Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Observation of a calcium-binding site in the $\gamma$-class carbonic anhydrase from *Pyrococcus horikoshii*

**Jeyaraman Jeyakanthan,[a,b] Sarani Rangarajan,[c] P. Mridula,[c] Shankar Prasad Kanaujia,[c,d] Yoshitsugu Shiro,[b] Seiki Kuramitsu,[b,e] Shigeyuki Yokoyama[b,f,g] and Kanagaraj Sekar[c,d]***

[a]National Synchrotron Radiation Research Center, 101 Hsin-Ann Road, Hsinchu Science Park, Hsinchu 30076, Taiwan, [b]RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo, Hyogo 679-5148, Japan, [c]Bioinformatics Centre (Centre of Excellence in Structural Biology and Bio-computing), Indian Institute of Science, Bangalore 560 012, India, [d]Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560 012, India, [e]Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, [f]Systems and Structural Biology Center, Yokohama Institute, RIKEN, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, and [g]Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Correspondence e-mail: sekar@serc.iisc.ernet.in, sekar@physics.iisc.ernet.in



Received 11 June 2008
Accepted 30 July 2008

**PDB References:** $\gamma$-carbonic anhydrase, 1v3w, r1v3wsf; 1v67, r1v67sf; 2fko, r2fkosf.

Carbonic anhydrases are zinc-containing metalloenzymes that catalyze the interconversion of carbon dioxide and bicarbonate. Three crystal structures of $\gamma$-class carbonic anhydrase (one of which is bound to a bicarbonate molecule) from the aerobic OT3 strain of the hyperthermophilic archeon *Pyrococcus horikoshii* have been solved by molecular replacement in space group $F4_132$. The asymmetric unit contains a monomer of 173 amino acids and a catalytic $Zn^{2+}$ ion. The protein fold is a regular prism formed by a left-handed $\beta$-helix, similar to previously reported structures. The active-site $Zn^{2+}$ ion located at the interface between the two monomers is bound to three histidyl residues and a water molecule in a tetrahedral fashion. In addition to the 20 $\beta$-strands comprising the $\beta$-helix, there is also a long C-terminal $\alpha$-helix. For the first time, $Ca^{2+}$ ions have been observed in addition to the catalytic $Zn^{2+}$ ion. It is hypothesized that Tyr159 (which corresponds to the catalytically important Asn202 in previously reported structures) utilizes C—H$\cdots\pi$ interactions to fulfill its functions. This study may shed light on the catalytic mechanism of the enzyme and throw open new questions on the mechanism of product removal in carbonic anhydrases.

## 1. Introduction

Carbonic anhydrases (CAs; EC 4.2.1.1) are zinc-containing metalloenzymes that have a variety of physiological roles including the facilitated diffusion of $CO_2$ and the reversible conversion of $CO_2$ to $HCO_3^-$ (Meldrum & Roughton, 1933). On the basis of amino-acid sequence similarity, carbonic anhydrases have been classified into three major groups, namely $\alpha$, $\beta$ and $\gamma$. A fourth category ($\delta$ class) of carbonic anhydrases has also been proposed (Tripp & Ferry, 2000). Sequence-homology studies involving CAs from bacteria and archaea show that the $\gamma$-class enzyme is likely to be the prototype of all carbonic anhydrases. Furthermore, phylogenic studies suggest that the date of evolution of the $\gamma$ class is near the origin of life (Smith *et al.*, 1999). The literature contains several structures of isozymes of $\alpha$-class carbonic anhydrases (Kannan *et al.*, 1984; Hakansson *et al.*, 1992; Eriksson & Liljas, 1993; Boriack-Sjodin *et al.*, 1995; Stams *et al.*, 1996) and $\beta$-class carbonic anhydrases (Mitsuhashi *et al.*, 2000; Covarrubias *et al.*, 2006; Cronk *et al.*, 2006; Sawaya *et al.*, 2006). Crystal structures of $\gamma$-class CAs have previously only been reported for a CA from the methanoarchaeon *Methanosarcina thermophila* (Kisker *et al.*, 1996; Iverson *et al.*, 2000).

Zn-Cam, the CA enzyme from *M. thermophila*, is a 70 kDa protein whose individual subunits contain 213 amino-acid residues (Alber & Ferry, 1994). The structure of the monomer consists of a left-handed parallel $\beta$-helix and an $\alpha$-helix. Each

face of the $\beta$-helix is comprised of parallel $\beta$-strands containing two to three residues each and is connected to the subsequent face by a 120° turn to form an equilateral prism (Kisker *et al.*, 1996). Three catalytic sites are arranged symmetrically in the trimer complex, in which $Zn^{2+}$ ions are bound by three histidines donated by His81 and His122 from one monomer and His117 from the neighbouring monomer. The binding of the $Zn^{2+}$ ion at the active site stabilizes its structure.

Even though CAs are found in living organisms ranging from bacteria to mammals, the $\gamma$ class of CAs are mainly present in archaea and some eubacteria. As an extracellular enzyme, $\gamma$-class CA is thought to play a role in acetate metabolism, especially when the substrate for growth is switched from methanol to acetate (Alber & Ferry, 1994; Hewett-Emmett & Tashian, 1996). As a cytosolic enzyme, it facilitates the uptake of acetate by supporting a $CH_3COO^-$/$HCO_3^-$ antiport system. In *M. thermophila*, energy is produced by the cleavage of the acetate C—C bond. Subsequently, the resultant methyl groups are reduced to methane using electrons derived from the oxidation of the carbonyl group to $CO_2$, which is then converted to bicarbonate with a turnover number of $6 \times 10^4$ $s^{-1}$ through the process of hydration (Alber *et al.*, 1999),

$$CO_2 + H_2O \longleftrightarrow HCO_3^- + H^+.$$

Outside the cell, this interconversion can facilitate either the uptake of acetate or the removal of $CO_2$ from the cytoplasm. Furthermore, it is involved in a variety of biosynthetic reactions including gluconeogenesis, the synthesis of certain amino acids (*via* pyruvate carboxylase), lipogenesis (*via* acetyl-CoA carboxylase), ureagenesis (*via* carbamoyl phosphate synthetase I) and pyrimidine-nucleotide biosynthesis (*via* carbamoyl phosphate synthetase II). These enzymes also play a significant role in processes such as respiratory gas exchange, photosynthesis, biosynthetic processes, pH homeostasis, bone resorption, tumorigenesis, calcification and ion transport (Chegwidden & Carter, 2000; Badger & Price, 1994; Tashian, 1989; Coleman, 1991). Thus, it is of significant interest to study the carbonic anhydrases, especially that from the aerobic strain OT3 of the hyperthermophilic archeon *Pyrococcus horikoshii*.

## 2. Materials and methods

### 2.1. Gene cloning, protein purification and dynamic light scattering

The carbonic anhydrase gene (PH1591) was amplified by the polymerase chain reaction (PCR) using *P. horikoshii* OT3 genomic DNA as a template. The recombinant plasmid was constructed using the super-rare-cutter system (Kanagawa *et al.*, manuscript in preparation). *Escherichia coli* BL21-CodonPlus (DE3)-RIL cells were transformed with the recombinant plasmid and grown at 310 K in LB medium containing 50 μg $ml^{-1}$ ampicillin for 20 h. The cells were harvested by centrifugation at 6500*g* for 5 min and suspended in 20 m*M* Tris–HCl pH 8.0 (buffer *A*) containing 0.5 *M* NaCl and 5 m*M* 2-mercaptethanol. The cell lysate was obtained by sonication and was clarified by centrifugation (15 000*g*, 30 min). The supernatant was heated at 363 K for 11.5 min and denatured proteins were removed by centrifugation (15 000*g*, 30 min). The supernatant was used as a crude extract for purification. It was desalted using a HiPrep 26/10 desalting column (Amersham Biosciences) and applied onto a Super Q Toyopearl 650M column (Tosoh) equilibrated with buffer *A*. The protein was eluted with a 0–0.3 *M* linear gradient of NaCl, desalted using a HiPrep 26/10 desalting column with buffer *A* and subjected to a Resource Q column (Amersham Biosciences) equilibrated with buffer *A*. The protein was eluted with a linear gradient of 0–0.3 *M* NaCl in buffer *A*. The buffer of the fractions containing the protein was exchanged to 10 m*M* sodium phosphate pH 7.0 using a HiPrep 26/10 desalting column. It was applied onto a Bio-Scale CHT-20-I column (Bio-Rad) equilibrated with the same buffer and the protein was eluted with a 10–200 m*M* linear gradient of sodium phosphate pH 7.0. The fractions containing protein were pooled, concentrated by ultrafiltration (Vivaspin, 5 kDa cutoff) and loaded onto a HiLoad 16/60 Superdex 200pg column (Amersham Biosciences) equilibrated with buffer *A* containing 0.2 *M* NaCl. The purified protein was homogeneous and was localized as a single band on SDS–PAGE.

A dynamic light-scattering experiment was performed on a DynaPro MS/X instrument from Protein Solutions (Lakewood, New Jersey, USA). Measurements were made at 291 K on purified protein at 0.5–1 mg $ml^{-1}$ in solution containing 20 m*M* Tris–HCl buffer pH 8.0 and 200 m*M* NaCl.

### 2.2. Crystallization and data collection

The protein solution subjected to crystallization contained 20 mg $ml^{-1}$ protein in 0.02 *M* Tris pH 8.0 buffer and 0.2 *M* NaCl. Single crystals were grown using the sitting-drop vapour-diffusion method; the buffer was composed of 26% PEG 4000 in HEPES–NaOH pH 7.5. For the native structure, 0.2 *M* $CaCl_2$ was used as the precipitating agent, while for the bicarbonate-bound structure (Zn-Cap-$HCO_3^-$) 15 m*M* sodium bicarbonate was additionally added to the precipitating agent. The third crystal was obtained using a buffer consisting of 22.5% PEG 4000 in MES–NaOH pH 5.8 with 0.1 *M* lithium chloride as the precipitating agent. Although crystallization with various buffers containing different chloride-ion-containing solutions ($ZnCl_2$, $MgCl_2$, $MnCl_2$, $NiCl_2$, $CoCl_2$ and LiCl) of varying concentrations was attempted, no crystals were obtained from any condition apart from the 0.1 *M* LiCl solution. A drop size of 2 μl, made up of 1 μl protein solution and 1 μl precipitant, was used for crystallization. For cryo-cooling, all crystals were soaked in a reservoir containing all the crystallization compounds along with 30%(*v*/*v*) glycerol. The concentration of glycerol was raised stepwise from 10 to 30% over a period of 8 h in order to avoid crystal cracking, reduce the mosaicity and improve the resolution.

The diffraction intensities of the native crystal were collected at 100 K on beamline BL26B1 at SPring-8 using a

**Table 1**
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

| | Zn-Cap | Zn-Cap-$HCO_3^-$ | LiCl-soaked |
|---|---|---|---|
| Wavelength (Å) | 1.0000 | 1.5418 | 1.5418 |
| Temperature (K) | 100 | 100 | 100 |
| Unit-cell parameters ($a = b = c$) (Å) | 156.16 | 158.51 | 157.41 |
| Space group | $F4_132$ | $F4_132$ | $F4_132$ |
| Resolution range (Å) | 20.0–1.5 (1.55–1.5) | 20.0–2.3 (2.38–2.3) | 20.0–1.85 (1.92–1.85) |
| No. of observations | 120843 | 96990 | 495389 |
| Unique reflections | 26357 (2602) | 7997 (765) | 14772 (1406) |
| Completeness (%) | 98.7 (99.8) | 99.5 (98.6) | 99.6 (98.0) |
| $R_{merge}$† (%) | 7.1 (11.1) | 7.1 (24) | 9.8 (28.8) |
| $V_M$ (Å$^3$ Da$^{-1}$) | 2.1 | 2.2 | 2.1 |
| Solvent content (%) | 40.9 | 44.0 | 42.6 |
| $Z$ | 1 | 1 | 1 |
| Wilson $B$ factor (Å$^2$) | 15.0 | 21.3 | 18.2 |
| $R_{work}$ (%) | 18.5 | 18.9 | 19.2 |
| $R_{free}$ (%) | 20.3 | 23.8 | 20.7 |
| PDB code | 1v3w | 1v67 | 2fko |

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of reflection $i$ and $\langle I(hkl)\rangle$ is the average intensity of multiple observations.

MAR Research image-plate scanner. The X-ray wavelength was set to 1 Å. The images were processed, scaled and reduced to 1.5 Å resolution using *DENZO* and *SCALEPACK* from the *HKL*-2000 package (Otwinowski & Minor, 1997). Diffraction data for the two remaining crystals were collected using an R-AXIS VII diffractometer. Data processing was carried out in a similar manner to that used for the native crystal.

### 2.3. Structure determination and refinement

Structure solution was achieved by molecular replacement using *AMoRe* (Navaza, 1994). For the native structure (Zn-Cap), the carbonic anhydrase from *M. thermophila* (PDB code 1qq0; Iverson *et al.*, 2000) with a sequence identity of 28% was used as the search model. A solution was obtained with this search model (*R* factor = 40%; correlation factor = 44%) using the resolution limits 15.0–6.0 Å. 5% of the reflections were set aside for the calculation of $R_{free}$ (Brünger, 1992). Structure refinement was carried out using *CNS* v.1.1 (Brünger *et al.*, 1998). Initially, 50 cycles of rigid-body refinement were carried out, followed by 50 cycles of positional refinement. Without the mutated residues, the $R_{work}$ dropped to 47% ($R_{free}$ = 50.5%) for all reflections in the resolution range 20–1.5 Å. The difference electron-density map was sufficient to mutate the residues into the electron-density map and the model was then subjected to simulated annealing by employing a slow-cooling protocol followed by 100 cycles of positional refinement, which lowered $R_{work}$ to 28.2% ($R_{free}$ = 30.5%). At this stage, using difference electron-density maps, one glycerol molecule in the active site and three ethylene glycol molecules were added to the model. Furthermore, the difference maps revealed electron density for the functionally important $Zn^{2+}$ ion in the active site. Interestingly, five more strong peaks were found in the difference maps (peak heights of >8–10$\sigma$). With the exception of $CaCl_2$, which was present in the crystallization buffer, no other metal ions were used during purification and crystallization. Six ligands were observed around the strong peak with a reasonable coordination distance (<2.6 Å). Thus, the strong peaks could be interpreted as either $Ca^{2+}$ ions or $Cl^-$ ions. In order to assign either of these two ions to the Fourier difference electron density, a systematic computational analysis was carried out on the structures available in the PDB to determine the ligand-binding distances of $Ca^{2+}$ and $Cl^-$ ions. This information was used to assign either a $Ca^{2+}$ or $Cl^-$ ion to the peaks. At this stage individual *B*-factor refinement was started; the ion–ligand distances were not restrained and water O atoms were picked and included in the refinement. Finally, OMIT maps (omitting 25 residues at a time) were calculated and used to correct or check the quality of the final protein model. The program *PROCHECK* (Laskowski *et al.*, 1993) was used to check the sterochemical values.

For both the remaining crystals, the three-dimensional atomic coordinates of the above refined native structure were used as the starting model since the molecular orientation and position of the model in the unit cell were isomorphous to those in the native structure. A total of 40 cycles of rigid-body refinement was followed by 100 cycles of positional refinement using *CNS* v.1.1 with the maximum-likelihood formalism target. For the bicarbonate-bound complex (Zn-Cap-$HCO_3^-$), $R_{work}$ was lowered to 25.5% ($R_{free}$ = 29.5%). For the LiCl-soaked crystal, $R_{work}$ was lowered to 27.5% ($R_{free}$ = 30.5%). The difference electron-density maps calculated at this stage were used to locate the catalytically important $Zn^{2+}$ ion in the active site. In addition, in the structure in which clear unambiguous density appeared for the bicarbonate molecule, the coordinates of the bicarbonate in the Cam complex (PDB code 1qrl; Iverson *et al.*, 2000) were used to fit the model into the electron density. Water O atoms were located and added to the model, using peaks with height greater than 2.5$\sigma$ in $|F_o - F_c|$ and 1.0$\sigma$ in $2|F_o - F_c|$ maps. Refinement converged at an $R_{work}$ and $R_{free}$ of 19.8% and 25.5%, respectively, for the bicarbonate-bound structure and at an $R_{work}$ and $R_{free}$ of 19.2% and 20.7%, respectively, for the LiCl-soaked crystal. Finally, OMIT maps were used to remove model bias. The stereochemical quality of the two structures was validated by using *PROCHECK* (Laskowski *et al.*, 1993).

### 2.4. *In silico* analysis

Several internet computing engines and software programs were used to carry out analysis of the crystal structures. Trimers and the unit cell were generated from the coordinates and space-group information of the monomer in the asymmetric unit using the online package *SEM* (*Symmetry*

**Table 2**
Statistics of the three crystal structures.

| | Zn-Cap | Zn-Cap-$HCO_3^-$ | LiCl-soaked |
|---|---|---|---|
| Protein atoms | 1362 | 1332 | 1336 |
| Water molecules | 193 | 142 | 189 |
| Bound $Zn^{2+}$ ions | 1 | 1 | 1 |
| Bound $Ca^{2+}$ ions | 4 | 1 | — |
| Bound chloride ions | 1 | — | — |
| Glycerol molecules | 1 | — | — |
| Bicarbonate molecules | — | 1 | — |
| Ethylene glycol molecules | 3 | — | 2 |
| R.m.s. deviations from | | | |
| Ideal bond lengths (Å) | 0.006 | 0.009 | 0.015 |
| Ideal bond angles (°) | 1.7 | 1.50 | 2.50 |
| Ideal dihedral angles (°) | 26.30 | 25.60 | 26.70 |
| Ideal improper angles (°) | 0.90 | 0.91 | 0.82 |
| Ramachandran statistics (%) | | | |
| Most favoured regions | 87 | 83.6 | 86.3 |
| Allowed regions | 12.3 | 15.7 | 13.7 |
| Generously allowed regions | 0.7 | 0.7 | 0 |

*Equivalent Molecules*; Hussain *et al.*, 2003). *PSAP* (*Protein Structural Analysis Package*; Balamurugan *et al.*, 2007) was used to find the various interactions in the structure, including the ion pairs and water-mediated ion pairs. Superimposition of the three-dimensional structures was carried out and invariant water molecules were identified using the WWW server 3*dSS* (Sumathi *et al.*, 2006). Perl scripts were used to find C—H$\cdots\pi$ interactions in the proteins. *PROMOTIF* (Hutchinson & Thornton, 1996) was used to identify the secondary-structural information from the PDB coordinates and the molecular-graphics program *PyMOL* (DeLano, 2002) was used to generate the figures.

## 3. Results and discussion

### 3.1. Crystallographic analysis

Three crystal structures of Zn-bound $\gamma$-class carbonic anhydrase (Zn-Cap) from the aerobic OT3 strain of *P. horikoshii* were solved by the molecular-replacement method and contained a monomer of 173 amino-acid residues in the asymmetric unit. The native structure of the carbonic anhydrase additionally bound to $Ca^{2+}$ was solved at a resolution of 1.5 Å (Zn-Cap; PDB code 1v3w) and a bicarbonate-bound structure (Zn-Cap-$HCO_3^-$; PDB code 1v67) was solved at a resolution of 2.3 Å. Under different crystallization conditions (involving soaking in LiCl solution), a third crystal was formed and yielded a structure (Zn-Cap; PDB code 2fko) at a resolution of 1.85 Å. The X-ray diffraction data and relevant statistics, including the Matthews coefficient (Matthews, 1968), are given in Table 1 and the details of the models that were built from this data, including the root-mean-square deviation (r.m.s.d.) from the ideal geometry for bond length, bond angles and dihedral angles (Engh & Huber, 1991), and analysis of the Ramachandran plot, are given in Table 2. To summarize, the final model of the native structure had a crystallographic $R_{work}$ of 18.5% ($R_{free}$ = 20.3%) for all reflections in the resolution range 20–1.5 Å and consisted of 1362 non-H atoms from 173 amino acids, 193 water molecules, an active-site $Zn^{2+}$ ion, four $Ca^{2+}$ ions, three ethylene glycol molecules, one glycerol molecule and one $Cl^-$ ion. The final refined protein model of the crystal structure of Zn-Cap-$HCO_3^-$ ($R_{work}$ = 18.9% and $R_{free}$ = 23.8%) consisted of 173 amino-acid residues, 142 water molecules, one active-site $Zn^{2+}$ ion, one $Ca^{2+}$ ion and a bound bicarbonate molecule. The model for the third crystal structure was refined to an $R_{work}$ of 19.2% ($R_{free}$ = 20.7%) and contained 1336 non-H atoms, 189 water molecules, two ethylene glycol molecules (ethylene glycol was added as the cryoprotectant at the time of data collection) and one catalytic $Zn^{2+}$ ion.

**Table 3**
R.m.s.d.s between various $\gamma$-class carbonic anhydrase structures.

| | R.m.s.d. (Å) | | |
|---|---|---|---|
| PDB code | Subunit *A* | $\beta$-Helix† | $\alpha$-Helix‡ |
| 1v3w§ | Fixed | Fixed | Fixed |
| 1v67 | 0.292 | 0.224 | 0.176 |
| 2fko | 0.226 | 0.183 | 0.108 |
| 1qq0 | 0.980 | 1.197 | 0.826 |
| 1thj | 1.083 | 1.238 | 0.899 |

† Each turn of the helix has been taken, including the loops connecting the turns; *i.e.* residues 2–134 for Zn-Cap and residues 10–169 for Zn-Cam. ‡ Residues 147–170 for Zn-Cap and 194–210 for Zn-Cam. § The high-resolution (1.5 Å) structure was used as a fixed molecule for all superposition calculations.

### 3.2. Overall structure

The overall fold of the protein is a regular prism, the cross-section of which appears to be an equilateral triangle. The present crystal structures reveal a monomer comprised of 19 parallel $\beta$-strands forming a left-handed $\beta$-helix leading to an antiparallel $\beta$-strand (residues 137–143) through a $\beta$-hairpin (Fig. 1). Along with a loop, this $\beta$-strand acts as a connector between the $\beta$-helix and the long $\alpha$-helix at the C-terminus (residues 147–170). On comparing the monomers by means of superimposition (Table 3), it can be seen that the Zn-Cap monomers (PDB codes 1v3w, 1v67 and 2fko) and Zn-Cam monomers (PDB codes 1qq0 and 1thj; Iverson *et al.*, 2000; Kisker *et al.*, 1996) show a low r.m.s.d. of about 1 Å. Furthermore, the structure-based sequence alignment of Zn-Cap with Zn-Cam (Fig. 2*a*) and the r.m.s.d. of the monomers without the loops (Table 3) indicate that the overall fold of Zn-Cap is similar to that of the Zn-Cam structure. In fact, most of the large differences between the two structures (high r.m.s.d. values in the graph shown in Fig. 2*b*) arise from tighter loops in Zn-Cap compared with Zn-Cam. However, one peak corresponds to the C-terminal $\alpha$-helix of Zn-Cap, which is relatively 'pushed away' from the axis of the $\beta$-helix (compared with Zn-Cam) owing to the difference in the connecting region. While Zn-Cam has a flexible loop and small $\alpha$-helix, Zn-Cap has a more rigid $\beta$-strand containing a classic $\beta$-bulge. However, the axis of the $\alpha$-helix later becomes antiparallel to the $\beta$-helix. This distortion at the base of the helices forms a closed pocket in the structure that is large enough to accommodate small molecules or ions. Furthermore, there are no significant overall structural differences between the

bicarbonate-bound (PDB code 1v67) and free (PDB codes 1v3w and 2fko) structures of Zn-Cap (Table 3).

All the structures have a conserved catalytic $Zn^{2+}$ ion along the $\beta$-sheets comprising the side of the prism-shaped monomers. In addition, a second metal-binding site is observed in the Zn-Cap structure at the $\beta$-strands that replace the loops in Zn-Cam. For example, one $Ca^{2+}$-bound site, the longest loop of Zn-Cap (residues 65–72), corresponds to a small $\alpha$-helix and a strand in Zn-Cam, which hinders the movement of the metal ion along that face. Thus, it is possible that $Ca^{2+}$ binding at this position is only possible in Zn-Cap and not in Zn-Cam. Having a trimer as the biological unit is consistent with the results of the dynamic light-scattering experiment and the native molecular weights of the previously reported structures (Fig. 3). The buried surface area of the generated trimer is 18 775 Å$^2$, which is less closely packed than the previously reported trimer (23 352 Å$^2$; Iverson *et al.*, 2000). In fact, unlike the packing of the earlier structures, which formed a hexamer, the trimers formed by the present structures are arranged in a tetrahedral fashion. Since most of the side chains of the $\beta$-helix point outwards, the interior of the $\beta$-helix is aliphatic and the surfaces of the centre of the trimer are hydrophobic. However, the interfaces have a number of common water molecules and a large number of interfacial residues form hydrogen bonds. Two conserved ion pairs (Asp46 OD2 with His166 NE2 and Asp60 OD2 with Arg44 NH2) and a number of water-mediated interactions are detected between the subunits. Both $Zn^{2+}$ and $Ca^{2+}$ ions (Figs. 4*a* and 4*b*) are coordinated by atoms from adjacent subunits of the trimer. Thus, a number of different forces contribute to the stability of the $\gamma$-class carbonic anhydrase trimer.

### 3.3. Active-site architecture

The active form of the Zn-Cap structure is a trimer, with each monomer contributing one $Zn^{2+}$ ion at the interface between adjacent monomers of the trimer. The superimposition of the active-site residues including the $Zn^{2+}$ ion between Zn-Cap and Zn-Cam is only 0.1 Å, indicating that the three-dimensional architecture of the active site is highly conserved. The $Zn^{2+}$ ion shows nearly the same coordination as observed in Zn-Cam and in structures from the other classes of carbonic anhydrases. The active-site $Zn^{2+}$ ion is coordinated to the side-chain atoms of three histidyl residues (His65 and His87 from subunit *A* and His82 from subunit *B*) and two functionally important water molecules. Literature studies indicate that the zinc-bound water is more important to the catalytic mechanism of carbonic anhydrases than the second water molecule, which is further away from the active site. This second water molecule forms a hydrogen bond to the side-chain OG atom of Thr66

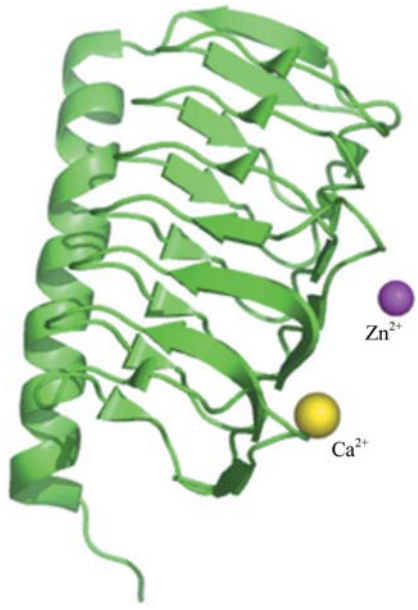


**Figure 1**
Three-dimensional representation of the Zn-Cap structure. The $Zn^{2+}$ and $Ca^{2+}$ metal ions are shown as magenta and yellow spheres.

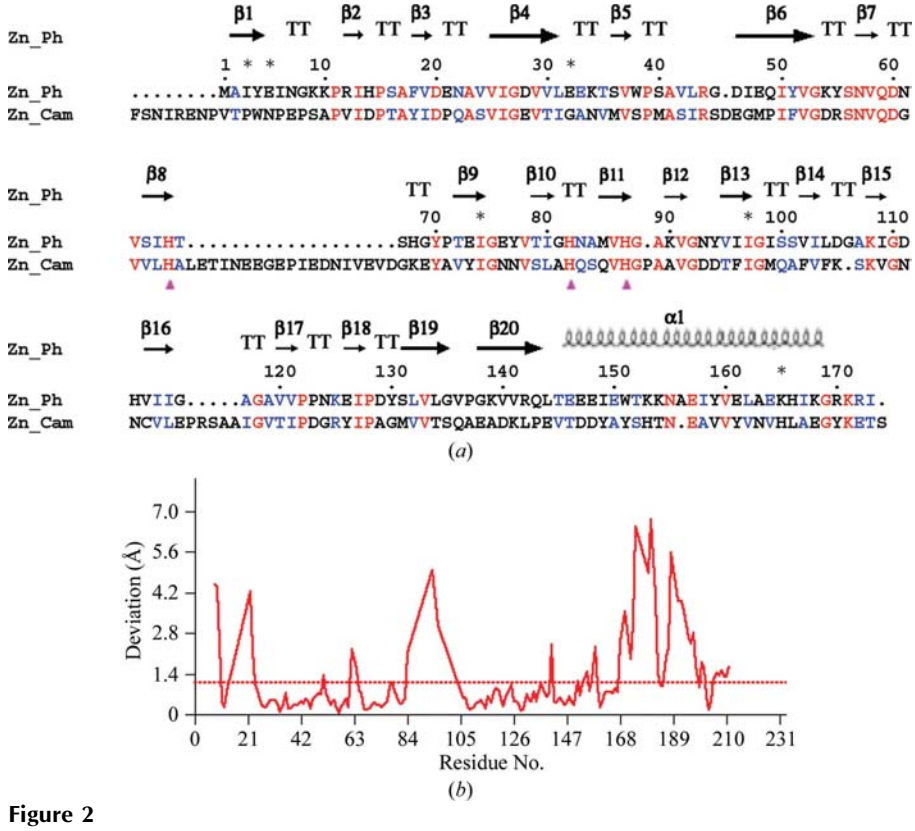


**Figure 2**
(*a*) Sequence alignment of Zn-Cap from *P. horikoshii* OT3 and Zn-Cam from *M. thermophila*: the secondary-structure elements are shown above the alignment and active-site residues are coloured. (*b*) Graph showing the r.m.s. deviation on superposition of the Zn-Cap and Zn-Cam monomers.

and is also involved in the formation of a network of hydrogen bonds around the active-site side-chain atoms. The possible mechanism of proton transfer may be as follows. Analysis of the three-dimensional crystal structures of Zn-Cap and Zn-Cam reveals that His68 is involved in the proton-transfer process in the case of Zn-Cap. However, the same role is carried out by Glu84 in the Zn-Cam structure (Tripp & Ferry, 2000). The loop residues Glu88 and Glu89 which have been shown to operate the proton transfer minimally in Zn-Cam and the corresponding loop are not found in the Zn-Cap structure. Interestingly, in the case of human carbonic anhydrase it has been observed that the similar process of proton transfer is carried out by His64 (Liang *et al.*, 1993).

The active sites of all the uncomplexed carbonic anhydrase structures are occupied by molecules used in the crystallization soup. In the present native structure, a glycerol molecule (which was used as a cryoprotectant during the data-collection procedure) binds near the active site and is also found to show clear electron density. In fact, the O atoms of the carboxyl groups of the glycerol molecule are coordinated to both the catalytically important zinc-bound water molecules. As a result, the active-site $Zn^{2+}$ ion and two water molecules are arranged in a distorted tetrahedral geometry in the native structure (Iverson *et al.*, 2000). On the other hand, in the bicarbonate-bound structure the active site is occupied by the substrate, the bicarbonate molecule. The three histidyl coordination contacts do not differ significantly between the free and ligand-complexed forms of Zn-Cap. While a previous study reported that two active-site water molecules are replaced by the bicarbonate, in this case the two O atoms of bicarbonate are coordinated in a bidenate fashion to the $Zn^{2+}$ ion, replacing only one water molecule (Fig. 4*a*) in Zn-Cap. One O atom forms hydrogen bonds to the side-chain atom (OD1) of Asn155 from the adjacent subunit, while another forms a bidentate hydrogen-bonding interaction with both the side-chain polar atoms of Gln59 (OE1 and NE2). It is interesting to note that the second water molecule (HOH109) interacts with the third bicarbonate O atom and a carboxylate O atom (OD2) of Asp105. A summary of the geometric parameters of $Zn^{2+}$ coordination in the free and complexed structures is given in Table 4.

### 3.4. C—H$\cdots\pi$ interactions in the active site

Interestingly, in both the native and complexed forms, the phenolic hydroxyl of Tyr159 provides coordination to $Zn^{2+}$ of the adjacent monomer and forms an ion pair with the NE2

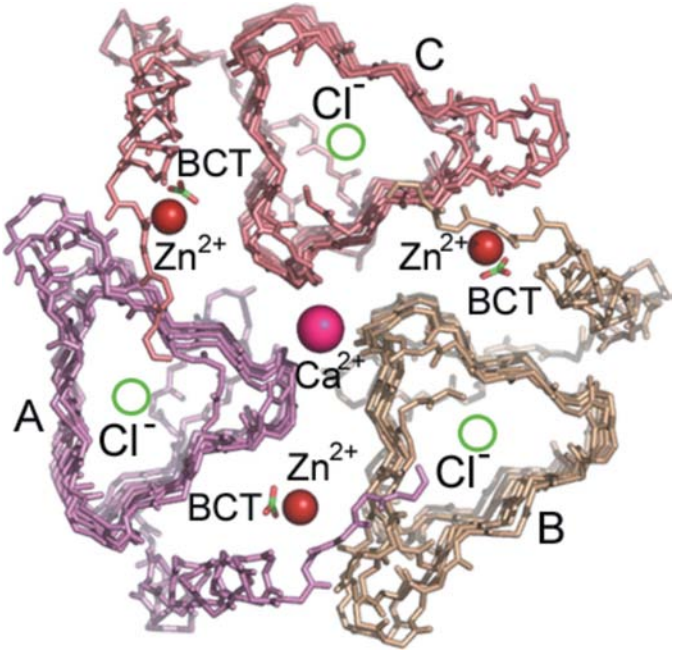


**Figure 3**
In this representation of the Zn-Cap trimer (generated from 1v67), metal ions are shown as spheres. The central $Ca^{2+}$ ion is in magenta, the $Zn^{2+}$ ions in brick red and the positions of the $Cl^{-}$ ions from a 1v3w trimer are shown as green circles.

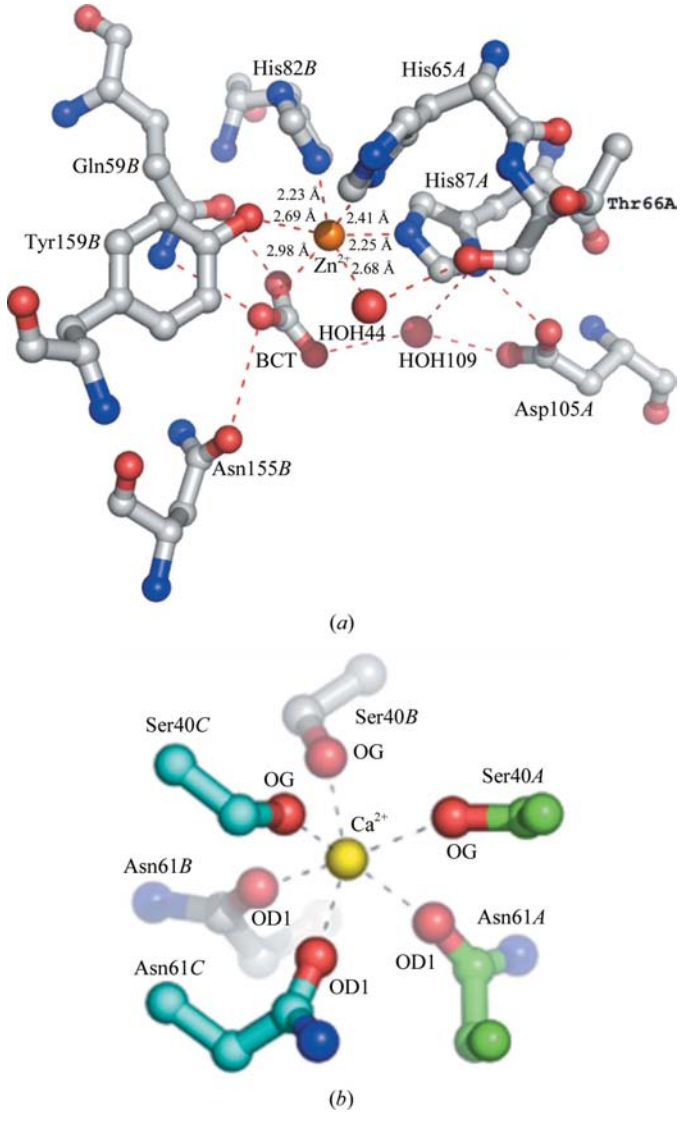


**Figure 4**
Metal-ion coordination. (*a*) $Zn^{2+}$ is coordinated to His65*A* ND1, His87*A* NE2, His82*B* NE2, HOH44 and Tyr159*B* OH. The OH is at a distance of 2.71 Å from $Zn^{2+}$. (*b*) The $Ca^{2+}$ ion is coordinated to Ser40 OG and Asn61 OD1 from all three chains.

**Table 4**
Active-site geometry.

(*a*) Zn-Cap.

| Atom (*X*) | Distance to Zn (Å) | *X*—Zn— second water (°) | *X*—Zn— His65 (°) | *X*—Zn— His87 (°) | *X*—Zn— His82 (°) |
|---|---|---|---|---|---|
| Zinc-bound water | 1.83 | 72.36 | 109.40 | 135.56 | 102.13 |
| Second water molecule | 2.82 | | 78.39 | 81.08 | 173.79 |
| His65 | 2.23 | | | 98.7 | 101.23 |
| His87 | 2.01 | | | | 105.07 |
| His82 | 2.17 | | | | |

(*b*) Zn-Cap-$HCO_3^-$.

| Atom (*X*) | Distance to Zn (Å) | *X*—Zn— BctO1 (°) | *X*—Zn— His65 (°) | *X*—Zn— His87 (°) | *X*—Zn— His82 (°) |
|---|---|---|---|---|---|
| BctO3 | 2.98 | 97.32 | 158.86 | 120.80 | 114.99 |
| BctO1 | 3.41 | | 159.65 | 98.40 | 84.51 |
| His65 | 2.23 | | | 75.76 | 76.03 |
| His87 | 2.01 | | | | 88.25 |
| His82 | 2.17 | | | | |

side-chain atoms of Gln59 and His82 of the same monomer. Furthermore, the OE1 side-chain atom of Gln59 (which is coordinated to the bicarbonate) also forms a $\pi$ interaction with the His82 residue from the bicarbonate-bound subunit. Additionally, the $\pi$-electron cloud of Tyr51 is involved in an O—H$\cdots\pi$ interaction with the bicarbonate in Zn-Cap-$HCO_3^-$ and a $\pi$-stacking interaction with Tyr159. Thus, it is likely that Tyr159, which corresponds to Asn202 in previously reported structures (Zimmerman & Ferry, 2006), serves the same functions through a different set of interactions, namely utilizing the extensive hydrogen bonding as well as C—H$\cdots\pi$ interactions.

### 3.5. $Ca^{2+}$ ion at the interface between subunits

A $Ca^{2+}$ ion has been observed for the first time at the interface between the monomers among the different classes of carbonic anhydrases (Kannan *et al.*, 1984; Hakansson *et al.*, 1992; Eriksson & Liljas, 1993; Boriack-Sjodin *et al.*, 1995; Stams *et al.*, 1996; Mitsuhashi *et al.*, 2000). In Zn-Cap, four $Ca^{2+}$ ions and a chloride ion are bound to the interface between the monomers, while in the bicarbonate-bound structure only one $Ca^{2+}$ ion is found. Although any metal could theoretically bind to this site, no lithium was found in the structures even for the crystal that was soaked in LiCl. Thus, the size of the metal atom may act as a filter for binding to this second site. The presence of $Ca^{2+}$ and chloride ions in the enzyme could arise from the use of $CaCl_2$ in the crystallization buffer. In order to differentiate between the two ions, a computational study was carried out (as described in §2) and the ions were assigned based on the coordinate distances (calcium–ligand distances are between 2.26 and 2.52 Å and chloride–ligand distances are between 3.16 and 3.26 Å).

The $Ca^{2+}$ ion that is invariant between the native and bicarbonate structures is found in a central position of the trimer (Fig. 3) on the threefold axis and is hexacoordinated to the side-chain atoms Ser40 OG and Asn61 OD1 (Fig. 4*b*) and a bridged water network. Since the invariant $Ca^{2+}$ ion is quite far from the enzyme active site, it is likely that the $Ca^{2+}$ ion is not involved in the catalytic mechanism. However, this $Ca^{2+}$ ion is bound to a polar valley (Fig. 5) and thus possibly plays a role in substrate entry and product exit. One possible mode of exit could be as follows: the $Zn^{2+}$-bound bicarbonate (stabilized by a Tyr51 O—H$\cdots\pi$ interaction) could move towards Tyr159, forming a hydrogen bond and a stronger C—H$\cdots\pi$ interaction. The $Ca^{2+}$ ion could then take over the tethering function of Tyr159 by forming electrostatic interactions with the bicarbonate and the product could then exit utilizing the hydrogen-bond network along the polar valley. In any case, the second metal-binding pocket and coordination of a second metal ion could contribute to the stabilization of the enzyme.

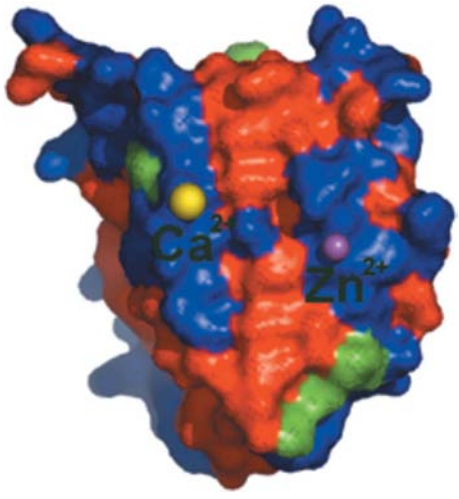


**Figure 5**
Surface of the native structure coloured by hydrophobicity. $Ca^{2+}$ (large yellow sphere) is bound to one polar valley and $Zn^{2+}$ (smaller magenta sphere) to another.

## 4. Conclusion

A second metal-binding site was found in the crystal structures of the $\gamma$-class carbonic anhydrase from *P. horikoshii*, which forms a regular prism-like structure. Although this is highly similar to the previously reported structure, a few differences are observed. The absence of loops probably allows the binding of the second metal ion. The interactions in the active site of the free and bound structures are almost the same. Owing to the replacement of the catalytically important Asn202 by Tyr159, an alternative set of interactions come into play for $HCO_3^-$ and $CO_2$ tethering that involve interaction with the $\pi$-electron cloud. A mechanism for the exit of the product is proposed based on the second metal-binding site and the $\pi$ interactions. The stability of the trimer and indeed the monomer is likely to be a consequence of the extensive formation of ion pairs and water-mediated interactions in the proteins.

The authors (RS, PM, SPK and KS) thank the Bioinformatics Centre (DIC), Bangalore, India and the Super-

computer Education and Research Centre (SERC), Bangalore, India. JJ thanks Professor T. H. Tahirov for useful discussion during the structure determination and the staff at beamlines BL26B1 and BL26B2 of SPring-8 for excellent facilities and timely assistance. This work was supported by the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

Alber, B. E., Colangelo, C. M., Dong, J., Stålhandske, C. M., Baird, T. T., Tu, C., Fierke, C. A., Silverman, D. N., Scott, R. A. & Ferry, J. G. (1999). *Biochemistry*, **38**, 13119–13128.
Alber, B. E. & Ferry, J. G. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 6909–6913.
Badger, M. R. & Price, G. D. (1994). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **45**, 369–392.
Balamurugan, B. *et al.* (2007). *J. Appl. Cryst.* **40**, 773–777.
Boriack-Sjodin, P. A., Heck, R. W., Laipis, P. J., Silverman, D. N. & Christianson, D. W. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 10949–10953.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, 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. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chegwidden, W. R. & Carter, N. D. (2000). *The Carbonic Anhydrases: New Horizons*, edited by W. R. Chegwidden, N. D. Carter & Y. H. Edwards, pp. 13–28. Basel: Birkhauser Verlag.
Coleman, J. R. (1991). *Plant Cell Environ.* **14**, 861–867.
Covarrubias, A. S., Bergfors, T., Jones, T. A. & Hogbom, M. (2006). *J. Biol. Chem.* **281**, 4993–4999.
Cronk, J. D., Rowlett, R. S., Zhang, K. Y. J., Tu, C., Endrizzi, J. A., Lee, J., Gareiss, P. C. & Preiss, J. R. (2006). *Biochemistry*, **45**, 4351–4361.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, Palo Alto, California, USA.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Eriksson, A. E. & Liljas, A. (1993). *Proteins*, **16**, 29–42.
Hakansson, K., Carlsson, M., Svensson, L. A. & Liljas, A. (1992). *J. Mol. Biol.* **227**, 1192–1204.
Hewett-Emmett, D. & Tashian, R. (1996). *Mol. Phylogenet. Evol.* **5**, 50–77.
Hussain, A. S., Kumar, C. K., Rajesh, C. K., Sheik, S. S. & Sekar, K. (2003). *Nucleic Acids Res.* **31**, 3356–3358.
Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.
Iverson, T. M., Alber, B. E., Kisker, C., Ferry, J. G. & Rees, D. C. (2000). *Biochemistry*, **39**, 9222–9231.
Kannan, K. K., Ramanadham, M. & Jones, T. A. (1984). *Ann. NY Acad. Sci.* **429**, 49–60.
Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G. & Rees, D. C. (1996). *EMBO J.* **15**, 2323–2330.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Liang, Z., Jonsson, B. H. & Lindskog, S. (1993). *Biochim. Biophys. Acta*, **1203**, 142–146.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Meldrum, N. U. & Roughton, F. J. W. (1933). *J. Physiol.* **80**, 113–141.
Mitsuhashi, S., Mizushima, T., Yamashita, E., Yamamoto, M., Kumasaka, T., Moriyama, H., Ueki, T., Miyachi, S. & Tsukihara, T. (2000). *J. Biol. Chem.* **275**, 5521–5526.
Navaza, J. (1994). *Acta Cryst.* A**50,** 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sawaya, M. R., Cannon, G. C., Heinhorst, S., Tanaka, S., Williams, E. B., Yeates, T. O. & Kerfeld, C. A. (2006). *J. Biol. Chem.* **281**, 7546–7555.
Smith, K. S., Jakubzick, C., Whittam, T. S. & Ferry, J. G. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 15184–15189.
Stams, T., Nair, S. K., Okuyama, T., Waheed, A., Sly, W. S. & Christianson, D. W. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 13589–13594.
Sumathi, K., Ananthalakshmi, P., Rosha, M. N. A. Md & Sekar, K. (2006). *Nucleic Acids Res.* **34**, W128–W132.
Tashian, R. E. (1989). *Bioessays*, **10**, 186–192.
Tripp, B. C. & Ferry, J. G. (2000). *Biochemistry*, **39**, 9232–9240.
Zimmerman, S. A. & Ferry, J. G. (2006). *Biochemistry*, **45**, 5149–5157.