PROTEINS: Structure, Function, and Genetics 16:29–42 (1993)

Refined Structure of Bovine Carbonic Anhydrase III at 2.0 Å Resolution

A. Elisabeth Eriksson[1] and Anders Liljas[2]
[1]*Department of Molecular Biology, Biomedical Center, Uppsala University, S-751 24 Uppsala, Sweden, and*
[2]*Department of Molecular Biophysics, Chemical Center, Lund University, S-221 00 Lund, Sweden*

ABSTRACT **The three-dimensional structure of bovine carbonic anhydrase III (BCA III) from red skeletal muscle cells has been determined by molecular replacement methods. The structure has been refined at 2.0 Å resolution by both constrained and restrained structure-factor least squares refinement. The current crystallographic R-value is 19.2% and 121 solvent molecules have so far been found associated with the protein. The structure is highly similar to the refined structure of human carbonic anhydrase II. Some differences in amino acid sequence and structure between the two isoenzymes are discussed. In BCA III, Lys 64 and Arg 91 (His 64 and Ile 91 in HCA II) are both pointing out from the active site cavity forming salt bridges with Glu 4 and Asp 72 (His 4 and Asp 72 in HCA II), respectively. However, Arg 67 and Phe 198 (Asn 67 and Leu 198 in HCA II) are oriented towards the zinc ion and significantly reduce the volume of the active site cavity. Phe 198 particularly reduces the size of the substrate binding region at the "deep water" position at the bottom of the cavity and we suggest that this is one of the major reasons for the differences in catalytic properties of isoenzyme III as compared to isozyme II.**

Key words: crystallography, molecular replacement, refinement, structure, muscle, carbonic anhydrase

INTRODUCTION

Carbonic anhydrase (CA, Mr = 29,500) is a zinc-containing enzyme that exists in a number of forms in the cytosol of different cell types of higher vertebrates. CA I and CA II can be found in large amounts in red blood cells as well as in many tissues, whereas the more recently discovered CA III is predominantly found in red skeletal muscle cells where it is the most abundant soluble protein.[1–4] The major function of CA I and CA II is to catalyze the reaction $CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$, thus facilitating CO_2 transport in the red blood cells, whereas the physiological function of CA III is more unclear.

The catalytic properties of CA III differ significantly from those of the other isoenzymes. The CO_2 hydration activity of BCA III is 3×10^3 s^{-1} at pH 8.5 and 25°C[5] which is only 0.3% of that of HCA II[6,7] and 1.5% of that of HCA I.[6] In addition, CA III has a lower esterase activity[1,4,8] and is not inhibited by most aromatic and heterocyclic sulfonamides[1,2,9,10] that otherwise are very specific inhibitors of CA I and CA II. These differences are probably due to amino acid substitutions in the active site which introduce three basic residues at positions 64*, 67, and 91 and one aromatic residue (Phe) at position 198. All three basic residues could be built into a model such that the proximity of their side chains to the zinc ion might affect the electrostatic environment of the zinc. Phe 198 is of special interest since it introduces a bulkier amino acid into the region of substrate binding.[11–14]

We have solved and refined the structure of bovine CA III (BCA III) at high resolution. Here we describe the new structure and compare it to the refined structure of human CA II (HCA II). In particular, we will discuss the effects of the differences in and near the active site on catalytic efficiency, substrate specificity, and inhibitor binding.

MATERIALS AND METHODS

Data Collection and Processing

BCA III, purified from bovine skeletal muscle cells[5] was kindly provided by Prof. Sven Lindskog, Umeå University and crystallized as described by Eriksson and Liljas.[16] The crystals have space group $P2_1$ with cell dimensions a = 50.6 Å, b = 44.7 Å, c = 56.9 Å and β = 90.3°. The crystals are very flat with an average dimension of 2.0–3.0 × 0.5 × 0.02 mm^3 and were therefore mounted on a plateau in specially prepared glass capillaries.[17]

Two separate X-ray diffraction data sets were col-

*The standard amino acid sequence for CA III starts with number 2, and number 126 is deleted from the sequence.

Received June 11, 1992; revision accepted October 12, 1992.
Address reprint requests to Dr. Anders Liljas, Department of Molecular Biophysics, Chemical Center, Lund University, Box 124 S-221 00 Lund, Sweden.

TABLE I. Data Collection Statistics for Native BCA III

	Film	Area detector	Film + area detector
Number of crystals	7	2	9
Number of measurements	24,662	12,159	36,821
Number of unique measurements	11,261	3,719	12,643
Number of unique measurements rejected (rejection ratio 0.4)*	513	114	646
Percentage of total data	62% at 2.0 Å	83% at 3.4 Å	69% at 2.0 Å
R_{merge} † (all measurements)	7.71	6.10	8.00
R_{merge} (rejection ratio 0.4)*	7.07	5.45	7.17

*Individual measurements are rejected if their relative deviation from the mean value ($|I-<I>|/<I>$) exceeds the specified (0.4) threshold limit.

† $$R_{merge} = \sum_{r}^{N}\left(\sum_{i}^{n(r)}|I(i,r) - (<Ir)>|\right)/\sum_{r}^{N}(n(r)*<I(r)>) \times 100$$

where N is the number of unique reflections, n(r) is the number of multiple measurements for the r-th reflection with the mean value <I(r)> and I(i,r) is the i-th individual measurements of the r-th reflection.

lected, one on film and one using an area detector. The film data set was collected at 2.0 Å resolution at room temperature on Enraf-Nonius (Delft, The Netherlands) Arndt-Wonacott oscillation cameras using Ni-filtered CuKα radiation from GX-20 and GX-6 rotating anodes (Elliott Bros. Ltd., England) and Reflex 25 (CEA AB, Sweden) X-ray-sensitive films. Films obtained from a 180° rotation around the b*-axis were scanned on a rotating drum densitometer (Optronics International Inc., Chelmsford, MA, U.S.A.) and evaluated.[18,19] One thousand one hundred and fifty-eight pairs of partially recorded reflections were merged and also included in the data set. All data handling and crystallographic calculations were performed with the program PROTEIN[20] unless otherwise stated. The second complementary data set was collected to 3.4 Å resolution on a Nicolet Xentronics area detector[21] using the software system described by Blum et al.[22]

Crystals used for some of the data had accidentally been soaked in a mounting solution containing azide ions. Azide is a strong inhibitor of BCA III[5,9,23] and therefore part of the final data set was collected on the BCA III-azide complex instead of the native structure. The consequences of this for the geometry of the zinc ion will be discussed later. Data collection statistics for the different data sets are given in Table I. The film data were used in molecular replacement calculations and CORELS refinement, whereas the merged film and area detector data were used in the continuing refinement with PROLSQ.

Molecular Replacement Calculations

The refined coordinates of HCA II[24] were modified and used as the target molecule in the molecular replacement calculations.[25] The amino acid sequences of HCA II and BCA III were compared and 113 nonidentical residues were replaced by either alanines or glycines. All solvent molecules were omitted from the HCA II structure. This modified model containing 1,736 atoms of a total of 2,056 protein atoms was then used in a Patterson search to determine the orientation of the BCA III molecule within the unit cell. The model was placed in a triclinic cell with orthogonal axes 80 Å in length. This size of the unit cell was chosen to prevent the occurrence of crossvectors between molecules in neighbouring cells. Triclinic structure factors were calculated from 7.0 to 3.0 Å resolution assuming an overall temperature factor of 20 $Å^2$. The correlation between the crystal and the model Patterson synthesis as a function of the orientation angles, ϕ, ω and κ (polar angles), was calculated using the fast rotation function POLARRFN written by W. Kabsch and included in the CCP4 program package.[26] Radial shells were varied between 15 and 20 Å and reflections between 7.0 and 3.0 Å were used. The orientation angles were varied in steps of 5.0° where ϕ = 0° to 360°, ω = 0° to 180° and κ = 0° to 180°, and subsequently in steps of 1.0° around the final peak. Calculated Eulerian angles were applied on the model using TRNSFMABG.[26] The location of the correctly oriented molecule was determined using the Patterson-like Q-function originally defined by Tollin and Cochran.[27,28]

Model Building and Refinement

The resulting coordinates from rotation and translation functions were further refined using the program CORELS[29], treating the whole molecule as a rigid body. The refinement then proceeded using CORELS where restraints over the main chain atoms were introduced. To reduce the buildup of restraint errors between peptides, the refinement was continued using PROLSQ.[30] At various stages during the refinement, the atomic structure was checked against new $2|F_{obs}|-|F_{calc}|$ or $|F_{obs}|-|F_{calc}|$ Fourier maps using model phases. Manual corrections of the atomic positions were made on an inter-

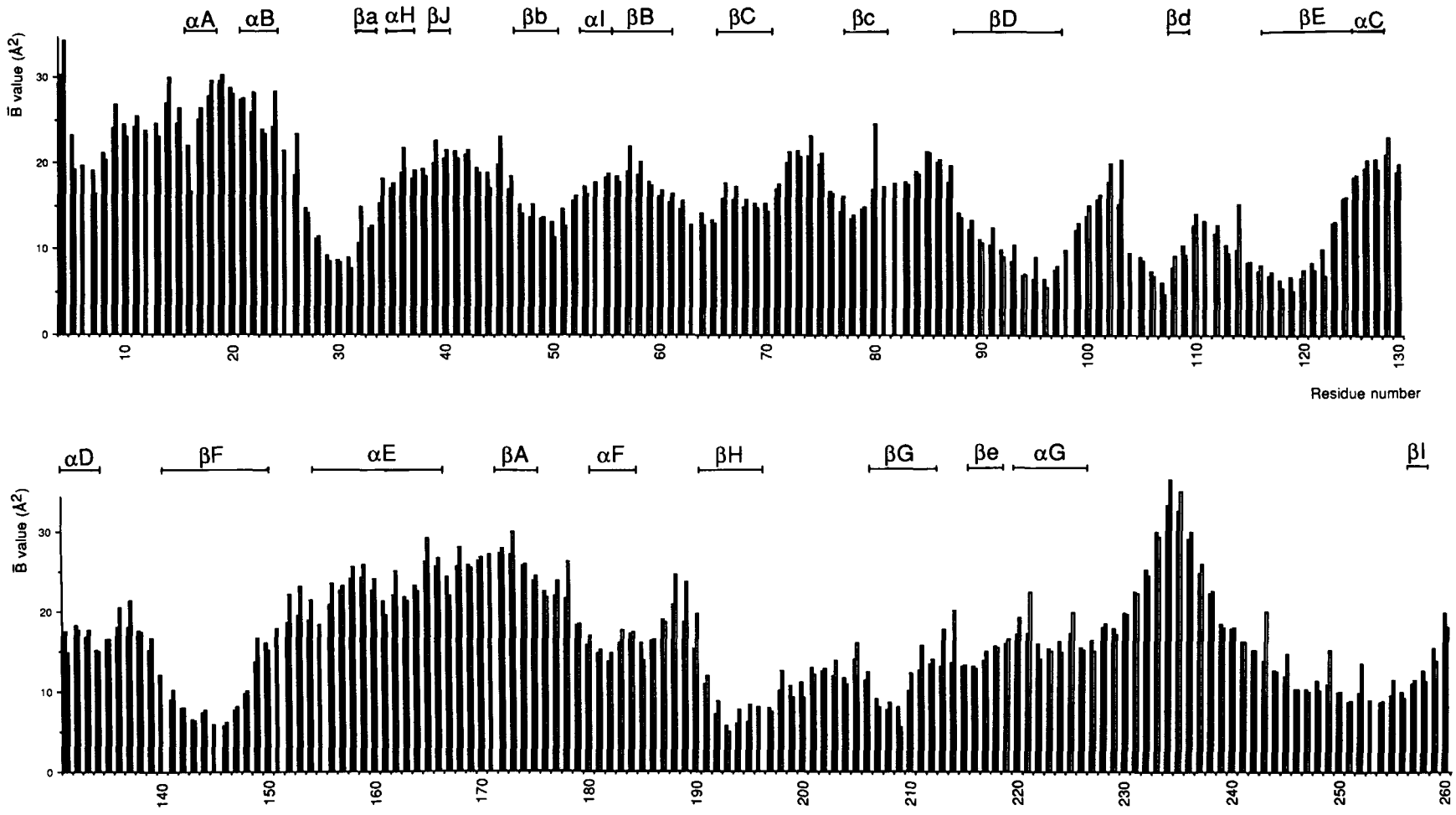


Fig. 1. Isotropic temperature factor (Å²) variation of the BCA III polypeptide chain plotted as a function of residue number. The average main chain and side chain B-values are represented by the filled and open bars, respectively.

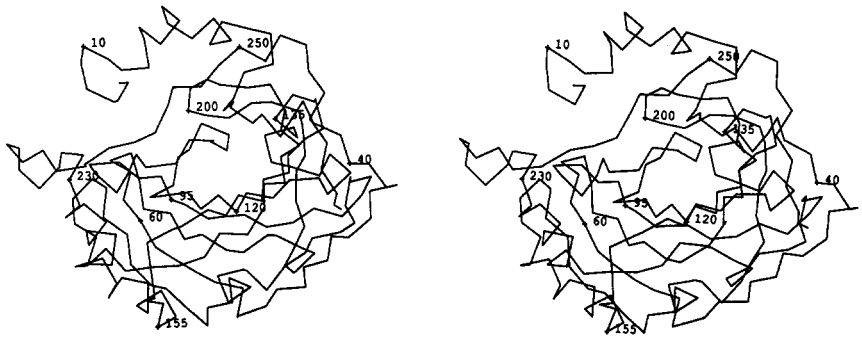


Fig. 2. Stereo drawing showing the α-carbon chain of BCA III.

TABLE II. Temperature Factors for Residue Side Chains and Water Molecules Within the Active Site*

Water molecule or residue number	B-value (Å^2)	Accessible surface (Å^2)
264 OHH	14.0	
265 OHH	13.2	
292 OHH	9.1	
303 OHH	21.8	
304 OHH	26.8	
335 OHH	14.6	
336 OHH	26.5	
345 OHH	31.8	
369 OHH	22.9	
389 OHH	7.3	
Tyr 7	16.4	35
Asn 62	15.4	37
Lys 64	12.7	19
Arg 67	17.1	66
Arg 91	12.3	77
Gln 92	9.0	10
His 94	6.9	15
His 96	5.5	1
Glu 106	6.8	0
Glu 117	7.3	0
His 119	5.0	2
Val 121	7.5	1
Tyr 131	14.9	41
Leu 135	16.5	33
Val 143	6.4	1
Phe 198	12.5	20
Thr 199	9.1	6
Thr 200	9.1	25
Glu 204	10.9	33
Ile 207	8.0	1
Trp 209	5.4	2
Zn	12.9	0.6

*The temperature factors vary between 4.7 Å^2 and 34.3 Å^2 for the side chain atoms in the protein structure. The table also includes solvent accessible surface[33] for the residues and the zinc ion.

active graphic system (Evans & Sutherland, PS 330 and PS 390) using FRODO.[31,32] The 113 alanines in the BCA III starting model were not replaced by their proper amino acid until their positions were clearly indicated in the electron density maps. All crystallographic calculations where done on VAX 11/750 and VAX 3600 (Digital Equipment, U.S.A.) computers except for most of the PROLSQ cycles that were done on a Convex/C1 (Convex Ltd., U.S.A.).

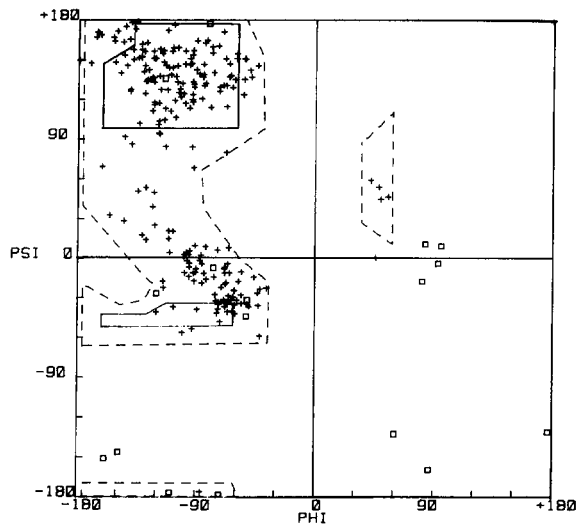


Fig. 3. Ramachandran plot of BCA III. Glycine residues are indicated by squares and nonglycines by crosses.

RESULTS AND DISCUSSION

Molecular Replacement Calculations

The modified model of HCA II was used to determine the orientation of the BCA III molecules within the crystal unit cell. The correct orientation was found in a rotational search with reflections between 6.0 and 3.0 Å. The major peak was 20% higher than any other peak. A corresponding search using reflections between 7.0 and 4.0 Å did not reveal the same peak. The obtained Eulerian angles of 48.1°, 115.6° and 18.1° were applied to the BCA III model. The following translation search showed a

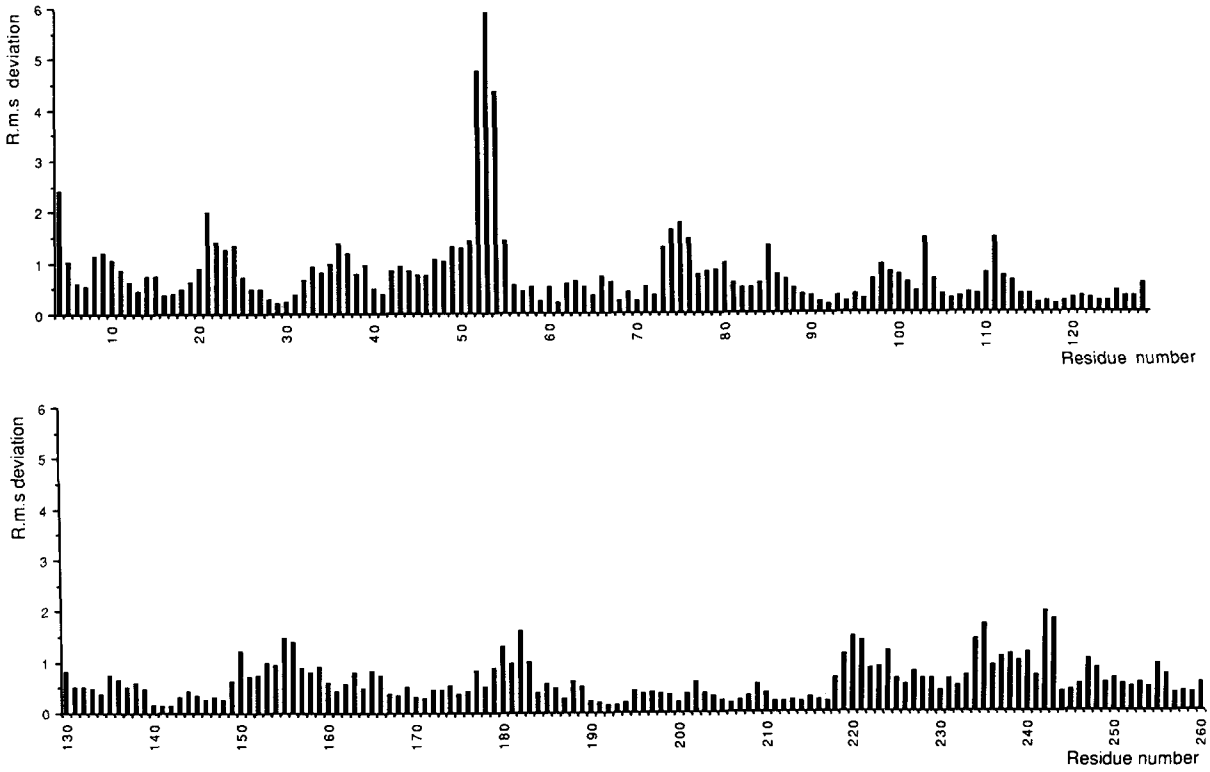


Fig. 4. The average rms deviations for the main chain atoms of HCA II and BCA III.

clear peak, 70% over noise level, which corresponded to a 1.5 Å translation in the x-direction and a 17.4 Å translation in z-direction. The final location of the BCA III molecules and the packing within the crystal lattice was checked on the graphic display using PACKIN (M. Bergdoll, Uppsala University). No severe interactions between the protein molecules were observed. The location of the molecule was also checked in a difference Fourier where the zinc ion had been deleted from the structure and did not contribute to the phase calculation. A significant peak was obtained at the same position as the deleted zinc ion.

Refinement and Stereochemistry of the Model

The starting model for refinement had a conventional crystallographic R-factor[†] of 48.6% including all reflections between 7.0 and 5.0 Å resolution. Rigid body refinement using CORELS, where 2 positional and 3 rotational parameters for the complete molecule were refined, translated the molecule 0.13 Å in the x-direction, 0.22 Å in the z-direction, and rotated it $-1.10°$ in ϕ, $-1.55°$ in θ and 0.75° in ρ which reduced the R-factor to 43.1%. The refinement proceeded using both CORELS and PROLSQ and the last refinement round included 2,177 atoms of which 121 where solvent molecules associated with the protein. Our current R-factor is 19.2% including data from 7.0 to 2.0 Å resolution. The rms[‡] departures from the restraints to ideal geometry in the last refinement cycle were 0.027 Å, 0.043 Å and 0.043 Å for bond, angle, and fixed dihedral angle distances, respectively; 0.011 Å for out of plane restraints, and 0.164 $Å^3$ for chiral volumes.[30] The deviation from the isotropic thermal factor restraints were 0.694 $Å^2$ for main chain bonds, 1.162 $Å^2$ for main chain angles, 0.723 $Å^2$ for side chain bonds and 1.205 $Å^2$ for the side chain angles. The rms parameter shifts in the last cycle were 0.010 Å for coordinates and 0.16 $Å^2$ for temperature factors.

General Features of the Molecule

The structures of BCA III and HCA II are almost identical. A complete description of the refined structure of HCA II at 2.0 Å resolution has been published[24] and recently extended to 1.55 Å.[14] In order to avoid repetition, only differences between the two structures will be discussed here and the reader is referred to our previous articles for a more thorough description of the structure.

The refined molecular model of BCA III is almost

[†]R-factor = $\Sigma||F_{obs}|-|F_{calc}||/\Sigma|F_{obs}|$.

[‡]rms = root-mean-square.

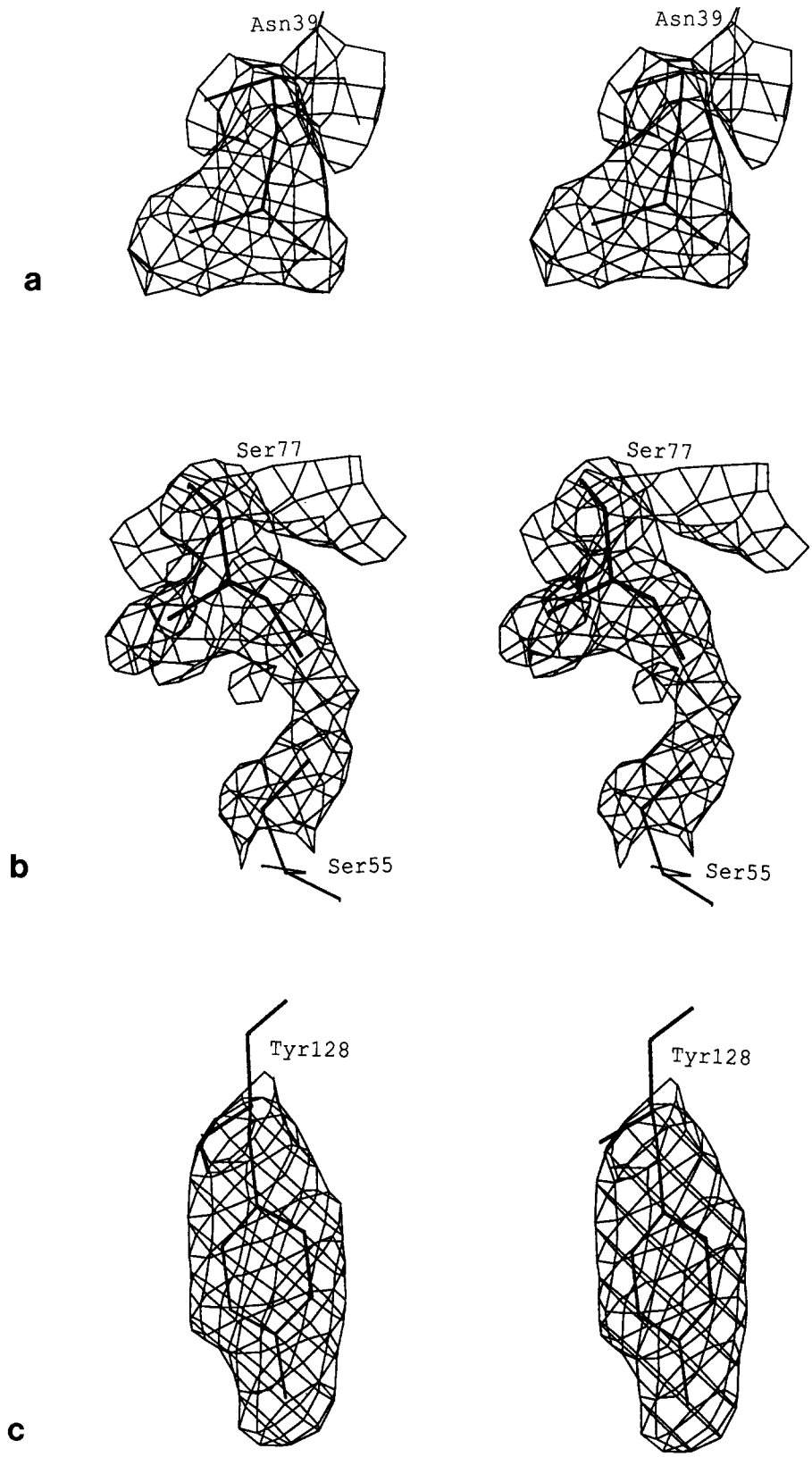


Fig. 5a–c. Legend appears on page 35.

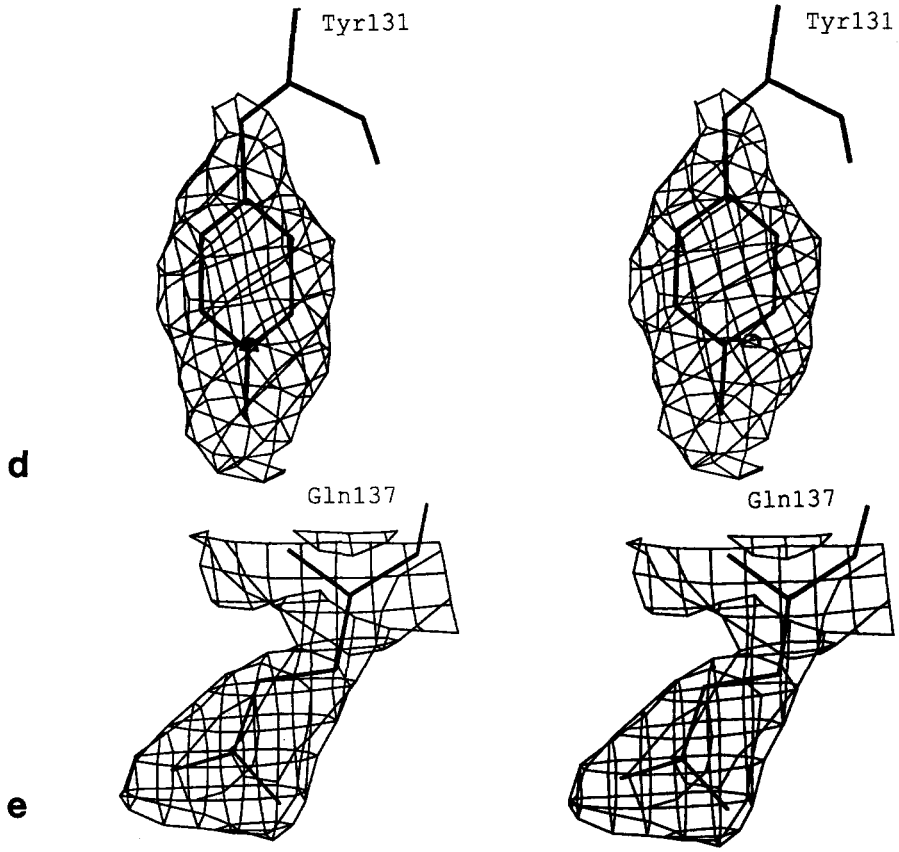


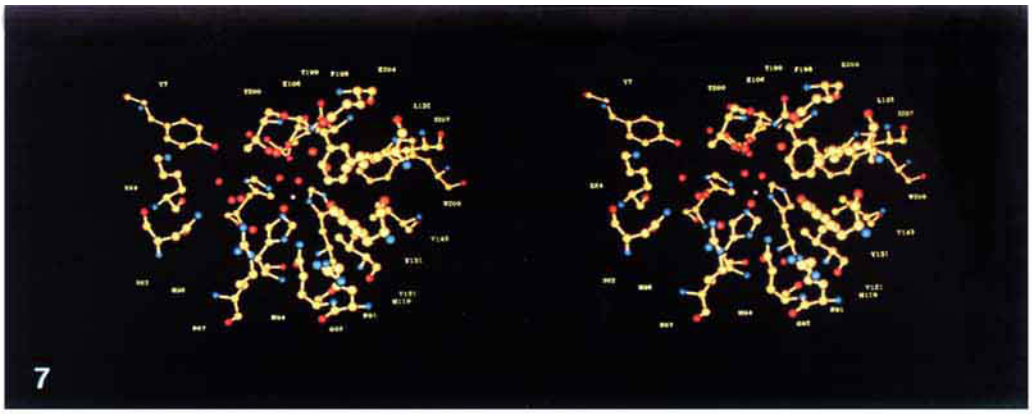
Fig. 5. Stereo views of the $2|F_{obs}|-|F_{calc}|$ difference map over the five amino acid corrections. **a:** Asn 39; **b:** Ser 77; **c:** Tyr 128; **d:** Tyr 131; and **e:** Gln 137.

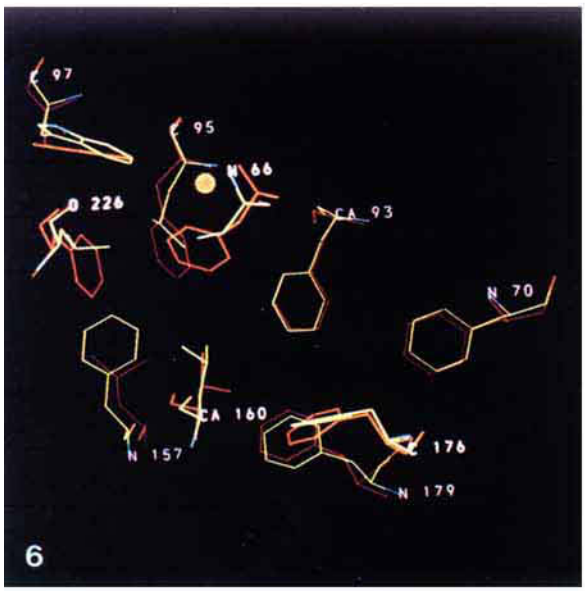
complete. However, the electron density is not defined at Leu 19–Phe 20 or Glu 234–Asn 235, regions situated at the surface of the molecule. The electron density is also weak at the side chains of Glu 4, Lys 24, His 103, and Arg 243. The distribution of temperature factors over the individual residues is shown in Figure 1. The average temperature factors for the main chain atoms for individual residues vary between 5.9 and 31.0 Å^2. Temperature factors for some residue side chains and water molecules within the active site are shown in Table II. This table also includes the solvent accessible surfaces for these residues.

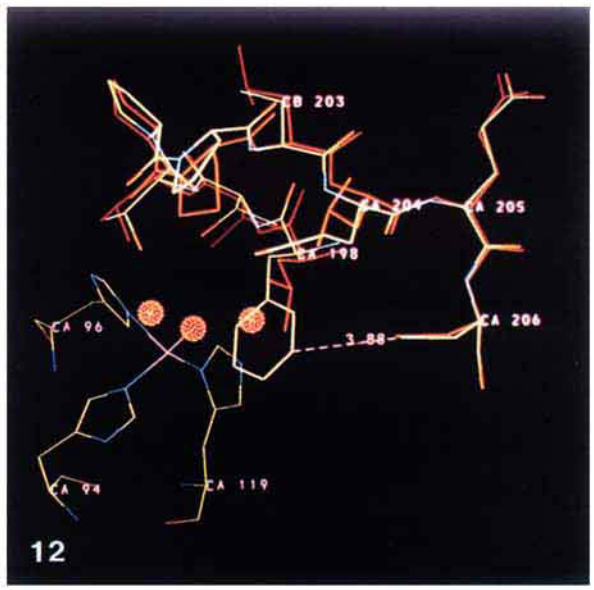
The α-carbon backbone of BCA III is shown in Figure 2 and secondary structure elements were defined using the program DSSP.[34] The same nomenclature as for HCA II is used. In addition to the secondary structure of HCA II, BCA III has two regions with a few residues in 3_{10}-helical conformation: [residues 35–37 (αH) and 53–55 (αI), see Fig. 1]. These regions in both proteins are situated on the surface of the molecule and interact with symmetry-related molecules in the unit cell.

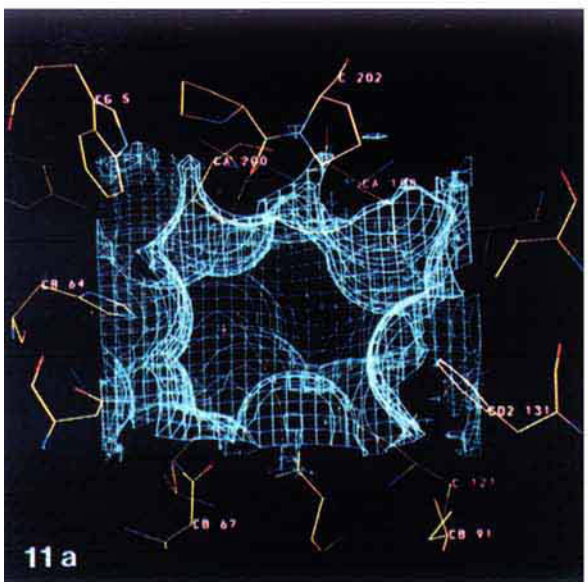
The distribution of the φ and ψ torsion angles which define the polypeptide backbone of BCA III is illustrated in Figure 3 as a Ramachandran plot.[35] All the amino acids except Glu 234 and a few glycines have allowed conformations. Glu 234 is situated in a poorly defined region of the molecule. The four additional nonglycine residues in left-handed helical conformation are all situated in different kinks and turns on the main chain. These four residues are Asp 110, Asn 178, Cys 203, and Lys 252. The structure also contains the conserved cis-prolines, Pro 30 and Pro 202.[15,24]

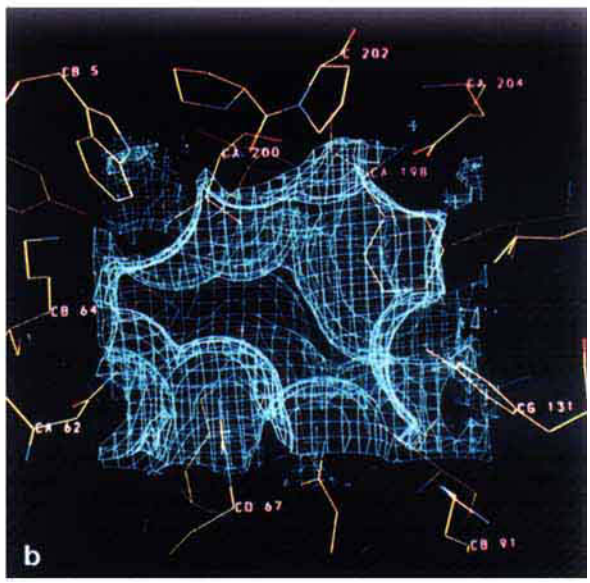
The structures of HCA II and BCA III were superimposed using program STRFIT (P. Kraulis, Uppsala University) which finds the best rotation[36] to relate two molecules using defined equivalent residues. The central β-sheet is highly conserved both in the amino acid sequence as in the protein structure. Thus, the main chain atoms for 95 residues situated in these β-sheet structures in HCA II and BCA III (numbers 39–41, 56–72, 89–98, 115–123, 140–149, 168–176, 187–218, and 256–260) were defined to be equivalent and the calculated rms fit was 0.39 Å including only these atoms. When all main chain atoms were included, the rms value was 0.92 Å. Figure 4 shows the rms deviation for the individual amino acids and as can be seen, the two polypeptide chains deviate significantly at residues 52–55. The two different conformations are probably due to two amino acid substitutions: Gln 53 and Lys 76 in HCA II are Pro and Arg, respectively, in BCA











Figs. 6, 7, 11a, b, and 12. Legends on page 37.

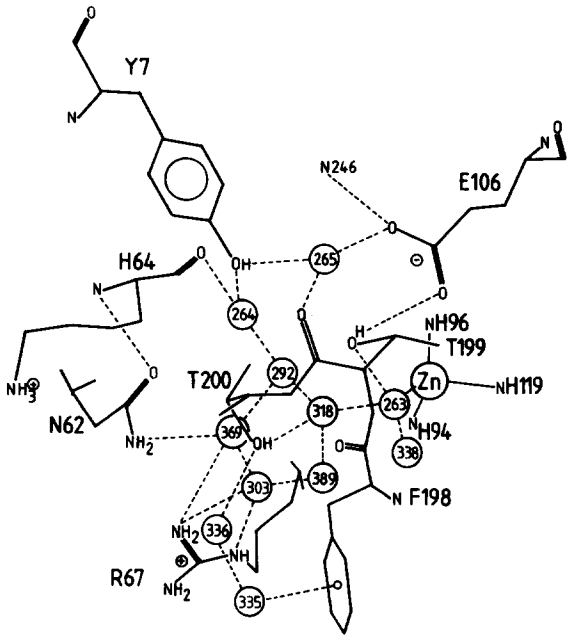


Fig. 8. Schematic drawing of the residues and water molecules in the active site of BCA III. Hydrogen bonds are indicated by dashed lines.

III. The conserved residue Asp 52 is situated in HCA II with its side chain pointing out from the molecule and hydrogen-bonded to the side chain of Gln 53. BCA III has a Pro at position 53 the side chain of which cannot form hydrogen bonds. After a conformational change of several Ångström, the Oδ1 atom of Asp 52 is instead hydrogen bonded to the Nε atom of Arg 76. This region interacts with symmetry-related protein molecules both in HCA II and BCA III and the two different conformations can be the effect of crystal packing and may be of no significance for the proteins in solution.

Amino Acid Sequence

The amino acid sequence of bovine CA III[37] was available from protein sequencing at the time of the structure determination with some corrections.[5] During the refinement we were very much aided by the amino acid sequence, but based on our electron density maps, we have suggested five additional corrections: Asp or Asn instead of Ser at position 39; Ser instead of Ala at position 77; Tyr instead of Phe at position 128; Tyr instead of Ile at position 131 and Glu or Gln instead of His at position 137 (see Fig. 5). Of these amino acid substitutions, Tyr 131 is of greatest interest since it is situated at the hydrophobic part of the active site and is a Phe in HCA II. All five amino acid changes were, in addition, consistent with the sequences of human CA III[38] equine CA III (ECA III[39]), mouse CA III (MCA III[40]) and rat CA III (RCA III[41]). Subsequent to the structure determination a revised sequence of bovine CA III has been reported.[37a] Here all our suggested modifications were confirmed. The modified sequence of BCA III as well as the sequences of the above-mentioned species are listed in Table III.

Aromatic Clusters

The two aromatic clusters in HCA I and HCA II[15,24] are also observed in BCA III with some modifications. The edge-to-face orientation which is a stabilizing feature in protein structures[42,43] is a dominating feature. The first smaller aromatic cluster at the N-terminal region including Trp 5, Tyr 7, Trp 16, and Tyr 20 is conserved. However, the packing of the second aromatic cluster situated just below the active site region is significantly different for BCA III. HCA I and HCA II have 7 conserved phenylalanines involved in this cluster (residues 66, 70, 93, 95, 176, 179, and 226), four of which are conserved in BCA III (70, 93, 176, and Phe 179). The other three have been substituted to nonaromatic residues (Cys 66, Leu 95, and Leu 226). It is noteworthy that the three substitutions are all situated at one end of the cluster leaving the other half of the cluster intact (see Fig. 6). Thus, the reduced sizes of the amino acids at positions 66, 95, and 226 make the packing less tight in this region. To some extent the introduction of Phe for Leu at position 157 and Leu for Val at position 160 compensates for this. The buried water molecule 295 which stabilizes secondary structure elements in HCA II does not exist in BCA III since the Sγ atom of Cys 66 occupies part of this space. The two β-strands (64–68 and 93–97) in BCA III seem to compensate for this missing water molecule by moving somewhat closer to each other (0.2–0.3 Å).

The Active Site

We refrain from a discussion of the zinc geometry or the positions of the active site water molecules situated closest to the zinc ion (numbers 263, 318, and 338 in HCAII) due to the partial binding of azide. However, we do not believe that the presence of azide ions in any way affects the positions of the

Fig. 6. The largest aromatic cluster in BCA III. The corresponding cluster in HCA II is shown in red.

Fig. 7. Stereo drawing showing the active site of BCA III. The photograph was taken on a SUN 3/110 using the program FRODO PAINT (Mats Kihlén, Dept. of Mol. Biology, Uppsala University, Sweden).

Fig. 11. Accessible surfaces of (a) HCA II and (b) BCA III computed in the form of an electron density layer (G. Vriend, Groningen University, The Netherlands).

Fig. 12. The loop region including residues 198–206 for BCA III and HCA II (of which the latter is shown in red). Also included are the zinc ion and its ligands (His 94, 96, 119, and water molecule 263), the water molecule 318 and the deep water, 338.

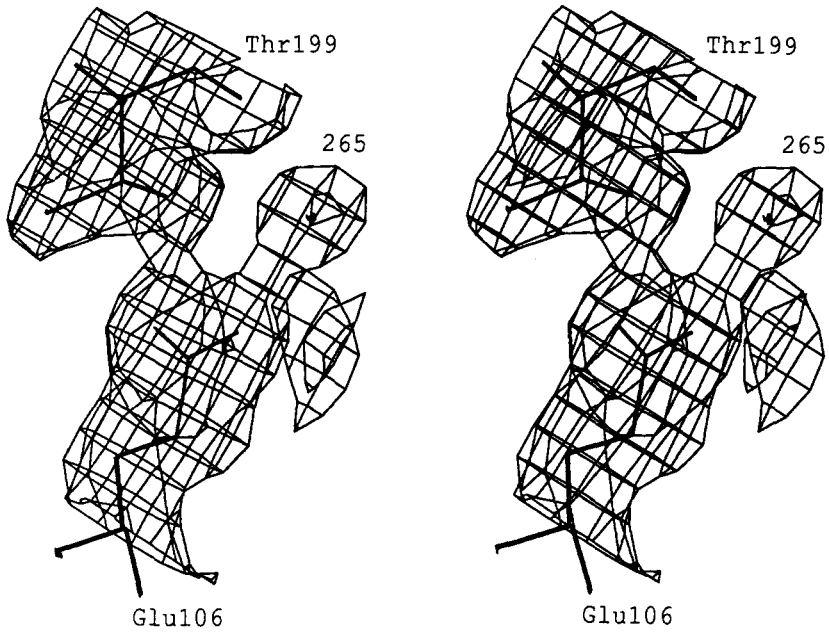


Fig. 9. Stereo views of the $2|F_{obs}|-|F_{calc}|$ difference map in part of the active site region of BCA III. Contours are drawn at a height of $+1\sigma$ where σ is the rms deviation throughout the unit cell.

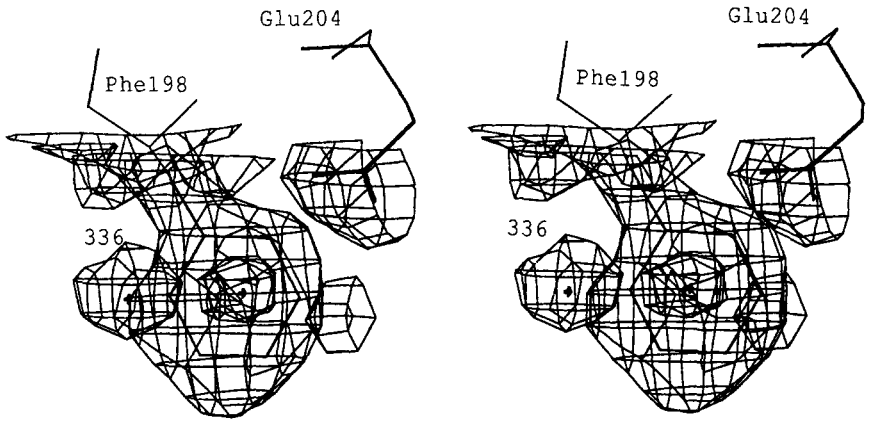


Fig. 10. Stereo views of the $2|F_{obs}|-|F_{calc}|$ difference map showing water molecule 335 hydrogen bonded to the π-electron system of Phe 198. The figure also includes Glu 204 and water molecule 336.

active site residues for the following reasons: first, since the azide ion in complex with HCA II binds to the zinc ion without change in conformation of the side chains in the active site (B. Jönsson et al., unpublished results), and secondly because the electron densities over all residues in the active site of BCAIII are very well-defined.

There are four major amino acid substitutions in the active site of BCA III compared to HCA II. BCA III has three basic residues, Lys 64, Arg 67, and Arg 91, located at the hydrophilic side and one aromatic residue, Phe 198, at the hydrophobic side. A stereo view of the active site is shown in Figure 7. The hydrogen bonding network, including some of the water molecules, is shown schematically in Figure 8 and the hydrogen bond distances are listed in Table IV. The quality of the electron density map can be judged from Figure 9.

The major conformation of His 64 in HCA II points towards the zinc ion.[14,44] However, this is not the case for Lys 64 in BCA III. Its Nζ atom is situated 12.3 Å from the zinc ion and probably forms a salt bridge to Glu 4. However, the electron density for the latter residue is not clearly defined. Likewise, Arg 91 is not in the proximity of the zinc ion. Its side chain is directed out of the cavity where its position is fixed by a salt bridge between its Nη1 atom and the Oδ1 atom of Asp 72. The Nη2 atom of Arg 91 has

TABLE III. The Primary Sequences of HCA III,[38] ECA III, [39] MCAIII,[40] RCA III,[41] and BCA III*

```
                1                  10                  20                  30
HCA  III        A K E W G Y A S H N G P D H W H E L F P N A K G E N Q S P
BCA  III        A K E W G Y A D H N G P D H W H E L F P N A K G E N Q S P
ECA  III        A K E W G Y A D H N G P D H W H E F Y P I A K G D N Q S P
MCA  III        A K E W G Y A R H N G P D H W H E L Y P I A K G D N Q S P
RCA  III        A K E W G Y A S H N G P E H W H E L Y P I A K G D N Q S P

                      *  40                  50                  60
HCA  III  I E L H T K D I R H D P S L Q P W S V S Y D G G S A K T I L
BCA  III  I E L N T K E I N H D P S L K P W T A S Y D P G S A K T I L
ECA  III  I E L H T K D I N H D P S L K A W T A S Y D P G S A K T I L
MCA  III  I E L H T K D I K H D P S L Q P W S A S Y D P G S A K T I L
RCA  III  I E L H T K D I R H D P S L Q P W S V S Y D P G S A K T I L

                         70                  *   80                  90
HCA  III  N N G K T C R V V F D D T Y D R S M L R G G P L P G P Y R L
BCA  III  N N G K T C R V V F D D T Y D R S M L R G G P L A A P Y R L
ECA  III  N N G R T C R V V F D D T Y D R S M L R G G P L T A P Y R L
MCA  III  N N G K T C R V V F D D T Y D R S M L R G G P L S R P Y R L
RCA  III  N N G K T C R V V F D D T F D R S M L R G G P L S G P Y R L

                        100                 110                 120
HCA  III  R Q F H L H W G S S D D H G S E H T V D G V K Y A A E L H L
BCA  III  R Q F H L H W G S S D D H G S E H S V D G V K Y A A E L H L
ECA  III  R Q F H L H W G S S D D H G S E H T V D G V K Y A A E L H L
MCA  III  R Q F H L H W G S S D D H G S E H T V D G V K L A A E L H L
RCA  III  R Q F H L H W G S S D D H G S E H T V D G V K Y A A E L H L

                      *  130 *               *  140                 150
HCA  III  V H W N P - K Y N T F K E A L K Q R D G I A V I G I F L K I
BCA  III  V H W N S - K Y N S Y A T A L K Q A D G I A V V G V F L K I
ECA  III  V H W N P - K Y N T Y G G A L K Q P D G I A V V G V F L K I
MCA  III  V H W N P - R Y N T F G E A L K Q P D G I A V V G I L L K I
RCA  III  V H W N P - K Y N T S E E A L K Q P D G I A V V G I F L K I

                        160                 170                 180
HCA  III  G H E N G E F Q I F L D A L D K I K T K G K E A P F T K F D
BCA  III  G R E K G E F Q L L L D A L D K I K T K G K E A P F N N F N
ECA  III  G R E K G E F Q L F L D A L D K I K T K G K E A P F T N F D
MCA  III  G R E K G E F Q I L L D A L D K I K T K G K E A P F T H F D
RCA  III  G R E K G E F Q I L L D A L D K I K T K G K E A P F N H F D

                        190                 200                 210
HCA  III  P S C L F P A C R D Y W T Y Q G S F T T P P C E E C I V W L
BCA  III  P S C L F P A C R D Y W T Y H G S F T T P P C E E C I V W L
ECA  III  P S C L F P T C R D Y W T Y R G S F T T P P C E E C I V W L
MCA  III  P S C L F P A C R D Y W T Y H G S F T T P P C E E C I V W L
RCA  III  P S C L F P A C R D Y W T Y H G S F T T P P C E E C I V W L

                        220                 230                 240
HCA  III  L L K E P M T V S S D Q M A K L R S L L S S A E N E P P V P
BCA  III  L L K E P I T V S S D Q I A K L R T L Y S S A E N E P P V P
ECA  III  L L K E P I T V S S D Q V A K L R S L F S S A E N E P P V P
MCA  III  L L K E P M T V S S D Q M A K L R S L F S S A E N E P P V P
RCA  III  L L K E P M T V S S D Q M A N V R S L F A S A E N E P P V P

                        250                 260
HCA  III  L V S N W R P P Q P I N N R V V R A S F K
BCA  III  L V R N W R P P Q P I K G R I V K A S F K
ECA  III  L V R N W R P P Q P L K G R V V R A S F K
MCA  III  L V G N W R P P Q P V K G R V V R A S F K
RCA  III  L V G N W R P P Q P I K G R V V R A S F K
```

*The BCA III sequence are from Tashian et al.[37] except for residues 65–67 and 173–208 which are from Engberg and Lindskog.[5] Our suggested amino acid changes are indicated by asterisks. Those changes have subsequently been verified.[37a]

TABLE IV. Hydrogen Bond Distances in the Active Site of BCA III

264 OHH	— 292 OHH	2.7 Å
264 OHH	— Oη Tyr 7	2.6 Å
264 OHH	— O Lys 64	3.2 Å
265 OHH	— Oη Tyr 7	3.0 Å
265 OHH	— Oε2 Glu 106	2.6 Å
265 OHH	— O Thr 199	2.8 Å
292 OHH	— 369 OHH	3.0 Å
303 OHH	— 369 OHH	3.0 Å
304 OHH	— Nδ2 Asn 62	2.8 Å
335 OHH	— 336 OHH	2.7 Å
335 OHH	— Oε1 Glu 204	2.8 Å
336 OHH	— Oγ1 Thr 200	3.1 Å
336 OHH	— O Pro 201	3.1 Å
345 OHH	— Nε2 Gln 92	2.6 Å
345 OHH	— Oη Tyr 131	3.2 Å
369 OHH	— Nδ2 Asn 62	3.0 Å
369 OHH	— Nη2 Arg 67	3.2 Å
389 OHH	— 303 OHH	2.8 Å
389 OHH	— Nε2 Gln 92	3.1 Å
Oδ1 Asn 62	— N Lys 64	3.3 Å
Oδ1 Asn 62	—O Thr 65	3.2 Å
Oδ1 Asp 72	— Nη1 Arg 91	3.1 Å
Oε1 Gln 92	— Nδ1 His 94	2.9 Å
Nδ1 His 96	— O Asn 244	2.7 Å
Oε1 Glu 106	— Oγ1 Thr 199	2.4 Å
Oε2 Glu 106	— N Arg 246	3.0 Å
Oε2 Glu 117	— Nε2 His 119	2.8 Å

no close neighbors and is situated 14.9 Å from the zinc ion. The only basic residue with a position relatively close to the zinc ion is Arg 67. Its side chain has no interaction with other residues and forms hydrogen bonds only to water molecules (numbers 269 and 303). The distance between the Nη1 atom of Arg 67 and the zinc ion is 8.7 Å. Several catalytic properties of CA III are suggested to be associated with one or both of the arginines in the active site. CA III has been shown to possess a low phosphatase activity unlike the other isoenzymes.[45] This activity can be abolished in pig CA III by modification with the arginine-specific reagent phenylglyoxal without affecting either the CO_2 hydratase or the esterase activity.[46] However, further modifications of arginine residues result in a decrease of the CO_2 activity.[46] On the other hand, modification of other CA III species with the arginine reagent butanedione stimulates the dehydration reaction for these enzymes.[47] Therefore, the function of the arginine residues in the active site of CA III needs further investigation. Since Arg 91 is situated at such a distant position, Arg 67 is more likely affecting the catalytic properties of the enzyme.

Phe 198 is situated at the hydrophobic side of the cavity. Interestingly a water molecule (number 335) is hydrogen bonded to the π-electron system of its benzene ring (see Fig. 10). The water is hydrogen-bonded at the flat side of the benzene ring and well-centered at 3.0–3.2 Å to all six ring carbon atoms. Weak polar interactions between the δ(+) edges and the δ(−) π-electron clouds of aromatic residues are generally occurring in protein structures[48,49] and results from infrared spectral studies indicate that the benzene ring also can function as a hydrogen bond acceptor.[50] Two hydrogen bonds to the same aromatic residue have been observed in the structure of BPTI.[51] Here one hydrogen bond is formed between the proton of the Nδ1 atom of Asn 44 and one side of the benzene ring of Tyr 35. The second hydrogen bond is formed between the proton of the peptide amide group of Gly 37 and the other side of the same tyrosine residue. However, to our knowledge, this is the first time a water molecule is observed hydrogen-bonded to an aromatic ring in a protein structure. The water molecule 335 forms two additional hydrogen bonds: one to the Oε1 atom of Glu 204 (which is a Leu in HCA II) and one to water molecule 336. The water molecule 336 is, in turn, hydrogen-bonded to the Oγ1 atom of Thr 200 resulting in a hydrogen bond network leading from Phe 198 into the proximity of the zinc ion (Fig. 8). The Oη atom of Tyr 131 is situated 3.3 Å from the edge of the Phe 198 ring. Thus, it seems as if a more polar area is created at the hydrophobic side of the active site of BCA III but whether this increased polarity is of significance for the catalytic properties at the zinc ion is not clear.

The introduction of Phe 198 seems important for the catalytic properties of CA III from another point of view. Crystallographic studies of HCA II have shown that the sulfonamide group of the inhibitors 3-acetoxymercuri-4-aminobenzenesulfonamide[11] acetazolamide[52] and aminobenzolamide[53] bind to the zinc ion at the deepest end of the active site which is presumably also the site for substrate binding.[11–14] When binding at the zinc ion, these inhibitors are within van der Waals' distances from Leu 198. In BCA III, Phe 198, together with Ile 207, significantly diminishes the space for inhibitor and substrate binding. The surface accessibilities in the active site cavities of HCA II and BCA III are illustrated in Figure 11. A smooth surface has been calculated which covers the protein at a distance of about 1.4 Å from nonhydrogen atoms. As can be seen in Figure 11, the volume of the active site cavity of BCA III is much reduced compared to HCA II primarily due to Phe 198 and Arg 67. From model building studies it can be concluded that Phe 198 prevents the aromatic part of the sulfonamide inhibitors from binding in the same manner as observed for HCA II.[11,53] This is consistent with the observation that most sulfonamides are only weak inhibitors of CA III. The aromatic sulfonamides, e.g., chlorzolamide[9] that inhibit CA III must therefore have other modes of binding. Mutation of Phe 198 to Leu in human CA III leads to improved sulfonamide binding as well as improved catalytic properties.[54]

However, the change of Leu 198 in HCA II to Phe does not lead to corresponding decreases in binding and catalysis.[55]

It has been reported that modification of one or several of the five cysteine residues in BCA III increases the CO_2 hydratase activity[56,57] but which of the cysteines associated with this activation could not be deduced. The strongest candidate seemed to be Cys 66 because of its location in the aromatic cluster just below the active site region (see Fig. 6). However, since its position is buried, no clear explanation can be made for its reactivity or the increase in the efficiency of the enzyme. We therefore think it is more likely that it is the modification of Cys 206 which affects the substrate binding region. The Sγ atom of Cys 206 is situated within van der Waals' distance from the edge of the Phe 198 benzene ring (see Fig. 12) and thus, modifying this residue might also affect the position of the Phe 198 side chain. It has been observed that modification with mercuric compounds leads to a conformational change of the Sγ atom of Cys 206 in HCA II.[11] In the native enzyme, the side chain is pointing towards the active site, whereas it rotates towards the surface of the protein in the mercury-bound complex. If the side chain of Cys 206 in BCA III is rotated, the benzene ring of Phe 198 moves out of the active site, thus reducing its sterical hindrance of the substrate binding region. It has been shown that modification of Cys 206 in HCA II does not affect the catalytic activity of the enzyme.[56] On the other hand, HCA II has a leucine at position 198 which does not have the same effect on the substrate binding site.

Monovalent anions are potent inhibitors of CA III isoenzymes[5,9] and our structure of BCA III is consistent with these observations. The SCN^- ion can be built in at the zinc, even though its orientation cannot be the same as observed in HCA II,[11] since the position of the sulfur atom has to be modified due to its proximity to the Cδ1 atom of Ile 207 and benzene ring of Phe 198.

In the absence of crystallographic inhibitor studies of BCA III, it seems as if the catalytic mechanism discussed for HCA II[11,14] could also apply for the muscle isoenzymes. The bicarbonate would be forced to adopt a slight difference in orientation in BCA III. The more restricted area in the active site as well as different and probably less than optimal geometry for the reactants could explain the reduced rate of catalysis seen in BCA III.

ACKNOWLEDGMENTS

We are thankful for help and advice with the structure determination from Professors Alwyn Jones and Hans Eklund University of Uppsala. Prof. Sven Lindskog, University of Umeå, is thanked for kindly providing us with purified BCA III enzyme as well as for continuous support. We are also grateful to Terese Bergfors for help with the manuscript. These studies have been supported by the Swedish Natural Science Research Council.

REFERENCES

1. Koester, M.K., Register, A.M., Noltmann, E.A. Basic muscle protein, a third genetic locus isoenzyme of carbonic anhydrase? Biochem. Biophys. Res. Comm. 76:196–204, 1977.
2. Holmes, R.S. Purification, molecular properties and ontogeny of carbonic anhydrase isozymes: Evidence for a A, B and C isozymes in avian and mammalian tissues. Eur. J. Biochem. 78:511–520, 1977.
3. Register, A.M., Koester, M.K., Noltmann, E.A. Discovery of carbonic anhydrase in rabbit skeletal muscle and evidence for its identity with "basic muscle protein". J. Biol. Chem. 253:4143–4152, 1978.
4. Carter, N., Jeffery, S., Shiels, A., Edwards, Y., Tipler, T., Hopkinson, D.A. Characterization of human carbonic anhydrase III from skeletal muscle. Biochem. Genet. 17:837–854, 1979.
5. Engberg, P., Millqvist, E., Pohl, G., Lindskog, S. Purification and some properties of carbonic anhydrase from bovine skeletal muscle. Arch. Biochem. Biophys. 241:628–638, 1985.
6. Khalifah, R.G. The carbon dioxide hydration activity of carbonic anhydrase. J. Biol. Chem. 246:2561–2573, 1971.
7. Steiner, H., Jonsson, B.-H., Lindskog, S. The catalytic mechanism of carbonic anhydrase. Eur. J. Biochem. 59: 253–259, 1975.
8. Tu, C., Thomas, H.G., Wynns, G.C., Silverman, D.N. Hydrolysis of 4-nitrophenyl acetate catalyzed by carbonic anhydrase III from bovine skeletal muscle. J. Biol. Chem. 261:10100–10103, 1986.
9. Sanyal, G., Swenson, E.R., Pessah, N.I., Maren, T.H. The carbon dioxide hydration activity of skeletal muscle carbonic anhydrase: Inhibition by sulfonamides and anions. Mol. Pharmacol. 22:211–220, 1982.
10. Pullan, L.M., Noltmann, E.A. Simultaneous and independent versus antagonistic inhibition of muscle carbonic anhydrase (CA III) by acetazolamide and cyanate. Biochem. Pharmacol. 33:2641–2645, 1984.
11. Eriksson, A.E., Kylsten, P.M., Jones, T.A., Liljas, A. Crystallographic studies of inhibitor binding sites in human carbonic anhydrase II: A penta-coordinated binding of the SCN^- ion to the zinc at high pH. Proteins 4:283–293, 1988.
12. Xue, Y., Vidgren, J., Svensson, L.A., Liljas, A., Jonsson, B.-H., Lindskog, S. Crystallographic analysis of Thr200→His human carbonic anhydrase II and its complex with the substrate, HCO3−. Proteins 15:80–87, 1993.
13. Lindahl, M., Svensson, L.A., Liljas, A. Metal poison inhibition of carbonic anhydrase. Proteins, 15:177–182, 1993.
14. Håkansson, K., Carlsson, M., Svensson, L.A., Liljas, A. The structure of native and apo carbonic anhydrase II, and some of its anion ligand complexes. J. Mol. Biol. 227:1192–1204, 1993.
15. Kannan, K.K., Ramanadham, M., Jones, T.A. Structure, refinement, and function of carbonic anhydrase isozymes: Refinement of human carbonic anhydrase I. Ann. N. Y. Acad. Sci. 429:49–60, 1984.
16. Eriksson, A.E., Liljas, A. Crystallization of and preliminary X-ray data for bovine carbonic anhydrase III. J. Biol. Chem. 261:16247–16248, 1986.
17. Åkervall, K., Strandberg, B. X-ray diffraction studies of the satellite tobacco necrosis virus. J. Mol. Biol. 62:625–627, 1971.
18. Rossmann, M.G. Processing oscillation diffraction data for very large unit cells with an automatic convolution technique and profile fitting. J. Appl. Cryst. 12:225–238, 1979.
19. Schmid, M.F., Weaver, L.H., Holmes, M.A., Grütter, M.G., Ohlendorf, D.H., Reynolds, R.A., Remington, S.J., Matthews, B.W. An oscillation data collection system for high-resolution protein crystallography. Acta Cryst. A37:701–710, 1981.
20. Steigemann, W. Ph.D. Thesis, Technische Universität Munich, 1974.
21. Durbin, R.M., Burns, R., Moulai, J., Metcalf, P., Freymann, D., Blum, M., Anderson, J.E., Harrison, S.C., Wiley, D.C. Protein, DNA, and virus crystallography with a fo-

cused imaging proportional counter. Science 232:1127–1132, 1986.
22. Blum, M., Metcalf, P., Harrison, S.C., Wiley, D.C. A system for collection and on-line integration of X-ray diffraction data from a multiwire area detector. J. Appl. Cryst. 20:235–242, 1987.
23. Kararli, T., Silverman, D.N. Inhibition of the hydration of CO_2 catalyzed by carbonic anhydrase III from cat muscle. J. Biol. Chem. 260:3484–3489, 1985.
24. Eriksson, A.E., Jones, T.A., Liljas, A. The refined structure of human carbonic anhydrase II at 2.0 Å resolution. Proteins 4:274–282, 1988.
25. "The Molecular Replacement Method. A Collection of Papers on the Use of Non-Crystallographic Symmetry." Rossmann, M.G. (ed.). New York, London, Paris: Gordon and Breach, Science Publishers, 1972.
26. CCP4 Program Suite, Rotation Function-POLARRFN, User documentation (Daresbury Laboratory, 1984).
27. Tollin, P., Cochran, W. Patterson function interpretation for molecules containing planar groups. Acta Cryst. 17: 1322–1324, 1964.
28. Tollin, P. On the determination of molecular location. Acta Cryst. 21:613–614, 1966.
29. Sussman, J.L., Holbrook, S.R., Church, G.M., Kim, S.-H. A structure-factor least-squares refinement procedure for macromolecular structures using constrained and restrained parameters. Acta Cryst. A33:800–804, 1977.
30. Hendrickson, W.A., Konnert, J.H. Incorporation of stereochemical information into crystallographic refinement. In: "Computing in Crystallography." Diamond, R., Ramaseshan, S., Venkatesan, K. (eds.). Bangalore: Indian Institute of Science, 1980:13.01–13.23.
31. Jones, T.A. A graphics model building and refinement system for macromolecules. J. Appl. Cryst. 11:268–272, 1978.
32. Jones, T.A. FRODO: A graphics fitting program for macromolecules. In: "Computational Crystallography." Sayre, D. (ed.). Oxford: Clarendon Press, 1982:303–317.
33. Lee, B., Richards, F.M. The interpretation of protein structures: Estimation of static accessibility. J. Mol. Biol. 55: 379–400, 1971.
34. Kabsch, W., Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637, 1983.
35. Ramachandran, G.N., Ramakrishnan, C., Sasisekharan, V. Stereochemistry of polypeptide chain conformations. J. Mol. Biol. 7:95–99, 1963.
36. Kabsch, W. A discussion of the solution for the best rotation to relate two sets of vectors. Acta Cryst. A 34:827–828, 1978.
37. Tashian, R.E., Hewett-Emmett, D., Stroup, S.K., Goodman, M., Yu, Y.-S.L. Evolution of structure and function in the carbonic anhydrase isozymes of mammals. In: "Biophysics and Physiology of Carbon Dioxide." Bauer, C., Gros, G., Bartels, H. (eds.). Berlin/Heidelberg/New York: Springer-Verlag, 1980:165–176.
37a. Hewett-Emmett, D., Tashian, R.E. Structure and evolutionary origins of the Carbonic anhydrase multigene family. In: "The Carbonic Anhydrases: Cellular Physiology and Molecular Genetics." Dodson, S.J., Tashian, R.E., Gros, G., Carter, N.E. (eds.). New York and London: Plenum Press, 1991:15–32.
38. Lloyd, J., McMillan, S., Hopkinson, D., Edwards, Y.H. Nucleotide sequence and derived amino acid sequence of a cDNA encoding human muscle carbonic anhydrase. Gene 41:233–239, 1986.
39. Wendorff, K.M., Nishita, T., Jabusch, J.R., Deutsch, H.F. The sequence of equine muscle carbonic anhydrase. J. Biol. Chem. 260:6129–6132, 1985.
40. Tweedie, S., Edwards, Y. Mouse carbonic anhydrase III: Nucleotide sequence and expression studies. Biochem. Genet. 27:17–30, 1989.
41. Kelly, C.D., Carter, N.D., Jeffery, S., Edwards, Y.H. Characterization of cDNA clones for rat muscle carbonic anhydrase. Biosci. Rep. 8:401–406, 1988.
42. Serrano, L., Bycroft, M., Fersht, A. Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles. J. Mol. Biol. 218:465–475, 1991.
43. Hunter, C.A., Singh, J., Thornton, J.M. π-π interactions: The geometry and energetics of phenylalanine-phenylalanine interactions in proteins. J. Mol. Biol. 218:837–846, 1991.
44. Nair, S.K., Christianson, D.W. Unexpected pH-dependent conformation of His-64, the proton shuttle of carbonic anhydrase-II. J. Am. Chem. Soc. 113:9455–9458, 1991.
45. Koester, M.K., Pullan, L.M., Noltmann, E.A. The p-nitrophenyl phosphatase activity of muscle carbonic anhydrase. Arch. Biochem. Biophys. 211:632–642, 1981.
46. Pullan, L.M., Noltmann, E.A. Specific arginine modification at the phosphatase site of muscle carbonic anhydrase. Biochemistry 24:635–640, 1985.
47. Tashian, R.E., Johansen, J.T., Christiansen, E., Chegwidden, W.R. Activation of mammalian skeletal-muscle carbonic anhydrase III by arginine modification. Biosci. Rep. 4:573–579, 1984.
48. Perutz, M.F., Fermi, G., Abraham, D.J., Poyart, C., Bursaux, E. Hemoglobin as a receptor of drugs and peptides: X-ray studies of the stereochemistry of binding. J. Am. Chem. Soc. 108:1064–1078, 1986.
49. Burley, S.K., Petsko, G.A. Amino-aromatic interactions in proteins. FEBS Lett. 203:139–143, 1986.
50. Engdahl, A., Nelander, B. A matrix isolation study of the benzene-water interaction. J. Phys. Chem. 89:2860–2864, 1985.
51. Tüchsen, E., Woodward, C. Assignment of asparagine-44 side-chain primary amide ^{1}H NMR resonances and the peptide amide N^{1}H resonance of glycine-37 in basic pancreatic trypsin inhibitor. Biochemistry 26:1918–1925, 1987.
52. Vidgren, J., Liljas, A., Walker, N.P.C. Refined structure of the acetazolamide complex of human carbonic anhydrase II at 1.9 Å. Int. J. Macromol. 12:342–344, 1990.
53. Vidgren, J., Svensson, L.A., Liljas, A. Refined structure of the aminobenzolamide complex of human carbonic anhydrase II at 1.9 Å and sulfonamide modelling of bovine carbonic anhydrase III. J. Macromol., in press, 1993.
54. LoGrasso, P.V., Tu, C., Jewell, D.A., Wynns, G.C., Laipis, P.J., Silverman, D.N. Catalytic enhancement of human carbonic anhydrase III by replacement of phenylalanine-198 with leucine. Biochemistry 30:8463–8470, 1991.
55. Ren, X., Jonsson, B.H., Lindskog, S. Some properties of site-specific mutants of human carbonic anhydrase II having active-site residues characterizing carbonic anhydrase III. Eur. J. Biochem. 201:417–420, 1991.
56. Engberg, P., Lindskog, S. Activation of bovine muscle carbonic anhydrase by modification of thiol groups. Eur. J. Biochem. 156:407–412, 1986.
57. Ren, X., Jonsson, B.-H., Millqvist, E., Lindskog, S. A comparison of the kinetic properties of native bovine muscle carbonic anhydrase and an activated derivative with modified thiol groups. Biochim. Biophys. Acta 953:79–85, 1988.