Sequence Analysis of the Genome of the Unicellular Cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence Features in the 1 Mb Region from Map Positions 64% to 92% of the Genome (Supplement)

Takakazu Kaneko,1 Ayako Tanaka,1 Shusei Satō,1 Hirokazu Kotani,1 Takashi Sazuka,1 Nobuyuki Miyajima,1 Masahiro Sugiura,2 and Satoshi Tabata1*

*Kazusa DNA Research Institute, 1532-3 Yanauchino, Kisarazu, Chiba 292, Japan 1*

and *Center for Gene Research, Nagoya University, Furho-cho, Nagoya 464-01, Japan 2*

(Received 16 August 1995)

The contiguous sequence of 1,003,450 bp from map positions 64% to 92% of the genome of *Synechocystis* sp. strain PCC6803 has been deduced, and potential ORFs, RNA coding genes and other unique sequences were analyzed by computer. The numerals on the horizontal lines with scale represent the sequence in kb starting from the map position 64%, and restriction sites for Asc I, Mlu I, Spl I, Not I and I-Ceu I are given by red vertical bars. Below the lines, the regions covered by cosmid clones (cs), λ clones (ls) and long PCR products (ps) which used for sequence analysis were indicated by blue bars, and their thick portions represent the regions where the sequences were actually deduced. The potential ORFs taken by the following principles are indicated by boxes with arrowheads indicating the reading direction: (i) All the ORFs longer than 100 codons and only shorter ones which show strong similarities to known genes were taken. (ii) If two ORFs partially or entirely overlapped on either strand, the longer one was chosen unless function of the shorter one was confidently predicted. (iii) Structural RNA genes were taken by preference. The result of similarity search is shown by the following color codes. Green; ORFs identical to reported genes, red; homologues of reported genes, blue; homologues of database-registered hypothetical genes, yellow; ORFs with weak similarities to reported genes, no color; ORFs with no apparent similarity to any reported genes or no significant protein motifs. The gene names are given to the ORFs identical to reported genes. The positions of three IS-like elements identified are shown by dotted boxes, and the rRNA operon and tRNA genes, by the closed box and the blue vertical bars, respectively. The sequence data reported here appear under the accession numbers D63999 (nucleotide positions 1-125,469), D64000 (125,470-271,740), D64001 (271,741-377,535), D64002 (377,536-488,443), D64003 (488,444-601,507), D64004 (601,508-735,706), D64005 (735,707-871,344), and D64006 (871,345-1,003,450).

* To whom correspondence should be addressed. Tel. +81-438-52-3933, Fax. +81-438-52-3934

Downloaded from https://academic.oup.com/dnaresearch/article-abstract/2/4/191/487250 by guest on 04 March 2019