Structural Mechanics of the pH-dependent Activity of β-Carbonic Anhydrase from Mycobacterium tuberculosis*

Adrian Suarez Covarrubias, Terese Bergfors, T. Alwyn Jones, and Martin Högbohm

From the Department of Cell and Molecular Biology, Uppsala University, S-751 24 Uppsala, Sweden

Tuberculosis is estimated to cause nearly 2 million deaths yearly. Multiple drug treatment for half a year or more can cure over 95% of patients; however, only a minority of patients who develop tuberculosis receive treatment. There is also a rapid development of multidrug-resistant strains representing over 10% of new tuberculosis cases in some areas (1–3). This study concerns the Mycobacterium tuberculosis Rv3588c gene, shown to be required for in vivo growth of the pathogen, encodes a β-carbonic anhydrase with a steep pH dependence of its activity, being active at pH 8.4 but not at pH 7.5. We have recently solved the structure of this protein, which was a dimeric protein with a blocked active site. Here we present the structure of the thiocyanate complexed protein in a different crystal form. The protein now forms distinct tetramers and shows large structural changes, including a carboxylate shift yielding the accessible active site. This structure demonstrated for the first time that a β-carbonic anhydrase can switch between the two states. A pH-dependent dimer to tetramer equilibrium was also demonstrated by dynamic light scattering measurements. The data presented here, therefore, suggest a carboxylate shift on/off switch for the enzyme, which may, in turn, be controlled by a dimer-to-tetramer equilibrium.

Carbonic anhydrases catalyze the reversible hydration of carbon dioxide to form bicarbonate, a reaction required for many functions, including carbon assimilation and pH homeostasis. Carbonic anhydrases are divided into at least three classes and are believed to share a zinc-hydroxide mechanism for carbon dioxide hydration. β-carbonic anhydrases are broadly spread among the domains of life, and existing structures from different organisms show two distinct active site setups, one with three protein coordinations to the zinc (accessible) and the other with four (blocked). The latter is believed to be inconsistent with the zinc-hydroxide mechanism. The Mycobacterium tuberculosis Rv3588c gene, shown to be required for in vivo growth of the pathogen, encodes a β-carbonic anhydrase with a steep pH dependence of its activity, being active at pH 8.4 but not at pH 7.5. We have recently solved the structure of this protein, which was a dimeric protein with a blocked active site. Here we present the structure of the thiocyanate complexed protein in a different crystal form. The protein now forms distinct tetramers and shows large structural changes, including a carboxylate shift yielding the accessible active site. This structure demonstrated for the first time that a β-carbonic anhydrase can switch between the two states. A pH-dependent dimer to tetramer equilibrium was also demonstrated by dynamic light scattering measurements. The data presented here, therefore, suggest a carboxylate shift on/off switch for the enzyme, which may, in turn, be controlled by a dimer-to-tetramer equilibrium.

Tuberculosis is estimated to cause nearly 2 million deaths yearly. Multiple drug treatment for half a year or more can cure over 95% of patients; however, only a minority of patients who develop tuberculosis receive treatment. There is also a rapid development of multidrug-resistant strains representing over 10% of new tuberculosis cases in some areas (1–3). This study concerns the Mycobacterium tuberculosis Rv3588c gene, shown to be required for in vivo growth of the pathogen, encodes a β-carbonic anhydrase with a steep pH dependence of its activity, being active at pH 8.4 but not at pH 7.5. We have recently solved the structure of this protein, which was a dimeric protein with a blocked active site. Here we present the structure of the thiocyanate complexed protein in a different crystal form. The protein now forms distinct tetramers and shows large structural changes, including a carboxylate shift yielding the accessible active site. This structure demonstrated for the first time that a β-carbonic anhydrase can switch between the two states. A pH-dependent dimer to tetramer equilibrium was also demonstrated by dynamic light scattering measurements. The data presented here, therefore, suggest a carboxylate shift on/off switch for the enzyme, which may, in turn, be controlled by a dimer-to-tetramer equilibrium.

Carbonic anhydrases (CAs)\(^2\) are Zn\(^{2+}\)-dependent enzymes that catalyze the reversible hydration of the membrane-permeable gas carbon dioxide to form the ionic compound bicarbonate (5–7). CAs are generally divided into three classes, α, β, and γ, although there is also a fourth (δ) form identified in one organism (8, 9). The α-class found in mammals, some bacteria, and algae is the most extensively studied (7, 10–12). The three classes have different folds and are believed to have evolved separately (10, 12–15). Crystal structures of α- and γ-CAs show that they share a three-histidine zinc ligation with at least one zinc-bound water molecule. Extended X-ray Absorption Fine Structure data of the δ-CA from Thalassiosira weissflogii also suggest the same type of zinc ligation (16). The β-CAs have a notably different zinc ligation comprising one histidine and two cysteines with the fourth coordination position occupied by either a water molecule or an aspartyl side chain (13, 17). A zinc-hydroxide mechanism for carbon dioxide hydration is well established for the α-class (7, 18). In this mechanism, a zinc-coordinated hydroxide makes an attack on carbon dioxide, forming bicarbonate. The product then exchanges for a water molecule, which is activated by the zinc coordination, lowering its pK\(_a\), leading to deprotonation (Reaction 1) (18). Kinetic studies suggest that this mechanism is shared also by the β-CAs (19–22).

β-CAs are broadly spread among the domains of life and are found in euacteria, archaea, algae, and plants, in which they play an important role in carbon fixation and metabolism (7, 10, 23–26). All β-CAs share an α/β fold with the core being a five-stranded β-sheet (27).

In all cases, the monomers combine to make homodimers (denoted essential dimer) to produce an extended β-sheet and complete the active site environment; the β-CA from Porphyridium purpureum constitutes an exception in which an apparent internal gene duplication produces a poly peptide that is twice as long and folds with a pseudo 2-fold symmetry, to mimic the dimer (13). All dimers and the P. purpureum monomer superimpose very well, and the dimer interaction geometry is conserved (27). Most β-CAs adopt even higher multimeric states as observed by biochemical data and indicated by crystal packing. The oligomeric states have been reported to be dimers, tetramers, hexamers, and octamers (7, 14, 19, 28–33). Tetramers seem to be the predominant variety, but this varies between species and reports.

The proposed CA reaction mechanism depends on the presence of a zinc-coordinated hydroxide as the reactive species. All observations of α-, γ-, and δ-CA active sites are consistent with this as the zinc is coordinated by at least one water molecule (or possibly hydroxide) in the active site structures that are not complexed by non-protein ligands, such as inhibitors. Conversely, the β-CAs display a considerable difference of the zinc ligation within the class. The β-CAs from the archaeon Methanobacterium thermoautotrophicum (17), the plant Pisum sativum (garden pea) (14), and Rv1284 from M. tuberculosis (27) all have three protein-derived zinc ligands; one histidine (His-104, M. tuberculosis Rv3588c numbering here and throughout); and two cysteines (Cys-51, Cys-107). The last open zinc coordination position is occupied by a water molecule, or in the case of P. sativum, an acetate ion from the crystallization mother liquor. This active site architecture is thus consistent with the proposed zinc-hydroxide mechanism of the α-class and...
Carboxylate Shift Regulation Mechanism in β-CA

![Diagram of carboxylate shift regulation mechanism](image)

TABLE 1

| Data collection and refinement statistics | Rv3588c tetrameric <sup>a</sup> |
|------------------------------------------|-------------------------------|
| **Data collection statistics from HKL2000** |                               |
| Beaml ine                                | I71                          |
| Wavelength (Å)                           | 1.094                        |
| Space group                              | P2                           |
| Cell parameters: (Å)                     | 67.9, 70.3, 84.4             |
| Resolution (Å)                           | 3.00–2.20                    |
| No. of unique reflections                | 40363                        |
| Redundancy                               | 3.7                          |
| Completeness (%)                         | 99.6 (97.7)                  |
| R <sub>work</sub> (%)                    | 5.8 (36.3)                   |
| R <sub>f</sub> (%)                       | 20.1 (2.7)                   |
| **Refinement statistics**                |                               |
| Resolution (Å)                           | 3.00–2.20                    |
| No. of unique reflections                | 40363                        |
| No. of reflection in test set            | 2016                         |
| R <sub>work</sub> (%)                    | 16.3                         |
| R <sub>f</sub> (%)                       | 21.4                         |
| No. of atoms                             |                               |
| Protein                                  | 6163                         |
| Metal ions                               | 11                           |
| Ligand                                   | 12                           |
| Solvent                                  | 312                          |
| r.m.s.d. <sup>ad</sup> from ideal values |                               |
| Bond distance (Å)                        | 0.013                        |
| Bond angle (°)                           | 1.37                         |
| Ramachandran outliers, %<sup>d</sup>    | 2.4                          |
| Average atomic B factors (Å<sup>2</sup>) |                               |
| Protein/solvent                          | 38.7/43.4                    |

<sup>a</sup> Values in parentheses refer to values in the highest resolution shell. 
<sup>b</sup> Ideal values from Ref. 56. 
<sup>c</sup> r.m.s.d., root mean square deviation. 
<sup>d</sup> Calculated using a strict-boundary Ramachandran plot (37).

will be denoted as “accessible” with regard to the protein zinc ligation. In contrast, the β-CAs from the red alga *P. purpureum* (13), *Escherichia coli* (*cyn* T2-*yadF* gene product) (34), and Rv3588c from *M. tuberculosis* (27) display four protein-derived zinc ligands, the histidine, the two cysteines, and an aspartic acid residue (Asp-53). This coordination setup (denoted “blocked”) leaves no room for a water or hydroxide ligand and seems incompatible with a zinc-hydroxide mechanism. This apparent inconsistency has led to alternative mechanistic proposals (13) and also the suggestion that the aspartate may be involved in activity regulation of the enzyme by somehow changing between the accessible and blocked states (34). However, no regulation of the enzyme by somehow changing between the accessible and blocked states (34). However, no β-CA has been observed to be able to adopt both forms of the active site, and a link to activity regulation has not been established.

We have recently solved the structures of two β-CAs from *M. tuberculosis*, Rv1284 and Rv3588c (27). As reported in the previous study, Rv3588c showed carbonic anhydrase activity at pH 8.4 but not at pH 7.5. The crystal structure at pH 7.0 had the blocked active site setup with four protein-derived zinc ligands and a dimeric structure with no tetramers formed in the crystal lattice. Here we present a structure of this protein in complex with the zinc ligand thiocyanate. Interestingly, the protein now displays the accessible active site conformation, providing the first direct evidence that a β-CA active site can change between the two states. Moreover, the protein now displays the accessible active site conformation, providing the first direct evidence that a β-CA active site can change between the two states. Moreover, the complex crystallizes in a different crystal form and forms a distinct 222 symmetric tetramer, arranged as a dimer of the basic dimers. The tetrameric arrangement is also verified in solution by dynamic light scattering (DLS). Based on the results presented here, we propose that the role of the conserved Asp-53 is to provide a carboxylate group to the tetramer conversion.

**MATERIALS AND METHODS**

Cloning, Protein Expression, and Purification—Cloning, expression, and purification of Rv3588c have been described previously (27). In short, the gene was PCR-cloned from *M. tuberculosis* strain H37Rv (35), and an N-terminal His<sub>6</sub> tag was introduced. The construct was ligated into the vector pCR<sup>®</sup>T7 and transformed into *E. coli* BL21/AI, in which protein was overexpressed by induction with 0.02% arabinose at 37 °C for 2 h. The protein was purified by nickel-nitrilotriacetic acid-agarose (Qiagen) and size-exclusion (Hiload<sup>®</sup> 16/60 Superdex<sup>®</sup> 75, Amersham Biosciences) chromatography. The protein was concentrated to 34 mg/ml in Vivaspin concentrators (Vivasience) with a molecular mass cut-off of 10 kDa and stored in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl, and 5% glycerol.

**DLS Studies**—DLS measurements were performed on a PD2000 DLS Plus system (Precision Detectors) and analyzed with the PrecisionDeconvolve software. Buffers were 25 mM HEPES, pH 7.5, or 25 mM TAPS pH 8.4; both buffers also contained 100 mM sodium sulfate to mimic the activity measurement conditions.

**Crystallization, Data Collection, and Processing**—Rv3588c was crystallized using the sitting drop vapor diffusion method (36). Final conditions for crystallization were as follows: 2 μl of stock protein solution (34 mg/ml) was mixed with 1 μl of an unbuffered reservoir solution containing 200 mM NaSCN, 20% polyethylene glycol 3350, and 10% glycerol. Monoclinic crystals grew to a size of about 0.15 × 0.07 × 0.07 mm<sup>3</sup> over a period of about 2 months. Crystals were cryo-protected by soaking for a few minutes in the mother liquor but with a total concentration of 20% glycerol and flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K at beam line I71 at the MAXII synchrotron in Lund, Sweden. The data were processed and scaled using HKL2000 (37); data collection statistics are shown in Table 1.

**Structure Determination and Refinement**—The structure was solved by molecular replacement with Phaser 1.2 (38), using a dimer of our previous structure (27) as the search model. The protein and water structure was rebuilt with Arp/Warp (39) and refined with Refmac5 (40). No or cut-off for the data was used in refinement. Manual adjustments were made in O (41). The free R-value (42) was calculated from 5% of the data and monitored throughout the refinement. Refinement statistics are shown in Table 1. Accessible surface areas were calculated with the CCP4 program arealMol. All figures were made with O and rendered with Molray (43). Atomic coordinates and structure factors have been deposited at the Protein Data Bank (44) with access code 2A5V.
Carboxylate Shift Regulation Mechanism in \( \beta\)-CA

**RESULTS**

**Overall Structure**—The structure of the thiocyanate complexed Rv3588c protein displays several interesting local and global differences when compared with the blocked enzyme (27). The most apparent difference is that the protein crystallized as a distinct tetramer in the asymmetric unit (Fig. 1A), the tetramer is arranged as a dimer of dimers, and the essential dimer geometry is the same as in the previous structure, as indeed it is in all \( \beta\)-CA structures solved to date (27). In the blocked structure, however, no tetramer was formed in the crystal lattice. There are also large backbone conformational differences throughout the molecule (Fig. 1B). The Ca root mean square deviation between the dimers is surprisingly large, 1.9 Å, for ordered residues. The changes are primarily due to changes within the protein subunits, and to a lesser extent, in dimer interface geometry. Pairwise monomer Ca RMSDs are \( \sim 1.6 \) Å between ordered residues in both structures. The pairwise differences between the four monomers in the tetrameric structure are much smaller, \( \sim 0.2 \) Å for all residues (excluding the ordered His tags). The changes between the two structures are concentrated at, but not limited to, the active site and the tetramer-forming interface in which the shifts cooperate to produce a rather flat surface for the interaction between the dimers. In the blocked structure, residues 32–41 are disordered and invisible in the electron density (27). In the present tetrameric structure, however, this sequence is ordered and forms an \( \alpha\)-helix (Fig. 1B, green) that packs on the other monomer of the essential dimer to help create a cavity and restrict access to the active site. As in the structure of Rv1284 \( \beta\)-CA (27), the active site cavity is very small and almost completely shielded from solvent (Fig. 2, A and B). This is remarkably different from the \( \alpha\)-CAs, which display a very open active site.

The ordering of three of the four N-terminal His tags is of structural interest but of no biological significance. Each His tag coordinates two metal ions, contributing two histidyl side chains per metal. The remaining metal coordinations are provided by surface histidines and carboxylates from a neighboring molecule in the crystal lattice. This type of ordering of His tags is rare but not unprecedented (45). We have modeled the metal ions as zinc because they refine to reasonable B-values, and the metals coordinated by the most well ordered His tag (chain A) have the same peak height in the electron density as the active site zinc ions. Metal analysis of the protein sample did not reveal any other metals than one equivalent of zinc per protein monomer, corresponding to the active site zinc (27).

**Active Site**—The single most interesting change from the previous, dimeric structure is the shift of Asp-53 from its zinc-coordinating position to become involved in a salt link with Arg-55 (Fig. 3A). This renders one of the zinc coordination positions open for coordination by exogenous ligands; in the present structure, this position is occupied by a direct coordination of the nitrogen atom of a thiocyanate ion from the mother liquor. This position is the one proposed to be occupied by a hydroxide in the CA zinc-hydroxide catalytic mechanism. The sulfur atom of the thiocyanate ion is located in a small and rather hydrophobic pocket in the larger part of the active site cavity (Fig. 2A, the left part of the cavity). It makes no tight interactions with any part of the protein; the closest atomic distances are on the order of 4 Å. This cavity may be suitable for binding and orienting the carbon dioxide substrate, as suggested previously by Kimber and Pai (14) for the \( \textit{P. sativum} \) enzyme in complex with the substrate analogue acetate.
The carboxylate shift of Asp-53 is accompanied by major backbone changes mainly of residues Ser-54 and Val-56 (Fig. 3B). The side chain of Ser-54 flips to point in the opposite direction with atomic shifts of up to 8 Å. In concert, Val-56 undergoes an equally dramatic change, whereas the side chain of Arg-55 undergoes relatively small shifts. The result of these molecular gymnastics is the creation of a salt link between Asp-53 and Arg-55.

Quaternary Structure—To verify the relevance of the tetrameric arrangement, we performed DLS measurements of the protein in the same buffers and pH as the activity measurements. At pH 7.5, the majority of the sample (84%) exists as a species slightly larger than 40 kDa, corresponding to the calculated molecular mass for the dimer of 43.6 kDa. A small amount (16%) exists as a larger species with a calculated molecular mass of slightly over 70 kDa, most likely corresponding to the tetrameric protein. At pH 8.4, the equilibrium is shifted toward the tetrameric form. The lower molecular weight species now represents 31% of the sample, and the larger represents 69%. The results from these experiments strongly supported the hypothesis that the protein is a dimer at pH 7.5, at which the protein is inactive, but that the majority of the protein exists as a tetramer at pH 8.4, at which the protein is active.

The structures that we have determined are dimers at crystallization conditions of pH 7 (27) and tetramers at pH 7.5 (although this number is somewhat uncertain since the protein was crystallized from an unbuffered mother liquor and the buffer concentration of the protein solution is low, only 20 mM). The tetrameric form may be influenced by the sodium thiocyanate concentration and the altered coordination of the active site zinc. For whatever reason, the conformational changes result in large and complementary tetramer-forming interaction surfaces. Since the same surface is employed for tetramer formation in other β-CA, we believe that this tetramer formation is not a crystallographic artifact.

The Tetramer Interaction—The tetramer-forming interface between the pairs of dimers is much less hydrophobic than the interface within the essential dimer. This is to be expected for a transient interaction in which the affinity should be controlled by external factors such as pH. The essential homodimers (corresponding to chains A-B and C-D) bury
FIGURE 4. Buried charged residues in the tetramer-forming interface. Chain A is in blue, chain B is in green, chain C is in red, and chain D is in gold. Tyr-126 and Asp-80 are indicated. Bubbles show hydrogen bonds and salt links.

~2200 Å² of solvent-accessible area per monomer. In the dimer-dimer interface, subunit A interacts mainly with subunit D, whereas B interacts mainly with C, burying ~850 Å²/monomer (Fig. 1). The A-C and B-D interaction is much smaller, burying only ~50 Å²/monomer. Nonetheless, it includes very interesting interactions made between tyrosines 126, situated centrally in the tetramer-forming interface. Tyr-126 forms a tight hydrogen bond with its counterpart Tyr-126D, as does Tyr-126D to Tyr-126C (Fig. 4). The rest of the interface is built up of extensive networks of salt links and hydrogen bonds, sometimes via water molecules. There are no less than 12 (three from each monomer) negatively charged (Asp-129, Glu-132, and Glu-159) and 20 (five/monomer) positively charged side chains involved in interactions. Interestingly, all 20 are arginines (Arg-128, Arg-133, Arg-142, Arg-143, and Arg-167), and no lysines or histidines are involved in the tetramer-forming interaction. Asp-129, Glu-132, Arg-133, and Arg-167, together with Tyr-126, form a 20-residue charged and buried ring surrounding a main interaction surface and are not in direct contact with each other (the closest distance between any two atoms of the two Asp-80 pairs is ~6 Å). The conformation of these aspartyl side chains is different in the two dimers; in the AB dimer, the carboxylates interact via a short H-bond and are thus likely protonated, whereas in the CD dimer, the side chains are rotated away from each other. There is, however, some residual positive and negative density around the Asp-80 side chains indicative of partial disorder. The remaining charged residues, not shown in Fig. 4, Glu-159, Arg-128, Arg-142, and Arg-143, are situated around the rim of the tetramer interaction surface and are solvent-exposed. Glu-159A - Glu-159D make direct salt links to Arg-128D - Arg-128A, whereas Arg-142A - Arg-142D H-bond to the carbonyl oxygen of Ile-117D - Ile-117A situated in the C-terminal part of the helix running from the active site Cys-107, thus functioning in charge capping of the helix dipole over the tetramer-forming interface.

DISCUSSION

The proposed reaction mechanism of the best studied CAs depends on the presence of a zinc-coordinated hydroxide as the reactive species. Structures of different β-CA proteins, however, show two distinct active site setups, one accessible that leaves room for a zinc-coordinated hydroxide (14) and one blocked in which all coordination positions are occupied by protein ligands (13). The blocked setup seems incompatible with the zinc-hydroxide mechanism, and this observation led to the proposal of a variation of the mechanism by which the observed zinc-bound aspartate (Asp-53 in Rv3588c) functions as a base that activates the water molecule (13). On the other hand, Cronk et al. (34) suggest that the aspartate may be involved in activity regulation of the enzyme by somehow changing between the accessible and blocked states. In the β-CAs with an accessible active site conformation, the aspartate corresponding to Asp-53 forms a salt link to the guanidinium group of an arginine (Arg-55) in all structures solved to date. Together with the three zinc ligands, these two residues are completely conserved in β-CAs despite an otherwise very low overall sequence similarity (23, 46). This suggests an active role for these residues, such as a catalytic and/or a regulatory one. However, until now, no β-CA has been observed to be able to adopt both an accessible and a blocked form of the active site.

Here we have shown that Asp-53 is indeed able to adopt different conformations in this protein and thus change the active site between the accessible and blocked conformations. The accessible and blocked conformations observed in Rv3588c are very similar to the ones observed in other β-CAs, including the dramatic local changes in the structure. Furthermore, the accessible conformation is concurrent with a tetrameric quaternary structure in the crystal, whereas the blocked conformation is not. To show that the tetramer formation is not an artifact of crystal packing, we also performed DLS experiments with the same buffer conditions as in the activity measurements. The results demonstrated that the protein exists mainly in a tetrameric form at the active pH 8.4 but is dimeric at pH 7.5, at which the protein is inactive. Taken together, these data strongly support a regulatory role for Asp-53 by functioning as an on/off switch, possibly controlled by the multimeric state of the enzyme.

The roles of the equivalents of Asp-53 and Arg-55 have been analyzed by site-directed mutagenesis of the M. thermoautoantricipic enzyme (47). The Asp → Ala mutation reduced the $k_{cat}/K_m$ but not to an extent that would suggest that the residue plays an essential role in the reaction. The Arg → Ala mutant, on the other hand, displayed a much larger reduction in $k_{cat}/K_m$ consistent with the aspartyl residue, now unable to make a salt link to the arginine, coordinating the zinc and blocking the active site. Also in line with this interpretation, the double mutant had a higher activity than the Arg → Ala single mutant. Use of an imidazole buffer increased activity of both the Asp → Ala and the double mutant, suggestive of an additional role for the aspartyl in proton transfer and not solely activity regulation by coordinating the zinc. The activity of the Arg → Ala mutant, however, could not be rescued by imidazole, consistent with an inactivating zinc coordination by the aspartyl. Carboxylate shifts of metal-coordinating glutamates and aspartates are common in enzymes with di-metal cofactors and are believed to provide mechanistic control by determining the accessibility of open metal coordination positions (48, 49). A carboxylate shift mechanism has previously also been observed in various zinc alcohol dehydrogenases (50–52). In this mechanism, however, the carboxylate coordinates from the opposite side of the zinc relative to the water ligand, and its coordination involves a flipping of the tetrahedral coordination geometry coupled to a 2 Å movement of the zinc. It seems that carboxylate shift control
mechanisms that restrict metal ion access in various ways are common in bioinorganic chemistry, and new mechanistic schemes continue to be discovered.

The rate-limiting step for α-CAs appears to be the inter- or intramolecular proton transfer from the active site (7, 8). Activity is commonly studied at a basic pH (above 8), at which all classes of CAs have their maximal activity due to the nature of the zinc-hydroxide mechanism. In the cases in which the pH dependence of β-CA activity has been studied, a somewhat more complicated scenario emerges. β-CAs have maximal activity at around pH 9 that decreases with lower pH. In some proteins, the activity remains significant at neutral pH, however, and unlike the canonical α-CAs, the pH dependence is poorly modeled by a titration curve including only a single effective pK_a. In some cases, the data could not be adequately modeled by even two or three pK_a values, and deviations from Michaelis-Menten kinetics at low pH have also been observed (20–22, 47, 53). Moreover, other β-CAs display a steep pH-dependent variation of activity suggestive of an inactive form at neutral pH (8, 27, 34). In addition to the intricate pH dependence of the activity, many other factors, including buffer and ion concentration, may influence the activity, which makes comparing activity measurements of CAs difficult. In any case, the complex activity behavior suggests that β-CAs are regulated in some way. In the present example, the key feature seems to be the opening and closing of the active site, possibly coupled to the pH-dependent tetramerization.

The charge bias of 20 positive versus 16 negative residues, including Asp-80, in the tetramer-forming interface could explain why a higher pH is advantageous for tetramer formation in this case. Charge neutrality in the interface could then be achieved by deprotonation of sites that are tuned to an appropriate pK_a by the surrounding residues. One example may well be the H-bonded tyrosine pairs that could delocalize a negative charge over both aromatic rings if a single proton is shared between the two tyrosines. The hydrogen-bonding distances in these pairs refine to unreasonably short distances of 2.0 and 2.2 Å, respectively, although the side-chain conformation is supported by the density (the density can be viewed at the Electron Density Server at Uppsala University) (54). The short distances may be indicative of so-called "low barrier” H-bonds (55). These H-bonds are usually found where the pK_a value of the H-bond donor and acceptor are very similar and the proton can be considered to be shared equally by the two groups. Moreover, the central part of the interaction surface is composed of 20 charged residues (Fig. 4) with intricate hydrogen-bonding and salt link patterns. It seems probable that this unusual setup produces an area that is tuned to be protonated/deprotonated at a suitable pH to promote the tetramer formation. As an example, the Asp-80 pairs show variation in conformation and H-bonding patterns between the two dimers, indicative of structural plasticity and possibilities for different levels of protonation.

The two tyrosines and the charged cluster are not conserved among the β-CAs, and the oligomeric states of the different β-CAs vary between species and reports. For example, the E. coli cynT β-CAs display a tetramer to dimer conversion in the presence of the product, bicarbonate (19).

The dramatic changes around Asp-53 in the active site (Fig. 1B, arrow 1) are most likely due to the movement of the aspartyl side chain. The large loop movement indicated in Fig. 1B by arrow 2 is located at the tetramer-forming interface and can be directly attributed to that interaction. In addition to this, there are conformational changes throughout the protein. It is tempting to speculate that the tetramerization is a way of controlling the overall protein structure, and in that way, facilitating the conversion between the accessible and blocked states. Based on the available data on this and other β-CAs, it also seems possible that this switch can be controlled by different factors in different β-CAs. In the M. tuberculosis Rv3588c, one way of controlling it is pH, perhaps giving the protein the possibility to work as a pH regulator because its rapid and acid-producing reaction is activated at high pH. In the E. coli cynT β-CA, on the other hand, bicarbonate leads to dissociation of the tetramer, which may be a mechanism for product inhibition in this case. Different modes of regulation could also explain the presence of several β-CAs in the same organism. For example, in addition to cynT, E. coli also has the cynT2 β-CA that has a similar pH dependence of its activity as Rv3588c (34).

The dimeric interaction geometry is conserved in all β-CAs solved to date. Dimerization is required to create the active site surroundings (Fig. 2) and, considering the nature of the dimer interface, most likely necessary for proper folding and solubility of the protein. In the cases in which the proteins display tetrameric arrangements in the crystal packing, however, there are some differences in quaternary structure. The dimer-of-dimers arrangement with 222 symmetry is preserved, and the interaction is also made with the same surface in all cases, but the relative orientations of the dimers change. The variation is best described as a rotation around the essential dimer 2-fold for one of the dimers relative to the other (in Fig. 1A, imagine the left dimer rotating relative to the right). The E. coli tetramer and the P. purpureum dimer (mimicking the tetramer) produce a dihedral angle between the strands of the central β-sheets in the dimers of ~45°. In the Rv3588c tetramer, one dimer is rotated more or less 45° from this orientation, and the angle between the strands is on the order of 90° (Fig. 1A). The Rv1284 tetramer is rotated even further, producing an angle of some 120°, using the same reference. In contrast, the P. sativum protein adopts an octameric arrangement with different interaction surfaces connecting the dimers than in the tetrameric proteins. The higher multimeric interactions thus display a larger variation than the essential dimers. Moreover, there is no clear correlation with the blocked/accessile active site structure and a dimeric/tetrameric arrangement in the crystal packing for all solved β-CA structures. For example, the E. coli structure has the blocked active site but displays a tetrameric arrangement in the crystal. This tetramer could simply be an effect of crystal packing, but it is also possible that this protein does not shift between the multimeric states as Rv3588c does. Furthermore, the accessible active site conformation is seen in the M. thermoaerotrophicum structure that has a dimeric arrangement in the crystal. In this case, a HEPES buffer molecule H-bonds directly to the aspartyl residue in one of the active sites, which may influence its conformation. Consequently, from the available structural data of β-CAs, it is not possible to conclude whether a coupled tetramerization and opening of the active site is a common feature. Nevertheless, based on the fact that the Asp/Arg pair is conserved in all β-CAs sequenced to date, it seems most likely that the Asp carboxylate shift regulation of activity is in effect in all β-CAs, although it may be regulated in different ways and by different factors.

The identification of the opening of the active site as a key regulatory feature in β-CAs is certainly interesting from an inhibitor design perspective. In this particular case, the active site cavity in the active state of the protein is very small and likely restricts access to molecules larger than the substrate. However, we have also shown that the protein possesses a large degree of flexibility, both in the active site and in the protein as a whole, and there may be conformational states of the enzyme that are more amenable to inhibitor binding. For example, an inhibitor targeting the inactive state with the disordered helix could potentially lock the protein in that conformation. In addition, the coupling of activity to a tetramerization event in the Rv3588c β-CA reveals a possible new point of attack for drug design against this protein that is
not present in mammals and has been shown to be required for M. tuberculosis growth in vivo.

Acknowledgments—We thank Stewart Cole of the Institut Pasteur for M. tuberculosis DNA and Daniel Ågren for help with DLS measurements.

REFERENCES

1. Frieden, T. R., Sterling, T. R., Munsiff, S. S., Watt, C. J., and Dye, C. (2003) Lancet 362, 887–899
2. Flynn, J. L., and Chan, J. (2001) Annu. Rev. Immunol. 19, 93–129
3. Flynn, J. L., and Chan, J. (2003) Curr. Opin. Immunol. 15, 450–455
4. Sassetti, C. M., and Rubin, E. J. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 12899–12994
5. Stadie, W. C., and O'Brien, H. (1933) J. Bacteriol. 23, 2323–2330
6. Meldrum, N. U., and Liljas, A. (1993)
7. Chegwidden, W. R., Carter, N. D., and Edwards, Y. H. (eds) (2000)
8. Tripp, B. C., Smith, K., and Ferry, J. G. (2001)
9. Meldrum, N. U., and Roughton, F. J. W. (1933)
10. Lindskog, S., and Liljas, A. (1993)
11. SoltesRak, E., Mulligan, M. E., and Coleman, J. R. (1997)
12. Lindskog, S., and Liljas, A. (1993)
13. Smith, K. S., and Ferry, J. G. (2000)
14. Kimber, M. S., and Pai, E. F. (2000)
15. Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., and Rees, D. C. (1996)
16. Christianson, D. W., and Cox, J. D. (1999)
17. Strop, P., Smith, K. S., Iverson, T. M., Ferry, J. G., and Rees, D. C. (2001)
18. Christianson, D. W., and Cox, J. D. (1999)
19. Chegwidden, W. R., Carter, N. D., and Edwards, Y. H. (eds) (2000)
20. Rowlett, R. S., Chance, M. R., Wirt, M. D., Sidelinger, D. E., Royal, J. R., Woodroffe, M.,
21. Guilloton, M. B., Korte, J. J., Lamblin, A. F., Fuchs, J. A., and Anderson, P. M. (1992)
22. Smith, K. S., Cosper, N. J., Stalhandske, C., Scott, R. A., and Ferry, J. G. (2000)
23. Smith, K. S., Jakubzick, C., Whittam, T. S., and Ferry, J. G. (1999)
24. Smith, K. S., and Ferry, J. G. (1999) J. Biol. Chem. 274, 6605–6613
25. Strop, P., Smith, K. S., Iverson, T. M., Ferry, J. G., and Rees, D. C. (2001) J. Biol. Chem. 276, 10299–10305
26. Christianson, D. W., and Cox, J. D. (1999) Annu. Rev. Biochem. 68, 33–57
27. Smith, K. S., and Ferry, J. G. (2000) Annu. Rev. Phys. Chem. 51, 571–594
28. Giege, R., and Ducruix, A. (1999) Annu. Rev. Biochem. 68, 6247–6253
29. Smith, K. S., and Ferry, J. G. (2000)
30. Ho¨gbom, M., Andersson, M. E., and Nordlund, P. (2001) Nature 407, 537–544
31. Tepass, U., and Graf, G. (1999) Crystallography of Nucleic Acids and Proteins, 2nd ed., Oxford University Press, New York
32. Toth, A., and Sturtevant, J. (1973)
33. Kandel, M., Gornall, A. G., Cybulsky, D. L., and Kandel, S. I. (1978) Nature 276, 571–581
34. Cronk, J. D., and Endrizzi, J. A., Cronk, M. R., O'Neill, J. W., and Zhang, K. Y. (2001) Plant Physiol. 117, 1341–1349
35. Rumeau, D., Cuine, S., Fina, L., Gault, N., Nicole, M., and Peltier, G. (1996) Planta 199, 79–88
36. Tobin, A. J. (1970) J. Biol. Chem. 245, 2656–2666
37. Kisiel, W., and Graf, G. (1972) Phytochemistry 11, 113–117
38. Sanghani, P. C., Boston, W. F., and Hurley, T. D. (2002) Acta Crystallogr. Sect. D Biol. Crystallogr. 58, 240–255
39. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
40. Myszyna, B., and Zablotowicz, J. (1992) Nature 355, 472–475
41. Harris, M., and Jones, T. A. (2001) Acta Crystallogr. Sect. D Biol. Crystallogr. 57, 1201–1203
42. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) Nucleic Acids Res. 28, 235–242
43. O'Neill, J. W., Kim, D. E., Baker, D., and Zhang, R. Y. J. (2003) Acta Crystallogr. Sect. D Biol. Crystallogr. 59, 480–487
44. Smith, K. S., and Ferry, J. G. (2000) FEBS Microbiol. Rev. 24, 335–366
45. Smith, K. S., Ingram-Smith, C., and Ferry, J. G. (2000) J. Bacteriol. 182, 4240–4245
46. Solomon, E. L., Brunold, T. C., Davis, M. I., Woodroffe, M., Wang, Y., and Ferry, J. G. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 12989–12994
47. Smith, K. S., and Ferry, J. G. (2000)
48. Solomon, E. L., Brunold, T. C., Davis, M. I., Woodroffe, M., Wang, Y., and Ferry, J. G. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 12989–12994
49. Ho¨gbom, M., Andersson, M. E., and Nordlund, P. (2001) J. Biol. Inorg. Chem. 6, 315–323
50. Esposito, L., Bruno, F., Sica, F., Raia, C. A., Giordano, A., Rossi, M., Mazzarella, L., and Zagari, A. (2003) Biochemistry 42, 14397–14407
51. Karslon, A., Al-Hmad, M., Jonsahn, K., Shafqat, J., Jornvall, H., and Ramaswamy, A. (2003) Biochemistry 42, 5371–5378
52. Wang, Y., and Ferry, J. G. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 12989–12994
53. Ho¨gbom, M., Andersson, M. E., and Nordlund, P. (2001) J. Biol. Inorg. Chem. 6, 315–323
54. Esposito, L., Bruno, F., Sica, F., Raia, C. A., Giordano, A., Rossi, M., Mazzarella, L., and Zagari, A. (2003) Biochemistry 42, 14397–14407
55. Karlsson, A., Al-Hmad, M., Jonsahn, K., Shafqat, J., Jornvall, H., and Ramaswamy, A. (2003) Biochemistry 42, 14397–14407
56. Sanghani, P. C., Boston, W. F., and Hurley, T. D. (2002) Biochemistry 41, 15189–15194
57. Karslon, A., Al-Hmad, M., Jonsahn, K., Shafqat, J., Jornvall, H., and Ramaswamy, A. (2003) Biochemistry 42, 14397–14407
58. Sanghani, P. C., Boston, W. F., and Hurley, T. D. (2002) Biochemistry 41, 15189–15194
59. Rowlett, R. S., Tu, C., Murray, P. S., and Chamberlin, J. E. (2004) Arch. Biochem. Biophys. 432, 25–32
60. Kleywegt, G. J., and Jones, T. A. (1996) Structure (Lond.) 4, 1395–1400