Expression of the chlI, chlD, and chlH Genes from the Cyanobacterium Synechocystis PCC6803 in Escherichia coli and Demonstration That the Three Cognate Proteins Are Required for Magnesium-protoporphyrin Chelatase Activity*

(Received for publication, January 22, 1996, and in revised form, March 21, 1996)

Poul E. Jensen§§, Lucien C. D. Gibson, Knud W. Henningsen, and C. Neil Hunter¶

From the 1 Krebs Institute for Biomolecular Research and the Robert Hill Institute for Photosynthesis, Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, United Kingdom, the Department of Ecology and Molecular Biology, Royal Veterinary and Agricultural University, Bulowsvej 13, DK-1870 Frederiksberg C, Denmark, and the 5 Department of Physiology, Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Copenhagen Valby, Denmark

Magnesium-protoporphyrin chelatase catalyzes the first step unique to chlorophyll synthesis: the insertion of Mg\textsuperscript{2+} into protoporphyrin IX. Genes from Synechocystis sp. PCC6803 with homology to the bchl and bchD genes of Rhodobacter sp. were cloned using degenerate oligonucleotides. The function of these genes, putatively encoding subunits of magnesium chelatase, was established by overexpression in Escherichia coli, including the overexpression of Synechocystis chlH, previously cloned as a homolog of the Rhodobacter bchH gene. The combined cell-free extracts were able to catalyze the insertion of Mg\textsuperscript{2+} into protoporphyrin IX in an ATP-dependent manner and only when the products of all three genes were present. The ChlH, ChlI, and ChlD gene products are therefore assigned to the magnesium chelatase step in chlorophyll a biosynthesis in Synechocystis PCC6803. The primary structure of the Synechocystis ChlD protein reveals some interesting features; the N-terminal half of the protein shows 40–41% identity to Rhodobacter Bchl and Synechocystis ChlI, whereas the C-terminal half displays 33% identity to Rhodobacter BchD. This suggests a functional as well as an evolutionary relationship between the “I” and “D” genes.

---

*This work was supported by Grants 13-5005-1 (to K. W. H.), and 9400619 (to P. E. J.) from the Danish Agricultural and Veterinary Research Council and by a grant (to C. N. H.) from the Biotechnology and Biological Sciences Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EBI Data Bank with accession number(s) X96599 and U35344.

To whom correspondence should be addressed. Tel.: 44-114-282-4240; Fax: 44-114-272-8697; E-mail: P.E.Jensen@Sheffield.ac.uk.

† The abbreviations used are: PCR, polymerase chain reaction; HPLC, high pressure liquid chromatography.

---

Magnesium-protoporphyrin chelatase catalyzes the insertion of Mg\textsuperscript{2+} into protoporphyrin IX, the first step unique to chlorophyll production, and lies at the branch point of the heme and chlorophyll production, and lies at the branch point of the heme and chlorophyll biosynthetic pathways (Fig. 1). Consequently, it is an essential first step in the development of photosynthetic membranes. Chelation of Mg\textsuperscript{2+} into protoporphyrin IX has been demonstrated in chlorophyll a-synthesizing organisms using isolated plastsids (1–3) and broken plastid systems (4–8). In these systems, ATP is absolutely required, and there is evidence that in pea (Pisum sativum) chloroplasts, two protein fractions participate in the enzymatic reaction and that ATP is required for both activation of the proteins and magnesium chelation (8). However, no specific proteins have been identified as components of magnesium chelatase.

In contrast, it has been possible to establish the identity of the genes encoding the magnesium chelatase enzyme in the purple bacterium Rhodobacter sphaeroides. Magnesium chelatase in this organism is encoded by the bchl, bchD, and bchH genes (9), and through expression of the bchl, bchD, and bchH genes in Escherichia coli, it has been shown that the three gene products are able to catalyze the insertion of magnesium into protoporphyrin IX in an ATP-dependent manner (9). In higher plants, nuclear genes corresponding to the Rhodobacter bchl and bchH genes have been identified (10–13). Furthermore, three loci required for insertion of Mg\textsuperscript{2+} into protoporphyrin IX have been identified in barley (Hordeum vulgare) (14, 15). Two of these loci correspond to the bchl and bchH genes of Rhodobacter, and the third locus is probably the plant homolog of the bacterial bchD gene (12). However, the third gene has not been identified and characterized, and more generally, magnesium chelatase has not been studied at the biochemical and molecular level in a chlorophyll a-producing organism.

The work described was carried out on the cyanobacterium Synechocystis sp. PCC6803 since this organism shares many features with higher plants. It synthesizes chlorophyll a and harbors an oxygen-evolving photosynthetic apparatus that, with exception of the peripheral light-harvesting complexes, is essentially identical to that of chloroplasts in plants and algae. At the same time, it retains the advantages of a prokaryote in terms of genome size and ease of genetic manipulation (see, for example, Refs. 16 and 17). Previously, we reported the cloning and sequencing of a gene (chlI) homolog to bchH of Rhodobacter (18). In this paper, we report the cloning of chlI and chlD and present biochemical evidence that the ChlI, ChlD, and ChlH polypeptides together compose the magnesium chelatase enzyme of the cyanobacterium Synechocystis PCC6803.

**EXPERIMENTAL PROCEDURES**

Cloning of chlI and chlD—Degenerate primers were designed from conserved regions of known CcsA/Ch42/ChlI sequences. One upstream primer (P1) is was made from the amino acid motif WHTVEREG (PI6). Two downstream primers were made from the motifs WNTVEREG (P16) and NPEEGELR (PI7), corresponding to 5'-CCCTC(G/C)(G/T)(G/C)(A/G)(A/G)(A/G)/(C/G)AGAT(S)AG-3' and 5'-C(A/C)(G)/CAG(10). Two downstream primers were made from the motifs WNTVEREG (P16) and NPEEGELR (P17), corresponding to 5'-CCCTC(G/C)(G/T)(C/C)(A/G)(A/G)(C/G)AGAT(S)AG-3' and 5'-C(A/C)(G)/CAG(10). Two downstream primers were made from the motifs WNTVEREG (P16) and NPEEGELR (P17), corresponding to 5'-CCCTC(G/C)(G/T)(C/C)(A/G)(A/G)(C/G)AGAT(S)AG-3', respectively. PCR\textsuperscript{1} was run in 16662
grown at 25°C in 50 ml of LB medium (20) containing 200 µg/ml of ampicillin/ml (pET3a) or 200 µg/ml of gentamicin/ml (pET9a) until the optical density of the cultures reached 0.6–1.0. Gene expression was induced by addition of isopropyl-β-D-thiogalactopyranoside to the cultures at a final concentration of 0.4 mM. After 4–9 h, the cells were harvested, and the cell pellets were stored at −20°C.

SDS-Polyacrylamide Gel Electrophoresis and Magnesium-protoporphyrin Chelatase Assays—SDS-polyacrylamide gel electrophoresis, purification of porphyrins, and magnesium-protoporphyrin chelatase assays were performed as described previously (9). The fluorescence emission spectra of enzyme assays were recorded between 550 and 650 nm on a Spex Fluorolog spectrofluorometer using an excitation wavelength of 420 nm.

HPLC Analysis of Pigments—For HPLC analysis, 200 µl of the assay mixture was extracted with 1.0 ml of 80% acetone containing 0.1 n NH₃, mixed with 400 µl of hexane, and centrifuged at 20,000 × g for 3 min. The acetone phase was analyzed for porphyrins by HPLC using a Beckman UltraspHERE ODS column (150 × 4.6 mm). The column was eluted with a 10-min linear gradient from 15% solvent B to 100% solvent B at 1 ml/min. Solvent A contained 0.05% (v/v) triethylamine in water; solvent B contained acetonitrile. Porphyrins were detected using a Waters in-line fluorescence detector. Magnesium protoporphyrin IX was detected using an excitation wavelength of 420 ± 5 nm and an emission wavelength of 595 ± 5 nm.

RESULTS

Cloning of chlI and chlD from Synechocystis Using PCR—Gene homologs to the Rhodobacter bchI and bchH genes encoding magnesium chelatase subunits have been identified in algae and plants. However, no gene corresponding to the Rhodobacter bchD gene had been identified in any chlorophyll-synthesizing organism, and it remained uncertain whether the magnesium chelatase of such organisms consists of three subunits as in Rhodobacter. The first stage in the cloning process relied upon the fact that there are some amino acid similarities between Rhodobacter capsulatus BchI and the amino-terminal region of BchD (see Fig. 3); in selected regions, the identity is 19–24% between BchD and the BchI/ChlI protein sequences from different organisms. Setting aside any possible functional significance, this homology provided a convenient location for the design of PCR primers, and in an attempt to exploit this conservation to clone a bchD homolog, we designed three oligonucleotides (P11, P16, and P17) (see Fig. 3) whose sequences were based on amino acid sequence identities in the BchI homologs from Arabidopsis thaliana (10), Euglena gracilis (22), Cryptomonas ph (EMBL accession no. Z21976), Olisthodiscus luteus (EMBL accession no. S32166), and R. capsulatus (23). Using these oligonucleotides as primers in PCR with Synechocystis PCC6803 genomic DNA as template, we obtained two different DNA fragments. With P11 and P16, a 483-base pair PCR fragment was obtained that was cloned (pSX) and sequencing, displayed only 39% identity to the BchI and Ch42 gene sequences of the governing magnesium chelatase subunit from each plasmid clone was used to screen genomic libraries, and several independent λ clones were isolated. Restriction mapping, subcloning, and sequencing revealed that clones isolated using the insert from pSX retained the sequence and orientation of the insert from pSX. A 2.3-kilobase region of λSI containing the entire open reading frame encoding the Synechocystis bchI homolog was sequenced (Fig. 2). This gene encodes a protein of 357 amino acids, with an estimated molecular mass of 39,629 Da and 51% overall identity to BchI of R. capsulatus and 69–73% identity to the plant and algal CcsA/Ch42 sequences. We propose that this gene should be called chlI. Of the genomic clones isolated with the insert of pSX, a 5.2-kilobase region was sequenced, and two major open reading frames were identified (Fig. 2). One of the
open reading frames encodes a 393-amino acid polypeptide (45,468 Da), and data base comparisons revealed that it had some homology to a putative heme-binding protein (Hbp) known from the chloroplast genome of liverwort and higher plants. The other open reading frame had the potential to encode a 676-amino acid polypeptide with an estimated molecular mass of 73,699 Da. This reading frame showed an overall amino acid identity of 32% when compared with BchD of Synecho-
cystis, but now the homology was not partic-
larly high in comparison with the identities between Synechocystis ChlI and ChlD and the corresponding R. capsulatus proteins, which are 51 and 41%, respectively. The similarity between the N-terminal regions of BchD and Bchl noted earlier was again seen with Synechocystis ChlD, but now the homology was much more marked (Fig. 3). Thus, the deduced ChlD polypeptide sequence shares similarity with both Bchl and BchD. The N-terminal part of ChlD (residues 1–326) displays 40–41% identity to Bchl and ChlI and only 18% identity to the corresponding region in BchD. The C-terminal part of ChlD (residues 327–676) displays 33% identity to the corresponding region of BchD, and in this region, several stretches of identical amino acids between the two proteins can be found, which makes the relation between the two proteins obvious. As for BchD, Synechocystis ChlD has a proline-rich region that divides the protein into two major domains. In fact, the similarity between the N-terminal parts of ChlD and Bchl ends just before this proline-rich region. Immediately following the proline-rich region of ChlD, a very acidic region is found; in a stretch of 40 amino acids, 19 residues are either aspartic acid or glutamic acid (Fig. 3).

Biochemical Role of ChlI, ChlD, and ChlH—The overall homology between BchD and ChlD would not, in itself, give confidence that ChlD serves the same function as part of a magnesium chelatase. Thus, to establish the biochemical function of the chlD gene from Synechocystis, this gene was overexpressed in E. coli, along with the Synechocystis chlI and chlH genes. DNA fragments containing the chlI, chlH, and chlD genes were cloned separately into pET3a or pET9a, yielding plasmids pET9a-ChlI, pET9a-ChlD, and pET3a-ChlH, respectively, and the E. coli cultures containing these constructs were induced with isopropyl-β-d-thiogalactopyranoside. The cells from the respective inductions were disrupted, and the crude extracts were centrifuged; the resulting cell-free supernatants were used for all subsequent work. The supernatants contained the proteins that gave rise to magnesium chelatase activity, suggesting that the ChlI, ChlD, and ChlH proteins accumulated in a soluble form. This was further supported by SDS-polyacrylamide gel electrophoresis analysis of the different fractions. We found that induction of protein production at 25°C for 4 h yielded significantly more soluble protein, especially of ChlI and ChlH, than induction at 37°C. The ChlI, ChlH, and ChlD proteins are clearly visible on the SDS-polyacrylamide gel, having apparent molecular masses of ~40, 70, and 150 kDa, respectively (Fig. 4). This is in good accordance with the molecular masses deduced from the nucleotide sequences, which are 39,629, 73,699, and 148,659 Da, respectively. In contrast to the expression of the R. sphaeroides bchD gene in E. coli (9), we observed significant expression of the Synechocystis chlD gene in E. coli. We estimate that overexpressed ChlI, ChlD, and ChlH compose ~30–40% of the total soluble protein in the E. coli cell after 4 h of induction.

Cell-free extracts prepared from E. coli strains containing ChlI, ChlD, and ChlH were mixed in various combinations and incubated with protoporphyrin IX, Mg2+, and ATP in a regenerating system in order to establish the combination of proteins that gave rise to magnesium chelatase activity. Formation of magnesium protoporphyrin IX in these assays was established by fluorescence emission spectroscopy. The fluorescence emission spectra of the assays after 30 min of incubation are shown in Fig. 5. Magnesium protoporphyrin IX has a characteristic emission maximum at 595 nm (Fig. 5, trace MgP), whereas the substrate has an emission maximum at 633 nm. The emission maximum at 595 nm is clearly present in the sample containing all three proteins (Fig. 5, trace I–D–H) and not in any of the other assays. The emission maximum at 595 nm is identical to that of the authentic magnesium protoporphyrin IX control (Fig. 5, trace MgP). In addition, the identity of the product formed in the ChlI, ChlD, and ChlH incubation was verified by HPLC; the retention times of authentic magnesium protoporphyrin IX and the product of the reaction are identical, demonstrating that the product formed is magnesium protoporphyrin IX (Fig. 6). An assay containing ChlI, ChlD, ChlH, Mg2+, and protoporphyrin IX, but no ATP, was also performed. In Fig. 7, the fluorescence emission spectrum of this assay is compared
with that of an assay containing ATP, and it is clear that the Synechocystis magnesium chelatase is ATP-dependent. Using the cell-free extracts and by roughly estimating the fraction of the overexpressed protein in these, we found that the molar ratio of the three components yielding optimal catalytic activity is approximately 10 ChlI:1 ChlD:10 ChlH.

**DISCUSSION**

The data presented in this paper provide the first direct evidence that magnesium insertion in a chlorophyll a-producing organism requires three different protein subunits. Thus, magnesium chelatase in Synechocystis PCC6803 consists of three components: the ChlI, ChlD, and ChlH proteins with deduced molecular masses of 39,629, 73,699, and 148,659 Da, respectively. When the genes are overexpressed in E. coli, the cell-free extracts contain significant levels of the overproduced proteins (Fig. 4). When the extracts containing these proteins are mixed, an active chelatase enzyme forms, and it has been demonstrated that this enzyme catalyzes the chelation of magnesium by protoporphyrin IX in an ATP-dependent manner.

It has also been demonstrated that there is an equivalent of the BchD protein, a subunit of the bacteriochlorophyll a-specific magnesium chelatase, in a chlorophyll a-synthesizing organism. The primary structure of the Synechocystis ChlD protein reveals some interesting features (Fig. 3). The amino-terminal 326 residues of ChlD displays 40–41% amino acid identity to the BchI and ChlI proteins of *R. capsulatus* and Synechocystis. This indicates an evolutionary relationship between "I" and "D" genes; it is possible that the D gene originated from a fusion between a duplicated I gene (the N-terminal part) and an ancestral D gene (the C-terminal part). The conserved motifs between ChlD and BchD/BchI probably represent residues with structural and/or functional importance.

The Synechocystis ChlD protein sequence is clearly divided into two major domains separated by the proline-rich sequence. A sequence of four or more prolines in a row adopts a single preferred conformation in solution, known as the polyproline II helix (26). This is an extended structure with three residues/turn, and in proteins, they generally form a rigid stem that protrudes from membranes or globular proteins (27). Although in many proteins the function of the proline-rich region remains uncertain, in some cases, it is involved in interaction with other proteins or cofactors; in other cases, the function is
structural, by separating the protein into two domains (reviewed in Ref. 27). In ChlD, both functions could be relevant; ChlD is divided into two major domains separated by the proline-rich region, and ChlD has to interact with at least one other protein component in the magnesium chelatase enzyme.

The amino acid sequence of ChlD has no long hydrophobic regions that could suggest membrane association. On the basis of their amino acid sequences, ChlI and ChlH also appear to be soluble, and it seems likely that magnesium chelatase is soluble. This is further supported by results obtained with pea and cucumber magnesium chelatases, which show no indication of being membrane proteins since the activity is readily solubilized by washing the total chloroplast membranes in buffer of low MgCl₂ content (28).

With respect to the number of different protein subunits, subunit sizes, and the absolute ATP requirement, the Synechocystis magnesium chelatase parallels the R. sphaeroides magnesium chelatase (9). Thus, the magnesium chelatase enzyme is conserved between purple bacteria and cyanobacteria. The Synechocystis ChlI, ChlD, and ChlH polypeptides are clearly related to BchI, BchD, and BchH of R. sphaeroides. However, it is possible that the subunit interactions and the stoichiometry between subunits are different for the two enzymes. Biochemical studies on magnesium chelatase from pea chloroplasts implied that the enzyme was of a multicomponent nature (8). The data presented here show that the magnesium chelatase of Synechocystis, and probably also of higher plants, consists of three different polypeptides and thereby demonstrate the multicomponent nature of this enzyme.

Acknowledgments—We are grateful to Dr. Hirokazu Kotani (Kazusa DNA Research Institute, Kazusa, J apan) for mapping the hbp, chlD, chlI, and chlH genes on the physical map of Synechocystis PCC6803. We are also grateful to Prof. Diter von Wettstein (Carlsberg Laboratory, Copenhagen) for stimulating discussions and critical reading of the manuscript. We thank Grant Naylor (University of Sheffield) for assistance in preparation of the figures. The use of the SEQNET facility of the Biotechnology and Biological Sciences Research Council (Daresbury, United Kingdom) is acknowledged.
REFERENCES

1. Castelfranco, P. A., Weinstein, J. D., Schwarz, S., Pardo, A. D., and Wezelman, B. E. (1979) Arch. Biochem. Biophys. 192, 592–598
2. Pardo, A. D., Chereskin, B. M., Castelfranco, P. A., Franceschi, V. R., and Wezelman, B. E. (1980) Plant Physiol. (Bethesda) 65, 956–960
3. Fuesler, T. P., Wright, L. A., Jr., and Castelfranco, P. A. (1981) Plant Physiol. (Bethesda) 67, 246–249
4. Richter, M. L., and Rienits, K. G. (1982) Biochim. Biophys. Acta 717, 255–264
5. Walker, C. J., and Weinstein, J. D. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 5789–5793
6. Lee, H. J., Ball, M. D., Parcham, R., and Rebeiz, C. A. (1992) Plant Physiol. (Bethesda) 99, 1134–1140
7. Walker, C. J., Hupp, L. R., and Weinstein, J. D. (1992) Plant Physiol. & Biochem. 30, 263–269
8. Walker, C. J., and Weinstein, J. D. (1994) Biochem. J. 299, 277–284
9. Gibson, L. C. D., Willows, R. D., Kannangara, C. G., von Wettstein, D., and Hunter, C. N. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1941–1944
10. Koncz, C., Mayerhofer, R., Koncz-Kalman, Z., Nawrath, C., Reiss, B., Redei, G. P., and Schell, J. (1990) EMBO J. 9, 1337–1346
11. Hudson, A., Carpenter, R., Doyle, S., and Coen, E. S. (1993) EMBO J. 12, 3711–3719
12. Jensen, P. E., Willows, R. D., Larsen-Petersen, B., Vothknechet, U. C., Stummann, B. M., Kannangara, C. G., von Wettstein, D., and Henningsen, K. W. (1996) Mol. & Gen. Genet. 250, 383–394
13. Gibson, L. C. D., Harrison, J. J., Leech, R. M., Jensen, P. E., Bassham, D. C., Gibson, M., and Hunter, C. N. (1996) Plant Physiol. (Bethesda) 111, 61–71
14. von Wettstein, D., Henningsen, K. W., Boynton, J. E., Kannangara, G. C., and Nielsen, O. F. (1971) in Autonomy and Biogenesis of Mitochondria and Chloroplasts (Bordmann, N. K., Linnane, A. W., and Smillie, R. M., eds) pp. 205–223, North-Holland, Amsterdam
15. Henningsen, K. W., Boynton, J. E., and von Wettstein, D. (1993) Biol. Skr. 42, 1–349
16. Kotani, H., Kaneko, T., Matsubayashi, T., Sato, S., Sugiura, M., and Tabata, S. (1994) DNA Res. 1, 303–307
17. Kotani, H., Tanaka, A., Kaneko, T., Sato, S., Sugiura, M., and Tabata, S. (1995) DNA Res. 2, 133–142
18. Jensen, P. E., Stummann, B. M., and Henningsen, K. W. (1996) Plant Mol. Biol. 30, 1075–1076
19. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987) Current Protocols in Molecular Biology, Greene Publishing Associates/Wiley-Interscience, New York
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. Rosenberg, A. H., Lade, B. N., Chui, D.-S., Lin, S.-W., Dunn, J. J., and Studier, F. W. (1987) Gene (Amst.) 56, 125–135
22. Orsat, B., Monfort, A., Chatellard, P., and Stutz, E. (1992) FEBS Lett. 303, 181–184
23. Armstrong, G. A., Alberti, M., Leach, F., and Hearst, J. E. (1989) Mol. & Gen. Genet. 216, 254–268
24. Zsebo, K. M., and Hearst, J. E. (1984) Cell 37, 937–947
25. Coomber, S. A., Chaundri, M., Connor, A., Britton, G., and Hunter, C. N. (1990) Mol. Microbiol. 4, 977–989
26. Cowan, P. M., and McGavin, S. (1955) Nature 176, 501–503
27. Williamson, M. P. (1994) Biochem. J. 297, 249–260
28. Walker, C. J., and Weinstein, J. D. (1995) Physiol. Plant 94, 419–424
29. Synechocystis Magnesium Chelatase 16667