CO₂ concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution

Murray R. Badger¹ and G. Dean Price

Molecular Plant Physiology Group, Research School of Biological Sciences, Australian National University, PO Box 475, Canberra City, ACT 2601, Australia

Received 11 April 2002; Accepted 10 October 2002

Abstract
Cyanobacteria have evolved an extremely effective single-cell CO₂ concentrating mechanism (CCM). Recent molecular, biochemical and physiological studies have significantly extended current knowledge about the genes and protein components of this system and how they operate to elevate CO₂ around Rubisco during photosynthesis. The CCM components include at least four modes of active inorganic carbon uptake, including two bicarbonate transporters and two CO₂ uptake systems associated with the operation of specialized NDH-1 complexes. All these uptake systems serve to accumulate HCO₃⁻ in the cytosol of the cell, which is subsequently used by the Rubisco-containing carboxysome protein micro-compartment within the cell to elevate CO₂ around Rubisco. A specialized carbonic anhydrase is also generally present in this compartment. The recent availability of at least nine cyanobacterial genomes has made it possible to begin to undertake comparative genomics of the CCM in cyanobacteria. Analyses have revealed a number of surprising findings. Firstly, cyanobacteria have evolved two types of carboxysomes, correlated with the form of Rubisco present (Form 1A and 1B). Secondly, the two HCO₃⁻ and CO₂ transport systems are distributed variably, with some cyanobacteria (Prochlorococcus marinus species) appearing to lack CO₂ uptake systems entirely. Finally, there are multiple carbonic anhydrases in many cyanobacteria, but, surprisingly, several cyanobacterial genomes appear to lack any identifiable CA genes. A pathway for the evolution of CCM components is suggested.

Key words: Carboxysomes, CO₂ concentrating mechanisms, cyanobacteria, evolution, genes, photosynthesis, transporters.

Introduction
Cyanobacteria have existed as oxygenic photosynthetic bacteria on earth for at least 2.7 billion years (Buick, 1992). During that time they have endured a changing gaseous environment where CO₂ has declined and O₂ has risen. This has imposed evolutionary pressure on them to evolve strategies for efficiently acquiring inorganic carbon for photosynthesis. In response to this they have developed an effective photosynthetic CO₂ concentrating mechanism (CCM) for improving the carboxylation by their relatively inefficient Rubiscos (Badger and Price, 1992; Price et al., 1998; Kaplan and Reinhold, 1999). This CCM is perhaps the most effective of any photosynthetic organism, concentrating CO₂ up to 1000-fold around the active site of Rubisco. In the past few years, there has been a rapid increase in the understanding of the mechanisms and genes involved in cyanobacterial CCMs. In addition, there has been a recent expansion in the availability of complete cyanobacterial genome sequences, thus increasing the potential to examine questions regarding both the evolution and diversity of components of the CCM across cyanobacterial species. This paper reviews current understanding of the mechanisms and genes underlying the operation of the cyanobacterial CCM, and takes the opportunity to employ comparative genomics to shed light on the evolution and diversity of the CCM among cyanobacterial species.

¹ To whom correspondence should be addressed. E-mail: murray.badger@anu.edu.au

Abbreviations: ABC, ATP binding cassette; CA, carbonic anhydrase; CCM, CO₂ concentrating mechanism; Ci, inorganic carbon; EZ, ethoxyzolamide; Fd, ferredoxin; NDH-1, NAD(P)H dehydrogenase; PQ, plastoquinone; Rubisco, ribulose 1–5–bisphosphate carboxylase-oxygenase EC. 4.1.1.39.
The cyanobacterial CCM model

A schematic diagram of a cyanobacterial cell with components of the CCM (see recent reviews by Price et al., 1998; Kaplan and Reinhold, 1999) is shown in Fig. 1. The basic concepts of this model are based on experiments with the model cyanobacterial species *Synechococcus* PCC7942, *Synechocystis* PCC6803 and *Synechococcus* PCC7002. Central to the functioning of the cyanobacterial CCM is the carboxysome, a protein micro-compartment within the cell that contains the Rubisco of the cell together with a carboxysomal carbonic anhydrase (CA). The CA functions to convert an accumulated cytosolic pool of HCO$_3^-$ into CO$_2$ within the carboxysome. The generation of CO$_2$ coupled with a diffusive restriction to the efflux from the carboxysome, possibly imposed by the protein shell, leads to the localized elevation of CO$_2$ around the active site of Rubisco within the carboxysome. The substrate for the carboxysome, HCO$_3^-$, is accumulated in the cytosol by the operation of a number of active CO$_2$ and HCO$_3^-$ transporters. These transporters are located on both the plasma membrane and the thylakoid membrane, and exist in both low affinity and high affinity transporter forms. Many cyanobacteria have the ability to improve their affinity for inorganic carbon (Ci) when grown at limiting Ci levels, and this acclimation is due primarily to the changes that occur in the synthesis and properties of various high and low-affinity Ci transporters associated with the cells.

The phylogeny of cyanobacteria

Attempts to divide cyanobacteria into phylogenetic groupings have used both photosynthetic pigment composition and ribosomal RNA subunit sequences. The prochlorophyte grouping was suggested as useful in grouping those cyanobacteria with and without light-harvesting phycobilisomes, however it has become obvious that there have been multiple evolutionary origins of so-called prochlorophytes within widely different cyanobacterial taxonomic groups (Urbach et al., 1992). Ribosomal RNA gene sequence analysis has become a more valid basis for establishing clearer evolutionary relationships (Urbach et al., 1998; Honda et al., 1999). Figure 2A shows one such analysis for 16S ribosomal RNA genes from a range of photosynthetic organisms. Cyanobacteria are clearly grouped within one radiation, with divisions within this primary radiation also being apparent. These analyses show them to be quite separate from β-proteobacteria, non-green algae, green algae, and higher plants. However, recent analysis of genomic sequences that have appeared for a number of cyanobacteria suggest that a division of cyanobacterial species based on their type of Rubisco may be much more appropriate (Badger et al., 2002) particularly with respect to the evolution of their CCMs.

A phylogenetic tree for Rubisco from various photosynthetic bacteria is also shown in Fig. 2B. The photosynthetic organisms are grouped according to their Form 1 (L$_8$S$_8$) Rubisco large subunit types, originally described by Delwiche and colleagues (Delwiche, 1999), with Form 1A, B, C, and D groups shown. Cyanobacterial species are contained within both the Form 1A and 1B domains, as initially noted by Tabita and colleagues (Tabita, 1999). This variation of cyanobacteria is associated with the divergence between β-proteobacterial-like cyanobacteria such as *Prochlorococcus marinus* species and *Synechococcus* WH8102 and cyanobacteria such as *Synechocystis* PCC6803. This has led to a recent suggestion that two primary groupings of cyanobacteria can be established based on their Rubisco phylogeny (Badger et al., 2002). These two grouping are referred to as α-cyanobacteria for those containing Form 1A Rubisco and β-cyanobacteria with Form 1B Rubisco.

The presence or absence of carboxysomes in various photosynthetic and chemoheterotrophic bacterial species is also indicated in Fig. 2B. Carboxysomes are protein bodies which are surrounded by a protein shell and contain the Rubisco of the cell. Hence there is potential significance between the nature of the carboxysome structure and the type of Rubisco that it contains. Sequencing of the
α-cyanobacterial genomes has revealed that α-cyanobacteria possess carboxysomes that are significantly different from the carboxysomes found in β-cyanobacteria. Based on these observations it is suggested that these carboxysomes should be termed either α-carboxysomes for those found in Form 1A Rubisco-containing photosynthetic bacteria, including α-cyanobacteria and proteobacteria such as *Thiobacillus* species (Shively et al., 1998a, b) while β-carboxysomes are those associated with Form 1B Rubisco in the β-cyanobacteria (Price et al., 1998). Little is known about the different physiological properties of α and β-carboxysomes or the potential ecological advantages that the possession of carboxysomes might confer. However, all cyanobacteria characterized to date have carboxysomes.

**Carboxysome structure and phylogeny**

Ideas about the structure and function of carboxysomes have been based on studies with experimental systems for both α and β-carboxysomes, although the α-carboxysomes studied have been those associated with proteobacteria such as *Thiobacillus* species rather than α-cyanobacteria (Cannon et al., 2002). However, the majority of studies on carboxysome function have been centred on β-carboxysomes in model laboratory species of β-cyanobacteria that are commonly used.

The carboxysome is a protein microbody, consisting of a protein coat, analogous to that surrounding a virus, and an interior soluble protein phase containing most, if not all, of the cellular Rubisco in cyanobacteria (Beudeker et al., 1980; Price et al., 1992; McKay et al., 1993). The Rubisco within these bodies appears to be packed into para-crystalline arrays (Shively et al., 1973; Holthuijzen et al., 1986). The composition of the protein shell has been most extensively characterized in α-carboxysomes from chemolithotrophic proteobacteria such as *Thiobacillus* (Cannon et al., 2001). Studies with these carboxysomes have shown the shell to consist of at least four different types of polypeptides. A number of small polypeptides (8–12 kDa), all related to each other, have been identified and named CsoS1, peptide A and peptide B. Frequently there are several members of each polypeptide type. Homologues of these peptides have been identified in β-cyanobacterial genomes and include CcmK, L, and O. Together, these proteins are related to each other by coding for proteins containing one or more regions of homology to bacterial micro-compartment domains (pfam 00936). These conserved structural domains have been identified by comparison of CcmK, CcmL, CcmO, or CsoS1-like proteins involved in carboxysome formation in α and β-cyanobacteria and β-proteobacteria (Price et al., 1998; Shively et al., 1998a, b; Cannon et al., 2001) as well as more recently discovered genes associated with enteric proteobacteria containing carboxysome-like micro-compartments specialized in both propanediol and ethanolamine metabolism and detoxification (Bobick et al., 1999; Kofoid et al., 1999). Figure 3 shows a phylogenetic tree highlighting the relationships of the different polypeptides. CcmK, CcmO and CsoS1 proteins form separate groups while CcmL, peptideA and peptideB form another. Figure 3 clearly indicates the differences in the polypeptide composition between α and β-cyanobacteria. CcmK, L and O genes are contained in β-cyanobacterial genomes while CsoS1 and peptide A and B are found in α-cyanobacteria.

Carboxysome shells also appear to contain two other larger polypeptides that bear no homology to each other. In α-carboxysomes there are the CsoS2 (80–90 kDa) and CsoS3 (55–65 kDa), while in β-carboxysomes these appear to be replaced by CcmM (55–70 kDa) and CcmN (26 kDa) (Price et al., 1998; Cannon et al., 2001). In general, these proteins have no functional homologies in other bacterial systems, except that the CcmM protein has domains within it that are homologous to both γ-CA and to the small subunit of the Form 1B Rubisco protein (Price et al., 1998; Ludwig et al., 2000). The functional significance of these homologies is unknown.

For CO2 fixation to occur in β-carboxysomes, it is envisaged that HCO3− diffuses through the proteinaceous shell of the carboxysome where a low activity of carbonic anhydrase inside the structure acts to catalyse the formation of CO2 from HCO3− at rates high enough to saturate the carboxylation reaction of Rubisco. The role of the carboxysome as a site for CO2 elevation was proposed by Reinhold et al. (1987, 1991). The models suggest that HCO3− diffuses into the carboxysome interior, and that CO2 is generated by a specialized CA. CO2 can be elevated within the carboxysome with the aid of some poorly understood diffusion barrier, such as the protein shell, that restricts CO2 diffusion out of the carboxysome. Further information on the role of carboxysomes in the cyanobacterial CCM can be found in recent reviews (Price et al., 1998; Kaplan and Reinhold, 1999; Cannon et al., 2001).

Recent comparative genome analyses among different cyanobacteria (Cannon et al., 2001, 2002; Hess et al., 2001; Badger et al., 2002) and the analysis in Fig. 3 have highlighted for the first time that α-cyanobacteria have α-carboxysomes rather than the β-carboxysomes previously studied in β-cyanobacteria. This suggests that these two carboxysome types have either been inherited or evolved in parallel, in association with the Form 1A and Form 1B Rubiscos found in each cyanobacterial group. Badger et al. (2002) have suggested the possibility that both Rubisco and carboxysome genes could be inherited by a single lateral gene transfer event as in many cyanobacteria they are often found organized on contiguous regions of the chromosome (Badger et al., 2002).

The presence of a specific carboxysomal CA enzyme is interesting in that such a CA gene product has only been
identified in β-carboxysomes from a number of β-cyanobacteria (Fukuzawa et al., 1992; Price et al., 1992; Yu et al., 1992). The emergence of complete genome sequences for a number of α and β-cyanobacteria has begun to confuse this picture. None of the α-cyanobacteria sequenced so far has a recognizable carboxysomal β-CA homologue, although one beta-CA (note that nomenclature is not related to α and β terminology used for carboxysomes and cyanobacteria) is present in the Synechococcus WH8102 genome (Badger et al., 2002). In addition, the recently published genomes for the β-cyanobacteria Trichodesmium erythraeum (http://genome.ornl.gov/microbial/tery/) and Thermosynechococcus elongatus (http://www.kazusa.or.jp/cyanobase/Thermo/index.html) have also indicated a lack of any recognizable CA genes. The absence of any clearly identifiable CA genes in either α or
β-cyanobacteria is intriguing and points to the potential for a different mode of carbonosome function in these cyanobacteria.

**Carbonic anhydrases**

An analysis of possible alpha, beta and gamma carbonic anhydrases (Smith and Ferry, 2000) in the cyanobacterial genomes shows that there is a wide diversity in carbonic anhydrase gene content (Badger et al., 1998). Being an ionic form of Ci, HCO$_3^-$ activity in the general cytosol (Volokita et al., 1994; Price and Badger, 1989). Regardless of which form of Ci is presented to the cell, CO$_2$ or HCO$_3^-$, the available evidence indicates that HCO$_3^-$ is the species accumulated within the bulk cytoplasm. Furthermore, HCO$_3^-$ and CO$_2$ are only slowly interconverted through the apparent absence of carbonic anhydrase activity in the general cytosol (Volokita et al., 1984; Price and Badger, 1989). Being an ionic form of Ci, HCO$_3^-$ is much less permeable to lipid membranes than the uncharged CO$_2$ molecule and only slowly leaks from the cell, unless leakage is facilitated through a channel. The best support for the view that HCO$_3^-$ is the accumulated species is that ectopic expression of human carbonic anhydrase within the cytoplasm of *Synechococcus* PCC7942 cells leads to a debilitating leakage of CO$_2$ from the cells due to rapid equilibration between HCO$_3^-$ and CO$_2$ in the cytosol (Price and Badger, 1989).

Current evidence indicates that there are 4–5 modes of Ci uptake identified from research with *Synechococcus* PCC7942 and *Synechocystis* PCC6803 (Price et al., 2002; Shibata et al., 2002a). The four modes so far identified are discussed in more detail in Figs 4 and 5, but briefly they are: (1) BCT1, an inducible high affinity HCO$_3^-$ transporter encoded by *cmpABCD* and belonging to the bacterial ATP binding cassette (ABC) transporter family (Omata et al., 1999); (2) an inducible medium affinity Na$^+$-dependent HCO$_3^-$ transport system (Shibata et al., 2002b); (3) a constitutive system for CO$_2$ uptake associated with a specialized form of a thylakoid-located NDH-1 complex, referred to as NDH-I$_4$ (Shibata et al., 2001; Maeda et al., 2002); and (4) a second specialized NDH-1 complex, named NDH-I$_3$ (Shibata et al., 2001; Maeda et al., 2002) that is inducible at limiting Ci conditions and exhibits a higher uptake affinity for CO$_2$.

**The BCT1 HCO$_3^-$ transporter**

The high-affinity HCO$_3^-$ transporter, BCT1, was the first cyanobacterial Ci transporter to have been conclusively identified and characterized. In *Synechococcus* PCC7942, BCT1 is encoded by the *cmpABCD* operon and is expressed under severe Ci limitation (Omata et al., 2002).
When the cmp operon was expressed under the control of a nitrate responsive promoter (nirA) the induced expression of the operon caused the appearance of high-affinity HCO$_3^-$ uptake under high-CO$_2$ growth conditions (Omata et al., 1999). BCT1, and its cyanobacterial orthologues, appear to be the first example of a primary transporter (uniporter) for HCO$_3^-$, but the transporter is clearly a member of the diverse subfamily of bacterial ABC transporters (Higgins, 2001). The cmpA gene codes for the precursor of the 42 kDa protein (CmpA) which has long been known to be induced under Ci limitation (Omata and Ogawa, 1986) and also under high-light stress (Reddy et al., 1989). The protein sequence of CmpA is closely related to NrtA (Omata, 1991) which in turn has been confirmed as the nitrate/nitrite binding lipoprotein for the nitrate/nitrite transporter (nrtABCD) in *Synechococcus* sp. PCC7942 (Maeda and Omata, 1997). Recently, it has been shown that CmpA is indeed a binding protein that is specific for HCO$_3^-$ with an affinity or $K_D$ of around 5 $\mu$M (Maeda et al., 2000).

Figure 4 shows a model for the operation of the BCT1 transporter. The cmpABCD genes code for four proteins:
the HCO$_3^-$ binding protein (CmpA; 42 kDa); the intrinsic membrane protein (CmpB); a large extrinsic membrane protein with an consensus ATP binding site (CmpC), and a smaller related protein, CmpD, also with an ATP site. CmpB most probably forms a dimer within the membrane by analogy to other ABC transporters. The stoichiometry of CmpA to the other components of the transporter is apparently much greater than the intrinsic proteins of the transporter (Omata and Ogawa, 1986), but this is often a feature of binding proteins (Higgins, 2001). CmpA is a lipoprotein and it appears that CmpA proteins act as an array of substrate collectors for the transporter and are able to diffuse in two dimensions with the N-terminus of the mature polypeptide attached to the plasma membrane.

Comparative genomic analysis shows that the cmpABCD operon is found in freshwater β-cyanobacteria (see Fig. 6), but is absent from the marine Synechococcus PCC7002 (D Bryant, personal communication) and Trichodesmium erythraeum (see http://genome.ornl.gov/microbial/tery/) and is also absent from all three marine α-cyanobacteria species sequenced to date (Badger et al., 2002).

### Sodium-dependent HCO$_3^-$ uptake

Cells of both Synechococcus PCC7942 and Synechocystis PCC6803 possess a Na$^+$-dependent HCO$_3^-$ uptake activity quite separate from BCT1 transport (Price et al., 2002; Shibata et al., 2002a). It has been suggested from physiological studies that cyanobacteria may possess a Na$^+$/HCO$_3^-$ symporter that is energized by a standing, inwardly directed Na$^+$ gradient energized by a Na$^+$/H$^+$ antipporter (Espie and Kandasamy, 1994). Recently, Shibata et al. (2002a, b) have isolated a gene from Synechocystis PCC6803 that appears to code for Na$^+$-dependent transport activity induced under Ci-limitation. The gene has been termed sbtA (SLR1512), and a mutant with lesions in sbtA/cmpB/ndhD3/ndhD4 genes was unable to grow under low CO$_2$ at pH 7 or 9. This transporter is also depicted in Fig. 4. Analyses suggest that the predicted gene product of sbtA codes for a protein of 374 amino acids, and has potential for 8–10 membrane spanning domains. It also has a 60 amino acid hydrophilic domain near the centre of the protein that may represent a membrane extrinsic region, possibly facing the cytoplasm.

---

**Fig. 5.** A speculative model for the functioning of a specialized NDH-1$_{3/4}$-type complex to hydrate CO$_2$ to HCO$_3^-$ The hydration is coupled to electron flow such as that supported by PSII cyclic electron transport. The model is based on the proposal of Price et al. (2002). It is postulated that ChpX and Y subunits are CO$_2$ hydration proteins bound to the cytoplasmic face of NDH-1$_{3/4}$ complexes. The Ndhd4/F3 and NdhdD4/D3 subunits form part of the proton translocation channel. This specialized NDH-1 complex is proposed to drive the net hydration of CO$_2$ to HCO$_3^-$ when coupled to photosynthetic electron transport. Here both NADPH and Fd$_{red}$ are depicted as being potential electron donors to the NDH-1$_{3/4}$ complexes. The energetics of CO$_2$ conversion may be further improved through the operation of a Q cycle similar to that operating in the b/f complex. Equations 1–3 describe the reaction sequence proposed to occur in conventional CA enzymes (E) (Silverman, 2000; Smith et al., 2000); the steps are (1) generation of a reactive hydroxyl group combined with spontaneous conversion of CO$_2$ to HCO$_3^-$, (2) binding a water molecule to Zn and (3) abstraction of a proton to a nearby His residue and then conduction along a proton wire to a buffer molecule (B) in the bulk medium. The essential part of this proposal is that the last step (3a) be coupled to electron-driven translocation of protons to the thylakoid lumen, making hydroxyl-mediated net hydration of CO$_2$ largely irreversible in the light.
Fig. 6. Diversity in components of the cyanobacterial CCM. Depicted are four types of cyanobacterial cells, currently identifiable from genome sequencing projects. The groups are differentiated by the presence of either α or β-carboxysomes, and the complement of Ci transporters present. The two β-cyanobacterial groups are freshwater and marine types. For the marine α-cyanobacteria, there are two groups typified by either *Synechococcus* WH8102 or *Prochlorococcus* species. The species used to derive the groups are listed on the figure.
Strong homologues of the sbtA gene are present in many β-cyanobacterial genomes, however, only weak homologues appear to present in the marine α-cyanobacteria (Badger et al., 2002). Surprisingly, there are no homologues in the β-cyanobacterium Thermosynechococcus elongatus (http://www.kazusa.or.jp/cyanobase/Thermol) and Trichodesmium erythraeum (http://genomeornl.gov/microbial/teryl/).

**NDH-1 genes and CO₂ uptake**

Previous studies have shown that the NDH-1 dehydrogenase complex is involved in enabling CO₂ uptake by cyanobacteria (Ogawa, 1992; Price et al., 1998; Klughammer et al., 1999; Ohkawa et al., 2000a, b). However, within β-cyanobacterial species there may be a number of distinct types of NDH-1 complexes with different roles within the cell. The complete genome database for Synechocystis PCC6803 revealed that there are a number of NDH-1 genes in Synechocystis that are present as single copies in the genome. These are: ndhA, ndhB, ndhC, ndhD3, ndhD4, and ndhF (3 homologues). Analysis of homology relationships indicated a large diversity in NdHD and NdHF proteins, with up to three groupings of NdHD being identifiable, as well as two groups of NdHF (Price et al., 1998).

The NdHD/D polypeptides together with NdHF1 probably are involved in forming a conventional respiratory NDH-1 complex, oxidizing NADPH and NADH and reducing plastoquinone, thus enabling cyclic electron transport around PSI (Ohkawa et al., 2000a, b). The NdHD3/D4, together with NdHD3 and F4 components are suggested as components of two forms of a specialized NDH-1 complex involved in catalysing active CO₂ uptake by converting CO₂ to HCO₃⁻ within the cell (Ohkawa et al., 2000a, b; Shibata et al., 2001; Maeda et al., 2002). The exact role of NdHD5/D6 polypeptides is unclear. In addition to this, two other genes/proteins are involved in enabling the CO₂ uptake activity of the NDH-1 complex, and these are referred to here as chpX and chpY (note that Shibata, Ogawa and colleagues have named these genes as cupA and cupB, while Price and co-workers have used chpX and chpY nomenclature; Shibata et al., 2001, 2002a; Maeda et al., 2002; Price et al., 2002). Recent evidence from both Ogawa and Price’s laboratories (Shibata et al., 2001; Maeda et al., 2002) have clearly indicated that the ndhF3/ndhD3/chpY genes code for polypeptides that are part of a high affinity CO₂ uptake NDH-1 complex, while the ndhF4/ndhD4/chpX genes code for a NDH-1 complex involved in low affinity CO₂ uptake.

Price et al. (2002) have speculated that the ChpX and ChpY polypeptides may be an integral part of the NDH-I CO₂ uptake complex (NDH-1₃/₄) and involved directly in the conversion of CO₂ to HCO₃⁻, linked to electron transport and proton translocation associated with the NDH-1 complex. Recent attempts to determine the localization of these complexes suggest that they may be restricted to the thylakoid membrane (Ohkawa et al., 2001) thus linking them directly to the photosynthetic electron transport chain. A model of the operation of such an NDH-1₃/₄ CO₂-uptake complex is shown in Fig. 5. Although it is apparent that the Chp proteins have no homologies with known CA protein families, multiple alignment of Chp proteins across ten α and β-cyanobacteria (Maeda et al., 2002; Price et al., 2002) show that it is possible to identify two conserved histidine residues and one conserved cysteine residue which could act as a potential Zn coordination site, as is found in conventional CA enzymes. The model shown in Fig. 5 is based on the theoretical potential of Chp proteins to bind Zn in a manner analogous to CA active sites, where Zn-OH⁻ can form and act as a strong nucleophile for CO₂ attack.

A potential reaction sequence for the conversion of CO₂ to HCO₃⁻ is shown in Fig. 5 and explained in the legend. Energization of CO₂ to HCO₃⁻ conversion would be achieved in a two-step process. Electron donation to the complex by donors such as NAD(P)H or Fdₜₘₜ produces a reduced intermediate within the NDH-1 complex. This intermediate acts as a base (B) to abstract protons from the Zn-H₂O to produce the nucleophilic Zn-OH⁻ species. In the second step, the abstracted proton is translocated across the membrane to the lumen via a proton shuttle pathway within the hydrophobic proton channel subunits. By analogy with the E. coli NDH-1 complex (Friedrich and Scheide, 2000), the NdD and F subunits are part of the proton channel and would be specialized for their interactions with either the ChpX or ChpY subunits of each complex. The net result of these reactions is the hydration of CO₂ to HCO₃⁻ energized by electron transport through the NDH-1₃/₄ complexes. This model is still speculative and needs further evidence on the NDH-1 location of Chp proteins and their ability to bind Zn in the manner described above.

There is significant diversity in NDH-1₃/₄ gene content within and between α and β-cyanobacteria. The genomes of five β-cyanobacteria contain single copies of ndhF, ndhD and chp genes coding for both low and high-affinity CO₂ uptake (Badger et al., 2002), however the marine β-cyanobacterium Trichodesmium erythraeum has only the low-affinity uptake forms. The marine α-cyanobacteria Synechococcus WH8102 has single copies of the ndhF4, ndhD4 and chpX genes coding for a putative low-affinity CO₂ uptake system while the two Prochlorococcus marinus species have no homologues of either the low-affinity or high affinity NDH-1₃/₄ specific genes (Badger et al., 2002). From what is known about β-cyanobacteria, this absence should mean that they lack the capacity for active CO₂ uptake, unless they possess another active CO₂ uptake system that is presently uncharacterized.
Diversity in CCM components between cyanobacteria and ecological adaptation.

Over the past two years, there has been a rapid increase in the number of complete cyanobacterial genome databases, and several more are due for completion in the next two years. This has provided the opportunity for comparative genomic analysis of the gene content and diversity of a range of evolutionarily distinct cyanobacteria from different environments. An analysis of existing genomes (Badger et al., 2002), and the summaries outlined above, have shown that at present at least four types of cyanobacterial CCM strategies can be classified, defined on the presence of the types of carboxysomes present in the cells, and the complement of Ci transporters employed to accumulate HCO$_3^\text{-}$.

A schematic summary of these four types is shown in the Fig. 6.

The presence of $\alpha$ and $\beta$-carboxysomes is perhaps the most striking variation between $\alpha$ and $\beta$-cyanobacteria. However, as nothing is known about the relative effectiveness of each type of carboxysome, it is difficult to speculate on the photosynthetic carboxylation advantages that each structure and its Rubisco may confer.

Most $\beta$-cyanobacterial genomes examined to date, except Trichodesmium, have genes correlated with both low and high affinity NDH-1 CO$_2$ uptake systems. This presumably implies that all these cyanobacteria are able to induce high affinity CO$_2$ uptake systems when inorganic carbon becomes limiting.

Common high and low-affinity HCO$_3^\text{-}$ transport systems may also be present in most freshwater $\beta$-cyanobacteria, if sbtA (SLR1512) is a Na$^+/\text{HCO}_3^\text{-}$ transport gene, although Thermosynechococcus elongatus lacks an sbtA homologue. However, it should be noted that although Synechocystis contains the cmpABCD operon, and it is expressed (Omata et al., 2001), there appears to be no associated high-affinity HCO$_3^\text{-}$ uptake in this species (Shibata et al., 2002b). The reason for this is unclear.

The marine $\beta$-cyanobacterium Synechococcus PCC7002 lacks the cmpABCD operon (see NCBI unfinished bacterial genomes: http://www.ncbi.nlm.nih.gov:80/cgi-bin/Entrez/framik?db=Genome&gi=5102), however, it still manages to induce high-affinity HCO$_3^\text{-}$ uptake when grown at low Ci (Sülttemeyer et al., 1995). There is a strong homologue of sbtA in this strain that may code for one of the HCO$_3^\text{-}$ transport activities. Interestingly, the important N$_2$-fixing marine $\beta$-cyanobacterium, Trichodesmium erythraeum, appears to lack both cmpABCD and sbtA and the molecular basis for any HCO$_3^\text{-}$ uptake in this species remains to be determined.

The Ci transport genes of marine $\alpha$-cyanobacteria may be quite different. Synechococcus WH8102 possesses genes that could code only for a CO$_2$ uptake system similar to the low-affinity CO$_2$ uptake system (NDH-1$_\alpha$) from $\beta$-cyanobacteria. However, these genes are absent from both Prochlorococcus species sequenced so far (Badger et al., 2002). Observations with Synechococcus WH7803 (a close relative of WH8102) that the cells were able to evolve CO$_2$ during photosynthesis using HCO$_3^\text{-}$ (Tchernov et al., 1997) may be consistent with either the absence of CO$_2$ uptake genes in this species or the presence of only the low affinity uptake system. Extensive physiological studies of $\alpha$-cyanobacterial species are lacking, but for Prochlorococcus, it would be expected that if these cyanobacteria do possess active HCO$_3^\text{-}$ transport then they should evolve CO$_2$ during active photosynthesis. However, the nature of bicarbonate transport systems in $\alpha$-cyanobacteria may be quite different from $\beta$-cyanobacteria. A strong homologue of sbtA appears to be absent (Badger et al., 2002). If $\alpha$ and $\beta$-cyanobacteria diverged in their evolution prior to the development of bicarbonate transport systems, as carboxysome differences may suggest, then different types of HCO$_3^\text{-}$ transport may have evolved independently.

The ecological significance of the differences in CCM gene content between $\alpha$ and $\beta$-cyanobacteria remains to be determined. However, it is clear that the marine $\alpha$ and $\beta$-cyanobacteria occupy quite different habitats compared with most freshwater $\beta$-cyanobacteria. The Prochlorococcus species, Synechococcus WH8102 and Trichodesmium erythraeum occur as dominant primary producers throughout oligotrophic oceanic waters (Moore et al., 1998; Partensky et al., 1999). In these environments it may be expected that inorganic carbon is never severely depleted and light and other nutrients may be major limiting factors. Many of the $\beta$-cyanobacteria occupy environments such as mats, films, estuarine situations, and alkaline lakes where higher population densities of organisms may occur, other nutrients may be more abundant and, overall, situations where inorganic carbon is a limiting resource may be much more common. This may explain why many $\beta$-cyanobacteria have the ability to induce various CO$_2$ and HCO$_3^\text{-}$ transport systems as their environmental conditions change. However oceanic $\beta$-cyanobacteria such as Trichodesmium species that live in more oligotrophic waters are obviously different. The marine oceanic $\alpha$-cyanobacteria may have developed a physiology where they may not have the ability to acquire or induce high affinity inorganic carbon transport systems, and in some species no active CO$_2$ uptake system may be present. Physiological studies of these species must be done to resolve this question.

Evolution of the cyanobacterial CCM

Considering current knowledge of cyanobacterial CCM genes and past evolutionary and climatic processes it is possible to speculate on the timing and evolution of CCMs in cyanobacteria. Figure 7 shows a chronological view of the development of photosynthetic cyanobacteria and
algae over the past 3.5 billion years. This is plotted together with deduced changes in CO2 and O2 over the past 540 million years (the Phanerozoic era).

Past atmospheric CO2 levels when cyanobacteria first arose were probably over 100-fold higher than present day conditions. This, combined with low O2 conditions, would have meant that the original cyanobacteria would not have needed a CCM to achieve effective photosynthesis. The initial development of a CCM in cyanobacteria would have been triggered by changes in CO2 and O2 that caused CO2 to be a limiting resource for photosynthesis and the Rubisco oxygenase reaction to become a significant problem. Clear records for changes in O2 and CO2 before about 600 million years ago are lacking, but is has been inferred that O2 was near present levels by the beginning of the Phanerozoic and CO2 may have been around 15–20 times current atmospheric levels. Given the properties of current cyanobacterial Rubiscos (Badger et al., 1998), these enzymes should have been able to achieve efficient photosynthesis under these conditions. About 400 million years ago, there was a large decline in CO2 levels and an almost doubling in the oxygen concentration. These changes would have placed significant pressures on both cyanobacterial and algal photosynthesis. It can be argued that this may have been the first time that major pressure was applied to photosynthetic organisms to develop CCMs (Raven, 1997a).

The first steps towards developing a cyanobacterial CCM may have been quite simple and speculation and has been offered previously (Badger et al., 2002) and is outlined in Fig. 8. In the initial stages of CO2 decline, the first step to developing a CCM would have been the evolution of a carboxysome structure for Rubisco. The cyanobacterial CCM is totally dependent on this structure and all other additions would have revolved around its presence. A carboxysome carbonic anhydrase would probably have been required at this stage as the rate of chemical conversion of HCO3 to CO2 would have been too slow. As CO2 limitation became more severe, the development of the NDH-1 based low and high-affinity CO2 hydration process would have maintained adequate internal HCO3 pools and provided adequate CO2 levels around Rubisco in the carboxysome. This process would have been based around the modification of an existing respiratory NDH-1 complex, and would have resulted in the efficient recycling of leaked CO2 as well as net acquisition of CO2 from outside the cell. Finally, as more extreme CO2 limitation was imposed, the evolution of low and high affinity bicarbonate transport systems and high affinity CO2-uptake would have been necessary.

Examining the genes involved in cyanobacterial carboxysomes and proteobacterial micro-compartments, it is obvious that the common components are the small ccmK, ccmO, ccmL-like and csoS1-bacterial micro-compartment genes (Fig. 3). The larger csoS2 and csoS3 genes are specific for Form 1A Rubisco α-carboxysomes, while the ccmM and ccmN genes are specific for Form 1B Rubisco β-carboxysomes. The view shown in Fig. 8 is that carboxysomes developed first in cyanobacteria and differentiated into both α and β carboxysomes. The appearance

Fig. 7. The evolution of photosynthetic cyanobacteria, algae, higher plants, and chemoautotrophic proteobacteria over the past 3.5 billion years. (drawn with reference to Raven, 1997b, and Badger et al., 2002). Also shown on the graph are the proposed changes in CO2 and O2 during the Phanerozoic period (Berner, 2001; Berner and Kothavala, 2001). CO2 is shown as the ratio of past to present levels (R_co2) while O2 is shown as percentage content. The shaded section indicates the period of CO2 limitation combined with O2 increase that may have initiated the development of CCMs in aquatic photosynthetic organisms.
of carboxysomes in β-proteobacteria could have occurred through processes of lateral gene transfer which is becoming seen as a major part of bacterial evolution (Eisen, 2000; Brown et al., 2001; Sicheritz-Ponten and Andersson, 2001).

The evolutionary divergence of α and β-cyanobacteria may have been a fairly ancient event and it has been argued that the types of cyanobacteria split before the advent of the primary endosymbiosis some 2 billion years ago (Tomitani et al., 1999). If this is the case, then it could be argued that both cyanobacterial groups developed CCM mechanisms independently of each other rather than from a common ancestor having CCM components.

A polyphyletic origin of CCMs in cyanobacteria and algae

An independent and polyphyletic origin of CCM mechanisms in cyanobacteria, algae has been previously argued (Raven, 1997b; Badger et al., 2002). A cornerstone to support this position has been the proposition that if CO2 limitation were not imposed on cyanobacteria until the Phanerozoic, then this would strongly imply that the cyanobacteria that were the basis for the original primary endymbiotic event(s) probably did not have CCMs. Thus the original Chlorophyte and Rhodophyte algae would also have lacked any CCM genetic elements in common with cyanobacteria, that could have aided their adaptation to falling CO2 levels in the Phanerozoic. Chlorophytes and Rhodophytes as well as the secondary and tertiary endosymbiont algae that arose during the CO2-limitation of the Phanerozoic would all have needed to develop independent strategies for adapting to low CO2. Indeed, it has been suggested that the development of secondary endosymbiont algae may have been driven by this decline in CO2, as absorption into an acidic vacular structure may have made CO2 more available by the conversion of HCO3- to CO2 (Lee and Kugrens, 2000).

A search of the existing higher plant and algal DNA databases indicates there are no homologues of cyanobacterial CCM genes to be found. Thus carboxysome genes are almost exclusively restricted to cyanobacteria and some proteobacteria. Likewise, the NDH-1,3 CO2 uptake genes and the bicarbonate transport genes are restricted to cyanobacteria. An exception to this generalization is the occurrence of carboxysome structures in the plastids of Glaucocystophytes such as Cyanophora paradoxa. However, the evolutionary position of these algae has been very confusing. They retain a plastid that obviously resembles a cyanobacterium and if the plastids of red and green algae are assumed to be monophyletic then these algae appear to be an outgroup (Löffelhardt and Bohnert, 1994). Perhaps their plastids were obtained in a later endosymbiotic event involving a cyanobacterium that had developed a carboxysome structure. The algal genome databases are limited at present and are dominated by algal chloroplast genomes and Chlamydomonas reinhardtii ESTs, and it is possible that further sequencing of other algal species may uncover some CCM homologues in the nuclear genome. However, it would seem reasonable at this stage to assume that there were multiple origins of aquatic CCMs in algae and the ancestors of higher plants.

References


Beuderke RF, Cannon GC, Kuenen JG, Shively JM. 1980. Relations between RuBP carboxylase, carboxysomes and CO2...
fixing capacity in the obligate chemolithotroph *Thiobacillus neapolitana* grown under different limitations in the chemostat. *Archives of Microbiology* 1241, 185–189.


Ludwig M, Sültemeyer D, Price GD. 2000. Isolation of ccmKLN genes from the marine cyanobacterium, *Synechococcus* sp. PCC7002 (Cyanobacteria), and evidence that CcmM is essential for carboxysome assembly. *Journal of Phycology* 36, 1109–1118.


Ludwig M, Sültemeyer D, Price GD. 2000. Isolation of ccmKLN genes from the marine cyanobacterium, *Synechococcus* sp. PCC7002 (Cyanobacteria), and evidence that CcmM is essential for carboxysome assembly. *Journal of Phycology* 36, 1109–1118.


Omata T, Ogawa T. 1986. Biosynthesis of a 42 kDa polypeptide in the cytoplasmic membrane of the cyanobacterium *Anacystis*
nidulans strain-R2 during adaptation to low CO2 concentration. Plant Physiology 80, 525–530.


