Carboxysomes are an essential part of the cyanobacterial CO2-concentrating mechanism, consisting of a protein shell and an interior of Rubisco. The β-carboxysome shell protein CcmM forms two peptides via a proposed internal ribosomal entry site (IRES) within the ccmM transcript in Synechococcus PCC7942. The abundant short form (35 kD, M35) consists of Rubisco small subunit-like repeats and binds Rubisco. The lower abundance long form (58 kD, M58) also contains a γ-carbonic anhydrase-like domain, which binds the carboxysomal carbonic anhydrase, CcaA. We examined whether these CcmM forms arise via an IRES or by other means. Mutations of a putative internal start codon (GTG) and Shine-Dalgarno sequence within ccmM, along with a gene coding for M35 alone, were examined in the high-CO2-requiring (HCR) carboxysomeless mutant, ΔccmM. Expression of wild-type ccmM in ΔccmM restored the wild-type phenotype, while mutation of putative start and Shine-Dalgarno sequences led to as much as 20-fold reduction in M35 content with no recovery from HCR phenotype. These cells also contained small electron-dense structures. Cells producing little or no M58, but sufficient M35, were found to contain large electron-dense structures, no CcaA, and had a HCR phenotype. Large subcellular aggregates can therefore form in the absence of M58, suggesting a role for M35 in internal carboxysome Rubisco packing. The results confirm that M35 is independently translated via an IRES within ccmM. Importantly, the data reveal that functional carboxysomes require both M35 and M58 in sufficient quantities and with a minimum stoichiometry of close to 1:1.
organizing components that may coordinate the initial nucleation of the carboxysome shell facets by binding Rubisco and the carboxysomal carbonic anhydrase (CA), CcaA, as well as organizing an outer layer of interlocking CcmK hexamers (Long et al., 2007). CcmM is characterized by two specific domains: an N-terminal γ-CA-like domain and a C-terminal domain of Rubisco small subunit (SSU)-like repeats (Price et al., 1993; Ludwig et al., 2000). The crystal structure of the CA-like domain of *Thermosynechococcus elongatus* BP-1 CcmM has recently been elucidated and shown to be an active CA in that and other species lacking *ccaA* (Peña et al., 2010). In addition, there is consistent evidence that CcmM exists as at least two forms in a number of β-cyanobacteria; our studies show that in *Synechococcus PCC7942* both the N terminus to be Ser217, possibly resulting from a possible post-translational modification (M35) that contains three SSU-like repeats (Price et al., 1998; Long et al., 2005).

In *Synechococcus PCC7942* M35 possibly results from a putative internal ribosomal entry site (IRES) within the *ccmM* gene transcript (Price et al., 1998; Ludwig et al., 2000; Long et al., 2005, 2007; Fig. 1). N-terminal sequencing of the short form of CcmM (M35) suggested the N terminus to be Ser217, possibly resulting from a putative GTG start codon downstream of the putative Shine-Dalgarno (SD) ribosomal binding sequence (GAGG), and subsequent N-terminal Met loss (Long et al., 2007). This corresponds to an equivalent peptide from the CcmM of *Nostoc* (Anabaenopsis) PCC7120 and matches likely short forms from a suite of related CcmM sequences (Long et al., 2007). In support of these findings, Cot et al. (2008) showed that CcmM from *Synechocystis* PCC6803 is present in carboxysomes as a number of peptides resulting from multiple putative IRESs and translation initiation codons within the gene transcript. In all cases, it appears that short forms of CcmM maintain complete Rubisco SSU repeats (Ludwig et al., 2000), which are responsible for Rubisco binding and playing an important role in carboxysome structure (Long et al., 2007). However, the true origin of CcmM short forms is unknown.

In a recent review we highlighted the potential to utilize components of the cyanobacterial CCM to enhance nitrogen- and water-use efficiency in C₃ crop plants (Price et al., 2008). A long-term part of this approach could incorporate carboxysome-like structures into C₃ chloroplasts to generate a compartmentalized Rubisco with an elevated CO₂ concentration. To make progress in this area, however, further comprehensive information on the interactions of cyanobacterial Rubiscos and their carboxysomal component proteins is required. Here we investigate the complicated expression of two important forms of CcmM to better understand their roles in carboxysome assembly. We show that an IRES and translation start codon are responsible for the production of two peptides from the *ccmM* gene and that there is an absolute requirement for both M35 and M58 in functional carboxysome formation. The results indicate specialized roles for M35 and M58 in both carboxysome structure and function.

**RESULTS**

The carboxysome production capabilities, requirement for high CO₂, and distribution of carboxysome proteins from each *ccmM* mutation (Fig. 1) are summarized in Table I. Mutation of the putative internal start codon or putative SD sequence in *ccmM*, in the carboxysomeless deletion mutant Δ*ccmM* (Woodger et al., 2005; Emlyn-Jones et al., 2006), led to a significant reduction in the CcmM short form (M35) in *Synechococcus PCC7942* cells (Table I; Fig. 2). Immunoblots of cell lysates and Mg²⁺ precipitable proteins from wild type and mutant cell types (Fig. 2) showed that both forms of CcmM were present in wild-type cell extracts and those of Δ*ccmM* + *ccmM*, Δ*ccmM* + gtc, Δ*ccmM* + NL, and Δ*ccmM* + SD mutants. However, M35 amounts were significantly diminished in

---

**Figure 1.** Wild-type and mutant *Synechococcus PCC7942 ccmM* gene sequence constructs. Partial wild-type *Synechococcus PCC7942 ccmM* gene sequence (WT) indicating the positions of the putative internal SD sequence (shaded) and putative internal start codon (underlined) is shown. Partial amino acid sequences of the resulting polypeptides (M58, M35, and M23) are shown with the confirmed N-terminal amino acid of M35 (boxed; Long et al., 2007). N-terminal sequence analysis suggests that the initial Met of M35 is lost posttranslationally (Long et al., 2007). The mutated read-through sequences, *ccmM*-NL and *ccmM*-gtc, are indicated with individual base changes from the wild-type sequence (underlined) and individual amino acid changes (bold). The resulting Ctg (Leu) codon in the *ccmM*-NL mutation is also a rare initiation codon in some instances (Missiakas et al., 1993; Sazuka et al., 1999; Tichy and Vermaas, 1999). The *ccmM*-FS mutation introduces a stop codon (shaded) while maintaining the M35 sequence by introducing a FS by removal of a single base (indicated by an arrowhead). The *ccmM*-SD mutation eliminates the putative internal SD sequence by a single codon mutation (underlined). The construct coding for M35 alone (*ccmM*-M35) has the translation start codon immediately downstream of the nirA promoter in plasmid pSE4 (Maeda et al., 1998; Price et al., 2004).
Table 1. Summary of carboxysome protein characteristics and CO$_2$ requirement in ccmM mutants

The presence of each of the carboxysome proteins identified in each Mg$^{2+}$ fraction is subjectively identified from western blots (Fig. 2) according to the following scheme: present but low abundance (+), present and high abundance (+++), present and equivalent abundance in each fraction (++), or absent (−). In wild-type cells, carboxysomes are predominantly associated with the Mg$^{2+}$ pellet fraction. HCR in each cell type is indicated.

| Mg$^{2+}$ Fraction | Cell Type       | Carboxysome Protein | Carboxysones | HCR | $K_{0.5}$ Ci$^+$ |
|-------------------|-----------------|---------------------|--------------|-----|-----------------|
| Wild type         | M58             | +                    | Present      | No  | 0.27 ± 0.08     |
| ΔccmM             | M58             | +                    | Present      | No  | 15.1 ± 1.39     |
| +ccmM             | M58             | +                    | Present      | No  | 0.23 ± 0.03     |
| +M35              | M58             | −                    | Absent       | Yes | 19.5 ± 0.83     |
| +ccmM-NL          | M58             | +                    | Small EDB$^b$ | Yes | 18.7 ± 1.62     |
| +ccmM-gtc         | M58             | −                    | Absent       | Yes | 19.9 ± 0.25     |
| +ccmM-FS          | M58             | −                    | Absent       | Yes | 19.7 ± 1.96     |
| +ccmM-SD          | M58             | −                    | Absent       | Yes | 10.1 ± 0.72     |
| Pellet            | Wild type       | +++                  | Present      | No  | 0.27 ± 0.08     |
| ΔccmM             | M58             | −                    | Absent       | Yes | 15.1 ± 1.39     |
| +ccmM             | M58             | −                    | Absent       | Yes | 19.5 ± 0.83     |
| +M35              | M58             | −                    | Absent       | Yes | 19.9 ± 0.25     |
| +ccmM-NL          | M58             | −                    | Absent       | Yes | 19.7 ± 1.96     |
| +ccmM-gtc         | M58             | −                    | Absent       | Yes | 10.1 ± 0.72     |
| +ccmM-FS          | M58             | −                    | Absent       | Yes | 19.9 ± 0.25     |
| +ccmM-SD          | M58             | −                    | Absent       | Yes | 10.1 ± 0.72     |

$^a$Photosynthetic half-saturation concentrations of Ci for wild-type and mutant cell types. Photosynthetic parameters were determined by membrane inlet mass spectrometry as described in the text. Data presented are means of triplicate measurements ±sd. Maximum photosynthetic rates measured were in the range 330 to 513 μmol O$_2$ evolved (mg Chl)$^{-1}$ h$^{-1}$. A representative set of mass spectrometric measurements of Ci-dependent O$_2$ evolution in wild-type Synechocystis PCC7942 and ΔccmM mutants is presented in Supplemental Figure S1.

$^b$EDB, Electron-dense bodies. The ΔccmM + ccmM-NL cells contain small electron-dense bodies with apparent planar surfaces (Fig. 3).

$\Delta$ccm+gtc, $\Delta$ccm+N-L, and $\Delta$ccm+S-D compared with quantities found in wild-type and ΔccmM+ccmM cells (Fig. 2).

Neither M58 nor M35 were present in ΔccmM cell extracts. M58 could not be detected in ΔccmM + M35 or ΔccmM + frame shift (FS) cells, where an introduced FS was designed to eliminate M58 production but maintain M35 via the uninterrupted IRES (Fig. 1). However, both of these cell types contained greater relative quantities of M35 than those found in wild-type cells (Fig. 2). The absence of M58 from ΔccmM + FS cells suggests successful termination of M58 translation at the introduced stop codon (TGA), while M35 production within these cells confirms independent translation of the shorter peptide via the putative GTG start codon.

While expression of the mutated forms of ccmM, or the M35 form alone, within the ΔccmM background did not alleviate the high-CO$_2$-requiring (HCR) phenotype, expression of the wild-type gene ($\Delta$ccmM + ccmM) recovered wild-type physiology (Table I). The ΔccmM + SD cell type displayed intermediate inorganic carbon (C) requirement for photosynthesis with a slightly lower $K_{0.5}$ Ci (concentration at which half maximum photosynthetic rate occurs) than the other ccmM mutants but it still required high CO$_2$ for growth (Table I; Supplemental Fig. S1).

Magnesium precipitation of carboxysome-related complexes from each type that showed the carboxysomal CA CcaA was undetectable in ccmM and ΔccmM + M35 cells (Fig. 2). However, CcaA was detected in all other cell types (Fig. 2). Nonetheless, all the cell types producing M58 or the 23-kD (M23) γ-Ca domain ($\Delta$ccmM + FS cells) M23 contained relatively low CcaA quantities compared to wild type (Fig. 2). Detection of M23 in ΔccmM + FS cells was not achieved by western-blot analysis due to both extremely poor immunogenicity of this region of CcmM in polyclonal antisera raised against M58, and probable low abundance of the peptide in this cell line (although, in the absence of specific proteases, M23 can be detected in high abundance in Escherichia coli expression cell lines; Supplemental Fig. S2). Because of the known interaction between CcaA and the γ-Ca domain of CcmM (Long et al., 2007; Cot et al., 2008), the detection of stable CcaA in ΔccmM + FS cells (Fig. 2) was presumed to indicate the presence of M23 at below-detectable quantities in these cells.  

Immunoblots of lysates, Mg$^{2+}$ supernatants, and pellets revealed that both RbcL and CcmK1 (the homolog of CcmK2 from Synechocystis PCC6803) were present in all cell types and were successfully precipitated from cell lysates with Mg$^{2+}$ in most cases (Fig. 2). Several additional bands cross-reacting with the CcmK1 antibody were observed in ΔccmM cell lysates but not in any other cell lines (Fig. 2), indicating potential proteolysis of this protein in this cell line; these protein bands appeared to be too small to be the homologs CcmK3 and CcmK4. With regard to CcmM, there was no evidence of any peptides smaller than...
35 kD that might arise from proteolytic cleavage of M58 into two distinct products, as evidenced by probing western blots with a polyclonal antibody specifically raised against the entire mature M58 polypeptide (data not shown). Specifically, no protein corresponding to a 23 kD or smaller M58 proteolysis fragment was observed. In ΔccmM + M35 and ΔccmM + FS cells, where M58 was absent, carboxysomal proteins were primarily found in the supernatant after Mg2+ treatment (Table I). This is not the case for ΔccmM cells, however. CcmK1 was confined to the supernatant fraction after Mg2+ precipitation of lysates from ΔccmM + SD, ΔccmM + FS, and ΔccmM + M35 cells (Fig. 2).

Electron-dense structures (possibly rhomboid and appearing rectangular in cross section) were observed in ΔccmM + FS and ΔccmM + M35 cells where M35 was the only CcmM polypeptide present (Fig. 3). Electron microscopy revealed that wild-type and ΔccmM + ccmM cell types contained carboxysomes while ΔccmM did not (Fig. 3). Cells of ΔccmM + NL, ΔccmM + gtc, and ΔccmM + SD mutants contained electron-dense structures although, for the most part, properly formed carboxysomes were not observed in sections of these cell types (Fig. 3). The ΔccmM + NL cells contained electron-dense regions with planar surfaces, reminiscent of carboxysome edges. The ΔccmM + gtc and ΔccmM + SD cells predominantly contained small dense structures that have the appearance of very small carboxysome structures and might also be carboxysome related (Fig. 3).

**DISCUSSION**

**M35 Arises from an IRES and GTG Start Codon in ccmM**

Mutations of a putative IRES and start codon within the *Synechococcus* PCC7942 *ccmM* gene have enabled us to determine a translational origin for the carboxysomal protein M35. Our data reveal a significant reduction in M35 production in the ΔccmM + gtc mutant (mutation of a GTG start codon to a frequently used but nonstart GTC codon), indicating this protein arises via a mechanism other than M58 proteolysis. Being a silent mutation, retaining the internal Val, ΔccmM + gtc should result in wild-type or at least ΔccmM + ccmM levels of both CcmM peptides if M35 results from M58 proteolysis. This is clearly not the case, with M35 approximately 20-fold lower in ΔccmM + gtc cells compared to wild type (Fig. 2). In addition, M35 production still occurs (albeit at low levels) in the *ccmM + NL* mutant that was designed to eliminate any potential proteolytic cleavage site within M58 by varying the amino acid residues at the known M35 N terminus, but retaining a known CTG start codon. Furthermore, immunoblots of cell extracts of all cell types using a polyclonal antibody (that recognizes the entire M58 polypeptide) was not successful in identifying any M58-like peptides (potential cleavage products) smaller than M35 (B.M. Long, unpublished data). This antibody shows relatively poor, yet sufficient, detection of the N-terminal (M23) γ-CA-like domain of CcmM when expressed in *E. coli*.

**Figure 2.** Carboxysomal proteins in wild-type Synechococcus PCC7942 and ccmM mutants. Western blots showing the presence or absence of carboxysomal proteins in the whole-cell lysates (L), Mg2+ supernatants (S), and Mg2+ pellets (P) from wild-type (WT), ΔccmM, and ΔccmM complementation mutant cell types. Cells were lysed and crude carboxysome-rich preparations formed by precipitation with 25 mM Mg2+ as described in the text. In all cases samples equivalent to 0.5 μg Chl in the lysate fraction were loaded onto SDS-PAGE gels. Separated proteins were transferred onto PVDF membranes and probed as described in the text. The proteins identified are indicated. NS signifies a nonspecific band that is routinely visible in CcmM western blots in our hands and corresponds to RbcL. This possibly results from M58 fragments comigrating with RbcL during electrophoresis. The proteins identified are indicated. NS signifies a nonspecific band that is routinely visible in CcmM western blots in our hands and corresponds to RbcL. This possibly results from M58 fragments comigrating with RbcL during electrophoresis. Relative CcaA amounts are expressed relative to the amount in wild-type cell lysates (set at 100%; SD (n = 3). NA. Not applicable; ND, not detected. A complete dataset is available in Supplemental Table S1.
Supplemental Fig. S2), suggesting that, if present, M23 is below detection limits in Synechococcus PCC7942 cells. Previous analysis by ourselves and others has not identified a peptide corresponding to the N-terminal domain of CcmM in β-carboxysomes (Long et al., 2005, 2007; Cot et al., 2008). Although N-terminal sequence analysis of M35 and the 36-kD homolog from Nostoc PCC7120 suggests that in both cases they might arise from a potential full-length CcmM cleavage at the carboxy side of an internal Val residue (Long et al., 2007), the data presented here clearly indicate a nonproteolytic origin for M35.

The ccmM gene of Synechococcus PCC7942, and related β-cyanobacteria, sits within a carboxysomal shell gene operon (ccmKLMNO; Omata et al., 2001) and codes for a 58-kD protein (Price et al., 1993). However, a short form of CcmM is routinely identified in greater quantities than the full-length form in both cell lysates and enriched carboxysome fractions from this and related cyanobacteria (Price et al., 1998; Long et al., 2005, 2007; Cot et al., 2008). A putative translation start site and IRES have been identified within the ccmM sequence (Ludwig et al., 2000; Long et al., 2007) that could account for independent M35 production. While IRESs are rare in cyanobacteria (the petH gene of Synechocystis PCC6803 the only known cyanobacterial example; Thomas et al., 2006), such complex translation processes have also been described in a number of organisms including E. coli (Sanatinia et al., 1995) and Salmonella typhimurium (Schoenhals et al., 1998).

In wild-type cells, M35 is produced in greater quantities compared with M58 (Fig. 2; Long et al., 2007), suggesting that the IRES and internal start codon of ccmM dominate translation of the ccmM mRNA on the ccmKLMNO transcript. A possible explanation for this is that IRES might result in both relatively high rates of M35 initiation as well as a greater rate of paused M58 translation (i.e. greater residence time of the ribosome at the IRES). Our analysis of local mRNA folding of

Figure 3. Ultrastructure of wild-type Synechococcus PCC7942 and ccmM mutant cells. Transmission electron micrographs showing the presence or absence of carboxysomes in wild-type and ccmM mutant cells of Synechococcus PCC7942. Cells were prepared as described in the text. In all cell types, where present, carboxysomes or related structures are indicated by arrowheads. Scale bars = 500 nm. A, Wild-type cells showing regular carboxysome structure. B, The carboxysomeless, ΔccmM mutant. C, ΔccmM + ccmM cells producing regular carboxysomes. D, ΔccmM + M35 cells contain electron-dense bodies that appear to be rectangular in cross section (possibly rhomboid). E, ΔccmM + ccmM-NL cells are characterized by dense regions within the cell that have planar surfaces, but regular carboxysomes are absent. F, ΔccmM + ccmM-gtc cells do not contain regular carboxysomes but produce numerous small electron-dense structures. G, ΔccmM + ccmM-FS cells produce electron-dense structures that are low in abundance and variable in size. H, ΔccmM + ccmM-SDΔccmM + ccmM-SD cells do not contain regular carboxysomes but produce numerous small dense structures.
hairpin loops at the transcription initiation regions of both M58 and M35 within the ccmM transcript give little indication as to which of these sites, if any, is likely to result in greater levels of translation (Supplemental Fig. S3; De Smit and Van Duin, 1994). It is worth noting that both out of context plasmid expression of ccmM (as in ΔccmM + ccmM) and expression from the wild-type operon lead to similar absolute quantities and ratios of M35 and M58 (Fig. 2). This confirms that competitive translation of the M58 or M35 regions of the ccmM message is independent of the ccm operon, but that processing signals are present within the ccmM gene.

Carboxysome Formation Minimally Requires Close to Equal Quantities of M35 and M58

The data presented here and in a previous study (Long et al., 2007) for ccmM mutants suggest that functional carboxysomes require M35 and M58 in at least an approximate 1:1 ratio to form functional carboxysomes. While M35 quantities are greater than M58 in wild-type cells (Fig. 2; Long et al., 2007), we have found that functional carboxysomes can form when M35:M58 stoichiometries are significantly perturbed by production of excess M58 (Long et al., 2007). However, the IRES-interrupting ΔccmM + NL, ΔccmM + gtc, and ΔccmM + SD mutants appear to contain far too little M35 to enable sufficient carboxysome formation. Notably, however, small electron-dense structures were found to form in ΔccmM + NL, ΔccmM + gtc, and ΔccmM + SD cells (Fig. 3), although these are apparently not fully functional carboxysomes (Table I). Relative quantitative analysis of M35 and M58 amounts in these mutants suggest approximately 3.5-fold and 20-fold decreases in M35 to M58 ratios in these cells (Table I).

In a previous study we found that the M35:M58 ratio of enriched carboxysome fractions from high-CO2-grown wild-type cells was approximately 4:1, while M35:M58 ratios in His6-tagged CcmM mutants, capable of producing functional carboxysomes, were in the range of approximately 1:1 to 11:1 (Long et al., 2007). Thus, approximately equal quantities, at least, of M35 and M58 are required for functional carboxysome formation. The ΔccmM + NL, ΔccmM + gtc, and ΔccmM + SD mutants in this study produce M35:M58 ratios that are substantially lower (0.1–0.6:1), yet M58 quantities are comparable with wild-type cells (Fig. 2). This suggests that M35 quantities are insufficient to allow functional carboxysome formation. Interestingly, through maintaining a HCR phenotype, the ΔccmM + SD mutant displayed intermediate C3 requirement for photosynthesis (Table I), had a M35:M58 ratio just below 1:1 (Fig. 2), and formed occasional electron-dense bodies (Fig. 3). In this case it is plausible that functional carboxysomes sometimes form due to near-adequate M35 to M58 ratios, but not sufficiently at a population level to overcome HCR for growth. Thus, appropriate absolute and relative quantities of M35 and M58 are essential for functional carboxysome formation. It follows, then, that fine control of the putative IRES and internal translation event within the ccmM transcript are important in the operation of the cyanobacterial CCM.

The absence of any M58 in the ΔccmM + M35 and ΔccmM + FS mutants (that did not alleviate the HCR phenotype; Table I) reveals an absolute requirement for the long form of CcmM in functional carboxysome formation. Notably, however, cells of these mutants contain large electron-dense structures (Fig. 3) that are for the most part similar to wild-type carboxysomes but relatively low in abundance (discussed below). While M35 provides an essential backbone for carboxysome formation, M58 must be present for CcaA recruitment and for organized shell facet formation (Long et al., 2007). In the ΔccmM + M35 mutant, not only is M58 absent but CcaA levels are essentially undetectable, as observed in ΔccmM cells (Fig. 2) and in the P-N carboxysomeless mutant (Price and Badger, 1991; Price et al., 1993). Thus, not only is carboxysome shell formation compromised in ΔccmM + M35, but CcaA is rapidly degraded, presumably to prevent extra-carboxysomal bicarbonate dehydration and CO2 loss. This indicates a specific proteolytic degradation of CcaA when it is not complexed with M58, a process that is likely to be relevant to the functioning of the cyanobacterial CCM but is outside the scope of this report.

M35 or M58 Alone Can Form Electron-Dense Subcellular Aggregates

In cell types producing M35 but no M58 (ΔccmM + M35 and ΔccmM + FS), relatively large electron-dense structures can be observed (Fig. 3) but cells still maintain a HCR phenotype (Table I). This suggests that M35 may form the structural basis for Rubisco holoenzyme interactions within the carboxysome and is not strictly limited to the shell layer as we have previously suggested (Long et al., 2007). Conversely, the small structures found in cell types producing predominantly M58 (ΔccmM + NL, ΔccmM + gtc, and ΔccmM + SD) could be carboxysome shell complexes with little internal Rubisco due to low M35 content. These hypotheses are supported by the regularly observed higher proportion of M35 over M58 (Fig. 2; Long et al., 2007) that would be required if M35 was distributed throughout the carboxysome and M58 confined to the shell layer. Indeed, analysis of absolute quantities of various carboxysomal component proteins from Synechococcus PCC7942 suggests that there is sufficient M35 to enable cross-linking of all Rubisco holoenzymes within the carboxysome and sufficient M58 to suggest that it may be confined to the shell layer (B.M. Long, unpublished data). Thus, specific roles for M35 as an internal carboxysomal Rubisco cross-linking protein and M58 as a shell-organizing scaffold protein can be hypothesized. This requires further investigation and may lead to im-

Copyright © 2010 American Society of Plant Biologists. All rights reserved.
portant advances in developing coordinated Rubisco complexes in C₃ plants (e.g. pseudocarboxysomes or pyrenoids).

IREs and Start Codons within ccmM Genes Are Conserved

Cot et al. (2008) found multiple CcmM polypeptides in Synechocystis PCC6803, each corresponding to putative IRES sequences and translation start codons, although the polyclonal antibody used was raised against the Synechococcus PCC7942 M58 protein (i.e. the same antibody used in this study). Analysis of ccmM from Synechococcus PCC7942 indicates additional putative IRES and translation initiation codons corresponding to doublet and singlet rbcS-like domains within the gene, corresponding to additional 23- and 12-kD polypeptides (Supplemental Figs. S3 and S4). However, unlike the findings of Cot et al. (2008), we have not yet found evidence for these smaller peptides being produced in Synechococcus PCC7942. Indeed, immunoblot analysis with polyclonal antibodies raised against both M35 and M58 polypeptides has provided no evidence for their existence thus far (B.M. Long, unpublished data). Nonetheless, the same putative GUG start codons and very similar SD sequences appear to be associated with these rbcS-like domains (Supplemental Figs. S3 and S4). The putative SD sequences at these sites are all commonly recognized sites in Synechocystis PCC6803 (Sazuka and Ohara, 1996). It is therefore possible that mRNA secondary or tertiary structure could play a role in preventing these sequences from being independently translated (De Smit and Van Duin, 1990; Boni, 2006), although the close proximity of the putative small SSU-3 SD sequence to the putative start may be less than optimal (Supplemental Fig. S4).

Our own attempts at determining secondary structure characteristics of the known and putative ccmM mRNA translation initiation regions provided no clear indication as to why one region may be favored over the others (Supplemental Fig. S3), suggesting that mRNA tertiary structure in particular may play a role. While it is plausible that the putative shorter CcmM forms (SSU-2 and SSU-3; Supplemental Fig. S4) arise from duplication and conservation of rbcS sequences, their independent translation would be redundant when triplet SSU-like repeats enable Rubisco cross-linking to occur. This is of particular relevance to the hypothesis that M35 plays a role in internalized carboxysomal Rubisco cross-linking, since three SSU-like domains would allow strong and uniform cross-linking both within and between Rubisco layers. Fewer domains would provide either weak cross-links (on only one plane) or none at all. Thus the three SSU-like domains in M35 might allow the formation of carboxysomal-like structures in the absence of M58 (such as those observed in the ΔccmM + M35 and ΔccmM + FS mutants), whereas fewer are less likely to. Why the putative IRESs and start codons of the addi-

tional SSU-like domains are therefore retained is unexpected, and the means by which one operates as an IRES but not the others is not clear.

Why Is CcmM a Two-Domain Protein?

It is well established that ccmM codes for two distinct protein domains (Price et al., 1993; Ludwig et al., 2000), and results presented here show that it can be translated into either a two-domain protein (M58) or a single-domain protein (M35) via an IRES. We have previously presented a model of how M35 and M58 may interact within the carboxysome shell to provide an essential structural framework: the N-terminal γ-CA domain of M58 forming an important structural trimer (Long et al., 2007). The N-terminal γ-CA domain of M58 is also required for binding the specific carboxysomal β-CA, CcaA (Long et al., 2007; Cot et al., 2008), and recruiting the enzyme into the carboxysome. In species that do not contain ccaA, it had been proposed that the N-terminal γ-CA domain of CcmM might act as the functional carboxysomal CA (Long et al., 2007; Cot et al., 2008; Cannon et al., 2010). A recent study of CcmM from T. elongatus BP-1, which does not have a ccaA gene, revealed that the N-terminal region of CcmM from this species is indeed an active trimeric CA under oxidizing conditions (Peña et al., 2010). The authors also showed that conserved regions of the T. elongatus polypeptide suggest this is the case in other ccaA-lacking β-cyanobacteria. Our model also suggests that M35 plays a structural role in binding Rubisco complexes to the shell layer via a cross-linking or Rubisco-organizing mechanism. This is based on our suggestion that multiple Rubisco SSU-like domains within a single M35 or M58 polypeptide enable cross-linking of Rubisco complexes (Ludwig et al., 2000; Long et al., 2007) and is supported by our experimental observation that M35 can bind to I₈ Rubisco cores in an E. coli expression system (Long et al., 2007).

The data presented here are consistent with M35 playing a specific role in Rubisco cross-linking within the carboxysome (perhaps within and between Rubisco layers). Thus the multiple SSU-like domains are required for Rubisco interaction and cross-linking. The fusion of the γ-CA domain and SSU-like domains therefore ensures the recruitment and close proximity of CcaA to Rubisco and carboxysome facet formation via M58. However, retention of the M35 gene start site allows independent SSU domain cross-linking with Rubisco where only structural interactions (and not CcaA recruitment) are required. Interestingly, while Synechococcus PCC7942 ccmM codes for a protein containing three Rubisco SSU-like domains, many β-cyanobacteria contain four or even five such repeats. While there is evidence for two CcmM forms in Nostoc PCC7120 (Long et al., 2007) and possibly up to four forms in Synechocystis PCC6803 (Cot et al., 2008), the number of CcmM short forms actually expressed in most β-cyanobacterial species is, as yet, unknown.
MATERIALS AND METHODS

SD Sequence, Start Codon, and Short Form Mutants of ccmM

Recombinant ccmM gene constructs (Fig. 1) were used to examine the putative IRES and internal start codon within the gene. Mutation of the putative internal start codon (GTC to Val216) to a nonstart GTC generates a silent, read-through mutation of the IRES is genuine. This mutation, ccmM + gtc, should result in one gene product (M58) if the IRES is genuine, or two gene products (M35 and M58) if a protease recognition site is required to produce both forms of CcmM. The ccmM + NL mutation is a GTC to CTG codon change at the putative start codon site. CTG is recognized as a rare start codon in cyanobacteria (Sazuka et al., 1999; Tichy and Vermaas, 1999) and was expected to result in a down-regulation of M35 production if CTG can operate as a start codon and if the IRES is genuine. The ccmM + NL mutation also contains an additional base change (ACC to AAC) that introduces T312Val to N312Ipeptide sequence change, aimed at disrupting any putative protease recognition site within M58 that may lead to M35 production. The ccmM + FS mutation introduces a FS and subsequent termination codon immediately prior to the putative M35 GTG start while maintaining the putative IRES (Fig. 1). This mutation was designed to allow M35 production in the absence of M58 if the IRES is genuine. It was also designed to produce the N-terminal yCA region of CcmM (M3, Fig. 1) independently of M35, thus enabling assessment of carboxysome formation without M58 but in the presence of both CcmM protein domains. Finally, a mutation in the putative SD sequence (GAGG to TCTG, ΔccmM + SD) was designed to test the hypothesis that ribosomal binding at this site is required for M35 translation. It was predicted that this mutation should lead to M35 production alone. All recombinant ccmM constructs were created via site-directed mutagenesis and expressed in the Synechococcus PCC7942 genome that GTG, CTG, and GTC are frequently used codons.

Cloning and Site-Directed Mutagenesis of ccmM

The full-length (M58) ccmM gene from Synechococcus PCC7942 was PCR adapted with primers: forward 5'-ttgcttctcaacaaccc (adding an NdeI site at the underlined codon) and reverse 5'-ctagtgcctagcggagttttggcaactaacgc (adding an in-frame Xhol site upstream of the underlined stop codon, and Xhol downstream). M35 was cloned with the forward primer 5'-aatctagaatgtgca (same reverse primer as used for M58 (above)) and the resulting M58 fragment was cloned into the NdeI and Xhol sites of a variant of the shuttle vector pSE4 (Maeda et al., 1998; Price et al., 2004) in which an NdeI site was removed from the plasmid backbone and one engineered into the multiple cloning site (pSE4-Nde-Del). The M35 fragment was cloned into the Ncol and Xhol sites of pSE4 and both sequences verified by DNA sequencing; these shuttle plasmids known as pSE4-ccmM and pSE4-M35. The shuttle vector possesses a nirA promoter from Synechococcus PCC7942 that is repressed in the presence of ammonium and derepressed in the presence of nitrate (Maeda et al., 1998; Price et al., 2004).

The NdeI/Xhol ccmM fragment (above) was cloned into pGEMT for site-directed mutagenesis using the QuikChange method (Stratagene) according to the manufacturer’s recommendations, as outlined (Fig. 1), and subsequently recloned into pSE4-Nde-Del. Primers for each of the ccmM mutations were as follows: GTC mutation, forward 5'-ctcaggagcagccagctgagttaac and reverse 5'-gttataagcgctcaatggctgctgctg; NL mutation, forward 5'-gttataagcgctcaatggctgctgctg and reverse 5'-ctcaggagcagccagctgagttaac and reverse 5'-gttataagcgctcaatggctgctgctg; FS mutation, forward 5'-gccaacagcagctgctgctgctg; SD mutation, forward 5'-ctcaggagcagccagctgagttaac and reverse 5'-ctcaggagcagccagctgagttaac.

Culture Conditions

Wild-type and mutant strains of Synechococcus PCC7942 were grown in modified BG-11 medium (Price and Badger, 1989a) containing 20 mM HEPES-KOH, pH 8.0 at 30°C and approximately 80 μmol photons m-2 s-1. Cultures were grown in 100-ml culture vessels and sparged with air enriched with 2% CO2 to circumvent the HCR phenotype. For mutants that harbor a spectroscopic mycin resistance marker antibiotics were added to liquid cultures at 7 μg/mL.

Carboxysome Isolation, Protein Electrophoresis, and Western Blotting

The presence of M35, M58, and other carboxysomal proteins was determined in cell extracts by western blotting. Cells from 3-d-old 50-mL cultures were collected by centrifugation (5,000g, 5 min) and lysed by lysozyme and French Press treatment as described (Long et al., 2005). Centrifugally clarified cell lysates were treated with 25 mM MgSO4 to generate carboxysome-enriched pellet fractions (Price and Badger, 1991; Price et al., 1998). Prior to SDS-PAGE, whole-cell lysate samples containing 500 ng chlorophyll (Chl) and equivalent volumes of resuspended Mg2+ pellets and Mg2+ supernatants were diluted in 4× SDS-PAGE sample buffer (Invitrogen), containing 4 μL urea, and boiled for 10 min. The use of 4.0 μL urea greatly enhanced the formation of monomeric CcmK1 that is otherwise observed predominantly as multimeric forms on SDS-PAGE in our hands. Proteins were separated on 4% to 12% NuPAGE Bis-Tris gels (Invitrogen) according to the manufacturer’s instructions, using the MES-based buffer system. In our hands, electrophoresis of monomeric CcaA on SDS-PAGE was also enhanced by addition of sodium bisulfite to a final concentration of 1 mM in the cathode buffer. Separated proteins were transferred to Immobilon-P polyvinylidene difluoride (PVDF) membrane (Whatman) using a semidyblot apparatus (1 mL/cm2, 60 min) and probed with polyclonal rabbit antiserum raised against Synechococcus PCC7942 Rubisco, M35, CcmK1 (the homolog of CcmK2 from Synechocystis PCC6803), and CcaA. In addition, blots were also probed using a polyclonal antibody raised against M58 in an attempt to identify cleavage products of the parent protein, other than M35, in the event that M35 may arise through proteolytic cleavage. Bound antibodies were detected using the AP-conjugate substrate kit (Bio-Rad) after secondary probing with a sheep-anti-rabbit alkaline phosphatase conjugated antibody (ICN).

Relative M35, M58, and CcaA quantities in mutant and wild-type cells were determined from western blots using the Attophos (Promega) detection system (ICN). Briefly, a dilution series of wild-type cell lysates (0.05–0.3 μg Chl on gel) and 0.15 μg Chl of each mutant was separated on SDS-PAGE gels. Proteins were blotted onto PVDF membranes as above and probed with polyclonal rabbit serum raised specifically against M35 (thus the quantitation of M35 and M58 were on an equivalent basis) or polyclonal rabbit serum raised against CcaA. Bound antibodies were detected with sheep-anti-rabbit alkaline phosphatase conjugated antibodies. Blots were visualized using a Bio-Rad versus doc system and relative quantities determined using QuantityOne software (Bio-Rad). For CcmM polypeptides, data were expressed relative to M58 found in wild-type extracts (set at 100%). Quantities of CcaA were expressed relative to that found in wild-type extracts (set at 100%).

Mass Spectrometric Measurements

Cells were prepared and analyzed in a membrane inlet mass spectrometer as previously described (Price and Badger, 1989b; Sulheimer et al., 1995). Cells were assayed at a Chl density of 2 μg mL-1 in BG-11 medium buffered with 50 mM BisTrisPropane-HCl (pH 7.9), containing 20 mM NaCl instead of NaNO3. Assays were performed in a thermostated (30°C) mass spectrometer cuvette allowing membrane inlet analysis of O2 (mass 32) and CO2 (mass 44). The irradiance used was 700 μmol photons m-2 s-1.

Transmission Electron Microscopy

Cells were prepared for electron microscopy essentially as described by Price and Badger (1989b). Stained sections were viewed using an Hitachi H-7100A TEM at 75 kV.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession number M96929.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. CcM-dependent photosynthesis in wild-type Synechococcus PCC7942 and ccmM mutants.

Supplemental Figure S2. Detection of M35 and M23 peptide products of ccmM-FS expressed in Escherichia coli.
**LITERATURE CITED**

Badger MR, Hanson D, Price GD (2002) Evolution and diversity of CO₂ concentrating mechanisms in cyanobacteria. Funct Plant Biol 29: 161–173

Badger MR, Price GD (2003) CO₂ concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot 54: 609–622

Badger MR, Price GD, Long BM, Woodger FJ (2006) The environmental plasticity and ecological genomes of the cyanobacterial CO₂ concentrating mechanism. J Exp Bot 57: 240–265

Boni IV (2006) Diverse molecular mechanisms for translation initiation in prokaryotes. Mol Biol 40: 658–668

Cannon GC, Heinitorh S, Bradburne CE, Shively JM (2002) Carboxysome genomics: a status report. Funct Plant Biol 29: 175–182

Cannon GC, Heinitorh S, Kerfeld CA (2010) Carboxysomal carbonic anhydrases: structure and role in microbial CO₂ fixation. Biochim Biophys Acta 1804: 382–392

Cot SSW, So AKC, Espie GS (2008) A multiprotein bicarbonate dehydration complex essential to carboxysome function in cyanobacteria. J Bacteriol 190: 936–945

De Smit MH, Van Duin J (1990) Secondary structure of the ribosome binding-site determines translational efficiency—a quantitative analysis. Proc Natl Acad Sci USA 87: 7608–7672

De Smit MH, Van Duin J (1994) Control of translation by mRNA secondary structure in Escherichia coli—a quantitative analysis of literature data. J Mol Biol 244: 144–150

Ellis RJ (1979) The most abundant protein in the world. Trends Biochem Sci 4: 241–244

Emlyn-Jones D, Woodger FJ, Andrews TJ, Price GD, Whitney SM (2006) A Synechococcus PCC7942 delta cm (Cyanophyceae) mutant pseudo-reverts to air growth without regaining carboxysomes. J Phycol 42: 769–777

Field CB, Behenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281: 237–240

Kerfeld CA, Sawaya MR, Tanaka S, Nguyen CV, Phillips M, Beemy B, Yeates TO (2005) Protein structures forming the shell of primitive bacterial organelles. Science 309: 936–938

Long BM, Badger MR, Whitney SM, Price GD (2007) Analysis of carboxysomes from Synechococcus PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J Biol Chem 282: 29323–29335

Long BM, Price GD, Badger MR (2005) Proteomic assessment of an established technique for carboxysome enrichment from Synechococcus PCC7942. Can J Bot 83: 746–757

Ludwig M, Sülttemeyer D, Price GD (2000) Isolation of ccmKLMN genes from the marine cyanobacterium, Synechococcus sp PCC7002 (Cyanobacteria), and evidence that CcmM is essential for carboxysome assembly. J Phycol 36: 1109–1118

Maeda S, Kawaguchi Y, Ohe T, Omata T (1998) cis-acting sequences required for ntcB-dependent, nitrate-responsive positive regulation of the nitrilase activation operon in the cyanobacterium Synechocystis sp. strain PCC7942. J Bacteriol 180: 4080–4088

McKay RML, Gibbs SP, Espie GS (1993) Effect of dissolved inorganic carbon on the expression of carboxysomes, localization of Rubisco and the mode of inorganic carbon transport in cells of the cyanobacterium Synechococcus UTEX 625. Arch Microbiol 159: 21–29

Missiakas D, Georgopoulous C, Raina S (1993) The Escherichia coli heat shock gene htpY: mutational analysis, cloning, sequencing, and transcriptional regulation. J Bacteriol 175: 2613–2624

Omata T, Gohta S, Takahashi Y, Harano Y, Maeda S (2001) Involvement of a CbbR homolog in low CO₂-induced activation of the bicarbonate transporter operon in cyanobacteria. J Bacteriol 183: 1891–1898

Peña KL, Castel SE, de Araujo C, Espie GS, Kimber MS (2010) Structural basis of the oxidative activation of the carboxysomal γ-carboxylyhyd- rase, CcmM. Proc Natl Acad Sci USA 107: 2455–2460

Price GD, Badger MR (1989a) Ethoxyzolamide inhibition of CO₂-dependent photosynthesis in the cyanobacterium Synechococcus PCC7942. Plant Physiol 89: 44–50

Price GD, Badger MR (1989b) Isolation and characterization of high CO₂-requiring-mutants of the cyanobacterium Synechococcus PCC7942: two phenotypes that accumulate inorganic carbon but are apparently unable to generate CO₂ within the carboxysome. Plant Physiol 91: 514–525

Price GD, Badger MR (1991) Evidence for the role of carboxysomes in the cyanobacterial CO₂-concentrating mechanism. Can J Bot 69: 963–973

Price GD, Badger MR, Woodger FJ, Long BM (2008) Advances in understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): functional components, θt transporters, diversity, genetic regulation and prospects for engineering into plants. J Exp Bot 59: 1441–1461

Price GD, Coleman JR, Badger MR (1992) Association of carboxylase anhydrase with carboxysomes isolated from the cyanobacterium Synechococcus PCC7942. Plant Physiol 100: 784–793

Price GD, Howitt SM, Harrison K, Badger MR (1993) Analysis of a genomic DNA region from the cyanobacterium Synechococcus sp. strain PCC7942 involved in carboxysome assembly and function. J Bacteriol 175: 2871–2879

Price GD, Sültemeyer D, Klughammer B, Ludwig M, Badger MR (1998) The lncr-ontrolling the CO₂ concentrating mechanism in several cyanobacterial strains—a review of general physiological characteristics, genes, proteins, and recent advances. Can J Bot 76: 973–1002

Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L (2004) Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proc Natl Acad Sci USA 101: 18228–18233

Satanihi H, Kofoid E, Morrison T, Parkinson J (1995) The smaller of two overlapping chaA gene products is not essential for chemotaxis in Escherichia coli. J Bacteriol 177: 2713–2720

Sakuta T, Ohara O (1996) Sequence features surrounding the translation initiation sites assigned on the genome sequence of Synechocystis sp. strain PCC6803 by amino-terminal protein sequencing. DNA Res 3: 225–232

Sato T, Yamasuchi M, Ohara O (1999) Cyanoo2Database updated: linkage of 234 protein spots to corresponding genes through N-terminal microsequencing. Electrophoresis 20: 2160–2171

Schoenhals GJ, Kihara M, MacNab RM (1998) Translation of the flagellar gene fiio of Salmonella typhimurium from putative tandem starts. J Bacteriol 180: 2936–2942

Sültemeyer D, Amoroso G, Fock H (1995) Induction of intracellular carbonic anhydrases during the adaptation to low inorganic carbon concentrations in wild-type and ca-1 mutant cells of Chlamydomonas reinhardii. Planta 196: 217–224

Tanaka S, Sawaya MR, Phillips M, Yeates TO (2009) Insights from multiple structures of the shell proteins from the β-carboxysome. Protein Sci 18: 108–120

Thomas JC, Ughy B, Lagoutte B, Ajlani G (2006) A second isoform of the ferredoxin: NADP oxidoreductase generated by an in-frame initiation of transcription. Proc Natl Acad Sci USA 103: 18368–18373

Tichy M, Vermaas W (1999) Accumulation of pre-apocytochrome f in a Synechocystis sp. PCC 6803 mutant impaired in cytochrome c maturation. J Biol Chem 274: 32396–32401

Woodger FJ, Badger MR, Price GD (2005) Sensing of inorganic carbon limitation in Synechococcus PCC7942 is correlated with the size of the internal inorganic carbon pool and involves oxygen. Plant Physiol 139: 1959–1969

Yeates TO, Kerfeld CA, Heinitorh S, Cannon GC, Shively JM (2008) Protein-based organelles in bacteria: carboxysomes and related micro-compartments. Nat Rev Microbiol 6: 681–691