Molecular Cloning and Sequencing of ADP-Glucose Pyrophosphorylase from Synechocystis PCC 6803

Genichi Kakefuda, Yee-yung Charng, Alberto A. Iglesias, Lee McIntosh, and Jack Preiss

Department of Biochemistry (G.K., Y.-y.C., A.A.I., L.M., J.P.), and Michigan State University-Department of Energy Plant Research Laboratory (L.M.), Michigan State University, East Lansing, Michigan 48824–1319

ADPGlc PPase is the regulatory enzyme for synthesis of starch in plants and glycogen in bacteria (8, 9). Previous work on cyanobacterial ADPGlc PPase has shown the enzyme to have intermediate characteristics to that of the higher plant and bacterial enzymes (4). ADPGlc PPase from Synechocystis PCC 6803 is allosterically activated by 3-phosphoglycerate and inhibited by Pi, as are the higher plant enzymes. The homotetrameric structure of Synechocystis ADPGlc PPase is similar to the enteric bacterial enzymes, which is in contrast with the heterotetrameric nature of all higher plant enzymes studied. Here we report the nucleotide sequence of ADPGlc PPase from Synechocystis PCC 6803. The Synechocystis clone was isolated from a Synechocystis PCC 6803 genomic DNA library. The probe used for screening the library was derived from PCR amplification of genomic Synechocystis DNA.

Amino acid sequences that were highly conserved in both higher plant and bacterial ADPGlc PPase sequences were used to design degenerate primers for PCR amplification of cyanobacterial DNA (Table I). Primer 1, which had a degeneracy of 512, was designed from the conserved amino acid sequences of the *Escherichia coli* ADPGlc PPase FBP activator binding site. The activator binding site determined for the *E. coli* enzyme is conserved in higher plants (7, 10). The conservation occurs despite the fact that FBP does not activate higher plant ADPGlc PPases. Primer 2, which had a degeneracy of 256, was designed from the conserved amino acid sequences (10) of the 8-azido-ADP-glucose affinity labeling site previously determined in the *E. coli* enzyme (5). PCR amplification of genomic *Synechocystis* DNA with these primers generated a fragment of expected size. This fragment was used to isolate a clone from a genomic library.

The nucleotide and deduced amino acid sequence of *Synechocystis* ADPGlc PPase is shown in Figure 1. The first 39 N-terminal amino acids from the deduced sequence are in identity with the sequence determined by N-terminal sequencing of the purified protein from *Synechocystis* (our unpub-

---

1 Supported by grants from the National Science Foundation (NSF 86–10319), National Institutes of Health (AI22835), and U.S. Department of Agriculture/Division of Energy/National Science Foundation Plant Science Center Program (88–37271–3964).

2 Present address: American Cyanamid Company, Plant Biotechnology, P.O. Box 400, Princeton, New Jersey 08543–0400.

3 Abbreviations: ADPGlc PPase, ADP-glucose pyrophosphorylase; PCR, polymerase chain reaction; FBP, fructose 1,6-bisphosphate.

---

**Table I. Characteristics of ADP-Glucose Pyrophosphorylase Genomic DNA from Synechocystis PCC 6803**

| Organism: | Synechocystis sp. strain PCC 6803. |
| Gene Product/Pathway: | ADP-glucose pyrophosphorylase (EC 2.7.7.27); glycogen biosynthesis. |
| Techniques: | PCR: |
| PCR primers: | primer 1, 5'- GAAAGCGGCAAXCCNGCNGT 3' primer 2, 5' ATCACGNGTNCCCAAYCCA 3' |
| N = A + G + T + C, X = A or G, Y = A or T, Z = T or C. |
| A genomic library constructed in lambda fix II (Stratagene) was screened with a radiolabeled probe utilizing the PCR-amplified fragment as template. Other techniques included: restriction enzyme and unidirectional deletion subcloning, complete deoxy sequencing of both strands, computer analysis, comparison, and management of sequences data (5). |
| Method of Identification: | Sequencing of the N-terminal 39 amino acids of ADPGlc PPase purified from Synechocystis PCC 6803 is identical to that deduced from the nucleotide sequence. Sequence similarity of deduced amino acids to that of ADPGlc PPases from spinach leaf 51 kD (65%), rice endosperm (63%), and *E. coli* (37%). |
| Features of Gene Structure: | The start codon is GTG; a Shine-Delgarno sequence located 7 bases upstream of the start codon (Fig. 1, shaded); a sequence with homology to the *E. coli* -35 and -10 box (Fig. 1) is observed. |
| Codon Usage: | Codons not present: TCG, TAA, TGA, CCA, CCG, ACA. (G + C) Content: 48.6% in the coding region. |
| Structural Features of Protein: | Open reading frame 429 amino acids. |
| Calculated M, 48,180. |
| Amino acid sequences similarity was found to: the *E. coli* FBP allosteric activator site (Fig. 1, amino acids 9–32) (7); the *E. coli* 8-azido-ADP-glucose affinity labeling site (Fig. 1, amino acids 95–107) (5); the spinach leaf 51 kD 3-PGA binding site (Fig. 1, amino acids 412–429) (6). |
| Antibodies: | Not available. Cross-reaction occurs with antibodies against either the 51 kD subunit, 54 kD subunit, or the holoenzyme of spinach leaf ADPGlc PPase (4). |

GenBank Accession No: M83556.
1 ATCATACGAGCCACGGGAGCATTGTACTCAGGGGAGTTTCCGACCTTTGCCATTCTGGTTT
61 ATCCGGATCCCCCACTGAACTGACCCGATCCTCGGAAATCCCAACGCGACGATGCACG
121 TTGCTTGGGGCATTAAAACCCTGCTGATTAGCCGAAATTTCCGTCGAGATTTCCCTTCCAG
-10 181 ATGTCCCCCTCCGGTTCTAATACTTTGAGCTCGAGATGTTGTTGGCAGATCGAGACTCG
241 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
301 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
361 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
421 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
481 ACTGCTGATGCGGTACGGCAATACCTCTGCTGTTAGGGGAATGGGACGTAGATGAATAT
541 GAGCTGGGCTTAATGAAAATCGACGCCCAGGGCAGAATTCTGACTTTTCTGAAAAGCCC
601 CTTTATTCTGTCCGGGACCATCTCCGCGAGTTCCGTCGAGATTCCCTTCTGAGGAC
661 CGCGAAGCCATGCGGCAATAACCTTCTCGGTGCTCCGATGACTGAAAGGCACCC
721 GAGCTGGGCTTAATGAAAATCGACGCCCAGGGCAGAATTCTGACTTTTCTGAAAAGCCC
781 ACTGCTGATGCGGTACGGCAATACCTCTGCTGTTAGGGGAATGGGACGTAGATGAATAT
841 GAGCTGGGCTTAATGAAAATCGACGCCCAGGGCAGAATTCTGACTTTTCTGAAAAGCCC
901 CTTTATTCTGTCCGGGACCATCTCCGCGAGTTCCGTCGAGATTCCCTTCTGAGGAC
961 CGCGAAGCCATGCGGCAATAACCTTCTCGGTGCTCCGATGACTGAAAGGCACCC
1021 GAGCTGGGCTTAATGAAAATCGACGCCCAGGGCAGAATTCTGACTTTTCTGAAAAGCCC
1081 CTTTATTCTGTCCGGGACCATCTCCGCGAGTTCCGTCGAGATTCCCTTCTGAGGAC
1141 GAGCTGGGCTTAATGAAAATCGACGCCCAGGGCAGAATTCTGACTTTTCTGAAAAGCCC
1201 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
1261 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
1321 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
1381 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
1441 GAGCTGGGCTTAATGAAAATCGACGCCCAGGGCAGAATTCTGACTTTTCTGAAAAGCCC
1501 ATCATACGAGCCACGGGAGCATTGTACTCAGGGGAGTTTCCGACCTTTGCCATTCTGGTTT
1561 ATCCGGATCCCCCACTGAACTGACCCGATCCTCGGAAATCCCAACGCGACGATGCACG
1621 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
1681 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
1741 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
1801 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
1861 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
1921 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
1981 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
2041 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
2101 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
2161 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
2221 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
2281 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
2341 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
2401 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
2461 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
2521 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
2581 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
2641 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
2701 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
2761 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
2821 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
2881 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
2941 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
3001 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
3061 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
3121 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
3181 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
3241 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
3301 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
3361 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT
3421 ATCCCCGCTCAATTTGACATCCAGAATCTTAAAATTCTGCTCTTACCCGATT
3481 GAGTTTGTGGAGTCTCGCCGGGCTTTTTGCTGAGATTCCCTTCTGAGGAC
3541 CTCTGTGAAGCCTTCAATTGATTCTCTGCTGCTGGCTGTGATAAAACATCTTCTCCTGCTCT
3601 TTAACCAACTGAGCCACAAACCCTGACCTTGGGCGAAATGATCGCTCTACTGAT

Figure 1. Underlined amino acids indicate those that have been confirmed by N-terminal protein sequencing of ADP-Glc pyrophosphorylase purified from Synechocystis PCC 6803. Double underline indicates proposed −10 and −35 box sequences. Shaded areas indicate a proposed Shine-Delgarno prokaryotic ribosome binding sequence. Asterisk indicates a stop codon.
lished results). N-terminal sequencing also confirms that the *Synechocystis* ADP-Glc PPase consists of a single subunit and is not a heterotetramer with subunits of similar molecular mass. The deduced amino acid sequence of *Synechocystis* ADP-Glc PPase was compared with the sequences of rice seed (1), spinach leaf (10), and *E. coli* (2). Based on the percent identity of amino acid sequences (Table I), despite being homotetrameric, the *Synechocystis* enzyme is more similar to the higher plant enzymes than to the bacterial enzymes. Furthermore, the *Synechocystis* protein is more similar to the small than to the large subunit of higher plant ADP-Glc PPases.

**ACKNOWLEDGMENTS**

We would like to thank Mr. James Schilling and Ms. Kim Brown of the University of California, Davis Protein Structure Laboratory for the N-terminal sequence analysis of the *Synechocystis* PCC 6803 ADP-Glc PPase. Brian Smith-White is gratefully acknowledged for advice on molecular techniques.

**LITERATURE CITED**


