Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Refined structure of bovine carboxypeptidase A at 1.25 Å resolution

Alexandra Kilshtain-Vardi,[a] Meir Glick,[b] Harry M. Greenblatt,[c] Amiram Goldblum[b] and Gil Shoham[a]*

[a]Department of Inorganic Chemistry and the Laboratory for Structural Chemistry and Biology, The Hebrew University of Jerusalem, Jerusalem 91904, Israel, [b]The David R. Bloom Center for Pharmacy at the Hebrew University and the Department of Medicinal Chemistry and Natural Products, School of Pharmacy, Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem 91904, Israel, and [c]Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel

Correspondence e-mail: gil2@vms.huji.ac.il.

Received 8 July 2002
Accepted 8 January 2003

PDB Reference: bovine carboxypeptidase A at 1.25 Å, 1m4l, r1m4lsf.

The crystal structure of the bovine zinc metalloproteinase carboxypeptidase A (CPA) has been refined to 1.25 Å resolution based on room-temperature X-ray synchrotron data. The significantly improved structure of CPA at this resolution (anisotropic temperature factors, R factor = 10.4%, $R_{\rm free}$ = 14.5%) allowed the modelling of conformational disorders of side chains, improved the description of the protein solvent network (375 water molecules) and provided a more accurate picture of the interactions between the active-site zinc and its ligands. The calculation of standard uncertainties in individual atom positions of the refined model of CPA allowed the deduction of the protonation state of some key residues in the active site and confirmed that Glu72 and Glu270 are negatively charged in the resting state of the enzyme at pH 7.5. These results were further validated by theoretical calculations that showed significant reduction of the p$K_{\rm a}$ of these side chains relative to solution values. The distance between the zinc-bound solvent molecule and the metal ion is strongly suggestive of a neutral water molecule and not a hydroxide ion in the resting state of the enzyme. These findings could support both the general acid/general base mechanism, as well as the anhydride mechanism suggested for CPA.

1. Introduction

Zinc metalloproteinases comprise a large family of enzymes with a wide variety of biological roles. These enzymes are characterized by the use of zinc as an essential catalytic entity (Vallee & Galdes, 1984). One of the most widely studied zinc metalloenzymes is carboxypeptidase A (CPA), which catalyzes both exopeptidase and exoesterase reactions. The three-dimensional structure of CPA was determined several years ago at 1.54 Å resolution (Rees *et al.*, 1983) and recently two new more complete and improved structures of the native enzyme were also refined at 1.5 Å resolution (Teplyakov *et al.*, 1993; Greenblatt *et al.*, 1998).

CPA has been the subject of extensive research in the last 30 years (Coleman & Vallee, 1961; Neurath *et al.*, 1970; Shoham *et al.*, 1984, 1988; Christianson & Lipscomb, 1989; Feinberg *et al.*, 1993, 1995; Greenblatt *et al.*, 1998). However, the exact catalytic mode of action of this representative enzyme and the rest of the zinc proteinase family is still not completely clarified. Several catalytic pathways have been suggested for CPA, which may be divided into two major groups. The first general pathway starts from direct nucleophilic attack by the active-site carboxylic group of Glu270 on the carbonyl group of the target peptide bond of the substrate (Lipscomb, 1980). The second general pathway involves a

general acid/general base (GAGB) mechanism in which the water that attacks the peptide carbonyl is initially activated by the zinc/Glu270 system (Christianson & Lipscomb, 1989) or by the C-terminal carboxylic group of the substrate (Mock & Zhang, 1991). Nevertheless, there is no conclusive evidence as yet for either of these pathways, either from biochemical experiments (Makinen *et al.*, 1976; Breslow & Wernick, 1977; Auld *et al.*, 1984; Lee & Kim, 2000) or from theoretical simulations (Kilshtain-Vardi *et al.*, 2002).

In addition to specific interest in the detailed mechanism of catalysis of CPA, it is hoped that, once deduced, the mechanism will serve as a prototypical catalytic scheme for the other members of the zinc proteinase family. This family includes important biological and medical target enzymes such as angiotensin-converting enzyme (ACE) which plays an important part of the blood-pressure control mechanism and has been a major target for developing inhibitors to treat hypertension (Soffer, 1976; Ondetti & Cushman, 1982). Another important member of this family is enkephalinase, which is involved in the natural mechanism of pain relief *via* the hydrolysis of enkephaline (Malfroy *et al.*, 1978; Schwartz *et al.*, 1990). Also included in this family are the recently characterized matrix metalloproteinases (MMPs) that have been linked to a variety of chronic disease conditions including cancer (Summers & Davidsen, 1998), arthritis (Bottomley *et al.*, 1998), osteoporosis (Tezuka *et al.*, 1994; Ohishi *et al.*, 1995) and multiple sclerosis (Woessner, 1994). Accumulation of information about the exact mechanism of these enzymes is hence very important and expected to aid, among other things, the rational design of mechanism-based inhibitors that block (or control) the excessive enzymatic activity when necessary. As such, these inhibitors are expected to serve as useful drugs for the treatment of the various clinical disorders mentioned above.

The importance of determining high-resolution three-dimensional structures stems from the fact that as the resolution improves it may be possible to establish a more accurate and detailed view of the protein structure and of its functional implications. Recent examples in the literature (*e.g.* Teixeira *et al.*, 2001; Sauter *et al.*, 2001) demonstrate that anisotropic refinement of high-resolution structures can help to determine the protonation state of key side chains, to locate disorder in side-chain positions and to establish a more complete representation of hydrogen-bond networks and solvent water networks in these structures. In addition, it has been shown that the error in atomic positions decreases with the improvement of the angular resolution of the diffraction data and with decreases in the crystallographic $R_{\rm factor}$ (Luzzati, 1952). An estimation of the structural error scale is extremely important for critical evaluation of small conformational changes in a series of slightly different macromolecular structures. The significance of structural differences is especially important if the results provide a basis for conclusions concerning a detailed mechanism of action or a detailed mode of substrate binding. Many recent advances in diffraction data measurement, including high-brilliance synchrotron-radiation sources, efficient X-ray optics and sensitive two-dimensional detection technologies, have enormously improved the possibilities of obtaining such high-resolution data with reasonably high quality and with a relatively short experimental time.

In this paper, we present a new and significantly improved structure of CPA which is based on room-temperature synchrotron data to 1.25 Å resolution and which was refined for the first time with anisotropic temperature factors for all non-H atoms of the model. Although this structure does not yet enable a unique determination of all bond types and proton positions, it can be used to deduce the protonation states of some side chains in the enzyme and also to obtain an improved and more accurate view of side-chain conformations, the positions of water molecules and a more exact geometry of the active-site zinc ion and its ligands. It should also improve our understanding of the catalytic mechanism of CPA, with possible implications for the general family of zinc proteinases. The structure could also serve as a better starting point for subsequent molecular-modelling studies.

2. Methods

2.1. Crystallization and data collection

CPA was purchased from Sigma Co.; the protein was further purified and crystallized according to previously described procedures (Shoham *et al.*, 1984, 1988). The fully grown crystals of native CPA were cross-linked prior to X-ray data measurement in order to impart greater mechanical strength. For this procedure, the buffer was changed to 0.02 *M* Veronal, 0.1 *M* LiCl and the cross-linking of the crystals was carried out with 0.025% glutaraldehyde for about 12 h. The buffer was subsequently changed back to 0.02 *M* Tris, 0.1 *M* LiCl pH 7.5, which was then used for the following structural studies.

The X-ray diffraction data collection was performed at room temperature on a rather large single crystal of native CPA ($\sim$0.3 $\times$ 0.3 $\times$ 0.8 mm). Diffraction measurement was performed on the X12B beamline, NSLS synchrotron facility, Brookhaven National Laboratories (New York, USA) using a MAR Research (MAR-300) imaging-plate area detector at a fixed wavelength (λ = 1.085 Å). In order to minimize radiation-damage effects, the crystal was translated along its long axis during the data collection, enabling a full diffraction data measurement from three different parts of the crystal. Images were integrated and scaled with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) to give the final merged intensity data set. Diffraction extended to 1.25 Å, but completeness in the highest resolution shell (1.25–1.28 Å) was only 36%. In lower resolution shells, the completeness was significantly higher and crossed the 80% level in the 1.35–1.38 Å resolution shell (83%). The full resolution range was used in the refinement. Representative measurement details and statistics for these data are listed in Table 1.

2.2. Structural refinement

The CPA crystals used for the present study proved to be completely isomorphous to previously studied native CPA crystals; hence, initial phases for the current refinement were

Table 1
Data measurement and refinement statistics for the present structure of native CPA (1.25 Å), compared with the corresponding PDB structures of CPA (1yme and 2ctb).

	1.25 Å	1yme	2ctb
Unit-cell parameters			
a (Å)	51.69	51.62	51.60
b (Å)	60.27	60.17	60.27
c (Å)	47.45	47.17	47.25
β (°)	97.47	97.40	97.27
No. of unique reflections	64976	34352	37803
Resolution (Å)	30–1.25	30–1.53	30–1.5
Completeness (last shell) (%)	86.3 (36.3)	79	88.7
$\langle I/\sigma(I)\rangle$	13.5	–	–
$R_{\rm sym}$ (last shell) (%)	6.6 (12.7)	3.6	–
No. protein atoms	2438	2438	2438
No. solvent atoms	375	192	218
R factor ($R_{\rm free}$) (%)	10.4 (14.51)	14.8	15.1
R.m.s.d. from ideal geometry			
Bond lengths (weight†) (Å)	0.015 (0.020)	0.017 (0.015)	0.014 (0.015)
Bond angles (weight†) (°)	0.030 (0.040)	0.036 (0.030)	0.035 (0.030)
Chiral volumes (weight†) (Å^3)	0.076 (0.100)	0.202 (0.200)	0.139 (0.200)

† The weights correspond to $1/\sigma^2$.

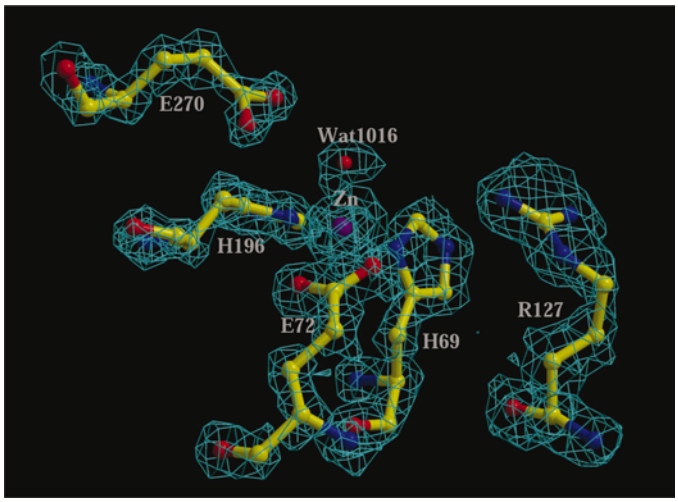

Figure 1
Electron-density 'omit' map around the active site of native CPA at 1.25 Å resolution. Electron-density contour levels are at +4.5σ (cyan). Superimposed on the density is the final 1.25 Å model of CPA (ball-and-stick representation, common atom colour codes), featuring the zinc and its ligands, as well as the catalytic residues Arg127 and Glu270.

obtained from the model of the native CPA structure which was reported recently at 1.5 Å resolution (1yme; Greenblatt *et al.*, 1998). A starting electron-density difference map [using ($F_o - F_c$) coefficients] generated from this model showed very good correlation between the initial model of the enzyme and the experimental data.

The first stage of refinement was performed using the *TNT* refinement program (Tronrud, 1992). Initial refinement restraints were loosened to allow the model to relax (five cycles), followed by a second round of refinement (nine cycles) with tight restraints until the refinement converged. Electron-density difference maps were generated after every round of refinement in order to follow the progress of the refinement and to confirm the improvement of the model.

The final model obtained from the *TNT* refinement was subsequently submitted to further refinement with *SHELX*97 (Sheldrick & Schneider, 1997). 5% of the diffraction data were randomly selected and omitted during the refinement for cross-validation analysis by means of the free R factor ($R_{\rm free}$; Brünger, 1992). Each step consisted of 10–15 cycles of CGLS minimization. Default values of the program for distance, planarity and chiral restraints and a diffuse solvent model were used throughout the refinement. New ordered solvent molecules were found in the newly generated difference Fourier maps [with peaks > 4σ, $B < 60$ Å^2] and incorporated in the model for the next cycle. Electron-density maps with ($2F_o - F_c$) and ($F_o - F_c$) coefficients were calculated after each step in order to examine the correlation between the experimental density and the emerging model. The model (protein and solvent) was checked and rebuilt using the program *CHAIN* (Sack, 1988). The R factor and $R_{\rm free}$ values calculated after the first *SHELX*97 run using isotropic temperature parameters (B factors) were 16.79 and 19.54%, respectively. Refinement of anisotropic displacement parameters (ADPs) of all non-H protein atoms dramatically decreased both R and $R_{\rm free}$ to 11.25 and 15.57%, respectively. ADPs of solvent molecules were restrained to be approximately isotropic. At this stage, clear electron density was observed near several amino-acid side chains, which were subsequently modelled in two alternative (disordered) conformations. The occupancy factors of solvent molecules were not refined, except those of water molecules associated with multiple protein conformations or those with nearly overlapping atomic positions ($d < 2$ Å).

When convergence of the CGLS minimization was achieved, all reflections used to calculate $R_{\rm free}$ were again included in the refinement, leading to a final crystallographic R factor of 9.94% for 58 913 measured reflections with $F_o > 4\sigma(F_o)$ and a final R factor of 10.4% for all data with $F_o > 1\sigma(F_o)$ (64 976 reflections) in the resolution range 30–1.25 Å. Table 1 shows the final refinement statistics and the final deviations from target stereochemistry for the *SHELX*97 refinement. At the end of the refinement, a final run of blocked-matrix refinement was performed in order to estimate individual coordinate errors. To this end, the protein was split into 20 overlapping blocks of 15 protein residues and one block of seven residues, with an overlap of two residues between successive blocks. Positional parameters and ADPs were refined in each of these blocks. The quality of the structure was assessed using the programs *PROCHECK* (Laskowski *et al.*, 1993) and *SHELXPRO* (Sheldrick & Schneider, 1997).

2.3. Proton assignment by theoretical methods

Even with the present 1.25 Å resolution structure of CPA, unambiguous determination of the protonation states of many key side chains is not directly possible based on X-ray results alone. However, a combination of high-resolution experimental coordinates with novel theoretical methods can help to

position polar protons and to provide estimates for the pK_a values of some crucial side chains, as described below.

H atoms were added to the current model of CPA at 1.25 Å by an energy-based stochastic method suggested by Glick & Goldblum (2000). This method is essentially a global optimization approach that has been recently extended for the prediction of rotameric states of side chains in proteins (Glick *et al.*, 2003). The algorithm assumes that the coordinates of the 'heavy' (non-H) atoms are known. First, non-rotatable H atoms, such as those found on the peptide backbone, are added according to standard geometric premises (bond lengths and angles). In the next stage, rotatable H atoms, such as those found on water molecules and the polar side chains of serine, threonine and tyrosine, are added according to this global optimization method. Unlike 'self-consistency' algorithms (*e.g.* Brünger & Karplus, 1988), the stochastic method optimizes the coordinates of all the polar H atoms simultaneously and not in a stepwise manner which might fail to detect the global energy minimum. Indeed, the stochastic approach achieved better accuracy than other methods when tested on five high-resolution crystal structures in which the protons positions were experimentally determined (Glick & Goldblum, 2000). The proton-addition program has an interface to an in-house program that calculates the pK_a values and charges of all appropriate residues for any given pH value (Seroussi, Glick & Goldblum, unpublished results). The pK_a values vary as a result of a modified Tanford–Kirkwood (Havranek & Harbury, 1999) calculation with all partial charges of the protein atoms. This calculation was performed using a Mehler-type dielectric (Mehler & Guarnieri, 1999) with screening and with proper weights for partially ionized residues (*i.e.* if the pK_a and the pH are close, both the neutral and ionized partial charges serve in calculating the electrostatic energy in the required proportions). H atoms are initially positioned at low pH and the pH is then gradually increased while charges are added or eliminated for acids and bases, respectively.

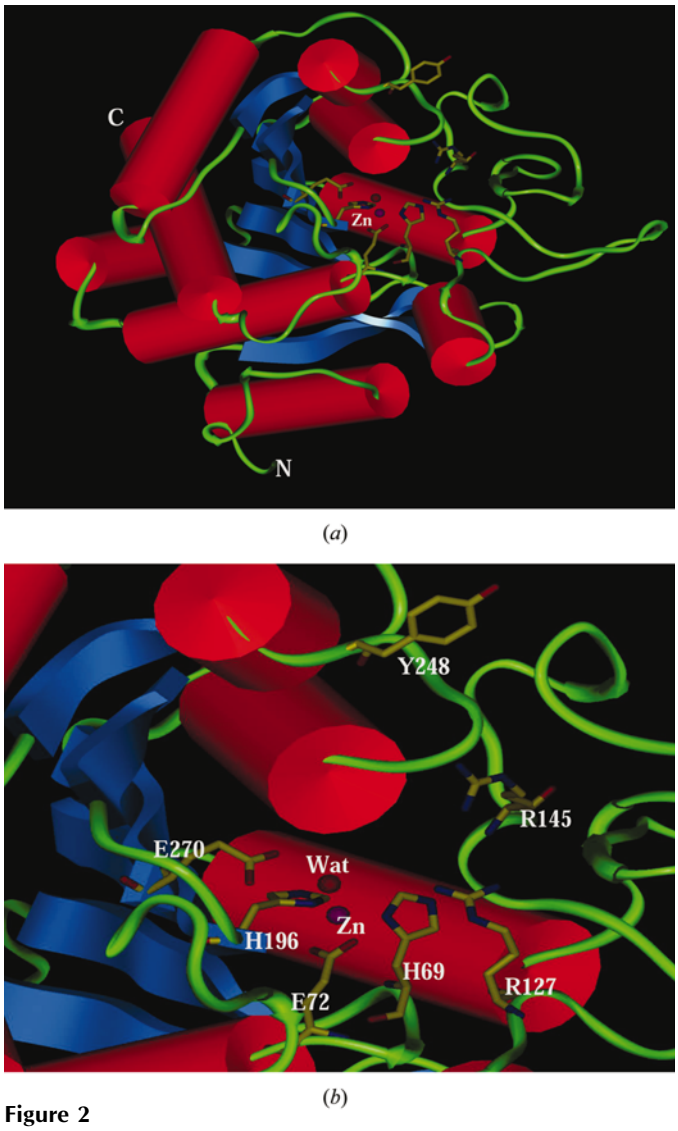

Figure 2
Secondary-structure diagram of the final structure of native CPA at 1.25 Å resolution, showing α-helices as cylinders (red), β-strands as ribbons (blue) and turns and random coil as thin tubing (green). Each secondary-structure element has directional cues: indented ends of helices represent the N-terminus of the helix, while conical ends represent the C-terminus; ribbons and turns have arrows at the C-terminal end. The active-site zinc cation (purple sphere), the zinc-bound water (dark-red sphere, Wat) and several key residues (stick representation) are shown in more detail. (*a*) An overall view of the enzyme. (*b*) A close-up view of the active site.

Input to the proton-addition program consisted of all the protein atoms of the final refined model of CPA at 1.25 Å. Since the program also adds H atoms to water molecules, the complete set of water molecules would have made the calculation too complex. This is the reason that only 38 water molecules (*i.e.* O atoms) were chosen within a radius of 4 Å around residues whose side chains may form hydrogen bonds with water molecules near the active site (serine residues Ser57, Ser102, Ser199, Ser158, Ser134, Ser135; threonine residues Thr78, Thr210; tyrosine residues Tyr169, Tyr248, Tyr240, Tyr198). This model served as a basis for non-polar and polar proton positioning. The output of the proton-addition program was then used as input for the pK_a calculations. The combination of the hydrogen-addition program together with the pK_a calculation program was employed in order to obtain more accurate results.

3. Structural results and discussion

3.1. The overall structure of CPA at 1.25 Å resolution

The final model of native CPA at 1.25 Å resolution consists of 2807 non-H atom positions representing 307 protein residues, one Zn^{2+} ion (at the active site) and 375 crystallographically determined water molecules. The final crystallographic R factor was 10.4% (at 1σ cutoff) and the last R_{free} value calculated before the inclusion of all the data was 14.51%. Both of these parameters indicate a consistent and reliable refinement of the structure. The final electron-density map is of excellent clarity, with the vast majority of the protein molecule displaying well defined electron density in the $(2F_o - F_c)$ map, even when contoured at the 4.5σ level (Fig. 1). The overall standard uncertainties for the positional

parameters, estimated from the blocked-matrix least-squares refinement, are between 0.01 and 0.05 Å for 80% of the protein atoms. Although it was recently suggested that this method underestimates the uncertainties by about 10% (Merritt *et al.*, 1998), the assignment of uncertainty values on a per-atom basis is a significant improvement relative to previous models, where it was only possible to obtain an estimate of the overall coordinate error. Final refinement parameters for the current CPA model (1.25 Å resolution, Fig. 2*a*) are summarized in Table 1 and compared with the corresponding
parameters in the previously reported structures of native CPA (PDB codes 1yme and 2ctb).

The Ramachandran plot calculated for the final CPA model with the program *PROCHECK* (Laskowski *et al.*, 1993) shows that 88.6% of the residues are in the 'most favoured' regions of the φ, ψ plane, 11.4% of residues are in the 'additionally allowed' regions and only one amino-acid residue (Ser199) appears in a 'disallowed' region. The abnormal conformation of Ser199 is explained by the presence of a *cis* peptide bond between Ser199 and Phe198, an unusual *cis* conformation which is probably required for the correct folding of the protein.

Although the significance of anisotropic temperature-factor refinement of a macromolecule at the current resolution can be borderline, the large drop (∼5%) in the R and $R_{\rm free}$ values following its implementation completely justifies this procedure. After taking ADPs into account, it became clear from inspection of the electron density that about 3% of the residues have more than one conformation. These multiple conformations were assigned and refined as discussed below.

Another improvement in the anisotropically refined model of CPA is the large number of water molecules, which is almost double (375 molecules) that found in previous models of native CPA [192 water molecules in the 1yme structure (Greenblatt *et al.*, 1998) and 218 water molecules in the 2ctb structure (Teplyakov *et al.*, 1993)]. As expected, most of the newly determined water molecules are located around the outer shell of the enzyme and have significantly higher temperature factors.

3.2. Amino acids in multiple conformations

As mentioned above, the high-resolution data and the ADP refinement of CPA lead to a sharp improvement in the electron-density map, indicating, among other things, multiple electron densities for some of the amino-acid side chains (and the corresponding solvent molecules). In nine of these cases (Arg2, Ser3, Leu24, Ile47, Ser57, Gln92, Ser194, Arg276 and Glu302), two alternative conformations were clearly visible and were modelled in subsequent refinement steps (Table 2).

The two alternative conformations for the side chain of Ser194 observed for the 1.25 Å structure were clearly observed in the two previous structures of CPA (2ctb; Teplyakov *et al.*, 1993; 1yme; Greenblatt *et al.*, 1998). In contrast, the structure of the complex of CPA with the inhibitor benzylsuccinate bound in the active site (Feinberg *et al.*, 1993, 1995) showed the side chain of Ser194 exclusively in a single conformation. It therefore apears that binding of affectors (substrates, analogues or inhibitors) in the active site causes some of the side chains (*e.g.* Ser194) to favour one of the two alternative conformations observed in the uncomplexed native CPA. As such, these cases of 'locked conformations' may be involved in further stabilization of the enzyme–affector complex and/or with the enhancement of binding specificity.

Table 2
Amino-acid residues with multiple conformations in the 1.25 Å CPA structure.

Residue	Partial occupancy	Specific side-chain atoms
Arg2	0.43/0.57	C_δ, N^ε, C^ζ
Ser3	0.39/0.61	O^γ
Leu24	0.62/0.38	$C^{\delta1}$, $C^{\delta2}$
Ile47	0.62/0.38	$C^{\delta1}$
Ser57	0.72/0.28	All atoms
Gln92	0.35/0.65	C^δ, $O^{\varepsilon1}$, $N^{\varepsilon2}$
Ser134	0.85/0.15	All atoms
Arg276	0.56/0.44	C^δ, N^ε, C^ζ
Glu302	0.84/0.16	O^γ

3.3. The protonation state of groups in the active site

Since the acidity of a given functional group should depend on its surroundings, perturbation is expected for the pK_a of side chains buried inside proteins or in cavities near the surface (Kaslik *et al.*, 1999; Dyson *et al.*, 1997; Gladysheva *et al.*, 1996). Knowledge of the actual protonation state of side chains that are located in the active site of an enzyme may be difficult to ascertain, yet it may be crucial for understanding the exact catalytic mechanism of the enzyme. Since certain side-chain functional groups undergo statistically significant changes in bond lengths upon deprotonation, accurate crystal structure models could provide information about the protonation states of various functional groups in the active site. This kind of correlation between local structural parameters and protonation states is best demonstrated in the crystal structure of 'small molecules', where the crystal structures are determined at a very high resolution and with a relatively large ratio of observations to parameters for the refinement. Based on such structures, the mean bond lengths of the C—OH and C═O bonds in the side-chain carboxylate group of a neutral glutamic acid or aspartic acid are 1.308 ± 0.019 and 1.214 ± 0.019 Å, respectively (Lide, 1991). Deprotonation of these carboxylate groups results (according to the accurate crystal structure of small molecules) in two equal partial double bonds of about 1.25–1.26 Å. Analysis of the CPA structure reported here shows that the side-chain carboxylate group of Glu270 has bond distances of 1.25 ± 0.03 Å for C^δ—$O^{\varepsilon1}$ and 1.24 ± 0.03 Å for C^δ—$O^{\varepsilon2}$. Similar distances were found for the side-chain carboxylate of Glu72: 1.24 ± 0.02 Å for C^δ—$O^{\varepsilon1}$ and 1.24 ± 0.02 Å for C^δ—$O^{\varepsilon2}$. These results are comparable to the mean values observed for deprotonated carboxylic acids, as discussed above, and therefore suggest that the side-chain carboxylic acid groups of

Glu72 and Glu270 are deprotonated in the native uncomplexed enzyme at pH 7.5. In order to make sure that the results were not affected by the initial model used for the refinement, we performed a parallel refinement in which the side-chain carboxylate groups of Glu72 and Glu270 were modelled initially as neutral groups with unequal bond lengths. At the end of this procedure, similar values were obtained for the two $C^{\delta}-O^{\varepsilon}$ bond lengths of the side chains of both residues (~1.24 Å), confirming the final structural assignments discussed above. These results are also consistent with theoretical calculations (as detailed below), which predict ionized states for both Glu72 and Glu270.

Of even greater relevance to the catalytic mechanism of CPA is the exact protonation state of the water molecule bound to the active-site zinc ion. Although binding of water molecules to metal ions is known to reduce their pK_a (Wooley, 1975), the actual pK_a obtained is very sensitive to the ligands involved in the zinc coordination, as shown, for example, by Kimura *et al.* (1990). In their work, the observed pK_a values of a zinc-bound water ranged from 7.3 to 9.8 depending on the immediate zinc coordination. In our case, the interactions of two negatively charged side chains (Glu72, Glu270) with the zinc ion and the zinc-bound water molecule would presumably mitigate, to a certain extent, the acidification of this molecule. Based on these considerations, it is generally assumed that the state of the zinc-bound water in the uncomplexed enzyme at physiological pH is neutral (H_2O) and it is deprotonated only at higher values of pH. The alkaline part of the pH/activity profile of the enzyme (with pK_a at about pH 9) is thus attributed to the deprotonation of this water molecule. Experimental structural support for this assignment comes from crystallographic studies of native CPA under several pH conditions, showing a significant decrease in the distance between the zinc ion and the bound water molecule with increasing alkalinity in the pH range 8.5–9.5 (Shoham *et al.*, 1984). The acidic pK_a of the activity profile of CPA around pH 6.5 has been similarly attributed to the protonation of Glu270 (Suh & Kaiser, 1976 and references therein). Others, however, have argued that in the active form of the enzyme the water molecule is deprotonated (hydroxide) and that the reduction in activity below pH 6.5 represents protonation of the hydroxide to water (Mock & Zhang, 1991; Mock & Cheng, 2000).

In the current structure, the refined distance between the zinc-bound water molecule and the zinc ion is 1.961 ± 0.003 Å. This interatomic distance is significantly shorter than the average value of 2.06 Å (Harding, 1999) found for analogous small molecules in the Cambridge Structural Database (CSD). Yet there are several examples in the CSD of water molecules (as confirmed by experimentally observed H atoms) which are bound to a single Zn atom, where the interatomic distance is equal or even smaller than 1.96 Å (for example, CSD structure LIVVUH; Bergquist & Parkin, 1999; CSD structure TURQEM; Potočňák *et al.*, 1997). On the other hand, the interaction between a single zinc cation and a hydroxide anion generally gives a distance of ~1.85 Å, as shown by Bergquist & Parkin (1999). The longest value for a zinc–hydroxide distance that we found in the CSD was 1.944 Å (CSD structure SIDMUN; Kimura *et al.*, 1990), which is still shorter than the zinc–water distance in the current structure of CPA. It should be noted also that the SIDMUN structure contains a unique feature in that the three hydroxide ions come from a three-fold-related group with an inter-oxygen distance of 2.256 Å. It is not clear, therefore, that this structure provides a true representation of zinc–hydroxide distances and significantly shorter distances are hence more representative and realistic.

It would appear, therefore, that the current crystal structure supports the proposal that in the resting state of CPA at neutral pH and at room temperature the solvent molecule bound to the active-site zinc ion is a neutral water molecule. Owing to the fine pK_a sensitivity discussed above, this situation can change as a result of small changes in the environment and/or the active site of the enzyme and especially upon binding of substrates and analogues.

Table 3
Representative interatomic distances (Å) between main groups in the active site of CPA.

Parameters from the present structure (1.25 Å) are compared with the previously reported structures of CPA (1yme and 2ctb).

		1.25 Å†	1yme	2ctb
His196 N^{δ}	Zn	2.04 (0.01)	2.08	2.10
His69 N^{δ}	Zn	2.03 (0.01)	1.94	2.06
Glu72 $O^{\varepsilon 1}$	Zn	2.27 (0.01)	2.21	2.20
Glu72 $O^{\varepsilon 2}$	Zn	2.25 (0.01)	2.29	2.36
Wat‡ O	Zn	1.96 (0.01)	1.94	1.94
Glu270 $O^{\varepsilon 1}$	Wat O	3.37 (0.02)	3.46	3.38
Glu270 $O^{\varepsilon 2}$	Wat O	2.64 (0.02)	2.66	2.66

† Experimentally estimated errors in the final distances (Å) are given within parentheses. ‡ 'Wat' indicates the zinc-bound water molecule (W1016 in the 1.25 Å structure).

3.4. The geometry around the active-site zinc

As shown previously in lower resolution structures of the enzyme, the catalytic zinc in CPA is coordinated to the $N^{\delta 1}$ atoms of His69 and His196, to the two carboxylic O atoms of Glu72 and to a water molecule (Fig. 2). This zinc coordination resembles a slightly distorted T_5 trigonal bipyramidal geometry. The more accurate zinc–ligand distances in the present structure are listed in Table 3, where they are also compared with the corresponding distances in the previously reported structures 1yme and 2ctb.

Geometric data for the metal centres of metalloproteins have recently been extracted by Harding (1999) from relevant small molecules in the Cambridge Structural Database (CSD). The structures were selected on the basis of a low crystallographic R factor ($\leq$0.065) and include molecules where the zinc-bound ligands are analogues of amino-acid side chains which are commonly found in proteins. A comparison between the coordination distances in the current 1.25 Å model of CPA and the values from the CSD (Table 4) interestingly shows that the distances between the zinc and the $N^{\delta 1}$ atom of His69 and His196 are in a better agreement with the Zn—N distance in the CSD than in the previous models of

Table 4
Representative zinc–ligand distances in proteins and small molecules compared with the relevant distances in the current structure of CPA (1.25 Å).

Average distances are given in Å; standard deviation or the range of the observed results is given in parentheses; the number of observations of the specific geometric parameter in the data set is shown in square brackets. The Zn—N(h) distance indicates that the coordinating nitrogen comes from a histidine residue. The Zn—O distances have been subdivided into two categories, depending on the type of oxygen-coordinating ligand (water or protein), labeled as Zn—O(w) and Zn—O(p), respectively. The literature results are based on surveys which have been performed on structural entries from the PDB and the CSD. In the Zn—O(p) category, the letter 'm' indicates a monodentate bond.

	Zn—N(h)	Zn—O(w)	Zn—O(p)
Native CPA (1.25 Å)	2.04 (2.034–2.045), [2]	1.96 [1]	2.26 (2.25–2.27), [2]
Proteins from PDB with resolution < 2.0 Å (Alberts *et al.*, 1998)	2.11 (2.01–2.22), [80]	2.09 (1.88–2.30), [16]	2.16 (2.00–2.32), [53]
Proteins from PDB with resolution < 1.6 Å (Harding, 2001)	2.00 (±0.08), [10]	2.09 (±0.02), [18]	2.04 (±0.09), [14(m)]
Small molecules from CSD (Harding, 1999)	2.00 (1.97–2.04), [2]	2.06 (1.97–2.28), [31]	2.00 (1.86–2.180), [∼35(m)]

Table 5
Representative zinc–ligand angles (°) of zinc-coordination sites with a T_5 trigonal bipyramidal geometry, compared with the corresponding angles in the current structure of CPA.

	Native CPA (1.25 Å)	T_5 trigonal bipyramidal zinc geometries (PDB†)
Ax—Zn—Ax (180)		
His196 $N^{\delta1}$—Zn—$O^{\varepsilon2}$ Glu72	151.30 (0.56)‡	154 (8)
Eq—Zn—Eq (120)		
His69 $N^{\delta1}$—Zn—O Wat§	118.36 (0.62)	117 (12)
Wat O—Zn—$O^{\varepsilon1}$ Glu72	114.96 (0.24)	121 (13)
Glu72 $O^{\varepsilon1}$—Zn—$N^{\delta1}$ His69	119.47 (0.75)	117 (12)
Eq—Zn—Ax (90)		
His196 $N^{\delta1}$—Zn—$O^{\varepsilon1}$ Glu72	94.81 (0.58)	102 (12)
His196 $N^{\delta1}$—Zn—O Wat	100.84 (0.44)	102 (12)
His196 $N^{\delta1}$—Zn—$N^{\delta1}$ His69	101.25 (0.82)	94 (20)
Glu72 $O^{\varepsilon2}$—Zn—$O^{\varepsilon1}$ Glu72	56.52 (0.19)	80 (19)
Glu72 $O^{\varepsilon2}$—Zn—O Wat	93.57 (0.27)	80 (19)
Glu72 $O^{\varepsilon2}$—Zn—$N^{\delta1}$ His69	93.62 (0.61)	102 (12)

† Based on Alberts *et al.* (1998). ‡ Standard deviations are given in parentheses. § 'Wat' indicates the zinc-bound water molecule (W1016 in the 1.25 Å structure).

CPA (Table 3). The distances between the zinc and the O^{ε} atoms of Glu72 are almost equal, suggesting that Glu72 is bound to the zinc in a symmetrical bidentate fashion. Zinc–carboxylate complexes in the CSD show an almost continuous range of M—O distances from a state where the two coordinating O atoms (O1 and O2) are nearly equidistant [$d(M$—O1) $\simeq$ $d(M$—O2) $\simeq$ 2.2 Å] to those non-symmetrical states where $d(M$—O1) > 3 Å and $d(M$—O2) $\simeq$ 2.0 Å, in which the interaction is practically monodentate. This flexibility in ligand geometry around the Zn atom may well be significant and relevant to its catalytic activity. Structures of CPA complexed with different ligands (Feinberg *et al.*, 1993, 1995; Kim & Lipscomb, 1991) demonstrated such flexibility in the Glu72–Zn interactions when different ligands are bound in the active site. For example, in the series of complexes between CPA and ketomethylene inhibitors (considered to be transition-state analogues; Feinberg *et al.*, 1995), it was shown that the coordination of Glu72 shifts from a clear bidentate coordination in the native enzyme to an almost monodentate coordination as a result of binding of these analogues. These results may suggest that the coordination mode of Glu72 is dependent, to some extent, on the mode of coordination of the zinc ion with various substrates, inhibitors and analogues, and that the binding of such affectors is coupled with a reduction in the coordination interactions of the zinc with Glu72.

Alberts *et al.* (1998) recently examined the high-quality crystal structures of zinc proteinases deposited in the Protein Data Bank with a resolution of better than 2 Å and a crystallographic R factor of better than 20%. On the basis of such data, they analyzed the average distances (and their standard deviations) between the catalytic zinc in these enzymes and its coordinating atoms in the T_5 trigonal bipyramidal geometry (see Table 4 for details). A similar study was reported recently by Harding (2001) based on more accurate structures of zinc proteinases deposited in the Protein Data Bank with a crystallographic resolution of better than 1.6 Å. A comparison of the zinc–ligand bond lengths reported in these two PDB studies with the current 1.25 Å structure of CPA shows a good general agreement in the geometry and distances around the catalytic zinc (Table 4), confirming the relevance of the zinc coordination in the 1.25 Å CPA structure to the wider family of zinc proteinases. A similar comparison was performed for the coordination angles around the catalytic zinc in the current structure of CPA with the corresponding values reported by Alberts and coworkers for other relevant zinc sites in the PDB (Table 5). As expected, these angles are within the range of the average values reported, with the obvious exception of the angle Glu72 $O^{\varepsilon1}$—Zn—$O^{\varepsilon2}$ Glu72 (57.73°).

3.5. Comparison of the new 1.25 Å model with previous models of CPA

In order to evaluate the new information obtained from the current 1.25 Å structure (Fig. 2), it was compared with the most recent structures of CPA available from the Protein Data Bank (1yme; Greenblatt *et al.*, 1998; 2ctb; Teplyakov *et al.*, 1993). The comparison was performed using a least-squares superposition algorithm (LSQ) from the *INSIGHT-II* software package (Accelrys Inc.). For those amino-acid side chains which were refined with multiple conformations, only the conformation with the highest statistical occupancy was considered for the comparison. A superposition using all non-H atoms gave values of 0.23 Å root-mean-square deviation (r.m.s.d.) for the 1yme structure and 0.47 Å for the 2ctb structure. Using main-chain atoms alone, an r.m.s.d. of only 0.16 Å was found for 1yme and 0.13 Å for 2ctb, indicating, as expected, that the side chains are more flexible and hence less accurately determined in these theoretically identical structures. In the active site of CPA (including Zn308, His69, His196, Glu72, Glu270, Arg127 and the zinc-bound water), the observed deviations are quite small, with r.m.s.d.s of 0.11 Å for 1yme and 0.09 Å for 2ctb.

Fig. 3 shows a plot of the r.m.s.d. *versus* the residue sequence number between the present structure and these two reference structures of native CPA. A plot of the mean temperature factors per residue (B_{eq}) for main-chain atoms *versus* sequence number (Fig. 4*a*) and a similar plot for the mean B_{eq} for side-chain atoms (Fig. 4*b*) for the 1.25 Å structure show that most of the significant differences in r.m.s.d. correspond to areas of high *B* factors in the new structure. Such differences also originate from side chains whose positions could not be traced accurately in the previous structures, but were successfully traced in the current structure, probably owing to the significant improvement in resolution (see examples below). Most of the areas of high B_{eq} factors are located on the surface of the enzyme, reflecting the conformational flexibility of these side chains. Related areas of high B_{eq} factors are located in several inner parts of the protein, where specific side-chain atoms do not participate in any interactions in a 3.5 Å radius around them. Such high B_{eq} factors are observed, for example, for the side chains of Ala1, Gln31, Asn58, Gln92, Lys153 and Lys168. These side chains exhibit relatively large conformational flexibility, which is the source of a substantial part of the r.m.s.d. differences between the structures. Other sources for r.m.s.d. differences may be attributed to the higher accuracy of the current 1.25 Å model, especially the anisotropic refinement, that allows a better determination of the exact side-chain conformations.

The high r.m.s.d. values for the backbone atoms of residues 133–135 (Figs. 3 and 4) correspond to very high B_{eq} values for their main-chain and side-chain atoms, as demonstrated in Figs. 4(*a*) and 4(*b*). These amino acids belong to a short loop (turn) which is located on the surface of the protein and as such cannot make efficient hydrogen bonds to other parts of the protein. The electron-density map of these residues was somewhat ambiguous in the previous structures of CPA (Greenblatt *et al.*, 1998; Teplyakov *et al.*, 1993), but was significantly clearer in the present structure, despite the relatively high *B* factors (B_{eq} of 60–65 Å^2) obtained in the refinement. In addition to the improved electron density in this region of the protein, we have also located several other side chains whose positions could not be determined in previous models, such as Gln92 (two alternative conformations), Ser134 (two alternative conformations), Lys190, Lys216, Lys231 and Arg276 (two alternative conformations). Obviously, these new experimentally observed side-chain conformations make the current model of CPA more accurate and more complete and as such a better starting position for various modelling studies.

3.6. Protonation states in native CPA derived from theoretical calculations

The structure of any protein is not complete without its H atoms. This is especially true for H atoms on ionizable groups involved in the reaction mechanism. As previously discussed, the increased resolution of the current structure is still insufficient to observe H atoms. Nevertheless, the higher resolution structure and the additional side chains and water molecules should provide a more accurate starting point for the modelling of H atoms, especially for ionizable polar groups in the structure. Since the theoretical predication of ionizable protons is very sensitive to the presence of nearby water molecules and other groups, greater accuracy in the description of the structure is clearly beneficial.

A reliable prediction of proton positions by any kind of theoretical simulation and even on the basis of a very accurate

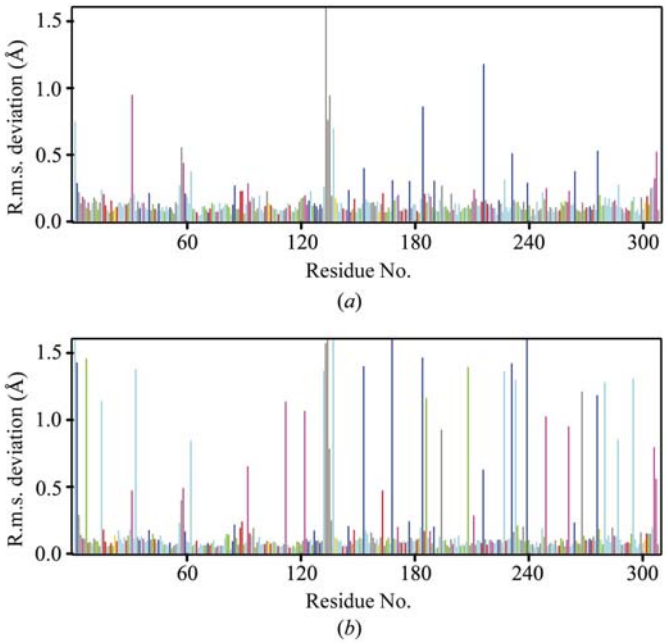

Figure 3
A plot of average r.m.s. difference (r.m.s.d.) per residue between main-chain atoms of the current 1.25 Å model of native CPA and the latest CPA models available from the PDB: (*a*) 1yme, (*b*) 2ctb. Residues are colour coded according to amino-acid type: yellow – Cys, Met; green – Phe, Tyr, Trp, His; cyan – Gly, Ala, Leu, Ile, Val, Pro; red – Glu, Asp; blue – Arg, Lys; purple – Gln, Asn; grey – Ser, Thr.

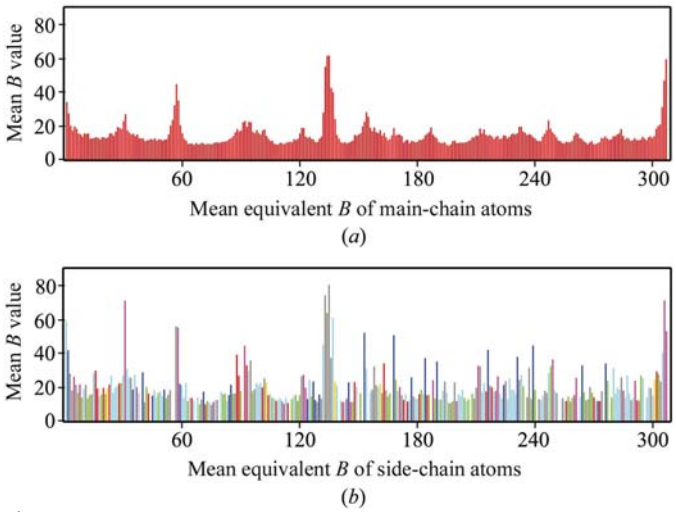

Figure 4
A plot showing the differences in mean equivalent *B* factors as a function of residue number in the final 1.25 Å structure CPA. (*a*) Mean values calculated for main-chain atoms. (*b*) Mean values calculated for side-chain atoms. Residues are colour coded according to amino-acid type: yellow – Cys, Met; green – Phe, Tyr, Trp, His; cyan – Gly, Ala, Leu, Ile, Val, Pro; red – Glu, Asp; blue – Arg, Lys; purple – Gln, Asn; grey – Ser, Thr.

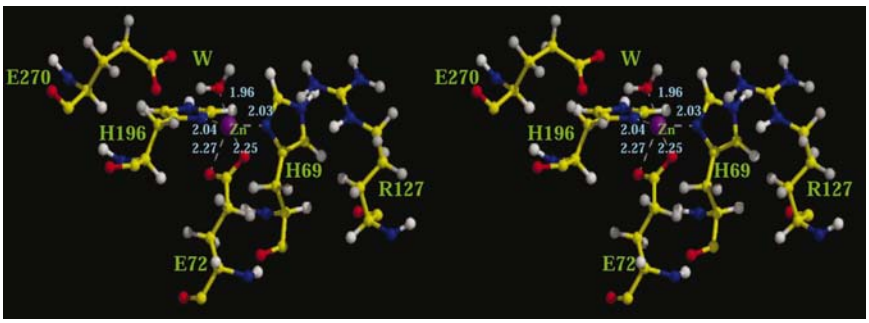

Figure 5
A stereoview of the active site of native CPA (stick model) following the proton-addition algorithm. Atoms are coloured according to conventional codes (C atoms, yellow; N atoms, blue; O atoms, red; H atoms, white; Zn atoms, purple). Zinc–ligand interactions are indicated (dashed lines) and their distances are displayed (Å). The zinc-bound water is indicated by 'W'.

Table 6
Calculated pK_a values of amino-acid side chains around the active site of CPA (based on the 1.25 Å structure).

Residue	Calculated pK_a
His69	2.04
His196	3.79
Glu72	1.82
Glu270	3.23
Arg127	11.55

protein structure is not trivial. It should be noted that all protons are positively charged to some extent, including those that are easy to position on tetrahedral C atoms in the main or side chains of proteins. H atoms are present in large numbers in the structure, comprising roughly half of the total number of atoms, and are thus a major source of electrostatic interactions. Polar protons are even more crucial and a single missing water molecule on the surface of the protein could result in an incorrect positioning of protons on side chains that interact with it. This could affect further positions of protons that are more distant from the surface of the protein. In order to minimize such complications, the H-atom positioning was performed in two steps where non-polar and polar protons were added initially, resulting in a model which was then subjected to pK_a calculations (see §2.3).

The results of the pK_a calculations for key residues in the active site of native CPA (at pH 7.5) are listed in Table 6. From this table, it is clear that the positively charged zinc ion increases the acidity of all the residues in its proximity. Although the distances between the zinc ion and its two histidine ligands are nearly identical (Fig. 5), the calculated pK_a values are quite different, with His69 being considerably more acidic than His196. This difference is presumably because of the residues which surround these two histidine side chains. His69 has three arginine residues in its vicinity (Arg71, Arg127 and Arg145), of which Arg127 is the closest. Although Asp142 makes a hydrogen bond to His69 (2.7 Å), the effect of its negative charge would be mitigated by its interaction with Arg145 (2.9 Å) and the main-chain N atom of Asn144 (2.9 Å).

One very significant result of these calculations is the predicted lowering of the pK_a of Glu270 (3.23) relative to its value in solution (4.3). This is inconsistent with the widely held assumption that Glu270 is responsible for the pK_a observed at ∼6.5 in the pH-activity profile, an assignment suggested by chemical modification experiments (Petra, 1971; Hass & Neurath, 1971*a*,*b*) and ^{35}Cl NMR studies (Stevens *et al.*, 1974). It was thought that the apparent increase in the pK_a of Glu270 might be a consequence of the hydrophobicity of its environment (Suh & Kaiser, 1976), an explanation forced by the presumption that the assignment was correct. An intuitive analysis, however, would propose that proximity of the zinc cation would cause acidification of Glu270. If this prediction of a lowered pK_a is correct, a new cause for the loss of activity below pH 6.5 must be found. While proponents of the reverse-protonation model (Mock & Zhang, 1991; Mock & Cheng, 2000) attribute this to the protonation of a zinc-bound hydroxide anion, this assignment is difficult to make for the reasons discussed above. It is quite feasible that some interactions removed from the active site but important in maintaining the overall structure of the enzyme are disrupted below pH 6.5. Thus, the loss of activity could reflect some structural incompetence on the part of the enzyme, rather than protonation of some active-site residue.

With respect to the proposed roles of Glu270 in the mechanism of the enzyme, a lower pK_a might at first glance be undesirable if the side chain were to play a general acid/base role. If one considers, however, that as attack of the zinc-bound water molecule on the target carbonyl carbon progresses the pK_a of the water proton that is closest to Glu270 would probably drop to below 0 (pK_a of alkyloxonium ions is less than −2), then some acidification of Glu270 presents no theoretical problem. As for a direct nucleophilic role for Glu270, the reduction in pK_a does not preclude this kind of potential behaviour.

Just as the zinc reduces the pK_a values of other residues that are close to it, it is likely to be involved in a reduction of the pK_a of the proximate water molecule as well. It is unclear, however, to what extent the pK_a of the zinc-bound water could be reduced. We assumed a neutral water molecule for all pK_a calculations of the relevant residues in the active site (as supported by the current high-resolution 1.25 Å structure). Had the water been ionized in the calculations, the pK_a values of the side chains reported above should change in the opposite direction, *i.e.* be increased relative to their values in Table 6.

4. Conclusions

The crystal structure of bovine CPA has been refined to a resolution of 1.25 Å using anisotropic thermal parameters for all non-H atoms. This presents a significant improvement in

the accuracy of the three-dimensional model of CPA compared with the structures of CPA available to date. The improvement in the experimental data allowed the modelling of conformational disorder (eight new disordered side chains), enhanced the description of the solvent network (375 water molecules compared with 192/218 in previous models) and provided a more accurate description of the interactions between the zinc and its ligands. The newly obtained metal–ligand distances are consistent with the values extracted from small-molecule crystal structures.

The calculation of standard uncertainties in individual atom positions allowed deduction of the protonation state of some key residues in the active site and confirmed that Glu72 and Glu270 are negatively charged in the resting state of the enzyme at pH 7.5. These results were further validated by theoretical calculations, which showed significant reduction of the pK_a values of these side chains relative to experimental solution values. In fact, all residues in the vicinity of the zinc ion showed an increased acidity, presumably owing to the local positive charge of the metal cation. The distance between the zinc-bound solvent molecule and the metal ion is strongly suggestive of a neutral water molecule and not a hydroxide ion in the resting state of the enzyme. Furthermore, the presence of a nearby carboxylate group (Glu270) would mitigate, to a certain extent, the depression of the pK_a value of this water molecule by the zinc ion. These findings could support both the GAGB mechanism and the anhydride mechanism suggested for CPA, but such conclusions should be verified by further investigations of the reaction mechanism (Kilshtain-Vardi *et al.*, 2002).

It is assumed that this highly accurate set of coordinates for the three-dimensional structure of CPA may serve now as a starting point for studying the detailed catalytic mechanism of this enzyme by a combination of further experimental and theoretical methods. Such studies, when performed, would hopefully allow us to distinguish between alternative reaction pathways for CPA and for related zinc metalloproteinases.

We thank the staff at the National Synchrotron Light Source (NSLS, X12B and X26C beamlines) of the Brookhaven National Laboratory for their helpful support in the X-ray synchrotron data measurement and analysis. This work was supported, in part, by grants from the Israeli Ministry of Science (MOS) and the Israel Academy of Sciences and Humanities - Israel Science Foundation (ISF).

References

Alberts, I. L., Nadassy, K. & Wodak, S. W. (1998). *Protein Sci.* **7**, 1700–1716.
Auld, D. S., Galdes, A., Geoghegan, K. F., Holmquist, B., Martinelli, R. A. & Vallee, B. L. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 5041–5045.
Bergquist, C. & Parkin, G. (1999). *J. Am. Chem. Soc.* **121**, 6322–6323.
Bottomley, K. M., Johnson, W. H. & Walter, D. S. (1998). *J. Enzyme Inhibit.* **13**, 79–101.
Breslow, R. & Wernick, D. L. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 1303–1307.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, A. T. & Karplus, M. (1988). *Proteins*, **4**, 148–156.
Coleman, J. E. & Vallee, B. L. (1961). *J. Biol. Chem.* **236**, 2244–2249.
Christianson, D. W. & Lipscomb, W. N. (1989). *Acc. Chem. Res.* **22**, 62–69.
Dyson, H. J., Jeng, M. F., Tennant, L. L., Slaby, I., Lindell, M., Kurpin, S. & Holmgren, A. (1997). *Biochemistry*, **36**, 2622–2636.
Feinberg, H., Greenblatt, H. M., Behar, V., Gilon, S., Cohen, S., Bino, A. & Shoham, G. (1995). *Acta Cryst.* D**51**, 428–449.
Feinberg, H., Greenblatt, H. M. & Shoham, G. (1993). *J. Chem. Inf. Comput. Sci.* **33**, 501–516.
Gladysheva, T., Liu, J. & Rosen, B. P. (1996). *J. Biol. Chem.* **271**, 33256–33260.
Glick, M. & Goldblum, A. (2000). *Proteins*, **38**, 273–287.
Glick, M., Rayan, A. & Goldblum, A. (2003). In the press.
Greenblatt, H. M., Feinberg, H., Tucker, P. A. & Shoham, G. (1998). *Acta Cryst.* D**54**, 289–305.
Harding, M. M. (1999). *Acta Cryst.* D**55**, 1432–1443.
Harding, M. M. (2001). *Acta Cryst.* D**57**, 401–411.
Hass, G. M. & Neurath, H. (1971*a*). *Biochemistry*, **10**, 3535–3540.
Hass, G. M. & Neurath, H. (1971*b*). *Biochemistry*, **10**, 3541–3546.
Havranek, J. J. & Harbury, P. B. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 11145–11150.
Kaslik, G., Westler, W. M., Graf, L. & Markley, J. L. (1999). *Arch. Biochem. Biophys.* **362**, 254–264.
Kilshtain-Vardi, A., Shoham, G. & Goldblum, A. (2002). *Int. J. Quant. Chem.* **88**, 87–98.
Kim, H. & Lipscomb, W. N. (1991). *Biochemistry*, **30**, 8171–8180.
Kimura, E., Shiota, T., Koike, T., Shiro, M. & Kodama, M. (1990). *J. Am. Chem. Soc.* **112**, 5805–5811.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–290.
Lee, M. J. & Kim, D. H. (2000). *Bioorg. Med. Chem.* **8**, 815–823.
Lide, D. R. (1991). *Handbook of Chemistry and Physics*, 72nd ed., pp. 9-2–9-15. Boca Raton/Ann Arbor/Boston: CRC Press.
Lipscomb, W. N. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 3875–3878.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Makinen, M. W., Yamamura, K. & Kaiser, E. T. (1976). *Proc. Natl Acad. Sci. USA*, **73**, 3882–3886.
Malfroy, B., Stewarts, J. P., Guyon, A., Roques, B. P. & Schwartz, J. C. (1978). *Nature (London)*, **276**, 523–526.
Mehler, E. L. & Guarnieri, F. (1999). *Biophys J.* **77**, 3–22.
Merritt, E. A., Kuhn, P., Sarfaty, S., Erbe, J. L., Holmes, R. K. & Hol, W. G. J. (1998). *J. Mol. Biol.* **282**, 1043–1059.
Mock, L. W. & Cheng, H. (2000). *Biochemistry*, **39**, 13945–13952.
Mock, L. W. & Zhang, J. Z. (1991). *J. Biol. Chem.* **266**, 6369–6400.
Neurath, H., Bradshaw, R. A., Petra, P. H. & Walsh, K. A. (1970). *Philos. Trans. R. Soc. London, Ser. B*, **257**, 159–176.
Ohishi, K., Fujita, N., Morinaga, Y. & Tsuruo, T. (1995). *Clin. Exp. Metastasis*, **13**, 287–295.
Ondetti, M. A. & Cushman, D. W. (1982). *Annu. Rev. Biochem.* **51**, 283–308.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–325.
Petra, P. H. (1971). *Biochemistry*, **10**, 3163–3170.
Potočňák, I., Heinemann, F. W., Rausch, M. & Steinborn, D. (1997). *Acta Cryst.* C**53**, 54–56.
Rees, D. C., Lewis, M. & Lipscomb, W. N. (1983). *J. Mol. Biol.* **168**, 367–387.
Sack, J. S. (1988). *J. Mol. Graph.* **6**, 224–225.
Sauter, C., Otalora, F., Gavira, J. A.,Vidal, O., Giege, R. & Garcia-Ruiz, J. M. (2001). *Acta Cryst.* D**57**, 1119–1126.
Schwartz, J. C., Gros, C., Lecomte, J. M. & Bralet, J. (1990). *Life Sci.* **47**, 1279–1297.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Shoham, G., Christianson, D. W. & Oren, D. A. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 684–688.
Shoham, G., Rees, D. C. & Lipscomb, W. N. (1984). *Proc. Natl Acad.*

Sci. USA, **81**, 7767–7771.
Soffer, R. L. (1976). *Annu. Rev. Biochem.* **45**, 73–94.
Stevens, R. S., Jentoft, J. E. & Bryant, R. G. (1974). *J. Am. Chem. Soc.* **96**, 8041–8045.
Suh, J. & Kaiser, E. T. (1976). *J. Am. Chem. Soc.* **98**, 1940–1947.
Summers, J. B. & Davidsen, S. (1998). *Annu. Rep. Med. Chem.* **33**, 131–143.
Teixeira, S., Leggio, L. L., Pickersgill, R. & Cardin, C. (2001). *Acta Cryst.* D**57**, 385–392.
Teplyakov, A., Wilson, K. S., Orioli, P. & Mangani, S. (1993). *Acta Cryst.* D**49**, 534–540.
Tezuka, K., Nemato, K., Tezuka, Y., Sato, T., Ikeda, Y., Kobori, M., Kawashima, H., Eguchi, H., Hakeda, Y. & Kumegawa, M. (1994). *J. Biol. Chem.* **269**, 15006–15009.
Tronrud, D. E. (1992). *Acta Cryst.* A**48**, 912–916.
Vallee, B. L. & Galdes, A. (1984). *Adv. Enzymol.* **56**, 283–430.
Woessner, J. F. (1994). *Ann. NY Acad. Sci.* **732**, 11–21.
Wooley, P. (1975). *Nature (London)*, **258**, 677–682.