Deutsche Forschungsgemeinschaft. We are indebted to Prof. Dr. W. Haehnel for the gift of *Synechococcus elongatus* DNA.

[1] S. S. Golden, J. Brusslan, and R. Haselkorn, EMBO J. **5,** 2789–2798 (1986).

[2] T. Yamaoka, K. Satoh, and S. Katoh, Plant & Cell Physiol. **19,** 943–954 (1978).

[3] L. K. Thompson, R. Blaylock, J. M. Sturtevant, and G. W. Brudvig, Biochem. **28,** 6686–6695 (1989).

[4] A. Trebst, Z. Naturforsch. **42c,** 742–750 (1987).