JOURNAL OF BACTERIOLOGY, Feb. 2004, p. 623–630 Vol. 186, No. 3
0021-9193/04/$08.00+0 DOI: 10.1128/JB.186.3.623–630.2004

A Novel Evolutionary Lineage of Carbonic Anhydrase (ε Class) Is a Component of the Carboxysome Shell

Anthony K.-C. So,[1] George S. Espie,[1] Eric B. Williams,[2] Jessup M. Shively,[2,3] Sabine Heinhorst,[2] and Gordon C. Cannon[2]*

Department of Botany, University of Toronto at Mississauga, Mississauga, Ontario, L5L 1C6, Canada[1]; Department of Chemistry and Biochemistry, The University of Southern Mississippi, Hattiesburg, Mississippi 39406-5043[2]; and Department of Genetics and Biochemistry, Clemson University, Clemson, South Carolina 29634[3]

Received 3 September 2003/Accepted 29 October 2003

A significant portion of the total carbon fixed in the biosphere is attributed to the autotrophic metabolism of prokaryotes. In cyanobacteria and many chemolithoautotrophic bacteria, CO_2 fixation is catalyzed by ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO), most if not all of which is packaged in protein microcompartments called carboxysomes. These structures play an integral role in a cellular CO_2-concentrating mechanism and are essential components for autotrophic growth. Here we report that the carboxysomal shell protein, CsoS3, from *Halothiobacillus neapolitanus* is a novel carbonic anhydrase (ε-class CA) that has an evolutionary lineage distinct from those previously recognized in animals, plants, and other prokaryotes. Functional CAs encoded by *csoS3* homologues were also identified in the cyanobacteria *Prochlorococcus* sp. and *Synechococcus* sp., which dominate the oligotrophic oceans and are major contributors to primary productivity. The location of the carboxysomal CA in the shell suggests that it could supply the active sites of RuBisCO in the carboxysome with the high concentrations of CO_2 necessary for optimal RuBisCO activity and efficient carbon fixation in these prokaryotes, which are important contributors to the global carbon cycle.

Many bacteria contain polyhedral protein microcompartments that are bounded by a 3- to 4-nm shell composed of several highly conserved polypeptides. However, the enzyme content of these structures varies depending on metabolic function and the bacterial strain in which they occur. For example, microcompartments found in a *Salmonella* sp. grown on propanediol contain propanediol dehydratase, while similar microcompartments appearing in *Escherichia coli* grown on ethanolamine contain ammonia lyase (12, 16). The widespread occurrence of such polyhedral bodies has led to the suggestion that their protein shell provides a catalytic advantage to the enzyme within.

The archetypal example of a prokaryotic microcompartment is the carboxysome, which is found in all cyanobacteria and many chemolithoautotrophic bacteria. Carboxysomes from the chemolithoautotroph *Halothiobacillus* (formerly *Thiobacillus*) *neapolitanus* measure 100 to 120 nm across and are filled with the CO_2-fixing enzyme of the Calvin-Benson-Bassham cycle, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO, form 1A), which is essential for autotrophic carbon metabolism (7, 27, 34). In addition to the large (CbbL) and small (CbbS) subunits of RuBisCO, up to seven additional polypeptides are associated with the carboxysome; five of these have been identified as shell proteins (3, 4, 11). The carboxysome polypeptides are encoded by genes within the *cso* gene cluster, an apparent operon that also includes *cbbL* and *cbbS* (29). A similar organization of carboxysome genes has been reported for three other members of the former genus *Thiobacillus* (6) and for several globally important cyanobacterial strains, which dominate the oligotrophic oceans (1, 8). Here we refer to these as cso-carboxysomes to distinguish them from ccm-carboxysomes found in several other marine and freshwater cyanobacteria such as *Synechocystis* sp. strain PCC6803 and *Synechococcus* sp. strain PCC7942 (1, 15). The components of ccm-carboxysomes appear to be encoded by the *ccmKLMN* gene cluster, which is typically located upstream of the genes coding for the large and small subunits of RuBisCO (form 1B) (15, 23). Although the putative shell proteins CcmK and CcmL share considerable sequence identity with CsoS1A, -B, and -C and with OrfA and OrfB, respectively, CcmM and CcmN have no readily apparent similarity to CsoS2 or CsoS3 (1, 8).

The number of carboxysomes per cell increases during CO_2 limitation in many organisms (5, 15, 19, 23). Deleterious mutations in the genes encoding carboxysome components result in a conditionally lethal phenotype that requires high concentrations of CO_2 or the induction of an alternate metabolic mode to permit survival and growth of the mutants (1, 2, 15, 23). Although the detailed biochemical mechanism by which carboxysomes enhance autotrophic CO_2 fixation is not well understood, the collective evidence suggests that the unique structural organization and the bounding shell of carboxysomes provide a distinct catalytic advantage that is essential to this process (7, 15).

Since the substrate for RuBisCO is CO_2, it has long been hypothesized that carboxysomes contain a carbonic anhydrase (CA; EC 4.2.1.1) (9, 15, 24). This enzyme catalyzes the reversible hydration of CO_2, and one model (25) predicts that CA would ensure rapid conversion of cytosolic HCO_3^- to CO_2 at a concentration that would support optimal RuBisCO activity. Indeed, CA activity has been reported to be associated with partially purified ccm-carboxysomes from two different cya-

* Corresponding author. Mailing address: Department of Chemistry and Biochemistry, The University of Southern Mississippi, Charles Lane Drive, Hattiesburg, MS 39406-5043. Phone: (601) 266-4221. Fax: (601) 266-6075. E-mail: gordon.cannon@usm.edu.

nobacterial strains (22, 31, 32). The enzyme has been shown to be an essential component of a cellular CO_2-concentrating mechanism (CCM) that is necessary for autotrophic growth under normal conditions (23, 32, 37), although its localization as an integral carboxysome component is still unclear. This CA (CcaA) is a member of the β class, one of three independent evolutionary lineages (α, β, and γ) that are widespread among the domains *Bacteria*, *Archaea*, and *Eucarya* (13). It was therefore surprising to find that the genomes of several prokaryotes encoding cso-carboxysomes lack genes for any identifiable member of the α, β, or γ class of CAs (10, 26).

In this study we demonstrate unequivocally that CsoS3, from the chemolithoautotroph *H. neapolitanus*, represents a new lineage of CA (ε class) and is a catalytically functional component of the shell of cso-carboxysomes. Homologues of CsoS3 from the marine cyanobacteria *Prochlorococcus* sp. and *Synechococcus* sp. are also ε-class CAs that likely reside in the carboxysome shell. The shell-localized CA is proposed to supply CO_2 to the active sites of RuBisCO in the carboxysome to support optimal CO_2 fixation activity.

MATERIALS AND METHODS

Carboxysome purification. *H. neapolitanus* was grown in a 3-liter chemostat at a dilution rate of 0.08 h^{-1} as previously described (9). Cells (6 to 8 g) were resuspended in 50 ml of TEMB (5 mM Tris-HCl [pH 8.0], 1 mM EDTA, 10 mM $MgCl_2$, 20 mM $NaHCO_3$) before they were ruptured by sonication and centrifuged (12,000 × *g*, 20 min). The supernatant was collected and centrifuged at 40,000 × *g* for 30 min. The resulting pellet was resuspended in 20 ml of a 33% (vol/vol) solution of CelLytic B-II (Sigma) in TEMB and centrifuged (40,000 × *g*, 30 min). The carboxysome-enriched pellet was resuspended in 3 ml of TEMB, clarified by centrifugation (3,000 × *g*, 1 min), and loaded onto 10 to 50% (wt/vol) linear sucrose density gradients. After centrifugation (105,000 × *g*, 30 min) in a Beckman JS24.38 rotor, the gradients were fractionated and the prominent carboxysome band near the middle of the gradient was collected and recentrifuged at 100,000 × *g* for 90 min. The resulting pellet of carboxysomes was resuspended in 1 ml of TEMB before RuBisCO (19) and CA (see below) activities were assayed. Carboxysomes were negatively stained in 2% (wt/vol) uranyl acetate on carbon-coated grids and were observed with a Philips EM 300 electron microscope. Thin-sections of *H. neapolitanus* cells were also prepared and viewed. Purified carboxysomes were disrupted by freezing pelleted carboxysomes at −20°C for 30 min. The pellet was subsequently thawed, resuspended in 100 to 200 μl of TEMB at 4°C, and subjected to centrifugation at 14,000 × *g* for 30 min. The resulting pellet consisted mainly of carboxysome shell "ghosts," while the supernatant contained most of the RuBisCO that originally had been packaged in the intact particles.

CsoS3 expression in *E. coli*. Recombinant CsoS3 was produced in *E. coli* by using several different expression systems. An existing clone containing pcsoS3ProEx (4) was employed to express a histidine-tagged CsoS3 protein. All other constructs were generated by PCR amplification of the *csoS3* gene using the primers shown in Table 1. The nucleotide sequences, orientations, and reading frames of all cloned fragments were verified prior to protein expression analyses. Clones producing T7-tagged CsoS3 were constructed by amplification of the *csoS3* gene from *H. neapolitanus* genomic DNA, subsequent digestion with *Eco*RI, ligation into pET-21b (Novagen), and transformation of *E. coli* BL21(DE3) cells with the resulting expression construct. Homologues of *csoS3* were similarly amplified from genomic DNA preparations of *Prochlorococcus marinus* strains MED4 and MIT9313 and *Synechococcus* strain WH8102, which were kindly provided by Sallie W. Chisholm (Massachusetts Institute of Technology) and Bianca Brahamsha (Scripps Institution of Oceanography). The amplified fragments were restricted with *Hin*dIII and ligated into the corresponding site in pET-21b, and the resulting plasmids were used to transform *E. coli*. To provide recombinant CsoS3 without any purification tag, the intein-based IMPACT Protein Purification System (New England Biolabs) was employed. Briefly, the *csoS3* gene from *H. neapolitanus* was amplified from plasmid pTn1, which contains the entire *cso* gene cluster from *H. neapolitanus* (2), and inserted into the *Nde*I and *Sap*I sites of pTYB1 to generate plasmid pE32A. Recombinant CsoS3 protein was overexpressed as a fusion protein containing a C-terminal intein with a chitin binding domain and was purified by affinity chromatography on a chitin column. After thiol-induced intein self-cleavage, recombinant CsoS3 was eluted from the column.

TABLE 1. Oligonucleotides used for PCR amplification of *csoS3* genes[a]

Primer designation	Nucleotide sequence	Original sequence designation (nucleotide positions)
HNIF	GGTGGTCATATGAACACCCG	4628–4648
HNIR	GGTGGTTGCTCTTCCGCATGCGG	6168–6145
HNF	GGAATTCGATCATGAACACCC	4626–4646
HNR	GGAATTCGACCCATATCAGC	6245–6226
MEDF	TAAGCTTATGCCTTTGAGAGG	1524589–1524569
MEDR	GAAGCTTATGCTCGAATCCC	1522981–1523000
MITF	TAAGCTTATGGCCTATCGC	524697–524715
MITR	GAAGCTTAAGAACCTTGCAG	526277–526258
WHF	TAAGCTTATGCCTCTCCG	727050–727033
WHR	GAAGCTTCCGTCTTGCAC	725218–725235

[a] Primers were designed from DNA sequences reported for *H. neapolitanus* (designations begin with HNI for the IMPACT system and with HN for the pET system), *P. marinus* strains MED4 (MED) and MIT9313 (MIT), and *Synechococcus* strain WH8102 (WH). Nucleotide substitutions were made in primer sequences in order to introduce restriction sites (underlined) that would facilitate subsequent cloning and CsoS3 expression.

Recombinant CsoS3 expression was typically induced by addition of 0.6 to 1.0 mM isopropyl-D-thiogalactoside to transformed cells, which were grown at 29°C with shaking (175 rpm) to an optical density at 600 nm of 0.5. After 4 h, cells were harvested by centrifugation (at 4,000 × *g* for 5 min) and washed with 20 ml of assay buffer (100 mM EPPS [*N*-(2-hydroxyethyl)piperazine-*N*′-3-propanesulfonic acid]–NaOH [pH 8.0]–20 mM $MgSO_4$). Cells were resuspended in 2 ml of assay buffer prior to disruption (four 20-s pulses) with an Ultrasonics Sonicator (model W-220F; Branson). Cell lysates were maintained on ice for a minimum of 6 h to facilitate proper folding of polypeptides prior to mass spectrometric (MS) analyses.

CA assays. CA activity in protein samples was detected as acceleration in the rate of $^{13}C^{16}O_2$ appearance (m/z = 45), arising from the exchange of ^{18}O from $^{13}C^{18}O^{18}O$ (m/z = 49) to H_2O. Assay buffer (6 ml) was sealed in a cuvette and equilibrated to 30°C. $K_2{}^{13}C^{18}O_3$ (400 μM) was added, and the progress of the uncatalyzed exchange reaction was monitored by measuring m/z 49, 47, and 45 with an MS (MM14-80SC; VG Gas Analysis) (20). After 3 min, a protein sample (100 μl) was added, and the effect was recorded. Samples were incubated with classical CA inhibitors for 20 min at room temperature and 5 min at 30°C prior to the start of each assay (20).

CA activity was also measured electrometrically. The protein sample was added to 800 μl of 20 mM Tris-HCl, pH 8.3, and the reaction was initiated by adding 600 μl of ice-cold water saturated with CO_2. The pH change resulting from the hydration of CO_2 was measured by a combination microelectrode (Microelectrodes Inc.). One Wilbur-Anderson (WA) unit of activity is defined as $2(T_0 - T)T^{-1}$, where T_0 and T are the times required for the pH to change from 8.3 to 6.3 for the uncatalyzed and catalyzed reactions, respectively (36). Inhibitors were incubated with protein samples for 30 min on ice before the start of the assay.

Western blot analysis. After protein transfer (at 100 V for 1 h) onto nitrocellulose membranes, blots were probed with a 1:2,500 (vol/vol) dilution of a rabbit antiserum raised against CsoS3 or with a 1:7,000 (vol/vol) dilution of a rabbit antiserum raised against CbbL from *H. neapolitanus*. The blots were subsequently developed with a 1:10,000 (vol/vol) dilution of alkaline phosphatase-conjugated anti-rabbit immunoglobulin G (Sigma) and 1-Step nitroblue tetrazolium–5-bromo-4-chloro-3-indolylphosphate (Pierce).

RESULTS

CA activity is associated with purified cso-carboxysomes. Carboxysomes from *H. neapolitanus* (Fig. 1A) were purified to homogeneity (Fig. 1B) and assayed for CA activity by using a sensitive MS assay (Fig. 2A). The assay measures the depletion of ^{18}O from $^{13}C^{18}O_2$ (m/z = 49) to H_2O that occurs by repeated hydration-dehydration, yielding $^{13}C^{18}O^{16}O$ (m/z = 47) as an intermediate and $^{13}C^{16}O_2$ (m/z = 45) as the final product

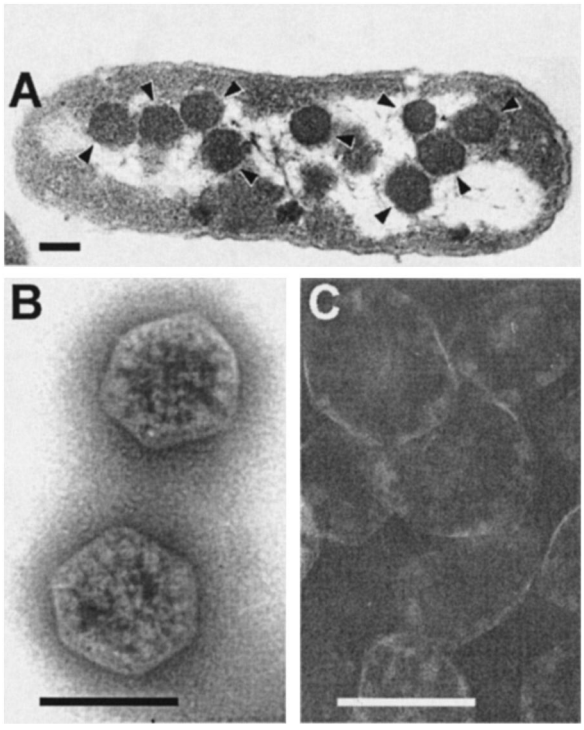


FIG. 1. Carboxysomes from *H. neapolitanus*. (A) Transmission electron micrograph of an *H. neapolitanus* cell containing numerous polyhedral carboxysomes (indicated by arrowheads). (B) Purified, negatively stained intact carboxysomes. (C) Negatively stained carboxysome shells after freeze-thaw treatment. Bars, 100 nm.

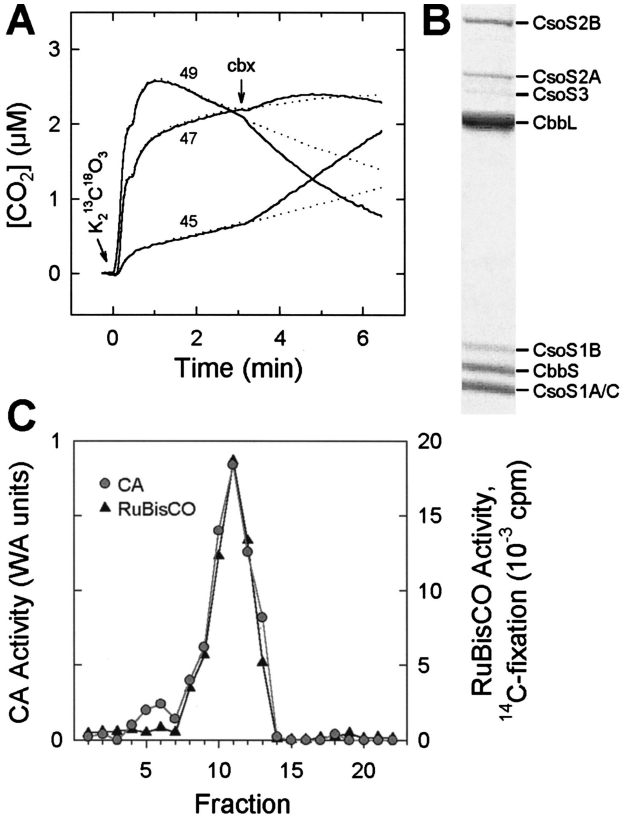


FIG. 2. CA activity associated with purified carboxysomes. (A) Measurements of ^{18}O exchange activity by MS. Experiments were initiated by the addition of 400 μM $K_2{}^{13}C^{18}O_3$ to the reaction vessel containing 100 mM EPPS-NaOH (pH 8.0) and 20 mM $MgSO_4$. The relative concentrations of $^{13}C^{18}O_2$ (m/z = 49), $^{13}C^{18}O^{16}O$ (m/z = 47), and $^{13}C^{16}O_2$ (m/z = 45) were monitored prior to and following the addition of purified carboxysomes (cbx; 120 μg of protein) from *H. neapolitanus*. The time course of the uncatalyzed isotopic exchange reaction in buffer alone is depicted by the dotted lines. (B) Electrophoretic separation of polypeptides from a homogeneous carboxysome preparation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Proteins were visualized following Coomassie staining. (C) CA and RuBisCO activities in fractions collected following carboxysome purification on sucrose density gradients. Partially purified carboxysomes were loaded onto a 10 to 50% sucrose density gradient and centrifuged. Fractions from the gradient were collected and assayed.

of this isotopic exchange reaction (20). CA activity was detected in cell lysates and all carboxysome-containing fractions obtained during purification (data not shown). When the purified carboxysomes (Fig. 2B) were added to the assay buffer containing $K_2{}^{13}C^{18}O_3$, a significant acceleration in the rate of loss of m/z 49 and a concomitant increase in the rate of appearance of m/z 45 was observed (Fig. 2A), consistent with the presence of CA activity in this fraction. CA activity assayed by the electrometric assay (36) was also found to comigrate with RuBisCO activity on sucrose density gradients (Fig. 2C), demonstrating a tight physical association between the two enzymes.

***csoS3* encodes a CA.** Purified *H. neapolitanus* carboxysomes are composed of nine polypeptides: CsoS1A, CsoS1B, CsoS1C, CsoS2, CsoS3, OrfA, OrfB, and the large and small subunits of RuBisCO (CbbL and CbbS) (Fig. 2B and 3A). These proteins were expressed individually in *E. coli*, and cell extracts of the bacteria were examined for CA activity. Only *E. coli* transformants expressing CsoS3 were found to possess CA activity (Fig. 3B). To determine if homologues of *csoS3* from the globally important cyanobacteria *P. marinus* MED4, *P. marinus* MIT9313, and *Synechococcus* strain WH8102 encode a functional CA, the *csoS3* genes from the three strains were cloned and expressed in *E. coli*. All three recombinant CsoS3 proteins also displayed substantial CA activity (Fig. 3C).

CsoS3 and CA activity reside in the carboxysome shell. A single freeze-thaw cycle releases much of the RuBisCO from purified carboxysomes and leaves the shells largely intact as empty "ghosts," as demonstrated previously (28). The two components can be separated by centrifugation, yielding a pellet consisting mainly of carboxysome shells (Fig. 1C) and a soluble fraction enriched in RuBisCO. When these fractions were tested for CA activity, on a protein basis, 95 to 100% of the activity found in the intact carboxysomes was observed in the shell fraction (Fig. 4A). Immunoblot analysis of the supernatant and pelletable fractions of broken carboxysomes from *H. neapolitanus* (Fig. 4B) confirmed the findings of the activity assays. Most of the protein that is recognized by an antibody directed against the CsoS3 polypeptide was confined to the carboxysome shell fraction. Conversely, the majority of CbbL detected by anti-CbbL was found in the supernatant. These results clearly demonstrate that the CA activity of CsoS3 is located in the carboxysomal shell or tightly associated with it.

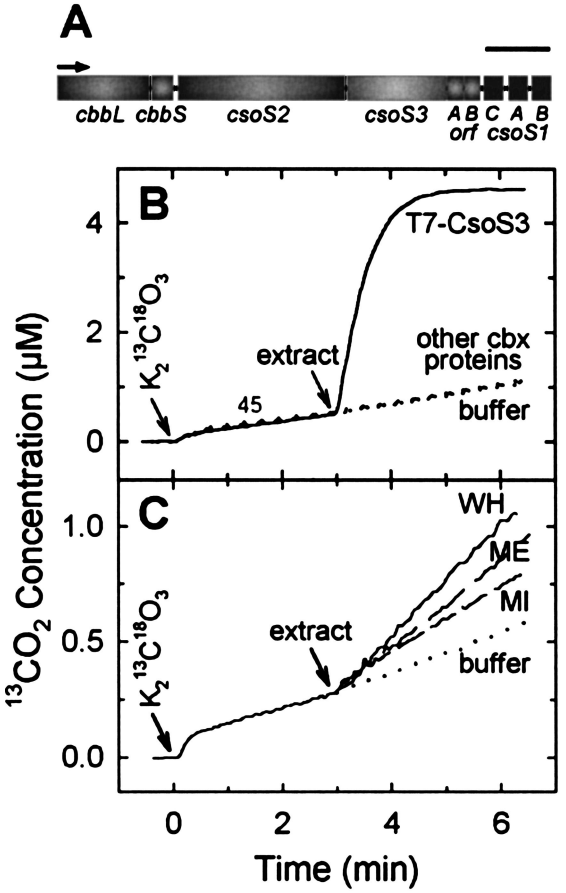


FIG. 3. Identification of CsoS3 as a CA. (A) Schematic of the *cso* gene cluster from *H. neapolitanus*. Bar, 1 kb. (B) Measurements of ^{18}O exchange in cell lysates (1.5 mg of protein) of *E. coli* expressing epitope-tagged CsoS3 and each of the other polypeptides encoded by the *H. neapolitanus cso* gene cluster. For clarity, only the *m/z* 45 traces are presented. (C) ^{18}O exchange activity in bacterial extracts (0.8 mg of protein) containing T7-CsoS3 from *Synechococcus* strain WH8102 and *P. marinus* strains MED4 and MIT9313.

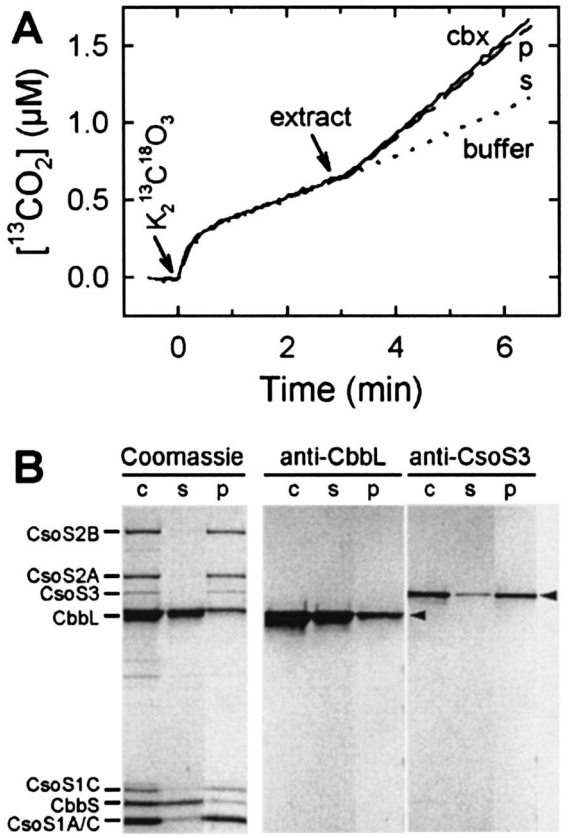


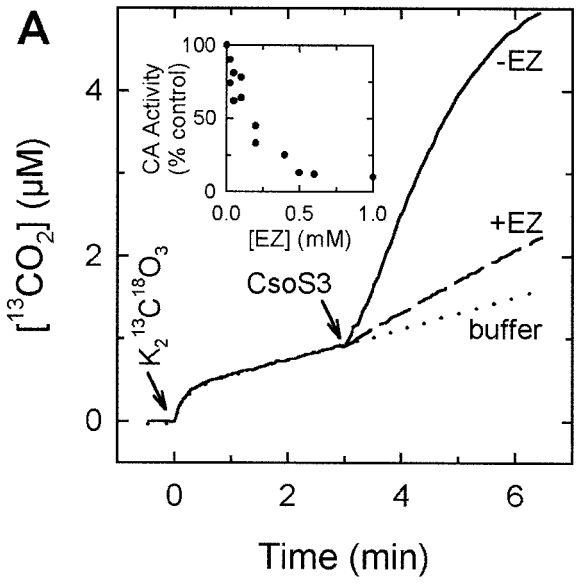
FIG. 4. Association of CsoS3 with the carboxysome shell. (A) ^{18}O exchange activity in intact (cbx) and broken carboxysome preparations. After freeze-thaw treatment of purified carboxysomes, pelletable (p) (71 μg of protein) and soluble (s) (49 μg of protein) fractions were assayed. (B) Coomassie staining of carboxysomal proteins and immunodetection of CsoS3 and CbbL. The faint, unassigned bands are incompletely denatured aggregates of shell proteins.

The localization of CsoS3 to the carboxysome periphery is also supported by immunogold electron microscopy (4).

Characterization of CA activity. To ensure that the observed CA activity was truly due to catalysis by the CsoS3 polypeptide, the *H. neapolitanus csoS3* gene was expressed as a fusion protein with a self-splicing intein (Impact protein purification system; New England Biolabs). Recombinant CsoS3 lacking any extraneous fusion tag was purified to homogeneity by affinity chromatography. This protein had a high level of CA activity (Fig. 5). By use of the electrometric assay (36), the specific activity of the recombinant CsoS3 was found to be approximately 30% of that measured with commercially obtained purified bovine erythrocyte CA (Fig. 5B). CsoS3 CA activity was inhibited by dithiothreitol, suggesting the involvement of cysteine residues in catalysis. Ethoxyzolamide (EZ), a classical inhibitor of CA that acts by interfering with the catalytically essential Zn^{2+} within the enzyme's active site, also inhibited the activity of the carboxysomal CA. Likewise, divalent metal chelators such as dipicolinic acid and nitrilotriacetic acid severely interfered with CA activity. These results are consistent with CsoS3 acting through a metal-mediated mechanism for the hydration of CO_2 and the dehydration of HCO_3^-, as is the case for other known CAs (30, 35).

CsoS3: a CA of distinct evolutionary origin. By using *H. neapolitanus* CsoS3 as a probe, searches of the National Center for Biotechnology Information nonredundant database (http://www.ncbi.nlm.nih.gov/), the Comprehensive Microbial Resource at The Institute for Genomic Research (http://www.tigr.org/), and unfinished genome sequences at the Joint Genome Institute (http://www.jgi.doe.gov/) revealed that the protein shares similarity with eight additional sequences, all found in organisms containing *cso* gene clusters (1, 8). These include the chemolithoautotrophs *Thiobacillus denitrificans*, *Acidothiobacillus ferrooxidans*, and *Thiomonas intermedia*, the three marine photolithoautotrophic cyanobacteria named earlier, *P. marinus* SS120 (or CCMP1375) (10), and *P. marinus* subsp. *pastoris* CCMP1378.

Alignment of the CsoS3 sequences demonstrated a high degree of sequence conservation, with up to 99% sequence identity observed between the MED4 and CCMP1378 polypeptides. Phylogenetic analysis (17) using nucleotide sequences indicated that *csoS3* could be separated into two closely related clades (Fig. 6), reflecting the chemolithoau-



B

Sample	CA Activity: Electrometric (WA units / mg protein)	CA Activity: MS (% control)
CsoS3 (control)	234	100
+ 1 mM DTT	29	14
+ 0.5 mM EZ	163	13
+ 2 mM NTA	103	-
+ 20 mM DPA	94	11
Boiled CsoS3	0	0
Purified, intact carboxysomes	14 (280)	10
Bovine CA	720	788

FIG. 5. Measurements of CsoS3 activity. (A) Effect of EZ (500 μM) on CsoS3-catalyzed ^{18}O exchange activity. (Inset) Graph of inhibition of CsoS3 by increasing concentrations of EZ. (B) Electrometric and MS measurements of CA activity. Recombinant, affinity-purified CsoS3 (38 μg) was derived from the IMPACT system. Electrometric assays were performed at 4°C for CsoS3 (1.75 μg), purified intact carboxysomes (100 μg), and bovine erythrocyte CA (1 μg). The value in parentheses represents the calculated specific activity of CsoS3, assuming that this protein constitutes 5% of the total carboxysomal protein. Alternatively, the activity of CsoS3 (6 to 20 μg) was determined at 30°C by MS as a percentage of the initial rate of *m/z* 45 appearance. Note that the solubility of EZ is greatly reduced at 4°C. DTT, dithiothreitol; NTA, nitrilotriacetic acid; DPA, dipicolinic acid.

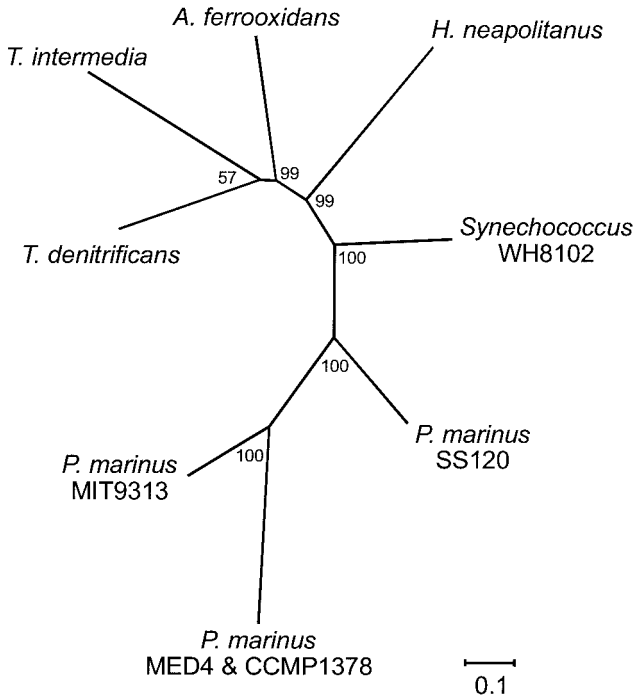


FIG. 6. Unrooted phylogenetic tree for *csoS3* homologues. Nucleotide sequence alignments were performed with ClustalX, version 1.81, and used to construct a neighbor-joining tree with MEGA, version 2.1. Bootstrap values are displayed at nodes as percentages of 1,000 replicates.

totrophic or photolithoautotrophic nature of the organisms from which the sequences were obtained. The high degree of relatedness of the *csoS3* genes from these otherwise phylogenetically distant autotrophic organisms is consistent with the notion that *csoS3* and the other members of the *cso* gene cluster have been laterally transferred as a genetic unit (1, 7).

BLAST searches using the most highly conserved regions of the CsoS3 amino acid sequence did not reveal other significant homologues. A search for remote homologues using the PROPSEARCH program (http://www.infobiosud.univ-montp1.fr/SERVEUR/PROPSEARCH/propsearch.html), which relies on analysis of protein properties, also failed to reveal any additional homologues, including any member of the α, β, or γ class of CAs. The overall three-dimensional folding patterns of the three classical CA types, as determined by X-ray crystallography, are unique and emphasize their independent evolutionary origins (30, 35). A CA (TWCA1) from the diatom *Thalassiosira weissflogii* may represent a fourth lineage (δ class) (35) of the enzyme, although some evidence suggests that it may instead be a distant homologue of the α-class CAs (18). Using the 3D-PSSM program (http://www.sbg.bio.ic.ac.uk/servers/3dpssm/), we attempted to identify common folding patterns between CsoS3 and other known proteins. No structural homology between CsoS3 and any of the recognized CA classes was evident. The collective data indicate, therefore, that CsoS3 is a hitherto undiscovered evolutionary lineage of CA, which we propose to designate the ε class.

All of the ε-class CA genes identified to date are found within *cso* gene clusters, flanked by *csoS2* and *orfA* in an arrangement similar to that observed for *H. neapolitanus* (Fig. 3A). The genes range from 1,503 to 1,602 nucleotides, with the exception of the *Synechococcus* strain WH8102 *csoS3* gene, which is 1,728 nucleotides. The predicted molecular masses of the CsoS3 polypeptides range from 55.2 to 63.4 kDa and are considerably greater than the masses (17 to 32 kDa) of typical α-, β-, and γ-class CA monomers. There are two potential translation start sites within the *Synechococcus* strain WH8102 *csoS3* gene. Only the larger of the two proteins (575 amino acids [M_r, 63,400] as opposed to 482 amino acids [M_r, 53,500]) was catalytically active (Fig. 3C). In all known instances, the

FIG. 7. Positions of candidate metal-binding residues in ε CAs and known zinc ligands of other CAs. (A) Alignment for a conserved portion of the CsoS3 polypeptides from *H. neapolitanus* (Hn), *T. denitrificans* (Td), *A. ferrooxidans* (Af), *T. intermedia* (Ti), *Synechococcus* sp. strain WH8102 (WH), and *Prochlorococcus* sp. strains CCMP1378 (CC), MED4 (MED), MIT9313 (MIT), and SS120 (SS). Amino acids conserved among all proteins are shaded. Potential zinc-binding residues are indicated by asterisks. (B) Zinc ligands (shaded) determined from the X-ray crystal structures for human CAII (α class), *P. purpureum* CA (β class), and *Methanosarcina thermophila* Cam (γ class). It must be pointed out that β- and γ-class CAs are multimeric enzymes with multiple Zn^{2+} binding sites. Amino acid numbering is indicated at the right of each sequence.

catalytically essential Zn^{2+} is coordinated to the CA apoprotein through three histidine residues or a combination of histidine, cysteine, and sometimes aspartate (Fig. 7) (35). There are six histidine, two cysteine, and six aspartate residues that are completely conserved across the nine CsoS3 polypeptides. Many (71%) of these candidate Zn^{2+} binding residues are located in a conserved 140-amino-acid stretch in the midportion of the protein (Fig. 7).

DISCUSSION

Electrometric and MS assays demonstrate that purified CsoS3 from *H. neapolitanus* catalyzes the reversible hydration of CO_2 and the exchange of ^{18}O between $^{13}C^{18}O_2$ and H_2O, reactions characteristic of the CAs (EC 4.2.1.1). Further support for this classification comes from the observation that catalysis may occur through a metal-mediated reaction mechanism that is inhibited by the classical CA inhibitor EZ. Recombinant CsoS3 from the marine cyanobacteria *Prochlorococcus* and *Synechococcus* also catalyze ^{18}O exchange and are, therefore, functional homologues of the *H. neapolitanus* enzyme. The CsoS3 monomer is considerably larger (>50%) than typical α-, β- and γ-class CA monomers. Functional CAs with monomeric sizes comparable to that of CsoS3 have been identified from the salt-tolerant green alga *Dunaliella salin*a (60 kDa) (11a) and the red alga *Porphyridium purpureum* (50 kDa) (20a). However, the increased size likely is the consequence of gene duplication, as the monomers contain either two α-class or two β-class CA domains. This is not the case for CsoS3. The larger size and mass of CsoS3 may reflect a dual role as a catalytically active enzyme and as a structural component of the carboxysome.

RuBisCO and CsoS3-associated CA activity colocalize to the carboxysome, with RuBisCO located within the core and CsoS3 tightly associated with the bounding shell. CsoS3 is a minor protein component of the carboxysome (Fig. 2B) and is therefore unlikely to form the major structural elements of the carboxysome shell. The close physical proximity of CsoS3-associated CA activity to RuBisCO within the cso-carboxysome interior, however, immediately suggests a direct role for this enzyme in the provision of CO_2 for autotrophic fixation. This suggestion is most compelling for *Prochlorococcus* spp. and *A. ferrooxidans*, since the genomes of these organisms do not contain recognizable genes for any member of the α, β, or γ class of CAs. *Synechococcus* strain WH8102 contains, in addition to *csoS3*, a single β-class CA gene whose product is most closely related to a noncarboxysomal β-CA thought to be associated with the cell surface in *Synechocystis* sp. strain PCC6803 (33). From a functional point of view, CsoS3 may be located at strategic positions throughout the carboxysome shell, where it can gain access to the intracellular HCO_3^- pool and channel CO_2 to RuBisCO. Direct measurements have established the presence of a cytosolic HCO_3^- pool in *H. neapolitanus* and many cyanobacteria (1, 14), but such mea-

surements are lacking for the marine cyanobacteria considered here. The colocalization of RuBisCO with a direct enzymatic source of its substrate, CO_2, provides an obvious catalytic advantage to autotrophic CO_2 fixation, but whether additional benefits are afforded by the unique organization of the carboxysome remains to be established. Knowledge of the spatial organization of CsoS3 and RuBisCO at the molecular level and of their kinetic characteristics is essential to furthering our understanding of the functioning of cso-carboxysomes.

Although CsoS3 catalyzes classic CA reactions, its primary sequence and predicted folding pattern are unlike those reported for any other class of CA. Phylogenetic analysis failed to find even remote similarities between CsoS3 and other CAs, leading to the suggestion that it represents a novel evolutionary lineage of the enzyme, the ε class. At present, ε-CA appears to be unique to the *Bacteria* and is narrowly distributed among several chemolithoautotrophs and certain marine cyanobacteria that contain cso-carboxysomes. Recent genome-wide analyses of *P. marinus* SS120 and MED4 indicate that these organisms may contain the minimal gene complement necessary for oxyphototrophic bacteria (10, 26). The perceived absence of a CA in these organisms has raised questions as to the necessity of this enzyme for autotrophic CO_2 fixation. Our work indicates that a catalytically functional ε-CA is encoded within these minimal genomes and that carboxysomal CA is an integral component of autotrophic metabolism.

From an ecological point of view, bacterial strains that contain ε-CA and cso-carboxysomes are globally distributed and, in the case of *Prochlorococcus* and *Synechococcus* spp., are abundant primary producers in the subtropical and tropical open oceans (21). Thus, through their photosynthetic and chemolithoautotrophic CO_2 fixation, carboxysomes and ε-CA influence the biogeochemical cycling of carbon in a globally significant way.

ACKNOWLEDGMENTS

This work was supported by the Ontario Graduate Scholarship Fund (to A.K.-C.S.), the Natural Sciences and Engineering Research Council of Canada (to G.S.E.), the National Science Foundation EPSCoR and MRSEC programs (to S.H. and G.C.C.), and the T. W. Bennett Distinguished Professorship Fund (to G.C.C.).

We thank S. W. Chisholm and B. Brahamsha for providing *Prochlorococcus* and *Synechococcus* genomic DNA.

REFERENCES

1. **Badger, M. R., D. Hanson, and G. D. Price.** 2002. Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. Funct. Plant Biol. **29:**161–173.
2. **Baker, S. H., S. Jin, H. C. Aldrich, G. T. Howard, and J. M. Shively.** 1998. Insertion mutation of the form I *cbbL* gene encoding ribulose bisphosphate carboxylase/oxygenase (RuBisCO) in *Thiobacillus neapolitanus* results in expression of form II RuBisCO, loss of carboxysomes, and an increased CO_2 requirement for growth. J. Bacteriol. **180:**4133–4139.
3. **Baker, S. H., S. C. Lorbach, M. Rodriguez-Buey, D. S. Williams, H. C. Aldrich, and J. M. Shively.** 1999. The correlation of the gene *csoS2* of the carboxysome operon with two polypeptides of the carboxysome in *Thiobacillus neapolitanus.* Arch. Microbiol. **172:**233–239.
4. **Baker, S. H., D. S. Williams, H. C. Aldrich, A. C. Gambrell, and J. M. Shively.** 2000. Identification and localization of the carboxysome peptide Csos3 and its corresponding gene in *Thiobacillus neapolitanus.* Arch. Microbiol. **173:**278–283.
5. **Beudeker, R. F., G. C. Cannon, J. G. Kuenen, and J. M. Shively.** 1980. Relations between D-ribulose-1,5-bisphosphate carboxylase, carboxysomes, and CO_2 fixing capacity in the obligate chemolithotroph *Thiobacillus neapolitanus* grown under different limitations in the chemostat. Arch. Microbiol. **124:**185–189.
6. **Cannon, G. C., S. H. Baker, F. Soyer, D. R. Johnson, C. E. Bradburne, J. L. Mehlman, P. S. Davies, Q. L. Jiang, S. Heinhorst, and J. M. Shively.** 2003. Organization of carboxysome genes in the thiobacilli. Curr. Microbiol. **46:** 115–119.
7. **Cannon, G. C., C. E. Bradburne, H. C. Aldrich, S. H. Baker, S. Heinhorst, and J. M. Shively.** 2001. Microcompartments in prokaryotes: carboxysomes and related polyhedra. Appl. Environ. Microbiol. **67:**5351–5361.
8. **Cannon, G. C., S. Heinhorst, C. E. Bradburne, and J. M. Shively.** 2002. Carboxysome genomics: a status report. Funct. Plant Biol. **29:**175–182.
9. **Cannon, G. C., and J. M. Shively.** 1983. Characterization of a homogeneous preparation of carboxysomes from *Thiobacillus neapolitanus.* Arch. Microbiol. **134:**52–59.
10. **Dufresne, A., M. Salanobaut, F. Partensky, F. Artiguenave, I. M. Axmann, V. Barbe, S. Duprat, M. Y. Galperin, E. V. Koonin, F. Le Gall, K. S. Makarova, M. Ostrowski, S. Oztas, C. Robert, I. B. Rogozin, D. J. Scanlan, N. Tandeau de Marsac, J. Weissenbach, P. Wincker, Y. I. Wolf, and W. R. Hess.** 2003. Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome. Proc. Natl. Acad. Sci. USA **100:** 10020–10025.
11. **English, R. S., S. C. Lorbach, X. Qin, and J. M. Shively.** 1994. Isolation and characterization of a carboxysome shell gene from *Thiobacillus neapolitanus.* Mol. Microbiol. **12:**647–654.

11a. **Fisher, M., I. Gokhman, U. Pick, and A. Zamir.** A salt-resistant plasma membrane carbonic anhydrase is induced by salt in *Dunaliella salina*. J. Biol. Chem. **271:**17718–17723.

12. **Havemann, G. D., E. M. Sampson, and T. A. Bobik.** 2002. PduA is a shell protein of polyhedral organelles involved in coenzyme B_{12}-dependent degradation of 1,2-propanediol in *Salmonella enterica* serovar Typhimurium LT2. J. Bacteriol. **184:**1253–1261.
13. **Hewitt-Emmett, D., and R. E. Tashian.** 1996. Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families. Mol. Phylogenet. Evol. **5:**50–77.
14. **Holthuijzen, Y. A., E. F. M. VanDissel-Emiliani, J. G. Kuenen, and W. N. Konings.** 1987. Energetic aspects of CO_2 uptake in *Thiobacillus neapolitanus.* Arch. Microbiol. **147:**285–290.
15. **Kaplan, A., and L. Reinhold.** 1999. CO_2 concentrating mechanisms in photosynthetic microorganisms. Annu. Rev. Plant Physiol. Plant Mol. Biol. **50:** 539–570.
16. **Kofoid, E., C. Rappleye, I. Stojiljkovic, and J. Roth.** 1999. The 17-gene ethanolamine (*eut*) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. J. Bacteriol. **181:**5317–5329.
17. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:**1244–1245.
18. **Lane, T. W., and F. M. M. Morel.** 2000. Regulation of carbonic anhydrase expression by zinc, cobalt, and carbon dioxide in the marine diatom *Thalassiosira weissflogii.* Plant Physiol. **123:**345–352.
19. **McKay, R. M. L., S. P. Gibbs, and G. S. Espie.** 1993. Effect of dissolved inorganic carbon on the expression of carboxysomes, localization of Rubisco and the mode of carbon transport in cells of the cyanobacterium *Synechococcus* UTEX 625. Arch. Microbiol. **159:**21–29.
20. **Miller, A. G., C. Salon, G. S. Espie, and D. T. Canvin.** 1997. Measurement of the amount and isotopic composition of the CO_2 released from the cyanobacterium *Synechococcus* UTEX 625 after rapid quenching of the active CO_2 transport system. Can. J. Bot. **75:**981–997.

20a. **Mitsuhashi, S., and S. Miyachi.** 1996. Amino acid sequence homology between N- and C-terminal halves of a carbonic anhydrase in *Porphyridium purpureum*, as deduced from the cloned cDNA. J. Biol. Chem. **271:**28703–28709.

21. **Partensky, F., W. R. Hess, and D. Vaulot.** 1999. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. Rev. **63:**106–127.
22. **Price, G. D., J. R. Coleman, and M. R. Badger.** 1992. Association of carbonic anhydrase activity with carboxysomes isolated from the cyanobacterium *Synechococcus* PCC7942. Plant Physiol. **100:**784–793.
23. **Price, G. D., D. Sultemeyer, B. Klughammer, M. Ludwig, and M. R. Badger.** 1998. The functioning of the CO_2 concentrating mechanism in several cyanobacterial strains: a review of general physiological characteristics, genes, proteins and recent advances. Can. J. Bot. **76:**973–1002.
24. **Reinhold, L., M. Zviman, and A. Kaplan.** 1987. Inorganic carbon fluxes and photosynthesis in cyanobacteria: a quantitative model, p. 6.289–6.296. *In* J. Biggens (ed.), Progress in photosynthetic research, vol. 4. Martinus Nijhoff, Dordrecht, The Netherlands.
25. **Reinhold, L., M. Zviman, and A. Kaplan.** 1989. A quantitative model for carbon fluxes and photosynthesis in cyanobacteria. Plant Physiol. Biochem. **27:**945–954.
26. **Rocap, G., F. W. Larimer, J. Lamerdin, S. Malfatti, P. Chain, N. A. Ahlgren, A. Arellano, M. Coleman, L. Hauser, W. R. Hess, Z. I. Johnson, M. Land, D. Lindell, A. F. Post, W. Regala, M. Shah, S. L. Shaw, C. Steglich, M. B. Sullivan, C. S. Ting, A. Tolonen, E. A. Webb, E. R. Zinser, and S. W. Chisholm.** 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature **424:**1042–1047. (First published 13 August 2003; 10.1038/nature01947.)
27. **Shively, J. M., F. Ball, D. H. Brown, and R. E. Saunders.** 1973. Functional

organelles in prokaryotes: polyhedral inclusions (carboxysomes) in *Thiobacillus neapolitanus.* Science **182:**584–586.
28. **Shively, J. M., F. L. Ball, and B. W. Kline.** 1973. Electron microscopy of the carboxysomes (polyhedral bodies) of *Thiobacillus neapolitanus.* J. Bacteriol. **116:**1405–1411.
29. **Shively, J. M., S. C. Lorbach, S. Jin, and S. H. Baker.** 1996. Carboxysomes: the genes of *Thiobacillus neapolitanus*, p. 56–63. *In* M. E. Lidstrom and F. R. Tabita (ed.), Microbial growth on C_1 compounds. Kluwer, Dordrecht, The Netherlands.
30. **Smith, K. S., and J. G. Ferry.** 2000. Prokaryotic carbonic anhydrases. FEMS Microbiol. Rev. **24:**335–366.
31. **So, A. K.-C., and G. S. Espie.** 1998. Cloning, characterization and expression of carbonic anhydrase from the cyanobacterium *Synechocystis* PCC6803. Plant Mol. Biol. **37:**205–215.
32. **So, A. K.-C., M. John-McKay, and G. S. Espie.** 2002. Characterization of a mutant lacking carboxysomal carbonic anhydrase from the cyanobacterium *Synechocystis* PCC6803. Planta **214:**456–467.
33. **So, A. K.-C., H. G. C. Van Spall, J. R. Coleman, and G. S. Espie.** 1998. Catalytic exchange of ^{18}O from $^{13}C^{18}O$-labelled CO_2 by wild-type cells and *ecaA*, *ecaB*, and *ccaA* mutants of the cyanobacteria *Synechococcus* PCC7942 and *Synechocystis* PCC6803. Can. J. Bot. **76:**1153–1160.
34. **Tabita, F. R.** 1999. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a different perspective. Photosynth. Res. **60:**1–28.
35. **Tripp, B. C., K. S. Smith, and J. G. Ferry.** 2001. Carbonic anhydrase: new insights for an ancient enzyme. J. Biol. Chem. **276:**48615–48618.
36. **Wilbur, K. M., and N. G. Anderson.** 1948. Electrometric and colorimetric determination of carbonic anhydrase. J. Biol. Chem. **176:**147–154.
37. **Yu, J.-W., G. D. Price, L. Song, and M. R. Badger.** 1992. Isolation of a putative carboxysomal carbonic anhydrase gene from the cyanobacterium *Synechococcus* PCC7942. Plant Physiol. **100:**794–800.