Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of human carbonic anhydrase III

David M. Duda,[a] Craig Yoshioka,[a] Lakshmanan Govindasamy,[a] Haiqian An,[b] Chingkuang Tu,[b] David N. Silverman[b] and Robert McKenna[a]*

[a]Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610, USA, and [b]Department of Pharmacology and Therapeutics, University of Florida, Gainesville, FL 32610, USA

Correspondence e-mail: rmckenna@ufl.edu

Received 12 December 2001
Accepted 26 February 2002

Carbonic anhydrases catalyze the interconversion of carbon dioxide to bicarbonate. Human carbonic anhydrase isozyme III with a C-terminal hexahistidine tag was overexpressed in *Eschericha coli*, purified and crystallized. Diffraction data (93.4% completeness) were collected to 2.2 Å resolution on an in-house R-AXIS IV++ image-plate system with Osmic mirrors and a Rigaku HU-H3R CU rotating-anode generator operating at 50 kV and 100 mA. A 60° sweep of data were collected from a single crystal with a crystal-to-detector distance of 150 mm and a 0.5° oscillation angle per frame using an exposure of 60 s per frame at 293 K. The crystals were shown to conform to the Laue hexagonal crystal system $P6$, with unit-cell parameters a = 44.7, c = 222.5 Å and a scaling $R_{\rm sym}$ of 0.087 for 11 962 unique reflections. Using the known crystal structure of the rat form of carbonic anhydrase isozyme III, a molecular-replacement model was built. This model was used for rotation and translation searches and uniquely defined the space group as $P6_5$. Rigid-body refinement of the model was used to generate an initial phased electron-density map with an $R_{\rm work}$ of 31.17%.

1. Introduction

Carbonic anhydrases (CAs) catalyze the reversible hydration of CO_2 to HCO_3^-. The mammalian class (α-class) of CAs is comprised of 14 known isozymes (I–XIV) with varying tissue distributions and catalytic activity (Parkkila, 2000). Human CAII (hCAII) is the most efficient of the mammalian isozymes with a catalytic turnover number ($k_{\rm cat}$) of $10^6\ {\rm s}^{-1}$ and has a wide tissue distribution compared with human CAIII (hCAIII; Lindskog, 1997), which is primarily found in slow-twitch red muscle fiber where it is the major soluble protein (Sly & Hu, 1995). HCAIII is distinguished from other α-class isozymes by its lower catalytic activity, $k_{\rm cat} = 10^3\ {\rm s}^{-1}$, and its resistance to most sulfonamide inhibitors (Sanyal *et al.*, 1982). The function of hCAIII is unclear but it may play a role in facilitated diffusion of CO_2 to the tissue capillaries (Riley *et al.*, 1982). It has also been shown that adipocytes have a high concentration of hCAIII (Spicer *et al.*, 1990), although its role is not well understood. HCAIII is also expressed at lower levels in other tissues, including salivary glands, smooth muscle cells in the uterus, red cells, prostate, lung, colon, kidney and testis (Sly & Hu, 1995).

In this paper, we report the purification, crystallization and initial phasing of hCAIII with a C-terminal histidine tag. This structure will enable a more detailed analysis of the catalytic activity of this isozyme and may provide insight into its mechanism of function. The structure of hCAII has previously been determined by X-ray crystallography (Eriksson *et al.*, 1988; Hakansson *et al.*, 1992) and therefore a direct structural comparison between the active sites of hCAII and hCAIII may provide insight into understanding the 1000-fold difference in catalytic efficiency between the two human isozyme forms.

2. Materials and methods

2.1. Mutagenesis

An expression vector containing the hCAIII coding region as described by Tanhauser *et al.* (1992) was used. Six histidine residues were added at the carboxy-terminus by PCR and additional subclone steps. In addition, the construct contained the following replacements: Cys183 and Cys188 were both replaced by Ser. These are surface residues on the side of the enzyme opposite the active-site cavity and the replacements were designed to promote crystallization. The construct was verified by DNA sequencing of the entire coding region. The expression plasmid was transformed into *E. coli* BL21(DE3)pLysS (Studier *et al.*, 1990) and gave a level of protein expression of up to 10 mg l^{-1}.

2.2. Purification

Cells were lysed by the freeze–thaw method in 0.2 *M* Na_2SO_4 in 0.1 *M* Tris buffer pH 8.0. Supernatant from the cell lysate was then

syringe filtered with a 0.45 µ*M* filter. hCAIII was then purified using a chelating column with Ni^{2+} bound to the matrix of the column. Protein was bound to the nickel column with 0.5 *M* NaCl in 0.02 *M* Tris buffer pH 7.4. Once all unbound cytosolic protein was eluted from the column, 0.5 *M* NaCl in the presence of 1.0 *M* imidazole and 0.02 *M* Tris buffer pH 7.4 was used at 4% imidazole (40 m*M*) to elute weakly bound cytosolic protein from the column. Once all weakly bound protein was eluted from the column, 100% imidazole (1.0 *M*) was used to elute the tightly bound hCAIII. The protein was desalted, concentrated and buffer exchanged into 10 m*M* Tris buffer pH 8.0 by several cycles of low-speed centrifugation using a Centricon (10 kDa molecular-weight cutoff). The protein was further purified using an anion-exchange column. The protein was loaded onto the column in the presence of 10 m*M* Tris pH 9.0 and eluted from the column with 10 m*M* Tris pH 8.0/1 *M* NaCl. The elution fractions from the anion-exchange column were then run through another cycle of desalting, buffer exchange and concentration prior to crystallization.

2.3. Crystallization

Initial crystallization screening was performed using the hanging-drop method (McPherson, 1982) over a range of 25–40% polyethylene glycol molecular weight 8000 (PEG 8000) as the precipitant in 10 m*M* Tris buffer pH 8.0. Crystal drops were prepared by mixing 5 µl of 10 mg ml^{-1} enzyme solution with 5 µl of precipitant solution. The drops were equilibrated by vapor diffusion against 1 ml of precipitant solution at 277 K.

Based on the results of the initial crystallization screening, useful X-ray diffraction quality crystals of hCAIII were obtained using 6.7 mg ml^{-1} enzyme solution in 10 m*M* Tris buffer pH 8.0. Crystal drops were obtained by mixing 10 µl of enzyme with 4 µl of precipitant solution consisting of 30% PEG 8000. The drops were equilibrated by vapor diffusion against 1 ml of precipitant solution at 277 K.

2.4. Data collection

Data were collected using an R-AXIS IV++ image-plate system with Osmic mirrors and a Rigaku HU-H3R CU rotating-anode generator operating at 50 kV and 100 mA. A 0.3 mm collimator was used with a crystal-to-detector distance of 150 mm and the 2θ angle fixed at 0°. All frames were collected using a 0.5° oscillation angle with an exposure time of 60 s per frame at 293 K. The data set was indexed using *DENZO* and scaled and reduced with *SCALEPACK* (Otwinowski, 1992).

2.5. Rotation and translation search

A preliminary model of hCAIII for molecular replacement was derived from the structure of rat CAIII (rCAIII; PDB code 1flj; Mallis *et al.*, 2000). This was achieved by interactively mutating the 23 (9.0%) amino acids that differ between the two enzymes (Hewett-Emmet & Tashian, 1996) using the program *O* (Jones *et al.*, 1991). No attempt was made to model the C-terminal molecular tag of six histidine residues. This model of hCAIII was used as the initial search model for the rotation- and translation-function calculations using the software package *CNS* (Brünger *et al.*, 1998).

2.6. Sequence alignment

A pairwise sequence alignment of hCAIII, rCAIII and hCAII was performed using *CLUSTALW* (Thompson *et al.*, 1994) (Fig. 1). Residues 5–260 were included in the alignment. The C-terminal hexahistidine tag in hCAIII was not considered. 23 residues out of 255 (9.0%) were found to differ between rCAIII and hCAIII, giving an overall sequence identity of 91.0%. 102 residues out of 255 (40.0%) were found to differ between hCAII and rCAIII, giving a sequence identity of 60.0%. A similar result was obtained for the alignment between hCAII and hCAIII, with 104 residues differing (40.1%), giving a sequence identity of 59.9% (Fig. 1).

3. Results and discussion

3.1. Crystallization

The initial crystallization screens produced small needle-like crystals (Fig. 2*a*) with a significant amount of precipitation that increased over the equilibration period of the drop. These crystals diffracted X-rays to approximately 3.5 Å resolution, but the data showed some evidence of disorder and twinning in the crystal lattice. Scaling of this data gave unsatisfactory R_{sym} values in the range 20–25% and attempts to use molecular-replacement methods to solve the structure failed.

Further refinement of the crystallization conditions yielded suitable crystals for X-ray diffraction studies. These crystals appeared in 2 d of equilibration against precipitant solution and grew to full size in one week.

```
                10        20        30        40        50        60
                 |         |         |         |         |         |
hCAIII      WGYASHNGPDHWHELFPNAKGENQSPIELHTKDIRHDPSLQPWSVSYDGGSAKTIL
rCAIII      WGYASHNGPEHWHELYPIAKGDNQSPIELHTKDIRHDPSLQPWSVSYDPGSAKTIL
hCAII       WGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRIL
               **        **        *   **** *****    * *      *****
                     70        80        90       100       110
                      |         |         |         |         |
hCAIII      NNGKTCRVVFDDTYDRSMLRGGPLPGPYRLRQFHLHWGSSDDHGSEHTVDGVKYAA
rCAIII      NNGKTCRVVFDDTFDRSMLRGGPLSGPYRLRQFHLHWGSSDDHGSEHTVDGVKYAA
hCAII       NNGHAFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAA
               **** *   *  *** *      *   *       * **       **
                120       130       140       150       160       170
                 |         |         |         |         |         |
hCAIII      ELHLVHWNP-KYNTFKEALKQRDGIAVIGIFLKIGHENGEFQIFLDALDKIKTKGK
rCAIII      ELHLVHWNP-KYNTFGEALKQPDGIAVVGIFLKIGREKGEFQILLDALDKIKTKGK
hCAII       ELHLVHWNT-KYGDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDVLDSIKTKGK
                  *   **  * **             *  * *** * * *  *
                180       190       200       210       220
                 |         |         |         |         |
hCAIII      EAPFTKFDPSSLFPASRDYWTYQGSFTTPPCEECIVWLLLKEPMTVSSDQMAKLRS
rCAIII      EAPFNHFDPSCLFPACRDYWTYHGSFTTPPCEECIVWLLLKEPMTVSSDQMAKLRS
hCAII       SADFTNFDPRGLLPESLDYWTYPGSLTTPPLLECVTWIVLKEPISVSSEQVLKFRK
            * *       *  * * *         *    **  ** **    **   * ** * *
            230       240       250       260
             |         |         |         |
hCAIII      LLSSAENEPPVPLVSNWRPPQPINNRVVRASFKHHHHHH
rCAIII      LFASAENEPPVPLVGNWRPPQPIKGRVVRASFK------
hCAII       LNFNGEGEPEELMVDNWRPAQPLKNRQIKASFK------
               ** *  ****      *  *   ***
```

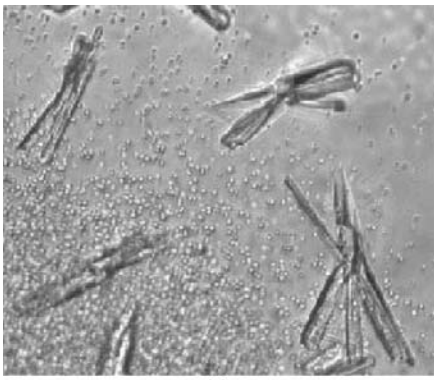
Figure 1
Pairwise sequence alignment of human carbonic anhydrase III (hCAIII), rat carbonic anhydrase III (rCAIII) and human carbonic anhydrase II (hCAII). Alignment was performed using *CLUSTALW* (Thompson *et al.*, 1994). Residues conserved in all three sequences are indicated by bold letters. Residues that are conserved between human carbonic anhydrase III and rat carbonic anhydrase III are indicated by a star. Amino-acid numbering is for the α-class of carbonic anhydrases.

The crystal habit was solid hexagons with approximate dimensions of 0.3 × 0.3 × 0.2 mm (Fig. 2*b*). All subsequent results were obtained from the hexagonal crystals.

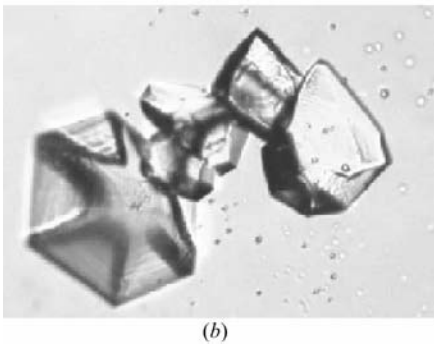
Table 1
Statistics of data collection.

Resolution shells (Å)	No. unique reflections	Completeness (%)	R_{sym}†
20.00–4.72	1241	94.0	0.044
4.72–3.76	1246	98.0	0.060
3.76–3.28	1239	96.1	0.072
3.28–2.98	1210	96.8	0.096
2.98–2.77	1208	94.2	0.128
2.77–2.61	1191	92.3	0.170
2.61–2.48	1150	91.2	0.220
2.48–2.37	1155	91.4	0.292
2.37–2.28	1168	90.2	0.364
2.28–2.20	1154	89.9	0.441
Total	11962	93.4	0.087

† R_{sym} is defined as $\sum |(I - \langle I \rangle)|/\sum I$, where I is the intensity of an individual reflection and $\langle I \rangle$ is the average intensity for this reflection; the summation is over all intensities.



(*a*)



(*b*)

Figure 2
Crystals of hCAIII. (*a*) Thin needle-like crystals from the initial crystallization screen. The precipitant for the initial screen was 25–40% polyethylene glycol molecular weight 8000 (PEG 8000) in 10 m*M* Tris pH 8.0. 5 µl of protein was mixed with 5 µl of precipitant to form the crystallization drop. (*b*) Hexagonal crystals from refined crystallization conditions. The refined precipitant was 30% PEG 8000 in 10 m*M* Tris pH 8.0. 10 µl of protein solution was mixed with 4 µl of precipitant to obtain the crystallization drop. The images were taken with a Zeiss Axioplan 2 microscope under polarized light.

3.2. Data collection

A total of 60° of data were collected (120 images) from a single crystal; a total of 88 415 reflections were measured to a maximum resolution of 2.2 Å. The data were initially processed with the Laue hexagonal crystal system *P*6, with unit-cell parameters a = 44.7, c = 222.5 Å. The data set was merged to a set of 11 962 independent reflections (93.4% complete, 89.9% in the outer resolution shell) resulting in an R_{sym} of 0.087 (0.441 in the outer resolution shell) (Table 1). On examination of the reflection file, no data was obtained along the c axis to allow assignment of a possible screw axis.

3.3. Rotation and translation search

From the hexagonal cell volume and the molecular weight of hCAIII, V_M (Matthews, 1968) values of 1.2 Å^3 Da^{-1} for three molecules per unit cell, 2.4 Å^3 Da^{-1} for six molecules per unit cell and 4.8 Å^3 Da^{-1} for 12 molecules per unit cell were calculated. Comparison with V_M values for other proteins indicates the most reasonable value is six hCAIII molecules per unit cell and suggests the space group to be either $P6_1$, $P6_2$, $P6_3$, $P6_4$ or $P6_5$, with one molecule per asymmetric unit and a solvent content of 47%. From considerations of the long c axis, the most probable packing of hCAIII is in space group $P6_1$ or $P6_5$.

A cross-rotation function was performed assuming the Laue group *P*6 and using the search model of hCAIII to determine the orientation of the reference molecule in the unit cell. This resulted in a single clear solution with θ_1 = 197.45, θ_2 = 47.20 and θ_3 = 156.43° and a peak of 7.2σ above the mean.

Using the rotation-function results, translation functions were performed assuming all six possible translational space groups of the Laue group *P*6. The space group $P6_5$ yielded the best solution, with translation components T_x = 14.65, T_y = 32.40 and T_z = 0.27 Å and a correlation coefficient of 0.731. The other space groups gave solutions with significantly lower correlation coefficients (Table 2).

The structure-factor file obtained from *SCALEPACK* was converted to *CNS* format and 5% of the data were randomly selected for R_{free} calculations. Using the obtained rotation and translation matrix, the hCAIII model was placed into the hexagonal $P6_5$ space group and rigid-body refinement was then used to generate an initial phasing data set with an R_{work} of 31.17%.

It is expected that the structure of hCAIII will be obtained relatively quickly and refined. This will allow direct structure–function comparison between the active sites of hCAII and hCAIII and may provide insight into the 1000-fold difference in catalytic efficiency between the two human isozyme forms. Specifically, this structure will allow a more detailed comparison of key active-site residues between the two isozymes.

3.4. Pairwise sequence alignment

The results of the sequence alignment were as expected. HCAIII and rCAIII showed the greatest sequence identity (91.0%); hCAII aligned with rCAIII and hCAIII showed less amino-acid conservation with 60 and 59.9% identity, respectively. It is expected that residues in the active site would have the most effect on the catalytic properties of these enzymes. The major amino-acid differences between the active sites of hCAII, hCAIII and rCAIII occur at positions 64, 67 and 198. In hCAII position 64 is occupied by a histidine residue found to be important for its ability to shuttle protons out of the active site (Eriksson *et al.*, 1988). The corresponding residue in hCAIII and rCAIII is lysine. Position 67 is occupied by an asparagine residue in hCAII. The corresponding residue in hCAIII and rCAIII at position 67 is an arginine. Finally, position 198 in hCAII is occupied by a leucine; the corresponding residue in human and rat CAIII is a phenylalanine. Previous work by LoGrasso *et al.* (1991) showed that replacing Lys64, Arg67 and Phe198 in hCAIII with His, Asn and Leu, respectively, enhanced the catalysis of CO_2 hydration by as much as 50-fold. Structural comparison of these residues between hCAII and rat and human CAIII may provide insight into the 1000-fold reduced catalytic activity of isozyme III with respect to isozyme II.

Table 2
Translation-function searches.

Space group	θ_1 (°)	θ_2 (°)	θ_2 (°)	T_x (Å)	T_y (Å)	T_z (Å)	CC†
*P*6	196.25	48.69	157.11	21.49	15.53	1.31	0.264
$P6_1$	197.25	48.15	156.00	14.46	32.04	0.50	0.334
$P6_2$	197.39	48.00	156.79	−7.56	19.71	0.39	0.397
$P6_3$	197.77	48.70	156.18	14.62	32.34	1.56	0.351
$P6_4$	197.22	47.69	157.25	2.38	28.57	0.13	0.262
$P6_5$	196.80	47.74	156.89	14.65	32.40	0.27	0.731

† $CC(r, \Omega) = [\langle |E_{obs}|^2 |E_m(r, \Omega)|^2 \rangle - \langle |E_{obs}|^2 \rangle \langle |E_m(r, \Omega)|^2 \rangle] \times \{(\langle |E_{obs}|^4 \rangle - \langle |E_{obs}|^2 \rangle^2)[\langle |E_m(r, \Omega)|^4 \rangle - \langle |E_m(r, \Omega)|^2 \rangle^2]\}^{-1/2}$. E_{obs} denotes the normalized observed structure factors and $E_m(r, \Omega)$ denotes the normalized structure factors of the search model oriented according to Ω.

The authors thank Mavis Agbandje-McKenna and Phillip Laipis for useful discussions, and Tim Vaught for taking the optical photographs of our crystals. This work was supported by grants from the National Institutes of Health GM25154 (DNS) and the University of Florida, College of Medicine start-up funds (RM).

References

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.

Eriksson, A. E., Jones, T. A. & Liljas, A. (1988). *Proteins Struct. Funct. Genet.* **4**, 274–282.

Hakansson, K., Carlsson, M., Svensson, L. A. & Liljas, A. (1992). *J. Mol. Biol.* **227**, 1192–1204.

Hewett-Emmett, D. & Tashian, R. E. (1996). *Mol. Phylogenet. Evol.* **5**, 50–77.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Lindskog, S. (1997). *Pharmacol. Ther.* **74**, 1–20.

LoGrasso, P.V., Tu, C., Jewell, D. A., Wynns, G. C., Laipis, P. J. & Silverman, D. N. (1991). *Biochemistry*, **30**, 8463–8470.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: Wiley & Sons.

Mallis, R. J., Poland, B. W., Chatterjee, T. K., Fisher, R. A., Darmawan, S., Honztko, R. B. & Thomas, J. (2000). *FEBS Lett.* **482**, 237–241.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Otwinowski, Z. (1992). *An Oscillation Data Processing Suite for Macromolecular Crystallography.* New Haven, CT, USA: Yale University.

Parkkila, S. (2000). *The Carbonic Anhydrases: New Horizons*, edited by W. R. Chegwidden, N. D. Carter & Y. H. Edwards, pp. 79–93. Basel: Birkhauser Verlag.

Riley, D. A., Ellis, S. & Bain, J. (1982). *Histochem. Cytochem.* **30**, 1275–1288.

Sanyal, G., Swenson, E. R., Pessah, N. I. & Maren, T. H. (1982). *Mol. Pharmacol.* **22**, 211–220.

Sly, W. S. & Hu, P. Y. (1995). *Annu. Rev. Biochem.* **64**, 375–401.

Spicer, S. S., Ge, Z. H., Tashian, R. E., Hazen-Marten, D. J. & Schulte, B. A. (1990). *Am. J. Anat.* **187**, 55–64.

Studier, W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzymol.* **185**, 60–89.

Tanhauser, S. M., Jewell, D. A., Tu, C. K., Silverman, D. N. & Laipis, P. J. (1992). *Gene*, **17**, 113–117.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.