

3-17-2006

The Structure of β -Carbonic Anhydrase from the Carboxysomal Shell Reveals a Distinct Subclass with One Active Site for the Price of Two

Michael R. Sawaya
University of California, Los Angeles

Gordon C. Cannon
University of Southern Mississippi, Gordon.Cannon@usm.edu

Sabine Heinhorst
University of Southern Mississippi, sabine.heinhorst@usm.edu

Shiho Tanaka
University of California, Los Angeles

Eric B. Williams
University of Southern Mississippi

See next page for additional authors

Follow this and additional works at: https://aquila.usm.edu/fac_pubs

 Part of the [Chemistry Commons](#)

Recommended Citation

Sawaya, M. R., Cannon, G. C., Heinhorst, S., Tanaka, S., Williams, E. B., Yeates, T. O., Kerfeld, C. A. (2006). The Structure of β -Carbonic Anhydrase from the Carboxysomal Shell Reveals a Distinct Subclass with One Active Site for the Price of Two. *Journal of Biological Chemistry*, 281(11), 7546-7555. Available at: https://aquila.usm.edu/fac_pubs/9056

This Article is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Faculty Publications by an authorized administrator of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.

Authors

Michael R. Sawaya, Gordon C. Cannon, Sabine Heinhorst, Shiho Tanaka, Eric B. Williams, Todd O. Yeates, and Cheryl A. Kerfeld

The Structure of β -Carbonic Anhydrase from the Carboxysomal Shell Reveals a Distinct Subclass with One Active Site for the Price of Two*

Received for publication, September 23, 2005, and in revised form, December 12, 2005. Published, JBC Papers in Press, January 10, 2006, DOI 10.1074/jbc.M510464200

Michael R. Sawaya^{†1}, Gordon C. Cannon[§], Sabine Heinhorst[§], Shiho Tanaka[‡], Eric B. Williams[§], Todd O. Yeates^{‡¶1}, and Cheryl A. Kerfeld^{¶||**2}

From the [†]Molecular Biology Institute, UCLA, Los Angeles, California 90095-1570, the [§]Department of Chemistry and Biochemistry, University of Southern Mississippi, Hattiesburg, Mississippi 39406-0001, the [¶]Department of Chemistry and Biochemistry, UCLA, Los Angeles, California 90095-1569, the ^{||}UCLA-Department of Energy (DOE) Institute for Genomics and Proteomics, Los Angeles, California 90095-1570, and the ^{**}Life Sciences Core, UCLA, Los Angeles, California 90095-1606

CsoSCA (formerly CsoS3) is a bacterial carbonic anhydrase localized in the shell of a cellular microcompartment called the carboxysome, where it converts HCO_3^- to CO_2 for use in carbon fixation by ribulose-bisphosphate carboxylase/oxygenase (RuBisCO). CsoSCA lacks significant sequence similarity to any of the four known classes of carbonic anhydrase (α , β , γ , or δ), and so it was initially classified as belonging to a new class, ϵ . The crystal structure of CsoSCA from *Halothiobacillus neapolitanus* reveals that it is actually a representative member of a new subclass of β -carbonic anhydrases, distinguished by a lack of active site pairing. Whereas a typical β -carbonic anhydrase maintains a pair of active sites organized within a two-fold symmetric homodimer or pair of fused, homologous domains, the two domains in CsoSCA have diverged to the point that only one domain in the pair retains a viable active site. We suggest that this defunct and somewhat diminished domain has evolved a new function, specific to its carboxysomal environment. Despite the level of sequence divergence that separates CsoSCA from the other two subclasses of β -carbonic anhydrases, there is a remarkable level of structural similarity among active site regions, which suggests a common catalytic mechanism for the interconversion of HCO_3^- and CO_2 . Crystal packing analysis suggests that CsoSCA exists within the carboxysome shell either as a homodimer or as extended filaments.

The efficiency of carbon fixation in cyanobacteria and many autotrophic bacteria is enhanced by a polyhedral microcompartment called the carboxysome. The carboxysome consists of a paracrystalline array of the carbon-fixing enzyme ribulose-bisphosphate carboxylase/oxygenase (RuBisCO)³ enveloped by a proteinaceous shell composed of at least five polypeptides (reviewed in Refs. 1–5). The structures of some of these shell proteins have been recently reported (6). It is proposed that car-

boxysomes enhance carbon fixation by elevating $[\text{CO}_2]$ up to 1000-fold around the active site of RuBisCO, effectively saturating it with substrate. The rate enhancement afforded by this mechanism is made evident when one considers the remarkably low affinity RuBisCO has for its substrate, CO_2 ($K_m > 150 \mu\text{M}$). The mechanism by which CO_2 is concentrated in the carboxysome is not clearly understood. It is proposed that HCO_3^- , which is actively transported and accumulated in the cytoplasm, diffuses through the proteinaceous shell and is dehydrated to CO_2 by a carboxysome-associated carbonic anhydrase. The close proximity of the carbonic anhydrase to encapsulated RuBisCO is thought to be a key factor in speeding the fixation of CO_2 . It is also possible that the carboxysomal shell provides a barrier to the escape of CO_2 from RuBisCO.

Carbonic anhydrases have been found associated with carboxysomes of the two known types: the ccm type, which appears to be limited to cyanobacteria, and the cso type, which occurs in non-photosynthetic autotrophs and in some cyanobacteria (2, 7). There are four known classes of carbonic anhydrases: α , β , γ , and δ . In the ccm-type carboxysomes, both β and γ classes were discovered. These are called CcaA and CcmM, respectively. CcaA activity co-purifies with carboxysome-enriched subcellular fractions (1–3, 8–10), and CcmM was found necessary for carboxysome formation (11). In cso-type carboxysomes, however, sequence comparisons could not discern a gene that might encode any of the four known carbonic anhydrase classes. Classical biochemical approaches revealed carbonic anhydrase activity in the carboxysomal shell protein CsoS3 (which we rename here CsoSCA, for carboxysome shell carbonic anhydrase) isolated from homogeneous carboxysome preparations from *Halothiobacillus neapolitanus* or prepared by recombinant techniques (12).

Contrary to earlier predictions that placed the carboxysome-associated carbonic anhydrase into the interior of the microcompartment, CsoSCA was shown to be an integral component of the shell (12, 13). The low abundance of the protein as compared with other shell proteins and the fact that a *H. neapolitanus* mutant carrying an in-frame kanamycin resistance cassette in the *csoS3* gene is able to form carboxysomes⁴ imply that CsoSCA is not a major structural determinant of the carboxysome shell. The protein is, however, crucial for carboxysome function since this *csoS3* mutant has a HCR (high CO_2 requiring) phenotype.

Because of its large size (57 kDa, approximately twice that of other carbonic anhydrases) and its lack of homology at the primary structure level, CsoSCA was assigned to a new (ϵ) class of carbonic anhydrase (12).

* This work was supported in part by grants from the U.S. Department of Agriculture, the National Institutes of Health, and the Biological and Environmental Research Program of United States Department of Energy Office of Science (to C. A. K and T. O. Y.), Grant MCB-0444568 from the National Science Foundation (NSF) (to G. C. C.), and Grant DMR 0213883 from the NSF Materials Research Science and Engineering Centers program (to G. C. C. and S. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2FGY) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¹ Supported by Howard Hughes Medical Institute.

² To whom correspondence should be addressed: UCLA-DOE Institute for Genomics and Proteomics, Box 951570, Los Angeles, CA 90095-1570. Tel.: 310-825-7417; Fax: 310-206-3914; E-mail: kerfeld@mbi.ucla.edu.

³ The abbreviations used are: RuBisCO, ribulose-bisphosphate carboxylase/oxygenase; r.m.s., root mean square.

⁴ G. C. Cannon, S. Dawn, E. B. Williams, J. M. Shively, and S. Heinhorst, manuscript in preparation.

TABLE 1

Data collection and refinement statistics

PIP, di-*m*-iodobis(ethylendiamine)diplatinum; PDB, Protein Data Bank; r.m.s.d., r.m.s. deviation.

Data collection	Native	HgCl ₂	Mersalyl	PIP	KI
Space group	P1	P1	P1	P1	P1
<i>a</i>	47.7	48.1	47.6	48.7	47.7
<i>b</i>	77.6	77.7	77.8	77.6	78.0
<i>c</i>	82.7	82.5	82.6	82.0	82.2
α	66.8	66.7	67.0	66.8	67.4
β	78.8	78.8	78.7	79.0	79.0
γ	79.2	79.0	79.2	78.2	79.4
Wavelength (Å)	1.0332	1.5418	1.5418	1.5418	1.5418
Resolution limit (Å)	2.2	3.3	3.0	2.9	3.9
R_{sym} (%) ^a	12.3	16.9	14.4	7.6	15.0
R_{sym} (%) (last shell)	43.0	47.6	50.3	29.3	34.6
I/σ (last shell)	2.5	2.4	2.4	3.8	3.2
Total observations	302,721	55,456	76,256	86,727	35,313
Unique reflections	52,789	15,845	21,490	23,643	9,567
Completeness (%)	98.0	98.0	97.7	97.4	98.3
Completeness (%) (last shell)	94.7	97.5	96.5	96.3	97.7
Phase determination					
R_{Cullis} ^b (% acentric, isomorphous)		81	83	94	92
R_{Cullis} ^c (% acentric, anomalous)	99	100	100	99	100
Phasing power ^d (acentric)		1.13	1.05	0.56	0.75
Number of sites	4	4	4	4	6
Mean overall figure of merit (before/after DM)			0.166/0.614		
Model refinement	Native				
R_{work} ^e (20 Å–2.2 Å)	17.1				
R_{free} ^f (20 Å–2.2 Å)	21.7				
PDB ID code	2FGY				
Number of residues (protein/water)	944/575				
Average B (main chain/side chain) (Å ²)	33.8/35.2				
Average B (waters) (Å ²)	40.7				
r.m.s.d. bonds (Å)	0.010				
r.m.s.d. angles (°)	2.0				
B-values (Å ² bonded)	1.8				

$$^a R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I.$$

$$^b R_{\text{Cullis}} = \sum \epsilon / \sum |F_{\text{PH}} - F_{\text{P}}|, \text{ where } \epsilon = \text{lack of closure.}$$

$$^c R_{\text{Cullis}} = \sum \epsilon / \sum |F^+ - F^-|, \text{ where } \epsilon = \text{lack of closure.}$$

$$^d \text{Phasing power} = \langle F_{\text{H}} / \epsilon \rangle.$$

$$^e R_{\text{work}} = \sum |F_{\text{o}} - F_{\text{c}}| / \sum F_{\text{o}}.$$

$$^f R_{\text{free}} = \sum |F_{\text{o}} - F_{\text{c}}| / \sum F_{\text{o}}, \text{ where all reflections belong to a test set of 5\% randomly selected data.}$$

Here, we describe the 2.2 Å crystal structure of CsoSCA from the chemoaototrophic bacterium *H. neapolitanus*, which suggests that the enzyme is a variant β -type carbonic anhydrase. The structure also provides clues to the catalytic mechanism of the dehydration reaction and reveals quaternary structure assemblies that may be biologically relevant in the context of carboxysome function.

EXPERIMENTAL PROCEDURES

Cloning, Protein Expression, Purification, and Crystallization—Cloning, overexpression, and purification of recombinant CsoSCA were performed as described for the pcsoS3ProEx-based system (12). For crystallization, purified CsoSCA with the uncleaved His tag was concentrated to 1.2 mg/ml in 10 mM Tris, pH 8.0. Crystals were obtained at room temperature by hanging drop (containing 1 μ l of protein plus 1 μ l of reservoir) vapor diffusion over a reservoir containing 0.1 M HEPES, pH 7.0, 0.2 M MgCl₂, and 10% ethanol. The crystals indexed in space group P1 ($a = 47.7$ Å, $b = 77.6$ Å, $c = 82.7$ Å, $\alpha = 66.8^\circ$, $\beta = 78.8^\circ$, $\gamma = 79.1^\circ$) with two molecules in the asymmetric unit.

Size Exclusion Chromatography—Purified recombinant CsoSCA (~1 mg/ml) was loaded at 1 ml/min over a Sephacryl S300 column (1 \times 16 cm, Amersham Biosciences) equilibrated in 10 mM Tris, pH 8.0, 50 mM KCl. Two peaks were observed, monitored by absorbance at 280-nm wavelength. The first peak (consisting of ~10% of the loaded protein) likely corresponds to aggregated species, lying completely within the exclusion volume ($M_r > 1.5 \times 10^6$). The second peak (having a symmetrical shape) corresponds to a dimer, eluting with an apparent M_r of 125,000. Measurements of carbonic anhydrase activity indicated

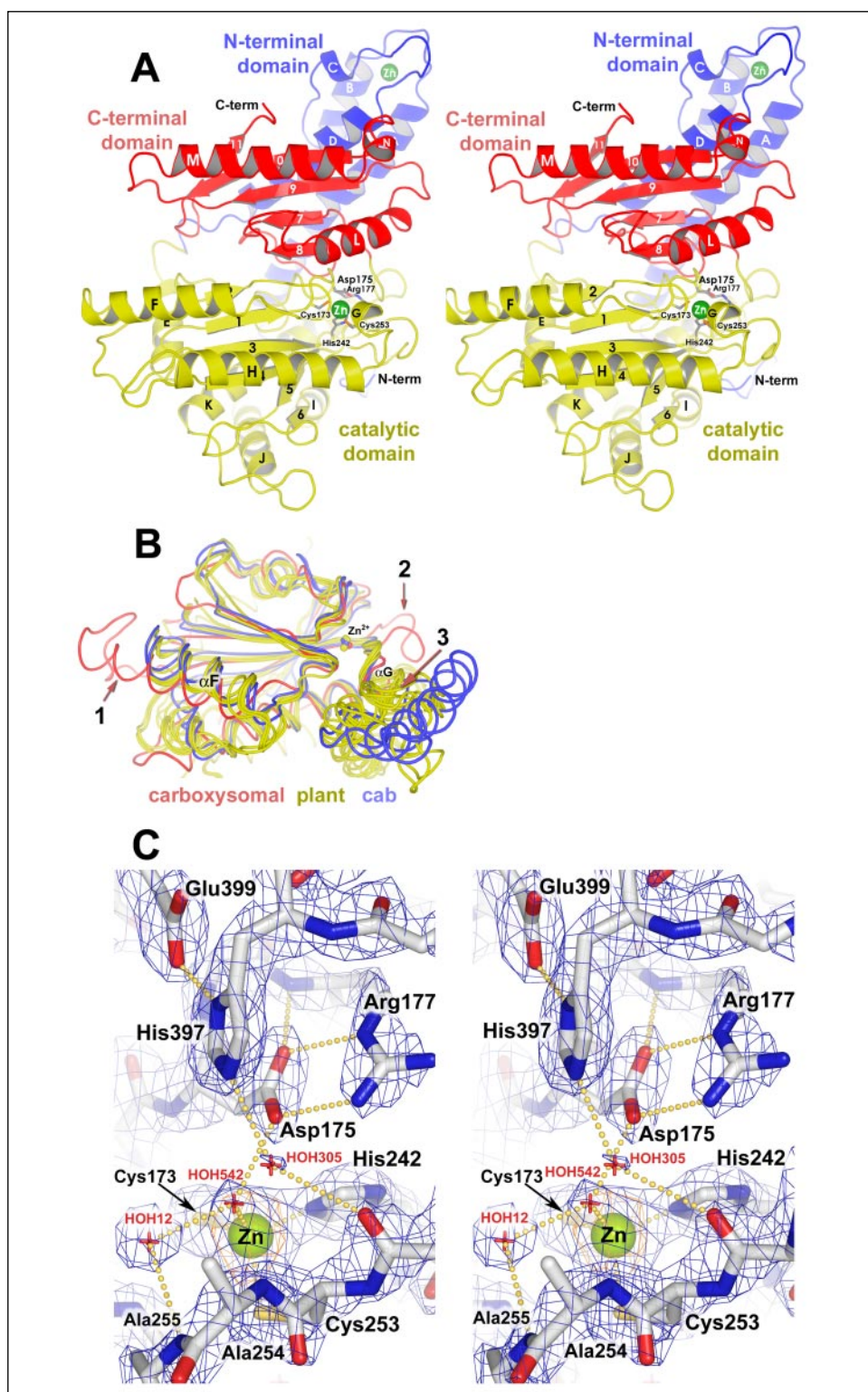
that the dimer peak was active, and no other carbonic anhydrase activity was detected throughout the elution pattern. It should be noted that solutions of CsoSCA with concentrations above 1.5 mg/ml accumulated precipitated aggregates when held at 4 °C over several days, and this aggregation roughly paralleled the loss of carbonic anhydrase activity, consistent with the observation that no activity was observed in the large molecular weight fractions eluted from the gel filtration column.

Structure Determination and Refinement—Potential heavy atom derivatives were screened for binding to CsoSCA using the native gel shift approach of Boggon and Shapiro (15). The following heavy atom compounds were identified as promising candidates for derivatization: HgCl₂, mersalyl, and di-*m*-iodobis(ethylendiamine)diplatinum. The compounds were dissolved at 2 mM concentration in a mixture of 65% reservoir, 35% ethylene glycol for cryo-protection. Crystals were soaked at room temperature for 1 h and then flash-cooled and stored in liquid nitrogen until data collection. An iodide derivative was also prepared by dissolving 8 mg of solid KI in 100 μ l of 65% reservoir, 35% ethylene glycol, and then soaking the crystal for 1 min in this solution.

The x-ray diffraction data set on the native crystal was collected at the Advance Light Source beamline 8.2.2, equipped with an ADSC quantum 315 CCD detector. Derivative data sets were collected on a Rigaku FR-D rotating copper anode x-ray generator equipped with an R-axis IV⁺⁺ imaging plate detector. Data were processed using DENZO/SCALEPACK (16) (Table 1). Mercury sites were identified by SHELXD using isomorphous differences between the native and HgCl₂ derivative data sets (17). Initial phases were calculated with MLPHARE and later improved by density modification and two-fold symmetry averaging

Structure of β -Carbonic Anhydrase from the Carboxysomal Shell

FIGURE 1. Ribbon diagrams of CsoSCA and β -carbonic anhydrase homologs. *A*, a stereo view of CsoSCA highlights its three domains: the N-terminal (*N-term*), catalytic, and C-terminal (*C-term*) domains. The catalytic domain (*yellow*) bears 5 residues that are strictly conserved among all β -carbonic anhydrases known to date (shown in *sticks*). They cluster around the catalytic zinc ion (*green sphere*). The C-terminal domain (*red*) is structurally similar to the catalytic domain but is smaller and lacks all the necessary zinc ligand residues; thus, no zinc is bound to it. The N-terminal domain (*blue*) bears a zinc ion, but this is thought to be non-physiological. *B*, structural superposition reveals a common fold among the three subclasses of β -carbonic anhydrases: plant (*yellow*, *P. purpureum*; *E. coli*; and *M. tuberculosis*, Rv3588c), cab (*blue*, *M. thermoautotrophicum* and *M. tuberculosis*, Rv1284), and carboxysomal (*red*, CsoSCA). The plant and cab subclasses are much more similar to each other than they are to CsoSCA. Numbered in the figure are three points of major structural difference between CsoSCA and the other two subclasses. The first region is an extension of helix F and its following loop (16-residue insertion). The second region is an insertion of an 8-residue loop preceding helix G. The third region corresponds to a severe foreshortening of helices G and H and intervening helices (30-residue deletion). *C*, the *blue mesh* corresponds to a $2F_o - F_c$ map of the catalytic site, contoured at 1.8 σ . The *orange mesh* centered on the zinc ion corresponds to anomalous difference Fourier density contoured at 8 σ . This stereo view of active site is similar to Fig. 4.



with DM (18). Isomorphous and anomalous difference Fourier methods revealed the positions of the heavy atoms in the other derivatives and the location of four zinc ions in the unit cell. Phasing information from all the derivatives was combined. The model was built manually using the graphics program O (19) and refined using conjugate gradient and simulated annealing algorithms as implemented by the program CNS (20). Strong non-crystallographic symmetry restraints were used throughout. Hydrogen bond restraints were helpful in the early stages of refinement

(21). This model was further refined with REFMAC (22), to introduce TLS parameters in the refinement. Data collection and refinement statistics are given in Table 1.

The geometric quality of the model was assessed with the following structure validation tools: ERRAT (23), PROCHECK (24), and WHATIF (25). PROCHECK reported that 92% of the residues fall in the most favored region of the Ramachandran plot and 8% of the residues were in additionally allowed regions. ERRAT reported an overall quality factor

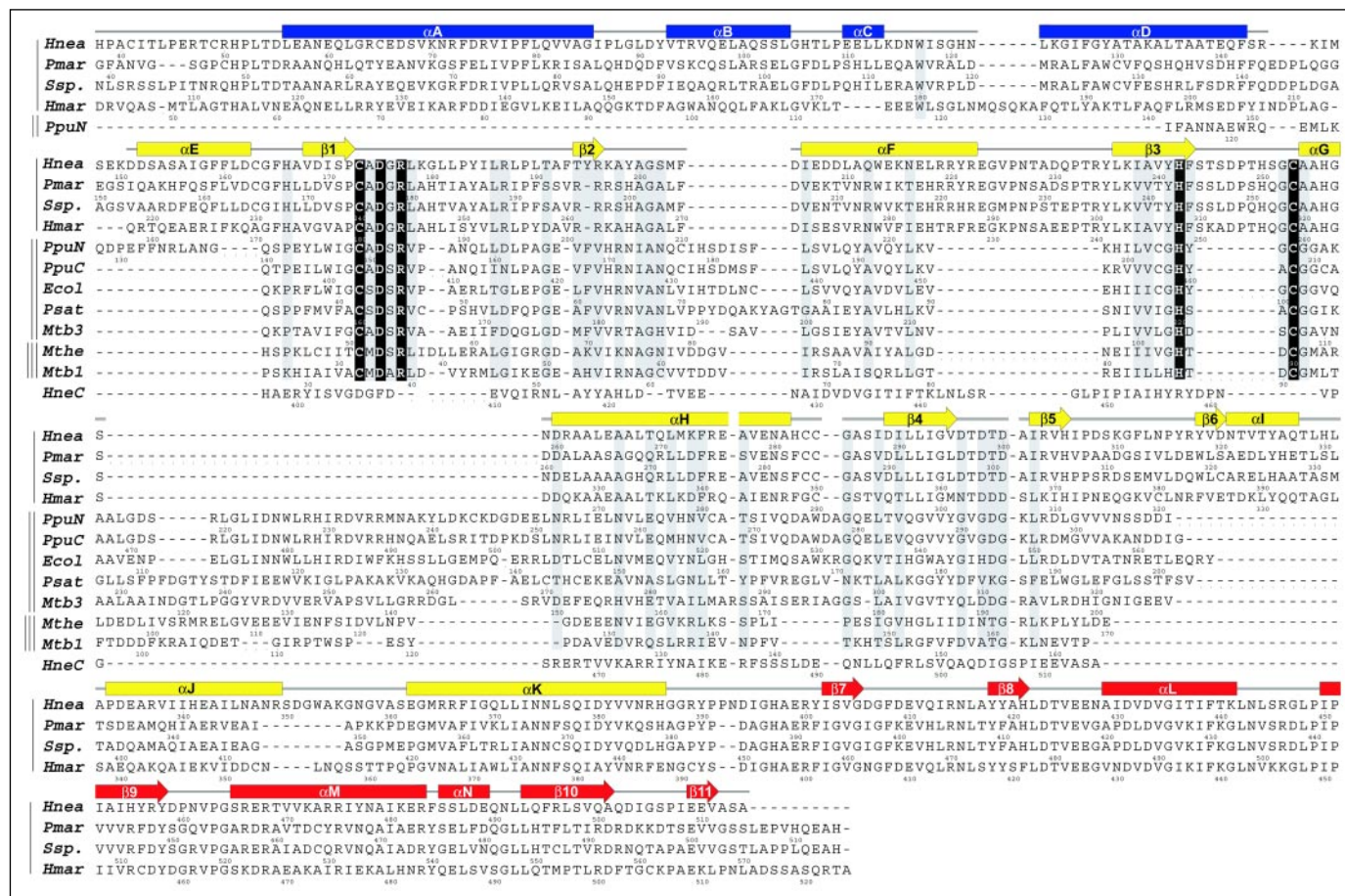


FIGURE 2. Structure based sequence alignment of CsoSCA and β -carbonic anhydrase homologs. Represented species are divided into three subclasses. A single vertical line in the left column signifies the carboxysomal subclass. A double vertical line specifies the plant subclass. A triple vertical line specifies the cab subclass. Species abbreviations are: *Hnea*, *H. neapolitanus*; *Pmar*, *Prochlorococcus marinus* str. MIT 9313; *Ssp*, *Synechococcus* sp. WH 8102; *Hmar*, *Hydrogenovibrio marinus*; *PpuN*, *Porphyridium purpureum* N-terminal domain; *PpuC*, *P. purpureum* C-terminal domain; *Ecol*, *E. coli*; *Psat*, *P. sativum*; *Mtb3*, *M. tuberculosis* Rv3588c; *Mthe*, *M. thermoautotrophicum*; *Mtb1*, *M. tuberculosis* Rv1284; *HneC*, *H. neapolitanus* C-terminal domain. Secondary structure assignments are color-coded to correspond to Fig. 1A.

of 96%. Protein structures were illustrated using the program PyMOL (26). Atomic coordinates and structure factors have been deposited in the Protein Data Bank, with ID code 2FGY.

RESULTS AND DISCUSSION

Domain Structure—CsoSCA from *H. neapolitanus* contains three domains (Fig. 1A): a novel N-terminal domain (residues 38–144) composed primarily of four α -helices; a catalytic (middle) domain (residues 151–397) bearing structural resemblance to the carbonic anhydrases of the β -class (Figs. 1B and 2) complete with a tightly bound zinc ion (B-factor of 35 \AA^2 at full occupancy) (Fig. 1C); and a C-terminal domain (residues 398–514) with weak structural similarity to the catalytic domain but, unlike the middle domain, completely devoid of catalytic metal ions. Only the N-terminal residues, 1–37, and residues 145–150 (connecting the N-terminal and catalytic domains) are missing from the structural model, presumably due to disorder.

N-terminal Domain—The structure of the N-terminal domain can be classified as an up-down four-helix bundle, but the individual helices are of disparate lengths (helix A, 29 residues; helix B, 12 residues; helix C, 4 residues; and helix D, 20 residues). As such, the domain bears no significant structural resemblance to the numerous helical bundles deposited in the Protein Data Bank. The closest structural similarity uncovered by VAST (27) is with a putative signal recognition particle (Protein Data Bank ID 1WGW) in which 43 residues align with an r.m.s. deviation of 2.4 \AA . An evolutionary relationship between the two could

not be established with any confidence. Related β -carbonic anhydrase structures from *Pisum sativum* (28) and *Escherichia coli* (29) also have N-terminal helical extensions in which the helices strengthen homodimer interactions by domain swapping. In CsoSCA, the lack of contacts between the N-terminal domain and other symmetry-related molecules seems to preclude a role in oligomerization. Furthermore, the N-terminal domain is located at an isolated extremity of the protein, suggesting the possibility that it might function to anchor CsoSCA to the carboxysomal shell or to molecules of RuBisCO.

Catalytic and C-terminal Domains Related by Gene Duplication and Fusion—The catalytic and C-terminal domains possess structural similarity to each other (Fig. 3C), suggesting that they arose from a gene duplication event. Sixty-one $C\alpha$ atoms from the catalytic domain can be superimposed on the C-terminal domain with an r.m.s. deviation of 1.5 \AA . The two domains each contain a central five-stranded mixed β -sheet with two helices flanking one side of the sheet and one helix flanking the other side. Interestingly, this latter helix, corresponding to helix K in the catalytic domain, is structurally analogous to helix D, which although tightly associated with the C-terminal domain, originates from the N-terminal part of the primary structure (Fig. 1A). The catalytic and C-terminal domains are related by an approximate two-fold symmetry axis perpendicular to the β -sheet (Fig. 3C).

The catalytic domain of CsoSCA bears structural similarity to numerous other carbonic anhydrases of the β -class (Fig. 1B). It can be

Structure of β -Carbonic Anhydrase from the Carboxysomal Shell

FIGURE 3. The loss of active site pairing in carboxysomal β -carbonic anhydrases. In *P. sativum* (A, upper panel), active site pairing is accomplished through homodimerization. Molecule A of the dimer is shown in yellow; molecule B is shown in red. Green spheres mark the location of the two identical active site zinc ions. An ellipse marks the location of a two-fold symmetry axis. Superimposed molecules A and B are shown in C, lower panel. Only the very C-terminal (C-term) tails are not superimposable (colored gray). In *P. purpureum* (B, upper panel), the same pairing is accomplished by a single polypeptide, not a dimer. The N-terminal (N-term) domain is shown in yellow; the C-terminal domain is shown in red. Pseudo-two-fold symmetry is still evident but is not exact (B, lower panel). Non-superimposable regions are shown in gray. Presumably, the *P. purpureum* carbonic anhydrase arose from gene duplication and fusion. Divergence appears minimal (70% sequence identity between domains). In A, upper panel, CsoSCA, like the *P. purpureum* enzyme appears to be the result of gene duplication and fusion. However, divergence between the two internal domains has progressed to the extent that the C-terminal domain has lost all active site residues that it presumably once contained. Superimposed catalytic and C-terminal domains (C, lower panel) show much larger areas of structural non-equivalence (gray).

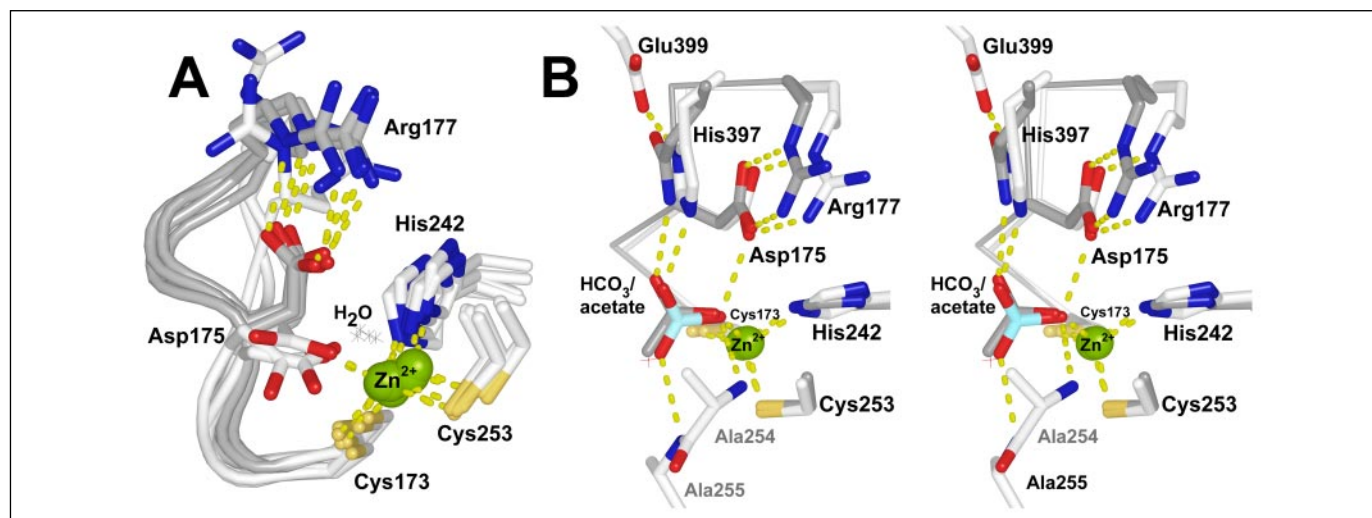
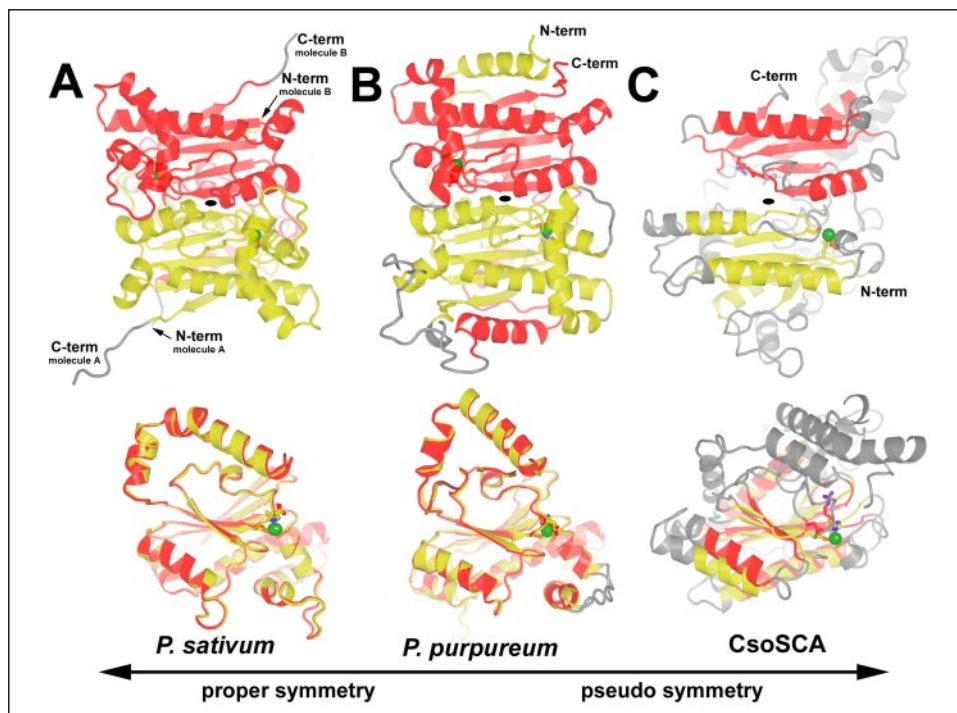


FIGURE 4. Conservation of active site structure and mechanism. A, currently available structures of β -carbonic anhydrases can be divided into two groups: those in which the conserved aspartate and arginine residues are hydrogen-bonded to each other (CsoSCA, *P. sativum*, *M. thermoautotrophicum*, and Rv1284) and those in which the hydrogen bonds are broken (*P. purpureum*, *E. coli*, and Rv3588c). These are represented in dark gray and light gray, respectively. In the former group, a water molecule serves as the fourth ligand to the zinc ion. In the latter group, the aspartate plays this role. The duality of conformations suggests a conformational flexibility that might play a role in catalysis. Residue numbering corresponds to CsoSCA. B, a stereo figure showing a superpositioning of CsoSCA (light gray) and *P. sativum* (dark gray) enzymes. A HCO_3^- ion was modeled into the active site, based on similarities with the position of the acetate ion found in the *P. sativum* structure. Small changes in orientation of the HCO_3^- allowed hydrogen bonds to form between the HCO_3^- and zinc, His-397, Asp-175, and backbone nitrogens of Ala-254 and Ala-255.

superimposed on carbonic anhydrases from *P. sativum* (28), *E. coli* (29), *Methanobacterium thermoautotrophicum* (30), and *Mycobacterium tuberculosis* (31) (Rv3588c is a *bona fide* carbonic anhydrase, and Rv1284 is a carbonic anhydrase-like protein) with the following r.m.s. deviations (*P. sativum*, 2.0 Å for 90 α C pairs; *E. coli*, 1.6 Å for 83 α C pairs; *M. thermoautotrophicum*, 1.5 Å for 85 α C pairs; *M. tuberculosis* Rv3588c, 1.4 Å for 70 α C pairs; *M. tuberculosis* Rv1284, 1.6 Å for 81 α C pairs). These structures all contain a central five-stranded β -sheet and at least three flanking helices. Each of these homologs is approximately half the mass of CsoSCA, or roughly the same mass as an isolated catalytic domain of CsoSCA.

In addition to the similarity in fold, the identity and positions of three zinc ligands are conserved in the active sites of all these β -carbonic anhydrases (Figs. 2 and 4A). Specifically, the zinc ion directly coordinates Cys-173, His-242, and Cys-253 in CsoSCA (Figs. 1C and 4B). An aspartic acid and an arginine are also structurally conserved in β -carbonic anhydrases (Asp-175 and Arg-177 in CsoSCA); these residues presumably assist in a number of catalytic steps, including substrate binding, proton shuffling, and product release.

The similarity between CsoSCA and other β -carbonic anhydrases extends beyond the active site. The homologs cited above all dimerize with a two-fold symmetry axis that coincides with the pseudo-symme-

try axis relating the catalytic and C-terminal domains of CsoSCA (Fig. 3). Apparently, the catalytic and C-terminal domains of CsoSCA have maintained their relative orientations following gene duplication and fusion, despite much subsequent divergence.

The carbonic anhydrase from *Porphyridium purpureum* bears special resemblance to CsoSCA in that it appears to have undergone a similar sequence of events in which two carbonic anhydrase domains are fused and arranged about a central two-fold axis of pseudo-symmetry (32) (Fig. 3B). However, the extent of the divergence between gene duplicates in CsoSCA is much greater than in *P. purpureum*; there is 70% sequence identity between the N- and C-terminal domains of *P. purpureum* carbonic anhydrase but only 11% identity between the catalytic and C-terminal domains of CsoSCA (Fig. 2).

CsoSCA Qualifies as a Distinct Subclass of the β -Carbonic Anhydrases—Although the structural similarities are sufficient to allow recognition of CsoSCA as a member of the β -class of carbonic anhydrases, three important differences merit qualification of CsoSCA as a member of a new, carboxysomal subclass of β -carbonic anhydrases. First, the sequence similarity between CsoSCA and the β -carbonic anhydrases is negligible; there is no statistically meaningful sequence similarity between CsoSCA and these proteins (less than 15% identity for aligned residues as compared with an average of 25% sequence identity among the remaining structurally determined β -carbonic anhydrases). The number of absolutely conserved residues among β -carbonic anhydrases decreases from 13 to 5 upon adding CsoSCA to the alignment (Fig. 2).

The numerous insertions and deletions in the CsoSCA primary structure as compared with other β -carbonic anhydrases offer a second distinguishing feature of carboxysomal carbonic anhydrases. These include: a 16-residue insertion extending helix F and its following loop; an 8-residue loop insertion between strand 3 and helix G; and severe shortening of helices G and H and intervening sequence (averaging a 30-residue deletion) (Figs. 1B and 2). This last region forms part of a dimer interface in some β -carbonic anhydrases (see below). Clearly, the structural differences between CsoSCA and the other β -carbonic anhydrases are far greater than those used to distinguish the two previously identified β subclasses, “plant” (*P. purpureum*, *E. coli*, *P. sativum*, *M. tuberculosis*, Rv3588c) and “cab” (*M. thermoautotrophicum*, *M. tuberculosis*, Rv1284). Indeed, the plant and cab subclasses superimpose with an average pairwise r.m.s. deviation of 0.9 Å for 132 α C pairs, whereas CsoSCA superimposes on these carbonic anhydrases with an average pairwise r.m.s. deviation of 2.4 Å over only 81 α C pairs.

Lastly, CsoSCA is distinct from the other structurally determined β -class carbonic anhydrases in that it alone has lost the two-fold symmetric pairing of active sites that most likely was once part of its heritage (Fig. 3). Active site symmetry is apparent in all species of β -carbonic anhydrases observed so far, except in CsoSCA. Active sites of carbonic anhydrases from *E. coli*, *P. sativum*, and *M. thermoautotrophicum* are organized as identical pairs through homodimerization of the carbonic anhydrase domain. The *P. purpureum* enzyme displays an intermediate degree of symmetry, in which two carbonic anhydrase domains have fused, and are organized with two-fold pseudo-symmetry (although only 70% sequence identity is maintained between domains). In CsoSCA, however, the C-terminal domain has diverged so greatly that it has lost the ligands for the catalytic zinc, including a loop on which one of the ligands (equivalent to Cys-253) would reside.

This loss of the catalytic potential of the C-terminal domain presents an evolutionary puzzle. In *P. purpureum*, one can speculate that by breaking the symmetry of its gene fusion product, the organism is able to gain some advantage by tuning the individual domains to have opti-

mum activity under different conditions. However, in the carboxysomal CsoSCA, the complete loss of an active site offers no apparent advantage gained in return for the cost of synthesizing the defunct domain. The fact that the C-terminal domain is about 130 residues smaller than the catalytic domain might suggest that by diminishing the size of the domain, the organism can partially mitigate the cost of its synthesis. Alternatively, one could propose a more purposeful role for the extreme divergence of the C-terminal domain, such as regulating the activity of the catalytic domain. The F_1 ATPase provides a well characterized analogy in which a non-catalytic α -subunit has evolved to regulate activity of the catalytic β -subunit (33). Still another possibility is that the C-terminal domain has evolved to interact with the carboxysome shell or RuBisCO. The observation of non-catalytic Gla and epidermal growth factor-like domains has led to similar proposals of their role in mediating protein-protein interactions (34).

A Layer of Dimers—CsoSCA protomers pack in layers within the crystal (Fig. 5A). The individual layers resemble egg crates, having two distinct faces: a flat face and a spiked face. The flat face of each layer originates from the tight-fitting lateral association of catalytic domains, whereas the spikes on the opposite face originate from the extremities of the N-terminal domains. Protomer-protomer contacts within the layer are numerous (see below), but layer-to-layer contacts are sparse, involving only one patch of 400 Å² between the N-terminal “spike” (Trp-117) and the flat catalytic domain surface (specifically, residues 345–349) of the next layer.

These layers obey plane group symmetry, p2, in which there are four unique two-fold symmetry axes normal to the plane (layer). The abundance of two-fold symmetry elements is not inconsistent with space group P1, the symmetry of the crystals used in this structure determination. The p2 layers apparently arise under the special circumstances observed here, in which the unit cell contains a homodimer having its non-crystallographic two-fold symmetry axis perpendicular to one of the unit cell planes, in this case perpendicular to the b-c plane.

The association of protomers within each layer is mediated by four distinct types of dimer interfaces, each centered on one of the four unique two-fold symmetry axes of plane group p2. Each interface is characterized by varying amounts of surface area buried and varying degrees of shape complementarity. In Fig. 5A, these are numbered 1–4 in order of decreasing amount of surface area buried by the two-fold symmetric interfaces (3304, 2188, 1030, and 584 Å², respectively). The shape complementarity of the four interfaces also decreases in the same order (0.717, 0.684, 0.584, <0.500).

The largest of these four dimer interfaces (Fig. 5A, interface 1) is most likely to be biologically relevant due to the immense amount of surface area it buries (3304 Å²). Also, the shape complementarity of this dimer interface is high (shape complementarity = 0.717), in the range of that seen for protease-inhibitor interfaces (35). Shape complementarity measures the closeness of fit between two molecular surfaces. Values can range from 0.0 to 1.0, where larger values indicate the absence of gaps between molecular surfaces. Lastly, 9 of the 29 residues lining this interface are absolutely conserved in an alignment of 10 carboxysomal carbonic anhydrase sequences. For these reasons, we term this dimer the principal dimer (Fig. 5, B and C). This dimer was chosen to represent the asymmetric unit in the coordinates deposited with the Protein Data Bank. A detailed description of the principal dimer interface is given in the following section.

The extent of the second largest interface (Fig. 5, interface 2) suggests that it may also be biologically relevant. With 2188 Å² of buried surface area, its size exceeds the threshold 1712 Å² used to discriminate between a true homodimer interface and simple crystal contacts between monomers (36). However, none of 19 residues are absolutely

Structure of β -Carbonic Anhydrase from the Carboxysomal Shell

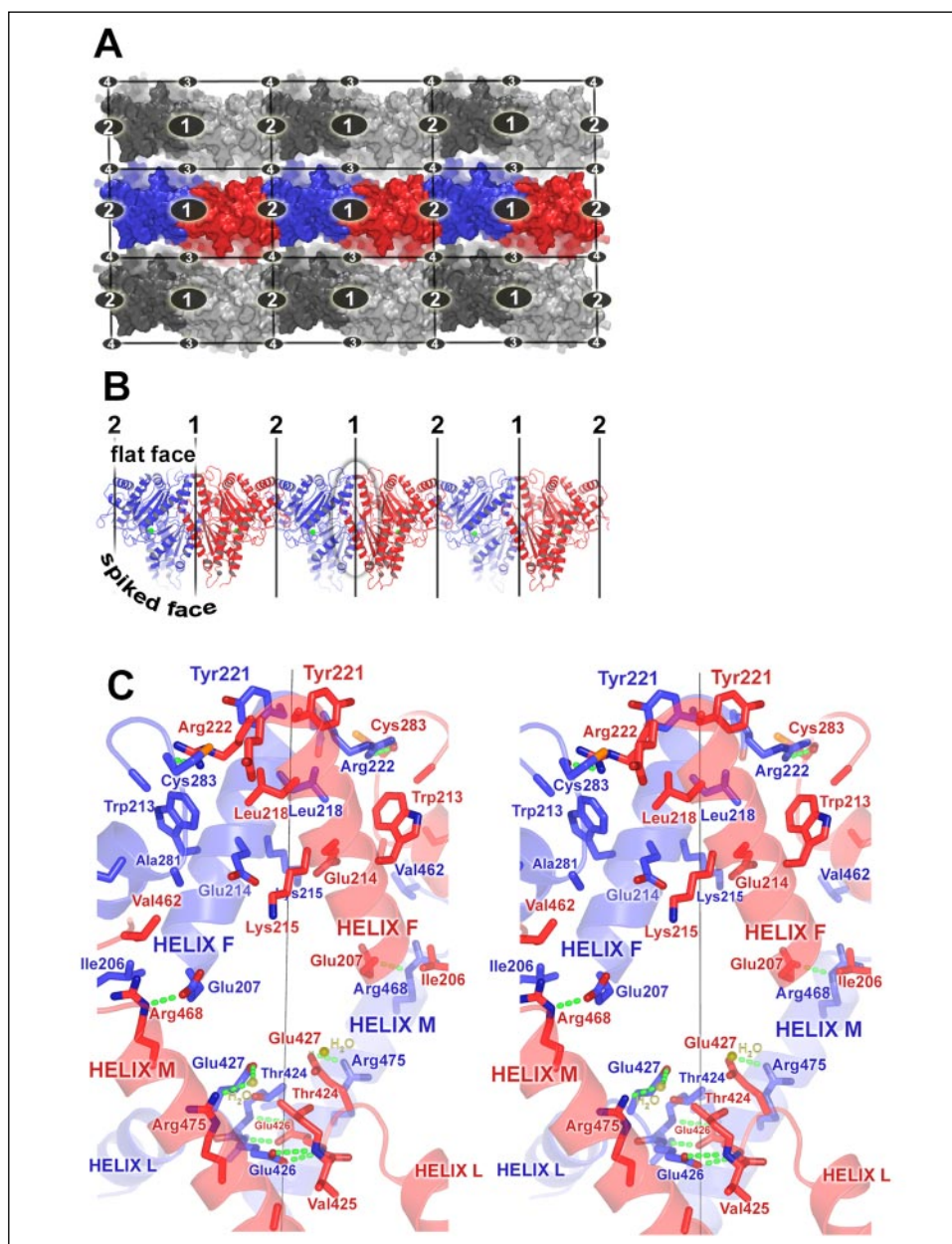


FIGURE 5. Oligomerization state of CsoSCA. *A*, crystal packing of CsoSCA reveals p2 layer symmetry when viewed perpendicular to the flat face of the layer. Ellipses mark the positions of two-fold symmetry axes, all of which project perpendicular to the plane of the page. There are four unique two-fold symmetry axes, numbered 1–4, each positioned at the center of a unique molecular interface. These axes, or interfaces, are numbered in order of decreasing amount of buried surface area and shape complementarity, *interface 1* being the strongest and *interface 4* being the weakest. Molecular surfaces of individual molecules are colored red, blue, dark gray, and light gray. The two strongest interfaces are juxtaposed to produce straight filaments (red and blue molecules). The other two interfaces are much weaker, revealing gaps of empty space between horizontal filaments. *B*, a single row of CsoSCA molecules is isolated from the layer and shown perpendicular to the view in *panel A*, as if the layer were viewed edge on. The distinction between the flat and spiked faces becomes evident. Molecules related by two-fold axes labeled 1 correspond to the principal dimer. An enlarged view of the principal dimer interface enclosed by the ellipse is shown in *panel C*.

conserved in this interface. The remaining interfaces in the layer (Fig. 5A, *interfaces 3* and *4*) fall below this threshold and so would most likely require the support of other carboxysomal proteins if they were truly present in the carboxysome.

The existence of p2 layer symmetry within the carboxysomal shell presents an attractive hypothesis about functional organization. In whole carboxysomes, CsoSCA is found tightly associated with the shell component, rather than the core component of RuBisCO molecules (12, 13). One might propose that interaction between CsoSCA and a shell protein such as CsoS1A might be facilitated by the complementarity of fit between their flat surfaces (6).⁵ If the N-terminal spikes were facing inward toward the center of the carboxysome, they could interact with RuBisCO molecules, perhaps to organize them with the proper periodicity for efficiency and compaction into paracrystalline arrays (38, 39). A possible role of CsoSCA in organizing RuBisCO molecules is implied

by analogy with the functionally related CcmM protein in which four repeats of the RuBisCO small subunit precede the carbonic anhydrase domain. It has been proposed that CcmM, anchored to the inside shell of the carboxysome, could bind to the large subunits of RuBisCO (11).

Three lines of evidence argue against the hypothesis that CsoSCA molecules are arranged with p2 symmetry in the carboxysomal shell. CsoSCA is a minor component of the carboxysome, appearing as a very faint band in a Coomassie Blue-stained SDS-polyacrylamide gel of purified carboxysomes (12, 13). Thus, it is unlikely that there is a sufficient number of molecules present to form a continuous layer in the shell. Secondly, the periodicity of molecules in the layer ($a = 90 \text{ \AA}$, $b = 48 \text{ \AA}$, $\gamma = 90^\circ$, Fig. 5A) does not match that of the CsoS1A shell component known from crystallographic studies ($a = b = 67 \text{ \AA}$, $\gamma = 120^\circ$).⁵ Specific interactions between sheets would be difficult to imagine if the layers do not share the same repeating dimensions. Lastly, examination of the protomer interfaces within the layer indicates that two of the four protomer interfaces in the layer are significantly larger and more enmeshed

⁵ Y. Tsai, M. R. Sawaya, F. Cai, S. Heinhorst, G. C. Cannon, C. A. Kerfeld, and T. O. Yeates, manuscript in preparation.

than the remaining two (Fig. 5A, compare *interfaces 1* and *2* with *interfaces 3* and *4*). Interestingly, the two strongest interfaces are juxtaposed to produce straight filaments (Fig. 5B). Perhaps it is these filaments, rather than sheets, that line the inside of the carboxysomal shell. Filaments have also been observed in the *M. thermoautotrophicum* β -carbonic anhydrase crystal structure (via an open-ended domain swap) and have been suggested to have functional importance (30). Although the extent of CsoSCA protomer assembly is uncertain, it is clear that the principal dimer is a key component. Indeed, size exclusion chromatography results suggest that a dimer is the principal species *in vitro*.

The Principal Dimer Interface—The interface of the principal dimer may be divided into two separate subregions; the first is centered on a pair of symmetry related helices F (residues 206–223), originating from the catalytic domains of protomers A and B (Fig. 5C). These two helices cross each other at an angle of 50° near their C-terminal ends. Hydrophobic interactions are observed near the crossing point where Leu-218 and Tyr-221 of protomer A contact equivalent residues in protomer B. Adjacent to these interactions, hydrogen bonds are reciprocated between Arg-222 (NH1) and Cys-283 (O) of opposing protomers, whereas salt bridges similarly link the side chains of Glu-207 (OE2) and Arg-468 (NE). Residues Glu-214 and Lys-215 are also paired across the dimer axis; however, these interactions are weaker (separated by 4 \AA), presumably because these side chains are exposed to solvent.

The second subregion of this dimer interface is also centered on the dimer symmetry axis but $\sim 20 \text{ \AA}$ away from the first subregion (Fig. 5C). It involves helix L of protomer A (from the C-terminal domain) and its two-fold symmetry-mate from protomer B. These two helices do not cross each other as the helices F (see above), but instead, their N-terminal ends abut, so their helical axes are roughly parallel. Such an interface would seem unfavorable due to the electrostatic repulsion arising from the proximity of the positive ends of the two-helix dipoles. However, Glu-426 at the very N-terminal end of the helix apparently neutralizes the charge repulsion. Hydrogen bonds are reciprocated between the side chain carboxylate of Glu-426 (O ϵ 2) and the amide nitrogen of Glu-426 (N) of the opposing protomer. The remaining carboxylate oxygen of Glu-426 (O ϵ 1) further participates in a second hydrogen bond with the amide nitrogen of Val-425 (N). This key residue, Glu-426, is absolutely conserved among carboxysomal carbonic anhydrases, as well as neighboring residue Thr-424, which forms van der Waals contacts between protomer A and its symmetry equivalent in protomer B. The interface is additionally strengthened by water-mediated hydrogen bonds between Glu-427 (O ϵ 2) and Arg-475 (NH $_2$) of opposing protomers.

The *P. purpureum* enzyme has a dimer interface analogous to CsoSCA, but more extensive (32). The *P. purpureum* enzyme uses roughly the same face of the molecule as a dimerization interface, but there is about 50° difference in the orientation of the monomers. The interface of the *P. purpureum* enzyme covers 8000 \AA^2 overall, as compared with 3304 \AA^2 in CsoSCA. A large segment of this interface is absent in CsoSCA and would correspond to a deletion of 30 residues between helices G and H (Fig. 2). Carboxysomal carbonic anhydrases appear to be the only subclass of β -carbonic anhydrases to have such an extensive deletion in this region. In *E. coli* (29) and Rv1284 (31), this region is used to form a tetrameric interface analogous to the dimeric interface of *P. purpureum*. However, in the octameric *P. sativum* (28), the filamentous *M. thermoautotrophicum* (30), and dimeric Rv3588c (31) enzymes, this region takes no part in any of its interfaces.

Thermodynamics of CO_2 Generation within the Carboxysome—For CsoSCA, as for all the structurally determined carbonic anhydrases to date, the enzyme-catalyzed reaction is reversible and is not driven by direct coupling to the hydrolysis of high energy compounds such as

ATP. That is, whether CsoSCA catalyzes a net gain in reactant (HCO_3^-) or product (CO_2) is determined by the difference in free energy (ΔG) between the two forms of inorganic carbon within the environment of the carboxysome shell. This energy difference must be sufficiently negative for CsoSCA to achieve its presumed biological role, the production of CO_2 for the fixation of carbon by RuBisCO inside the carboxysome (14). Since the standard free energy difference (ΔG°) contributes only a small negative term to ΔG , it is likely that ΔG (and the direction of the reaction) must be favorably influenced by the principle of mass action. That is, local concentrations of the reactant, HCO_3^- , are elevated by its active transport into the cell (1), whereas the local concentration of the product, CO_2 , is diminished by the close proximity of encapsulated RuBisCO, a CO_2 sink (37).

The sensitivity of the reaction direction to the environmental conditions is evidently key to the diverse roles played by β -carbonic anhydrase in biology. For example, in contrast to CsoSCA, the biological role of β -carbonic anhydrase in C_4 plants is not the production of CO_2 but rather the hydration of CO_2 to produce HCO_3^- , the substrate for phosphoenolpyruvate carboxylase (40). Thus, small changes in environmental conditions are presumably sufficient to reverse the biological role of β -carbonic anhydrases in the cell.

Active Site Similarities across Subclasses—Despite the level of sequence divergence that separates CsoSCA from the remaining two subclasses of β -carbonic anhydrases, there is a remarkable level of structural similarity in the active site region, which suggests a common catalytic mechanism for the interconversion of HCO_3^- and CO_2 . These active site residues fall into two main categories: zinc binding and substrate/product binding.

Zinc-mediated catalysis is common to all β -carbonic anhydrases studied to date. Crystal structures of these enzymes reveal that the zinc ion is ligated by two cysteines and a histidine with tetrahedral geometry (28–32). In CsoSCA, these ligands are Cys-173(S γ), His-242(N ϵ 2), and Cys-253(S γ). Indeed, zinc coordination geometry is remarkably conserved among β -carbonic anhydrase structures (Fig. 4A); however, the identity of the fourth zinc ligand can vary from structure to structure. In CsoSCA, it is a water molecule, analogous to those observed in β -carbonic anhydrase structures from *P. sativum*, *M. thermoautotrophicum*, and *M. tuberculosis* (Rv1284). In the remaining β -carbonic anhydrase structures, this ligand is a strictly conserved aspartate: Asp-151 and Asp-405 in *P. purpureum*, Asp-44 in *E. coli*, and Asp-53 in Rv3588c from *M. tuberculosis*. It is observed that whenever water occupies the position of the fourth zinc ligand, the aspartate (Asp-175 in CsoSCA) consequently takes a second conformation, fixed by two hydrogen bonds to a strictly conserved arginine (Arg-177 in CsoSCA) (Fig. 4A). The alternate conformations of this aspartate might play multiple roles in the catalytic mechanism (see below).

The composition of the HCO_3^- binding site might appear to be more subclass-specific than those involved in zinc binding; however, this impression might simply be due to the uncertainty of the location of HCO_3^- binding. Within the plant subclass of β -carbonic anhydrases, a conserved glutamine residue has been implicated in HCO_3^- binding. Specifically, in the *P. sativum* structure, Gln-151 is observed donating a hydrogen bond to substrate analog, acetate (28). In the carboxysomal carbonic anhydrases, this glutamine residue is substituted (conservatively) with a histidine (His-397 in CsoSCA). This histidine superimposes remarkably well with Gln-151 (Fig. 4B). The N ϵ 2 atom of His-397 is in proximity of an oxygen atom of acetate. His-397 is held in position by a hydrogen bond between N δ 1 and the relatively buried Glu-399. Backbone amide nitrogens of Ala-254 and Ala-255 (helix G) are in position to donate hydrogen bonds to the remaining HCO_3^- oxygens. Struc-

Structure of β -Carbonic Anhydrase from the Carboxysomal Shell

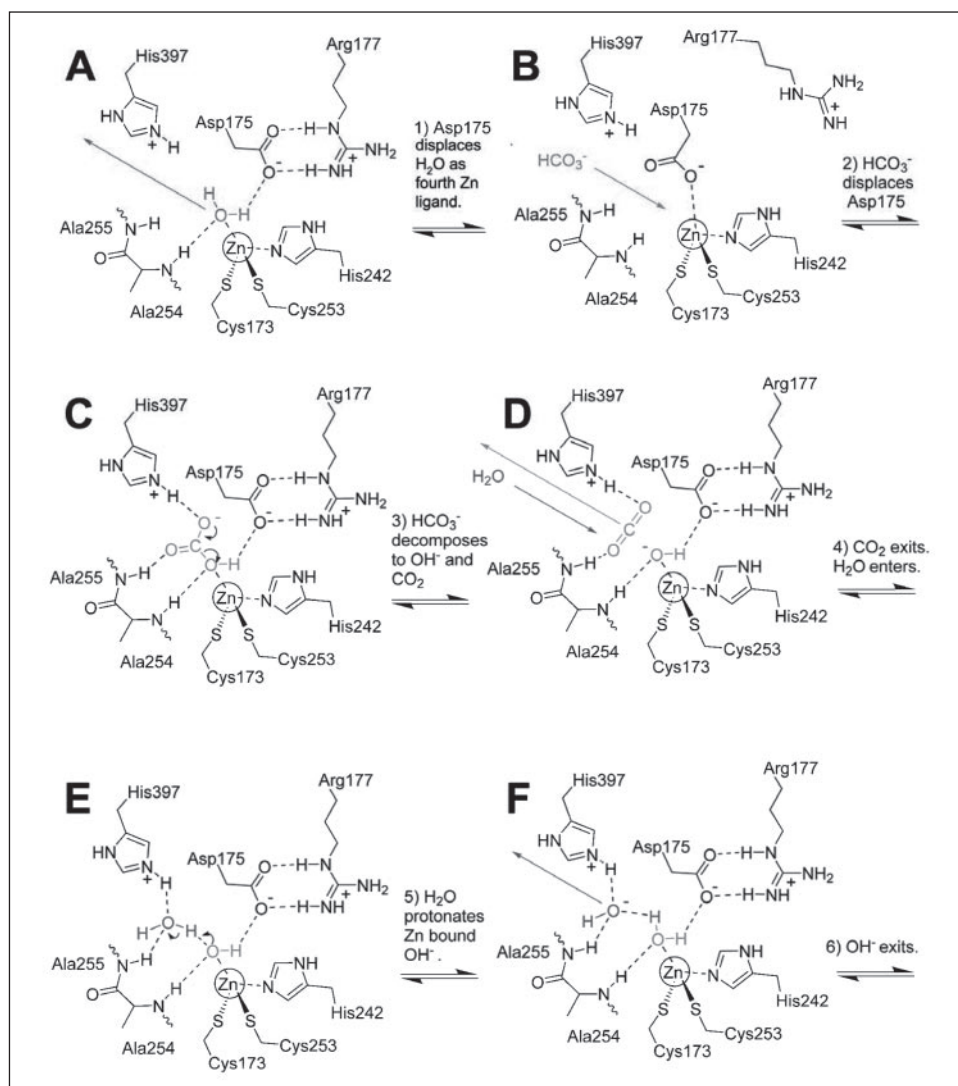


FIGURE 6. Schematic of the CsoSCA catalyzed dehydration of HCO_3^- .

turally equivalent amide nitrogens are present in all β -carbonic anhydrases, even the cab subclass, which appear to lack a residue equivalent to the Gln-151 of the plant. A model for HCO_3^- binding to CsoSCA is presented in Fig. 4B.

A Conserved Catalytic Mechanism—Structural similarity among β -carbonic anhydrases allows us to propose a mechanism for the dehydration of HCO_3^- that is analogous to that described for the reverse reaction (hydration of CO_2) in the literature (28–32). Here, we adopt the convention that dehydration is the forward direction since CO_2 is considered the biologically important product in the carboxysome. However, it should be understood that notions of forward or reverse directions are entirely a matter of convenience. Furthermore, we note that CsoSCA provides no advantage in directionality of the reaction as compared with other carbonic anhydrases, as enzymes have no effect on the equilibrium positions of the reactions they catalyze.

The catalytic cycle begins with water bound as the fourth ligand in the coordination shell of zinc, just as observed in the current crystal structure (Figs. 1C and 6, A and B). In preparation for HCO_3^- binding, conformational flexibility in Asp-175 might aid in dislodging the water from this site. In addition to being one of the products of the dehydration reaction, water is present at a high concentration and so might otherwise be difficult to displace. In support of this role, Asp-175 has been observed to successfully compete with water for binding to the zinc in

the β -carbonic anhydrase structures of *P. purpureum* (32), *E. coli* (29), and *M. tuberculosis* (Rv3588c) (31).

As HCO_3^- approaches the zinc site for coordination, Asp-175 further aids catalysis by selecting a catalytically competent binding geometry for HCO_3^- . Although all three oxygen atoms of HCO_3^- have affinity for zinc, it is important that the zinc coordinate only the protonated oxygen since this oxygen will form the leaving group. Selection is achieved by the close proximity of Asp-175 to the zinc-coordinated oxygen of HCO_3^- (Fig. 6C). Asp-175 is held in this conformation by a pair of hydrogen bonds with Arg-177 (as seen in the crystal structure, Fig. 1C). Since Asp-175 can serve only as a hydrogen bond acceptor, not as a donor, only the protonated oxygen of HCO_3^- can complete the hydrogen bond and coordinate zinc at the same time. An analogous “gatekeeper” role has been proposed for this strictly conserved aspartate in the hydration reaction (28). The backbone amide of Ala-254 donates a second hydrogen bond to the coordinated oxygen of HCO_3^- , and one of the two remaining HCO_3^- oxygen atoms accepts a hydrogen bond from the backbone amide of Ala-255.

Decomposition of HCO_3^- to CO_2 and OH^- follows (Fig. 6D). CO_2 diffuses from the active site, and a water molecule takes its place (Fig. 6E), binding where water 305 is observed in the crystal structure (Fig. 1C). The zinc-bound OH^- ion is likely protonated via this water molecule (Fig. 6F), which is in contact with bulk solvent.

Fortuitous Zinc Binding Site—In the furthest extremity of the N-terminal domain, a zinc ion was observed bound between surface loops (Fig. 1A). Specifically, the zinc ion directly coordinates His-92, Asp-115, and His-121. The electron density attributed to zinc could not have originated from water since a strong peak (8σ) was observed centered on this atom in an anomalous difference Fourier map. An anomalous scattering signal of this magnitude could only have arisen from a metal ion. Indeed, the metal site was used to improve the phases used to calculate subsequent electron density maps. The atomic element was identified as zinc due to the observance of a fluorescence maximum corresponding to the zinc peak wavelength. (Note: The same methods were used to verify the identity of the zinc ion in the catalytic domain).

The appearance of a zinc ion bound to the novel N-terminal domain (residues His-92, Asp-115, and His-121) gave the initial impression that this domain contains a catalytic site for the carbonic anhydrase reaction. However, it was later discovered that His-92, one of the zinc ligands, arose only as the result of a fortuitous point mutation in the clone used to produce the protein for crystallization. Since this mutant does not display increased activity as compared with the wild-type protein,⁶ it seems unlikely that this zinc atom remains intact in the wild-type protein, which contains a tyrosine at position 92. The tyrosine hydroxyl would likely displace the zinc ion due to the increased length of the side chain. Furthermore, if this apparently fortuitous metal binding site were truly advantageous, one would expect that the 3 residues directly coordinating the zinc site would be evolutionarily conserved. Sequence alignments clearly show this not to be the case (Fig. 2, and additional alignments not shown). We are therefore compelled to conclude that this zinc ion found in the N-terminal domain is not biologically relevant.

Conclusions—The similarities observed in the active site structure and likely catalytic mechanism between the carboxysomal carbonic anhydrase CsoSCA from *H. neapolitanus* and known β -carbonic anhydrases clearly place this enzyme into the carbonic anhydrase β -class. However, they must be considered as belonging to a distinct subclass since CsoSCA and closely related carboxysomal enzymes display an almost complete lack of primary structure similarity to other carbonic anhydrases and do not possess the characteristic paired active sites of other β -carbonic anhydrases. The significance of these structural deviations for the function of carboxysomal carbonic anhydrases is not fully understood, but they might be essential to the role of CsoSCA as a component of this primitive prokaryotic organelle.

Acknowledgments—We thank the staff at ALS beamline 8.2.2, especially Corie Ralston for collecting the native data set. We thank Duilio Cascio and Alex Lisker for technical assistance and Professor Jess Shively for helpful discussions.

REFERENCES

1. Price, G. D., Sultemeyer, D., Klughammer, B., Ludwig, M and Badger, M. R. (1998) *Can. J. Bot.* **76**, 973–1002

⁶ G. C. Cannon, F. Cai, E. B. Williams, D. Murin, J. M. Shively, and S. Heinhorst, manuscript in preparation.

2. Cannon, G. C., Bradburne, C. E., Aldrich, H. C., Baker, S. H., Heinhorst, S., and Shively, J. M. (2001) *Appl. Environ. Microbiol.* **67**, 5351–5361
3. Badger, M. R., and Price, G. D. (2003) *J. Exp. Bot.* **54**, 609–622
4. Kaplan, A., and Reinhold, L. (1999) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 539–570
5. Beudeker, R. F., Cannon, G. C., Kuenen, J. G., and Shively, J. M. (1980) *Arch. Microbiol.* **124**, 185–189
6. Kerfeld, C. A., Sawaya, M. R., Tanaka, S., Nguyen, C. V., Phillips, M., Beeby, M., and Yeates, T. O. (2005) *Science* **309**, 936–938
7. Cannon, G. C., Heinhorst, S., Bradburne, C. E. Shively, J. M. (2002). *Funct. Plant Biol.* **29**, 175–182
8. Price, G. D., Coleman, J. R., and Badger, M. R. (1992) *Plant Physiol.* **100**, 784–793
9. So, A. K.-C., and Espie, G. S. (1998) *Plant Mol. Biol.* **37**, 205–215
10. So, A. K.-C., and Espie, G. S. (2005) *Can. J. Bot.* **83**, 721–734
11. Ludwig, M., Sultemeyer, D., and Price, G. D. (2000) *J. Phycol.* **36**, 1109–1118
12. So, A. K.-C., Espie, G. S., Williams, E. B., Shively, J. M., Heinhorst, S., and Cannon, G. C. (2004) *J. Bacteriol.* **186**, 623–630
13. Baker, S. H., Williams, D. S., Aldrich, H. C., Gambrell, A. C., and Shively, J. M. (2000) *Arch. Microbiol.* **173**, 278–283
14. Cannon, G. C., and Shively, J. M. (1983) *Arch. Microbiol.* **134**, 52–59
15. Boggon, T. J., and Shapiro, L. (2000) *Structure (Lond.)* **8**, R143–R149
16. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276A**, 307–326
17. Sheldrick, G. M., and Schneider, T. R. (2001) in *Methods in Macromolecular Crystallography* (Turk, D., and Johnson, L., eds.) pp. 72–81. IOS Press, Amsterdam
18. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
19. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
20. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
21. Fabiola, F., Bertram, R., Korostelev, A., and Chapman, M. S. (2002) *Protein Sci.* **11**, 1415–1423
22. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255
23. Colovos, C., and Yeates, T. O. (1993) *Protein Sci.* **2**, 1511–1519
24. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291
25. Vriend, G., and Sander, C. (1993) *J. Appl. Crystallogr.* **26**, 47–60
26. DeLano, W. L. (2002). The PyMOL User's Manual, DeLano Scientific, San Carlos, CA
27. Gibrat, J. F., Madej, T., and Bryant, S. H. (1996) *Curr. Opin. Struct. Biol.* **6**, 377–385
28. Kimber, M. S., and Pai, E. F. (2000) *EMBO J.* **19**, 1407–1418
29. Cronk, J. D. Endrizzi, J. A., Cronk, M. R. O'Neil, J. W., and Zhang, K. Y. J. (2001) *Protein Sci.* **10**, 911–922
30. Strop, P., Smith, K. S., Iverson, T. M., Ferry, J. G., and Rees, D. C. (2001) *J. Biol. Chem.* **276**, 10299–10305
31. Covarrubias, A. S., Larsson, A. M., Högbom, M., Lindberg, J., Bergfors, T., Björkelid, C., Mowbray, S. L., Unge, T., and Jones, T. A. (2005) *J. Biol. Chem.* **280**, 18782–18789
32. Mitsuhashi, S., Mizushima, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Moriyama, H., Ueki, T., Miyachi, S., and Tsukihara, T. (2000) *J. Biol. Chem.* **275**, 5521–5526
33. Gao, Y. Q., Yang, W., and Karplus, M. (2005) *Cell* **123**, 195–205
34. Stenflo, J. (1999) *Crit. Rev. Eukaryotic Gene Expression* **9**, 59–88
35. Lawrence, M. C., and Colman, P. M. (1993) *J. Mol. Biol.* **234**, 946–950
36. Ponstingl, H., Henrick, K., and Thornton, J. J. (2000) *Proteins* **41**, 47–57
37. Reinhold, L., Kosloff, R., and Kaplan, A. (1991) *Can. J. Bot.* **69**, 984–988
38. Shively, J. M., Ball, F. L., and Kline, B. W. (1973) *J. Bacteriol.* **116**, 1405–1411
39. Holthuijzen, Y. A., Vanbreem, J. F., Konings, W. N., and Vanbrugg, E. F. (1986) *Arch. Microbiol.* **144**, 258–262
40. Hatch, M. D., and Burnell, J. N. (1990) *Plant Physiol.* **93**, 380–383

The Structure of β -Carbonic Anhydrase from the Carboxysomal Shell Reveals a Distinct Subclass with One Active Site for the Price of Two

Michael R. Sawaya, Gordon C. Cannon, Sabine Heinhorst, Shiho Tanaka, Eric B. Williams, Todd O. Yeates and Cheryl A. Kerfeld

J. Biol. Chem. 2006, 281:7546-7555.

doi: 10.1074/jbc.M510464200 originally published online January 10, 2006

Access the most updated version of this article at doi: [10.1074/jbc.M510464200](https://doi.org/10.1074/jbc.M510464200)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 38 references, 10 of which can be accessed free at <http://www.jbc.org/content/281/11/7546.full.html#ref-list-1>