Complete Sequence of One Copy of the *psbA* 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, Qb-Site, *psbA* Gene, *Synechococcus elongatus*

One copy of the *psbA* 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 Cys212 and D-2 Cys212 is responsible for the thermostability of the photosystem II reaction center from *Synechococcus elongatus*.

One of the three copies of the *psbA* gene in *Synechococcus elongatus* was sequenced completely and compared with the three copies of the *psbA* gene in the cyanobacterium *Anacystis nidulans* R2 (PCC 7942), formerly called *Synechococcus* R2 [1]. Like in *Anacystis nidulans* R2 there are also most probably three copies of the *psbA* gene in *Synechococcus elongatus* (unpublished results). An EcoRI-Apal restriction endonuclease fragment of 2.6 kb contains one complete copy of one *psbA* gene and in addition 330 base pairs of a second *psbA* gene. The third *psbA* gene is not located on this fragment. The complete *psbA* gene contains an open reading frame of 360 amino acids which is homologous to *psbA* genes of other organisms and encodes the photosystem II reaction center D-1 protein (Fig. 1). The *psbA* gene from *Synechococcus elongatus* has an *E. coli* like consensus promoter sequence at −10: TATATT but differs at the −35 sequence: TTTCAA.

---

* 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 GGC 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 TCC TTC ATC TAC CCC ATT GGT CAA
151 L A S A F A V F I Y P I G Q
V A A A T A V F I I Y P I G Q
* * * * * * * * * * * * * *
496 GCC AGC TTC TCT GAC GGG ATG CCC TCT GCC 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 TCT GTA GCC ATG ATA TTT CAC 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 GCT TCT GTA CTT GTG GGG GCG CTG
196 P F H Q L G V A G V F G C 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 CTC
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 ACC AAC TAC GGT TAC AAA TTT GTT
226 E T T T E T E S A N Y G Y K F G
E T 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
* * * * * * * * * * * * *
The predicted amino acid sequence of the coded D-1 protein was compared to the two different polypeptides coded by the three *psbA* 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 Ser212 to Cys in transmembrane helix IV. Based on the model of the photosystem II reaction center [4] it is feasible that D-1 Cys212 is connected via a disulfide bridge to D-2 Cys212 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.