Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of bovine carbonic anhydrase II at 1.95 Å resolution

Ryuta Saito, Takao Sato, Atsushi Ikai* and Nobuo Tanaka

Department of Life Science, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Japan

Correspondence e-mail: aikai@bio.titech.ac.jp

Received 21 September 2003
Accepted 10 February 2004

PDB Reference: bovine carbonic anhydrase II, 1v9e, r1v9esf.

Carbonic anhydrase (CA) is a zinc-containing enzyme that catalyzes the reversible hydration of CO_2 to HCO_3^-. In eukaryotes, the enzyme plays a role in various physiological functions, including interconversion between CO_2 and HCO_3^- in intermediary metabolism, facilitated diffusion of CO_2, pH homeostasis and ion transport. The structure of bovine carbonic anhydrase II (BCA II) has been determined by molecular replacement and refined to 1.95 Å resolution by simulated-annealing and individual *B*-factor refinement. The final *R* factor for the BCA II structure was 19.4%. BCA II has a C-terminal knot structure similar to that observed in human CA II. It contains one zinc ion in the active site coordinated to three histidines and one putative water molecule in a tetrahedral geometry. The structure of BCA II reveals a probable alternative proton-wire pathway that differs from that of HCA II.

1. Introduction

Carbonic anhydrases (CAs) are zinc metalloenzymes that can be grouped into three broad classes called α, β and γ, with very little or no amino-acid sequence homology between them (Hewett-Emmett & Tashian, 1996). The α class includes mammalian enzymes such as the bovine carbonic anhydrases (BCAs) used in this study. Class α CAs are monomeric proteins with a molecular weight of around 30 kDa. At least 14 α CA isozymes have been identified so far (Parkkila, 2000). These catalyse the same reversible reaction $CO_2 + H_2O \rightarrow HCO_3^- + H^+$, but differ in catalytic efficiency and inhibitor-binding properties. The kinetic properties of cytosolic CA II, one of seven isozymes (CA I to VII), show it to be one of the most efficient catalysts known, along with the catalases (Stainer *et al.*, 1975).

Since the report by Wong & Tanford (1973) of intermediate states in the transition from the native to the denatured state or *vice versa*, intensive investigations have focused on BCA II and human CA II (HCA II) as models for characterizing the partially folded or unfolded states of globular proteins (McCoy & Wong, 1981; Cleland & Wang, 1990; Hammarstorm *et al.*, 1997). BCA II is also a good model for its human homologue, with which its sequence homology is 79.5%. The structural stability of BCA II has also been investigated by atomic force microscopy. In this case, the protein was mechanically unfolded by pulling its N- and C-termini apart, revealing the presence of a tight core structure that showed a remarkable resistance towards tensile stretching (Alam & Ikai, 2001; Wang *et al.*, 2002).

An earlier X-ray structural study of HCA II revealed it to have a knot-forming topology in its C-terminal region (Eriksson *et al.*, 1988). More recently, crystal structures of murine CA isozymes IV and V and BCA isozyme III have also been determined (Stams *et al.*, 1998; Heck *et al.*, 1996; Eriksson & Liljas, 1993). The atomic structure of BCA II is required for full interpretation of the data from the single-molecule stretching experiments and to confirm the C-terminal knot structure. There have been reports of BCA II crystallization (Carlsson *et al.*, 1973; Kumar *et al.*, 1986) but no structural coordinates have been made available, even though the BCA II structure was determined at 2.8 Å resolution (Kumar *et al.*, 1989). Therefore, we set to work on solving the crystal structure of BCA II.

2. Materials and methods

2.1. Crystallization of commercial BCA II

Bovine erythrocyte BCA II was purchased from Biozyme Laboratories (London, England) and purified by gel-filtration chromatography on a Superdex-75 column (Amersham Biosciences, Buckinghamshire, England). BCA II crystals were grown by the batch method at 277 K using 17 mg ml^{-1} protein solution and a precipitant solution containing 50 m*M* Tris buffer pH 7.5 and 2.4 *M* ammonium sulfate. Crystals of typical dimensions 0.2 × 0.3 × 0.3 mm appeared after two weeks. These crystals belonged to the hexagonal space group $P6_1$, with unit-cell parameters $a = b = 66.7$, $c = 240.0$ Å and two BCA II molecules per asymmetric unit. The unit-cell

Table 1
Crystallographic data and refinement statistics.

Values in parentheses are for the outermost shell.

Unit-cell parameters	
$a = b$, c (Å)	66.7, 240.0
α, β, γ (°)	90, 90, 120
Space group	$P6_1$
Z value	12
V_M (Å^3 Da^{-1})	2.66
No. protein atoms	4114
No. heteroatoms	2
No. water molecules	573
Resolution range (Å)	20.0–1.95 (2.06–1.95)
Total No. measurements	213573
No. unique reflections	35487
$I/\sigma(I)$	6.4 (2.3)
Completeness (%)	90.6 (90.6)
R_{sym}	0.076 (0.284)
R factor	0.194
R_{free}	0.239
Average B factors (Å^2)	22.16
Zn atoms	15.79
Water atoms	29.79
R.m.s. deviation from ideal geometry	
Bond lengths (Å)	0.005
Bond angles (°)	1.42
Dihedrals (°)	24.7
Impropers (°)	0.832
Luzzati coordinate error (Å)	0.22
σ_A coordinate error (Å)	0.16
Ramachandran outliers	1

volume was 9.25×10^5 Å^3 which, on the basis of two molecules of molecular weight 29 kDa each, gives a V_M value of 2.66 Å^3 Da^{-1}.

2.2. Dynamic light scattering

12 samples of BCA II were prepared varying the pH and the precipitant and protein concentrations around the conditions described above and were examined as possible inducers of protein aggregation in mother liquor. Each sample was made up to 40 µl volume in a standard fluorescence cuvette and was kept at 281 K. To monitor the crystallization process in these solutions, a dynamic light-scattering system was used consisting of an ALV-5000 system (ALV, Langen, Germany) coupled with a He–Ni laser (λ = 633 nm). The total scattering intensities were measured and integrated five times in succession for each sample. Each autocorrelation function was analysed using the *CONTIN* program to obtain the average hydrodynamic radius and its distribution. In the evolution of a supersaturated solution of BCA II, monomers first assembled to form aggregates, which then led to crystal growth. Only when the aggregate size was larger than approximately 3000–4000 nm in diameter did nucleation of crystal growth follow.

2.3. Data collection and processing

Crystals of BCA II were flash-frozen by transfer into an artificial mother liquor containing 25%(*v*/*v*) glycerol as a cryoprotectant followed by freezing in an N_2 stream at 100 K. The diffraction data were collected using synchrotron radiation (λ = 1.00 Å) with an ADSC Quantum 4R CCD detector (ADSC, California, USA) on beamline 6A at the Photon Factory, Tsukuba, Japan. The data were processed, integrated and scaled with *MOSFLM* and *SCALA* from the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). Integrated intensities were converted into structure factors using *TRUNCATE* from the *CCP*4 suite. The data-collection and processing statistics are given in Table 1.

2.4. Structure solution and refinement

Initial phases for BCA II were obtained by molecular replacement using the *EPMR* program (Kissinger *et al.*, 1999). The coordinates of the HCA II structure (protein molecule only without inhibitor or water molecules; PDB code 2cbb) were used as an initial search model. The amino-acid sequences of HCA II and BCA II were compared and 52 non-identical residues were replaced with alanines or serines. The initial R factor and correlation coefficients for data in the resolution range 15.0–4.0 Å were 0.508 and 0.357, respectively, when a single 2cbb molecule was inserted and 0.412 and 0.609, respectively, when two molecules were inserted.

The structural model of BCA II was then refined within the resolution range 20.0–2.3 Å using *CNS* (Brünger *et al.*, 1998) by iterative cycles of simulated annealing and individual B-factor refinement, with manual rebuilding using *XtalView* (McRee, 1992). The model was initially refined with strict non-crystallographic symmetry (NCS) relating the two molecules in the asymmetric unit. The 52 alanines or serines in the initial BCA II model were replaced with their correct amino-acid residues when their positions were clearly indicated by electron density in $2|F_o| - |F_c|$ or $|F_o| - |F_c|$ difference Fourier maps. The final model contained residues 1–259 for two molecules, with no chain breaks. All residues except the 258th have main-chain dihedral angles that fall within the allowed regions of the Ramachandran diagram as defined by the program *PROCHECK* (Laskowski *et al.*, 1993). The asymmetric unit also contained two Zn^{2+} ions and 573 water molecules (Table 1). R and R_{free} were reduced to 0.194 and 0.239, respectively, in the resolution range 20.0–1.95 Å. The free R value was calculated with 5% (2515) of the total reflections.

Of the 573 solvent molecules identified in the electron-density map, all were within proper hydrogen-bonding distances of the protein (<3.4 Å or >2.4 Å) or other solvent molecules, with refined B factors of less than 50 Å^2. The average B factors of the main-chain atoms of individual residues varied between 20.2 and 21.9 Å^2, compared with 22.2 Å^2 for the overall molecule.

2.5. Semi-empirical quantum-mechanical calculation

To analyse the chemistry in the catalytic sites, theoretical computations were performed using v.3.5 of *WinMOPAC* Pro (Fujitsu Ltd, Tokyo, Japan), based on semi-empirical quantum-mechanical calculation. Optimizations were performed with the eigenvalue-following method (EF) and an SCF field was achieved. The model for the catalytic site was studied, including the central Zn^{2+} ion. The first coordinate shell ligands consisted of residues His93, His95 and His118 and the oxygen of water molecule 419 (W419), the second-shell coordination sphere consisted of Gln91, Glu105, Glu116, Thr197 and W33 and the outer shell consisted of Tyr6, Asn61, His63, Asn66, Thr198, Arg244 and nine water molecules. The system has an overall charge of +1 (MOPAC keyword CHARGE = +1). We performed EF geometry optimizations on all possible protonated forms and atomic charges (MOPAC keyword = AM1).

3. Results and discussion

3.1. Comparison of BCA II structure with HCA II

The refined structure of BCA II is a complete protein model that includes several residues that were not observed in HCA II (PDB code 1ca2; Eriksson *et al.*, 1988; PDB code 1moo; Duda *et al.*, 2003). The N-terminal residues Ser1 and His2 were clearly located in BCA II, whereas they were in flexible terminal regions and were difficult to locate in HCA II. The root-mean-square (r.m.s.) difference in C^α-atom positions between the two BCA II monomers was found to be 0.11 Å, showing that the two molecules in the asymmetric unit are almost identical. The number of bound water molecules was different, however, being 277 for the A monomer and 300 for the B monomer. The r.m.s. difference in backbone atom positions is 0.58 Å between HCA II and BCA II, showing that the structure of BCA II is very similar to that of HCA II. The typical motifs of HCA II are preserved: the central fold of BCA II consists of ten

strands of parallel or antiparallel β-sheet forming a core structure, as shown in Fig. 1(*a*).

BCA II was found to contain three *cis*-prolines, namely Pro29, Pro200 and Pro258, in 19 proline residues (Kumar *et al.*, 1989), compared with two *cis*-prolines in HCA II (Eriksson *et al.*, 1988). The extra *cis*-proline, Pro258, interacts with the side chain of *trans*-Pro41 on the turn intersection to the C-terminus.

The C-terminal region of HCA II is characterized by a unique 'knot' topology, meaning that the polypeptide chain crosses itself in such a manner that one chain segment goes 'below–above–below' the other segment. We have confirmed that the 'knot' topology is maintained in BCA II. In BCA II, some new interactions related to the 'knot' topology were found, including a salt bridge between Arg255 $N^{\eta1}$ and Asp40 $O^{\delta1}$ and a hydrogen bond between Pro258 and Pro41. We suggest that the 'knot' topology of BCA II is thus tighter and more stable than that of HCA II. The consequence of this 'knot' structure is that if the polypeptide chain is pulled from both ends, the molecule will become tightened rather than loosened and, after the structure is completely destabilized, the chain will remain as a straight chain with a single knot. This structure provides definite support for the interpretation of the experimental results of mechanical stretching in terms of the presumed 'knot' topology in BCA II (Wang *et al.*, 2001, 2002; Alam & Ikai, 2001; Alam *et al.*, 2002).

3.2. Active site

The active site of BCA II forms a funnel-shaped channel with an outer diameter of ~13 Å that extends from the molecular surface to the centre, with the zinc ion located at the bottom of the channel at a depth of ~10 Å from the surface. In Fig. 1(*b*) the active site Zn^{2+} ion is coordinated to three histidyl residues (His93, His95 and His118). The fourth ligand coordinated to the Zn^{2+} ion is a water molecule (W419 in molecule *A* and W420 in molecule *B*), the so-called 'zinc water' (Eriksson *et al.*, 1988). No electron density could be found for a water molecule corresponding to the so-called 'deep water' in the HCA II molecule, however.

The zinc coordination is almost tetrahedral and the zinc coordination distances, angles and deprotonation energies are shown in Table 2. The three histidines have very stable conformations, as indicated by their low thermal *B* factors and formation energies. After optimization of the structure by molecular mechanics, the ligand-binding affinity can be estimated, showing net gains of 8.8–82 kJ mol^{-1} per imidazole when the tetracoordinated aqua complex of Zn^{2+} ion is compared with structures containing one, two and three imidazoles. The spontaneous deprotonation of the zinc water is more exothermic than the deprotonation of the imidazole N atoms of His93, His118 and His95 in descending order. We think that this stability results from the fact that the three histidyl residues have their second imidazole N atoms hydrogen bonded to other residues: His93 $N^{\delta1}$ to Gln91 $O^{\varepsilon1}$, His95 $N^{\delta1}$ to Glu105 $O^{\varepsilon2}$ and His118 $N^{\varepsilon2}$ to Glu116 $O^{\varepsilon2}$.

The number of hydrogen bonds formed suggests that the carboxylic groups of Glu105 and Glu116 are negatively charged as predicted by *MOPAC*, with a partial charge of −0.53. With this assumption, it is possible to assign the donor–acceptor relationships of several hydrogen bonds in the active site. The $O^{\gamma1}$ proton of Thr197 is hydrogen bonded to Glu105 $O^{\varepsilon1}$; Thr197 $O^{\gamma1}$ can thus only be a hydrogen-bond acceptor for the proton of the zinc-bound hydroxyl ion.

In addition to the zinc-bound water molecule, nine and ten water molecules are found in the active sites of the two inde-

Table 2
Coordinate structure of the zinc environment in BCA II.

Active site	Ligand *X*	Deprotonation (kJ mol^{-1})	Bond distance (Å)	Bond angle (°)		
				X—Zn—His93	*X*—Zn—His95	*X*—Zn—His118
Zn	W204 OH^-	1000	2.19	101.3	123.0	109.2
	His93 $N^{\varepsilon2}$	−432	2.11	—	109.5	115.7
	His95 $N^{\varepsilon2}$	143	2.15	—	—	99.0
	His118 $N^{\delta1}$	−374	2.26	—	—	—

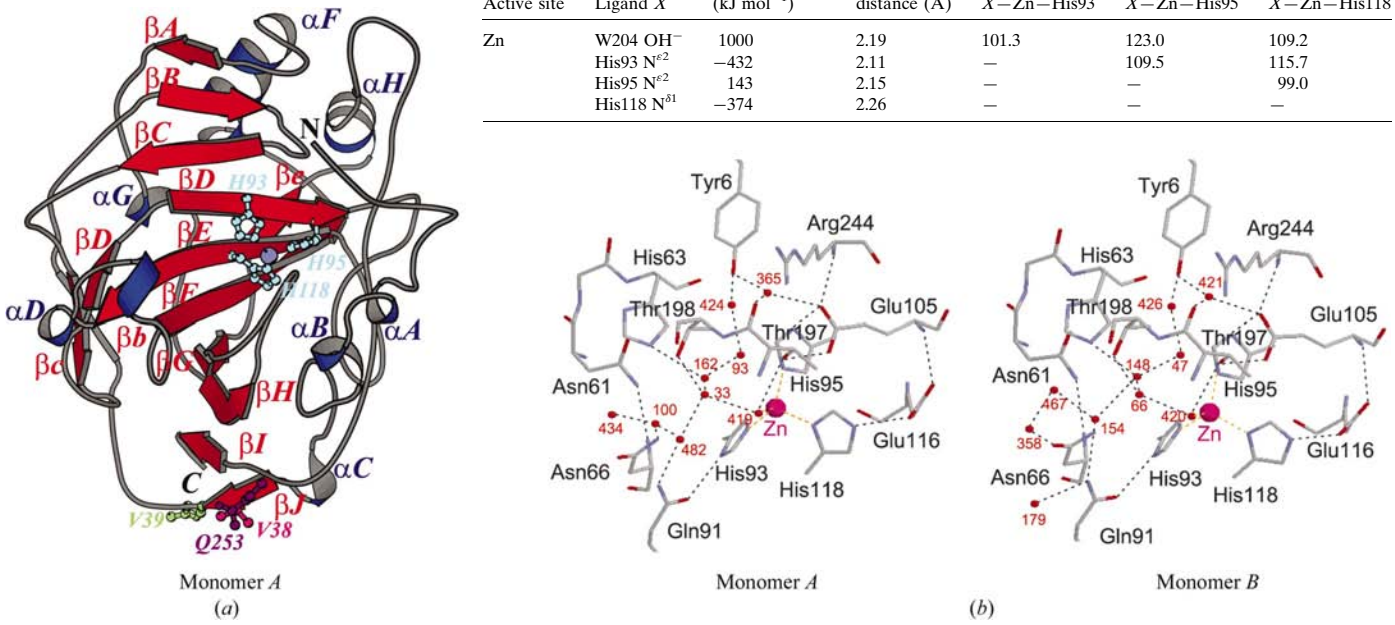

Figure 1
Crystal structure of BCA II molecule. (*a*) Ribbon diagram showing the BCA II monomer. The colour-coding of the secondary elements is as follows: red, β-strands; blue, α-helices; grey, coil. The position of the zinc ion (lilac) with three histidine ligands (sky blue) and the N- and C-termini are indicated. The core of the monomer consists of ten β-sheet strands. (*b*) Active-site residues and bound water molecules within 10 Å of the Zn^{2+} ion. Water molecules are shown as small red spheres, with hydrogen bonds shown by dotted lines. The zinc-bound water ligand is W419 in the *A* monomer and W420 in the *B* monomer. The hydrogen-bond network of water molecules differs near Gln91, with the number of bound water molecules in the active sites being nine and ten, respectively, for the *A* and *B* monomers. Figure created using *BOBSCRIPT* (Esnouf, 1997) and *RASMOL*.

pendent BCA II molecules, respectively, as shown in Fig. 1(*b*). These water molecules were predicted by *MOPAC* to have zero or slightly positive electric charges and to form the electric slope required for proton transfer. Only the oxygen of the zinc-bound hydroxyl ion is negatively charged, with a partial charge of −0.50 as predicted by *MOPAC* (see Fig. 2). Several potential proton wires can be envisaged leading from the zinc-bound H_2O. One leads to His63 *via* W33 and W162 (molecule *A*) or W66 and W168 (molecule *B*) in a manner similar to that in HCA II (Eriksson *et al.*, 1988). A second involves a water molecule bound to Gln91 (W100 in molecule *A*, W154 in molecule *B*) which is part of a potential proton wire consisting of W434-W100-W482-W33-W419. Switching may occur between the various proton wires during the proton-transfer process, making use of the extensive water network leading away from the zinc ion.

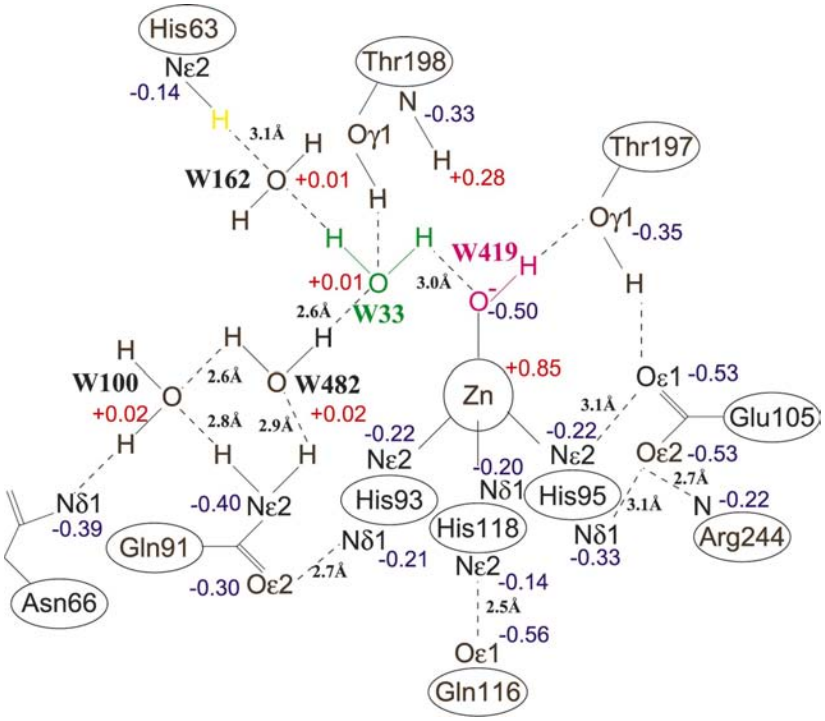

Figure 2
Schematic drawing of interactions and distances around the active site determined from the crystal structure of BCA II at pH 7.5. The net charges assigned by theoretical calculations using *MOPAC* are shown. The side-chain atom Gln91 $O^{\varepsilon 1}$ (net charge −0.30) accepts a hydrogen bond from of His93 $N^{\delta 1}$ (net charge −0.21) and contributes to His93 ligand stabilization. The side-chain atom of Gln91 $N^{\varepsilon 2}$, with a net charge of −0.40, has a dominant role in the binding of water molecule W482, with net charge of +0.02, in its slightly acidic or cationic form. This interaction is likely to be more hydrogen-bonding in character than the interaction between W162 (net charge +0.01) and His63 $N^{\varepsilon 2}$ (net charge −0.14) that has been a biological focus in the case of HCA II. This finding suggests that the dipole donor group of Gln91 may also participate in processes that require relatively rapid proton movement or release.

4. Conclusions

We report here the structure of BCA II at 1.95 Å resolution, representing the highest resolution structure of BCA II reported to date. As expected, the overall refined structure of this protein is very similar to that of HCA II and confirms the presence of the C-terminal knot structure in BCA II. This is important for interpretation of their mechanical responses. The high level of activity for BCA II is most likely to be a consequence of another water pathway in the active site forming a 'secondary' proton wire. We have also performed theoretical calculations on the active-site cavity and carried out automated docking studies of inhibitor binding. The structure gives insight into the existence of a secondary proton wire (W482, W100 and W434).

The present research was undertaken with the approval of the Photon Factory Advisory Committee, Japan. The authors wish to express their sincere thanks to the staff at the Photon Factory for their help with the data collection. This work was partly supported by grants-in-aid to AI from the Japan Society for the Promotion of Science (Research for the Future Program No. 99R16701) and the Japanese Ministry of Education, Culture, Sports, Science and Technology [Scientific Research on Priority Areas (B) No. 11226202] and by a grant from the 21st Century COE Program and Grant-in-Aid to TS for Scientific Research on Priority Areas (A) (No. 13033011) from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

Alam, M. T. & Ikai, A. (2001). *Appl. Phys. A*, **72**, S121–S124.
Alam, M. T., Yamada, T. & Ikai, A. (2002). *FEBS Lett.* **519**, 35–40.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gross, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Carlsson, U., Lindskog, S., Andersson, U., Lindqvist, O. & Olsson, G. (1973). *J. Mol. Biol.* **80**, 373–375.
Cleland, J. L. & Wang, D. I. (1990). *Biochemistry*, **29**, 11072–11078.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Duda, D., Govindasamy, L., Agbandje-McKenna, M., Tu, C., Silverman, D. N. & McKenna, R. (2003). *Acta Cryst.* D**59**, 93–104.
Eriksson, A. E., Jones, T. A. & Liljas, A. (1988). *Proteins*, **4**, 274–282.
Eriksson, A. E. & Liljas, A. (1993). *Proteins*, **16**, 29–42.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Hammarstrom, P., Kalman, B., Jonsson, K. P. & Carlsson, U. (1997). *FEBS Lett.* **420**, 63–68.
Heck, R. W., Boriack-Sjodin, P. A., Qian, M., Tu, C., Christianson, D. W., Laipis, P. J. & Silverman, D. N. (1996). *Biochemistry*, **35**, 11605–11611.
Hewett-Emmett, D. & Tashian, R. E. (1996). *Mol. Phylogenet. Evol.* **5**, 50–77.
Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999) *Acta Cryst.* D**55**, 481–491.
Kumar, V., Kannan, K. K. & Chidambaram, R. (1989). *Curr. Sci.* **58**, 344–348.
Kumar, V., Sankaran, K. & Kannan, K. K. (1986). *J. Mol. Biol.* **190**, 129–131.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
McCoy, L. F. Jr & Wong, K. P. (1981). *Biochemistry*, **20**, 3062–3067.
McRee, D. E. (1992). *J. Mol. Graph.* **10**, 44–47.
Parkkila, S. (2000). *The Carbonic Anhydrases: New Horizons*, edited by W. R. Chegwidden, N. D. Carter & Y. H. Edwards, pp. 79–94. Berlin: Birkhauser Verlag.
Stainer, H., Jonsson, B. H. & Lindskog, S. (1975). *Eur. J. Biochem.* **59**, 253–259.
Stams, T., Chen, Y., Boriack-Sjodin, P. A., Hurt, J. D., Liao, J., May, J. A., Dean, T., Laipis, P., Silverman, D. N. & Christianson, D. W. (1998) *Protein Sci.* **7**, 556–563.
Wang, T., Arakawa, H. & Ikai, A. (2001). *Biochem. Biophys. Res. Commun.* **285**, 9–14.
Wang, T., Arakawa, H. & Ikai, A. (2002). *Ultramicroscopy*, **91**, 253–259.
Wong, K. P. & Tanford, C. (1973). *J. Biol. Chem.* **248**, 8518–8523.