RESEARCH PAPERS

Carboxypeptidase A: Native, Zinc-Removed and Mercury-Replaced Forms

H. M. Greenblatt,[a] H. Feinberg,[a] P. A. Tucker[b] and G. Shoham[a]*

[a]*Department of Inorganic and Analytical Chemistry, Hebrew University, Jerusalem, Israel, 91904, and* [b]*Structural Biology Programme, European Molecular Biology Laboratory, Postfach 10.2209, Meyerhofstrasse 1, D69012 Heidelberg, Germany. E-mail: gil@vms.huji.ac.il*

Abstract

The crystal structure of the zinc-containing exopeptidase bovine carboxypeptidase A (CPA) has been refined to high resolution, based on a data set collected from a single crystal, incorporating new sequence information based on cloning of the bovine gene. In addition, new refined structures are available for the zinc-removed form of the enzyme, apo-CPA, as well as the mercury-replaced form, Hg-CPA. The native structure reveals that the zinc-bound water molecule does not appear to be more loosely bound than the rest of the zinc ligands, at least when B-factor values are considered. Nor is there any evidence for a secondary location of this water molecule. The apo-enzyme structure does not show any density in the place of the removed zinc ion. The only significant change appears to be a χ_2 rotation of one zinc histidine ligand to form an ion-pair interaction with a glutamic acid side chain. The structure of Hg-CPA reveals a solvent Tris molecule bound to the mercury cation, as well as an unidentified cation bound to Glu270. The location of this cation agrees with previous proposals for the binding site of inhibitory zinc. These observations may explain some of the differences in kinetics observed in metal-replaced 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 participant (Vallee & Galdes, 1984). One of the most widely studied zinc metalloenzymes is carboxypeptidase A (CPA), which catalyzes both exopeptidase and exoesterase reactions. These two types of reaction appear to proceed *via* the same catalytic pathway, although the initial binding modes are believed to differ (Christianson & Lipscomb, 1989).

The zinc ion of many zinc metalloenzymes, including CPA, can be removed and replaced by other divalent transition metals, such as nickel, manganese, cobalt, mercury, *etc.*, while still maintaining activity. Despite the fact that kinetic studies on metal-replaced carboxypeptidase A (CPA) were first carried out more than 30

years ago (Coleman & Vallee, 1961), no satisfactory explanation has been found for the observed changes in the rates of activity of metal-substituted CPA. Potential reasons for the changes in activities could be changes in metal-binding geometry, changes in substrate-binding geometry, electronic effects, or a combination of the above. The zinc-removed enzyme (apo-CPA) is also of interest, since it may provide a means of studying early binding of actual substrates, and so give clues as to the nature of the pre-catalytic complex.

Although the three-dimensional structures of apo-CPA (Rees & Lipscomb, 1983), and some metal derivatives were studied by X-ray crystallography (Rees *et al.*, 1986), the structures were at medium resolution, and were unrefined. Two derivatives (Ni^{II}-CPA and Co^{II}-CPA) were refined at high resolution (Hardman & Lipscomb, 1984), but no inhibitor binding studies were performed. In order to unambiguously establish the structure of apo-CPA and its various metallo derivatives, and in order to use these forms of CPA in substrate and inhibitor binding studies, we used methods designed to remove all traces of zinc from crystals of native CPA, and then collected high-resolution data sets which were subjected to rigorous refinement. In addition, we required an improved model for native CPA with better geometry. In this paper we present the results from the first stage of this work, namely the improved structure of native CPA and the refined structures of apo-CPA, and the Hg-CPA metal derivative. In addition to these results we also had the opportunity to compare data collected on different data-collection devices. Two devices used europium-doped barium halide imaging plates [R-AXIS IIC (Sato *et al.*, 1992), DIP100 (Tanaka *et al.*, 1990)]. The third was a TV-type detector, the FAST area detector (Arndt & Thomas, 1982; Arndt & in't Veld, 1988). The results of these limited comparisons are also presented.

2. Experimental

CPA was purchased from Sigma Co.; the protein was further purified and crystallized according the procedure described in Shoham *et al.* (1984).

The crystals were cross linked to impart greater mechanical strength, allowing the numerous manipulations subsequently performed on the crystals. The buffer was changed to 0.02 M veronal/0.1 M LiCl, and the cross linking carried out with 0.025% glutaraldehyde for 12 h. The buffer was then changed back to 0.02 M Tris/0.1 M LiCl, pH 7.5 for the structural and soaking studies.

Removal of zinc was effected by dialysis against hydroquinonesulfonic acid (HQSA; 1–5 mM). The dialysis solution was changed daily, with the concen-

tration of HQSA increased over a period of 3 weeks, until no zinc could be detected by atomic absorption measurements on solutions of dissolved CPA crystals. The presence of protein in these solutions was confirmed by absorption at 278 nm ($\varepsilon = 6.42 \times 10^4$).

Crystals of apo-CPA were soaked, over a period of 2 weeks, in increasing solutions (10^{-5}–10^{-3} M) of the inhibitor BOC-Phe-ΨPhe (BocFF). The crystals were allowed to soak at the highest concentration of inhibitor for at least 2 weeks before data collection. Intensity data were measured on a MAC Science DIP100 auto-
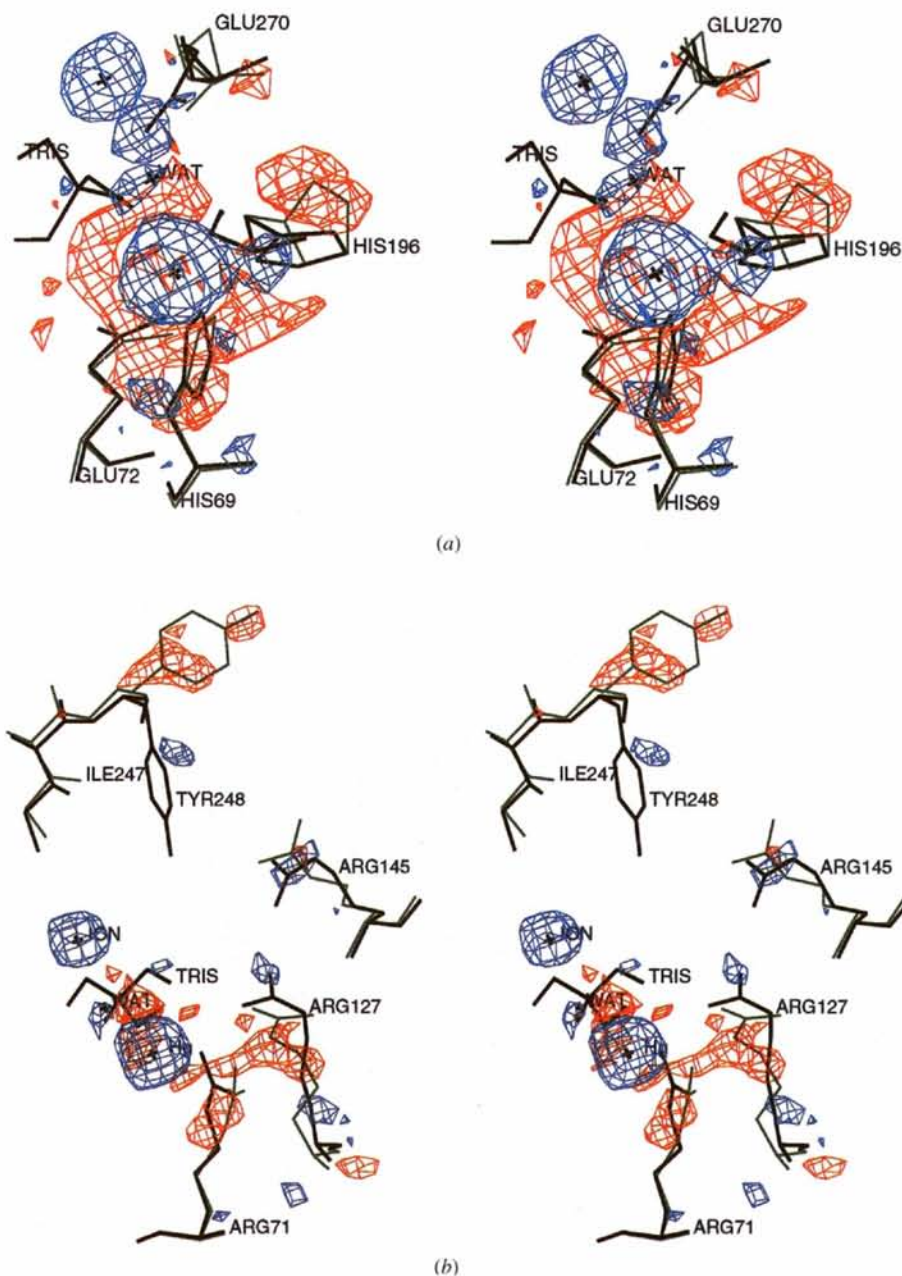

Fig. 1. (a) Difference map generated using model for apoCPA against data for Hg-CPA. Electron-density contour levels are at -3.5σ (red) and $+3\sigma$ (cyan). Superimposed on the density is the final model of Hg-CPA (black lines) and apo-CPA (green lines). Note that there is density for the N atom of the Tris molecule. The peak height for the ion bound to Glu270 is 13σ. (b) The same difference map as shown in (a) showing the three arginine residues in the active site of CPA as well as Tyr248.

matic data-collection system with an imaging plate (Tanaka *et al.*, 1990) on loan to EMBL, and mounted on a MAC Science MX18 rotating-anode generator operating at 45 kV and 100 mA. The Cu $K\alpha$ doublet was selected using a pyrolytic graphite monochromator. Data-collection details are given in Table 1(*a*). The crystal orientation and unit-cell parameters (see Table 1*a*) were determined and refined using the *ATIX* program (Tanaka *et al.*, 1990). Data were recorded and reflection shoeboxes were extracted with the *ADAC* program (Tanaka *et al.*, 1990). Integrated intensities were obtained using a profile-fitting routine (Tucker, 1990) and symmetry-equivalent reflections were scaled and averaged using the programs of the *BIOMOL* package from the University of Gröningen. Crystals of

apo-CPA were also soaked, over a period of two weeks in increasing solutions $(10^{-5}$–$10^{-3} M)$ of the CPA inhibitor benzyl-succinate. Again the crystals were allowed to soak at the highest concentration of the inhibitor for at least 2 weeks. Intensity data were measured on a FAST area detector system (Arndt & Thomas, 1982; Arndt & in't Veld, 1988) mounted on a GX21 rotating-anode generator operating at 44 kV and 90 mA with a 300 μm focus. The Cu $K\alpha$ doublet was selected using a pyrolytic graphite monochromator. Data-collection details are summarized in Table 1(*b*). Cell dimensions and crystal orientation were determined and refined with the *MADNES* program (Messerschmidt & Pflugrath, 1987), as summarized in Table 1(*b*). Intensities were measured and processed
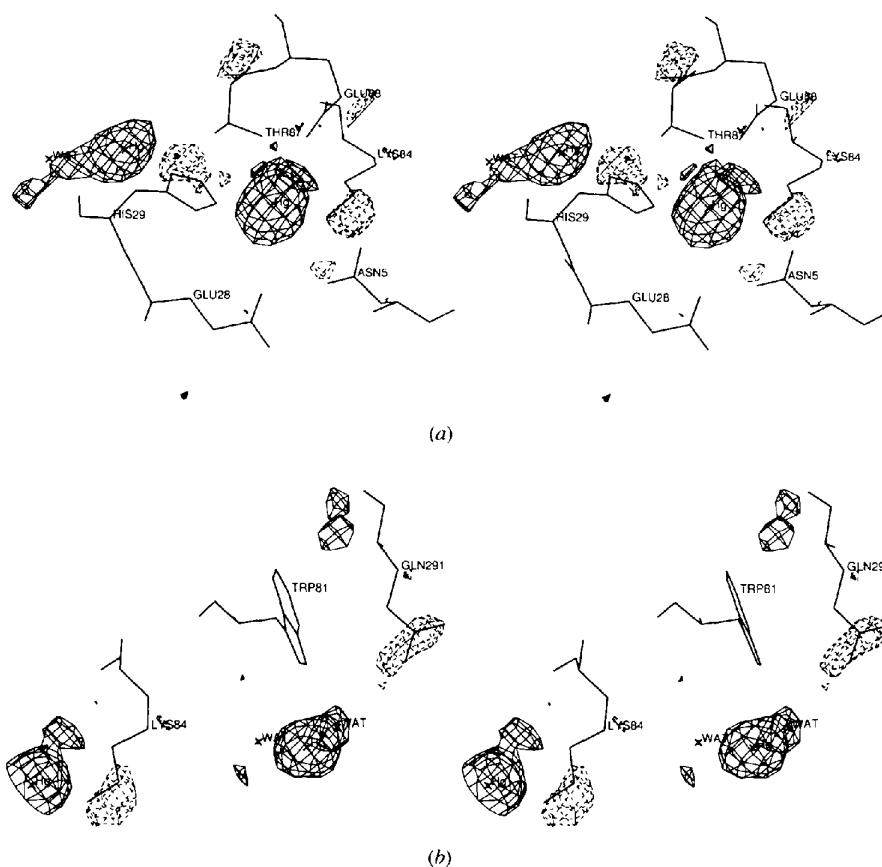

(*a*)

(*b*)

Fig. 2. (*a*) Stereoview of two Hg^{2+} binding sites on the surface of Hg-CPA, on either side of His29. Difference map $(F_o - F_c)$ was generated using starting model for Hg-CPA. Contours are at $\pm 3\sigma$, with positive contours drawn with solid lines and negative contours with dashed lines. Negative contours have a 1.0 Å cover radius, which shows only those contours within 1.0 Å of any atom. Positive contours have a 2.0 Å cover radius. The peak height for the cation between Lys84 and His29 is greater than 14σ, while the other peak height is only somewhat above 13σ. The distance between both cations and their respective nitrogen ligands on His29 is 2.2 Å. This is also the distance between Lys84 N^ζ and the cation, although the negative density around Lys84 N^ζ suggests a conformational change upon binding of Hg^{2+}. Note that there is negative density around His29 $N^{\delta 1}$ which suggests some movement of the side chain, although it is impossible to tell which binding site gives rise to this movement. (*b*) Same difference map as in (*a*) showing Hg^{2+} binding site in vicinity of Trp81. The peak height is between 13 and 14σ. Note that the cation does not appear to interact with any protein atoms directly, although there is some indication that the side chain of Gln291 does change its conformation, and may even occupy the lobe of density currently occupied by a water molecule. Note the close proximity of this binding site to the other two shown in (*a*). The distance between the two Hg^{2+} cations shown is 11 Å. The cation is 4.4 Å from the $N^{\varepsilon 1}$ of Trp81.

with the *MADNES* program and symmetry-equivalent reflections were scaled and averaged using programs of the *BIOMOL* package. In a third experiment, crystals of apo-CPA were soaked in a similar manner in increasing solutions (10^{-5}–10^{-3} M) of the substrate dansyl-Gly-L-Trp (DGT). Intensity data were measured on an R-AXIS IIC system mounted on an RU-300 fitted with a mirror/mirror optical system running at 50 kV and 100 mA; details of data collection and cell constants are summarized in Table 1(*c*). Determination and refinement of cell dimensions and crystal orientation, intensity measurement and data reduction were all carried out using the *Rigaku* software package (Sato *et al.*, 1992).

To prepare Hg-CPA, crystals of apo-CPA were soaked for at least two weeks in a 0.1 M LiCl/0.02 M Tris solution containing 10^{-3} M HgCl$_2$. These crystals were then soaked in increasing concentrations of benzyl-succinate, in the same manner described for the other inhibitor soaks above. Intensity data were measured on a MAC Science DIP100 system, as described above, and as summarized in Table 1(*a*).

Crystals of native CPA were cross linked and used for data collection on an R-AXIS IIC mounted on an RU-200 rotating-anode generator, running at 40 kV and 100 mA, fitted with graphite monochromator, and a 0.3 mm double collimator. Data-collection parameters and cell constants are summarized in Table 1(*c*). Cell-dimension and crystal-orientation determination,

refinement of these parameters, intensity measurements and processing were all carried out using the *Rigaku* software package (Sato *et al.*, 1992). In a parallel experiment, intensity data from a cross-linked crystal of native CPA were collected on an R-AXIS IIC device mounted on a RU200 rotating-anode generator, running at 40 kV and 100 mA equipped with a mirror/ mirror focusing system. Data-collection parameters and cell constants are summarized in Table 1(*c*). Determination and refinement of cell dimensions and crystal orientation, intensity measurement and data reduction were all carried out using the *Rigaku* software package (Sato *et al.*, 1992).

The restrained-parameter least-squares refinement program of Hendrickson & Konnert (1980), modified by one of the authors (HMG) to run on a four-processor Silicon Graphics 240GTX, was used for all the structures presented in this work. Initial restraints were made loose, in order to allow the model to adjust to the data and remove any bias from the starting model. The weighting scheme used for the structure factors changed during the course of *PROLSQ* refinement. Initial weights for F_o were constant with resolution, thus $\omega = 1/(\sigma_F)^2$, where $\sigma_F \simeq \langle |F_o| - |F_c| \rangle/2$. Once refinement had progressed enough that the value of $\langle |F_o| - |F_c| \rangle/2$ varied in a consistent fashion with resolution, the value of σ_F was fitted to the function $\sigma_F = C1 + C2[(\sin\theta)/\lambda - 1/6]$ where $C1$ and $C2$ are constants. The slope of the line, given by $C2$, was
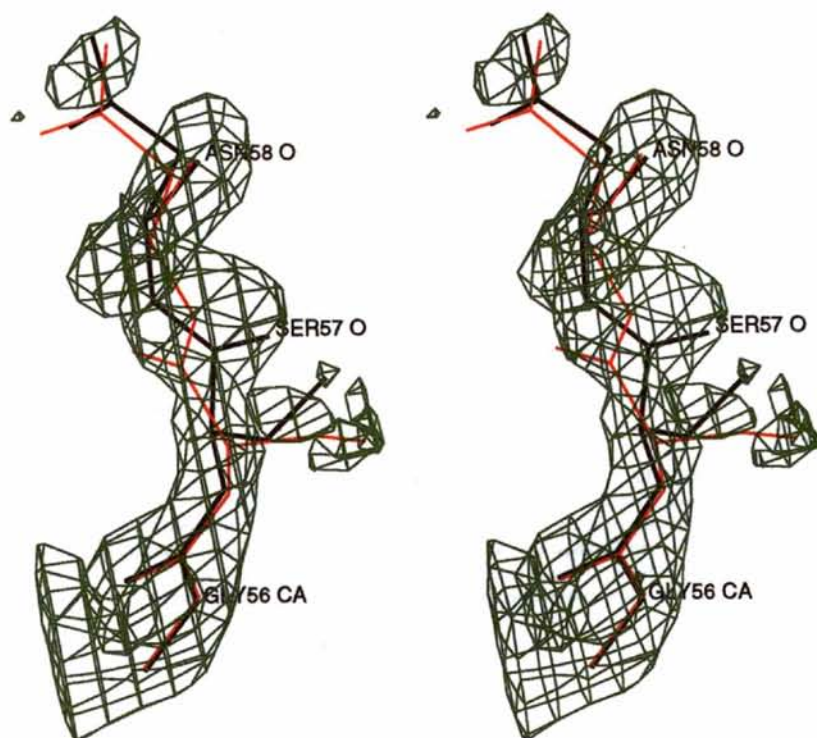

Fig. 3. Stereoview of the region surrounding Ser57 in native CPA. Shown in black is the structure of CPA after retracing and refinement with TNT. In red is the structure of CPA after refinement with *PROLSQ*. Electron-density map (green) was calculated after *PROLSQ* refinement, using the coefficients $2|F_o| - |F_c|$ (1σ contours). Density around the region of Thr133 was much poorer (not shown), and the new conformation was only a marginally better fit.

Table 1. *Data-collection details*

(a) DIP100

	Apo-CPA + BOC-Phe- Ψ-Phe	Hg-CPA + benzyl succinate
Crystal-to-plate distance (mm)	85	75
Frames measured	1–88 89–102	1–137
Frame width (°)	1.5 1.5	1.5
Frame measurement times (s)	1125 1250	1125
Number of observations	84305	107065
Number of unique reflections	21540	26225
Rejected observations	995	359
Overall R_{sym} (cutoff)/ completeness (%)	6.54(0σ)/89	9.08(0σ)/82
Maximum resolution (Å)	1.86	1.74
Cell dimensions (Å, °)		
a	51.70	51.70
b	60.32	60.32
c	47.20	47.20
β	97.39	97.39

(b) FAST

	Apo-CPA + benzyl succinate
Frame width (°)	0.2
Frame measurement time (s)	60
Detector swing angle (°)	25
Crystal-to-detector distance (mm)	55
κ-axis goniostat datum position	
ω	180.0 −122.97
κ	0.00 −134.74
φ	0.00 57.03
Scan axis	ω ω
Scan range (°)	−70 to 110 −25 to 25
Number of observations	46617
Rejected observations in integration	475
Number of unique reflections	23464
Rejected observations in scaling	117
R_{sym} (cutoff)/completeness (%)	5.93(0σ)/90
Maximum resolution (Å)	1.85
Cell dimensions (Å, °)	
a	51.70
b	60.32
c	47.20
β	97.39

(c) R-AXIS IIC

	Native CPA monochr.	Native CPA mirrors	Apo-CPA + DGT
Crystal-to-plate distance (mm)	55	64.8	70.4
Frames measured	45	60	45
Frame width (°)	2	2	2
Frame measurement times (s)	2700	2700	2700
Number of observations	58534	74056	48816
Number of unique reflections	34352	30920	26987
Rejected observations	3	48	0
Overall R_{sym} (cutoff)/ completeness (%)	3.6(1σ)/79	4.5(1σ)/87	2.8(1σ)/85
Maximum resolution (Å)	1.53	1.65	1.70
Cell dimensions (Å, °)			
a	51.62	51.74	51.69
b	60.17	60.30	60.23
c	47.17	47.27	47.29
β	97.40	97.47	97.34

Table 2. *Summary of amino-acid changes as a result of gene sequencing of bovine carboxypeptidase*

Original sequence	Gene sequence (atom change)
Gln28	Glu28 (N$^{\epsilon2}$ to O$^{\epsilon2}$)
Glu31	Gln 31 (O$^{\epsilon2}$ to N$^{\epsilon2}$)
Asn89	Asp 89 (N$^{\delta2}$ to O$^{\delta2}$)
Asn93	Asp 93 (N$^{\delta2}$ to O$^{\delta2}$)
Asn114	Asp114 (N$^{\delta2}$ to O$^{\delta2}$)
Glu122	Gln122 (O$^{\epsilon2}$ to N$^{\epsilon2}$)
Asn185	Asp185 (N$^{\delta2}$ to O$^{\delta2}$)

negative, giving more weight to the higher resolution data. All solvent molecules were initially refined with fixed occupancies, while variable occupancies were used in the later cycles of the refinement. After an initial round of refinement for all models, restraints were given tight values.

The refinement program *TNT* (Tronrud *et al.*, 1987; Tronrud, 1992) was used on three of the models after completion of *PROLSQ* refinement.

3. Structure solution and refinement

The crystals were isomorphous with native CPA and so initial phases were obtained from the model of the native structure (Rees *et al.*, 1983) available from the Brookhaven Protein Data Bank (entry code: 5CPA). For the apo-CPA structure, the side chains of the active-site residues involved in zinc coordination and substrate binding were removed in the starting model (His69, Arg71, Glu72, Arg127, Arg145, His196, Tyr248, Glu270). All solvent molecules were also removed. The residues comprising, and surrounding the three cis-peptide bonds (196–199, 204–207, 271–274) were also excluded from the starting model. Also incorporated were several corrections to the original sequence (Bradshaw *et al.*, 1969) as revealed by the bovine gene sequence (Shoham *et al.*, 1990; Le Huërou *et al.*, 1991), as summarized in Table 2. The significance of these changes with respect to the hydrogen-bond structure in the regions surrounding them is discussed below.

A starting difference map $(F_o − F_c)$ was generated from this model, using the DIP100 data, and all the missing residues were built into the difference density, using both 2σ and 3σ contours, where σ was the r.m.s. value of the map. No difference density could be attributed to a bound molecule of BocFF. Refinement was initialized once all the residues that had been removed were placed into density. Solvent molecules were not added until after seven cycles of restrained least-squares refinement (see below). Peak searches down to the 3σ level of a $F_o − F_c$ map were assigned water molecules if they were within 4.0 Å of a protein atom capable of hydrogen bonding (N, O atoms). The final round of refinement (cycles 18–32) used no σ cutoff for the data. After remodeling of two mobile

areas of the main chain (see below, native CPA) a further 15 cycles of *TNT* refinement were run. See Tables 3(*a*) and 3(*c*) for a summary of the refinement results. Coordinates for the final refined model of apo-CPA have been deposited with the Protein Data Bank.†

Refinement of the apo-CPA data from the FAST area detector began with the final model of apo-CPA obtained from the DIP100 imaging plate. Eight cycles of refinement were run with loose restraints in order to allow the model freedom to absorb errors and minimize model bias. Restraints were then tightened and an additional seven cycles of refinement were applied. The cutoff used for the data was 1σ on F for the first eight cycles, and 0σ for the last round, as was carried out for the apo-CPA data from the DIP100 (see Table 3*a*). No density was observed in the active site that was attributable to a bound molecule of benzyl-succinate.

The starting model for the Hg-CPA–benzyl-succinate crystals was taken from the apo-CPA structure after 31 cycles of refinement with 183 water molecules. A simple $F_o - F_c$ difference map using the apo-enzyme as the model showed the changes which occurred in the structure upon addition of mercury cation (Figs. 1*a*, 1*b*). Thus, all the active-site side chains removed to build the initial model of apo-CPA were also removed for Hg-CPA, as well as the active-site solvent molecules. The *cis*-peptide bond residues and the rest of the solvent molecules were retained from the apo-CPA model. An $|F_o| - |F_c|$ omit map generated using this partial model was used to actually fit the missing side chains. No bound inhibitor could be located in the active site, although Tyr248 displayed density in the 'down' position, and so was modeled as such (see §4). An extra cation, however, was observed bound to Glu270 and was tentatively identified as a calcium or similar cation, and was refined as a Ca^{2+} ion at full occupancy. Three mercury cations were located bound to surface residues of Hg-CPA, removed from the active site. These were refined at 40% occupancy. Two of these apparent binding sites are on either side of His29, although one site appears to be favoured (see Fig. 2*a*). The other site is close to Trp81, but does not have any direct contact with a protein side chain (Fig. 2*b*).

After seven cycles of refinement, a difference map clearly showed the presence of a Tris buffer molecule [$NH_2C(CH_2OH)_3$] bound to the mercury *via* the amine group. The presence of a Tris buffer molecule has been previously reported (Rees *et al.*, 1986). It is interesting to note that the interaction of the amine group of Tris with the mercury cation is similar to the interaction between the amine group of the bound dipeptide Gly-Tyr (Christianson & Lipscomb, 1986) and the zinc ion of

native CPA. The Tris molecule was refined at 30% occupancy, based on *B*-factor considerations. Since the presence of both the extra cation and the Tris molecule are probably mutually exclusive for steric reasons, the identification of the cation as Ca^{2+} was likely to be incorrect. In order to maintain the same B factor at only 70% occupancy, the cation was arbitrarily labeled a copper ion (see *Discussion*).

After several cycles of refinement, a number of side chains of leucine, isoleucine, valine and threonine showed great deviation from ideal chiral volumes. In these side chains, this is very often an indication that the side chain is in the 'opposite' conformation from what it should be. Thus, for example, a threonine may have the $O^{\gamma}1$ and $C^{\gamma2}$ switched around, and requires rotation about χ_1. This can often be confirmed by the presence of negative electron density around the current position of the terminal atoms of the side chain ($O^{\gamma1}$ and $C^{\gamma2}$ in the example of threonine) and positive density in their new location upon rotation. In the vast majority of cases, the new conformation significantly improved the refined chiral volume, and removed the negative and positive peaks of electron density in the difference maps. In certain cases the manual adjustment did not help, a likely indication that the side chain is distributed between both conformers in a roughly equal fashion.

Waters were assigned using peak searches as described above. At the end of refinement, all solvent molecules underwent extensive checking as described in Greenblatt *et al.* (1989) and similar to the method used by Fujinaga *et al.* (1985). Briefly, 50 solvents were removed at one time, the resulting model subjected to several cycles of refinement, and a difference map was calculated. Solvent molecules that had been identified that did not return with any 3σ density on the difference map were discarded. Borderline cases, where there was a 3σ peak at one or two grid points, were accepted only if the 2σ density was well defined and rounded. When this process was complete, the model was refined until the total r.m.s. positional shifts were close to 0.10 Å or less, indicating that the refinement had converged. Any solvent molecules that had B factors higher than 50.0 Å^2 at the end of all the refinement were also discarded.

As in apo-CPA (DIP100), two main-chain regions were remodeled after completion of the *PROLSQ* refinement. In addition two other changes were made. Electron-density maps calculated with the coefficients $2|F_o| - |F_c|$, showed that the side chain of Arg127 required adjustment. Furthermore, refinement with *TNT* had reduced the B factor of the N atom in the Tris molecule to 1.0, the minimum allowed by the program. To correct this, a water molecule was placed in the same location at 70% occupancy, and the refinement repeated. The final refinement parameters for Hg-CPA are summarized in Tables 3(*b*) and 3(*c*). Coordinates for the

Table 3. *Final refinement parameters*

(*a*) Apo-CPA

Data set	Apo-CPA	Apo-CPA	Apo-CPA
Data-collection device	DIP100	FAST	R-AXIS IIC (mirror/mirror)
Type of refinement	Restricted LS PROLSQ/HK	Restricted LS PROLSQ/HK	Restricted LS PROLSQ/HK
No. of cycles	32	15	19
No. of protein atoms (excluding metal)	2437	2437	2437
No. of water molecules	184	180	173
R.m.s. deviation (σ, target)			
Bond lengths	0.013 (0.015)	0.016 (0.015)	0.016 (0.015)
Bond angles 1–3	0.037 (0.030)	0.039 (0.030)	0.039 (0.030)
Dihedral angles 1–4	0.034 (0.030)	0.038 (0.030)	0.037 (0.030)
Plane restraints	0.012 (0.015)	0.013 (0.015)	0.013 (0.015)
Chiral volumes	0.235 (0.150)	0.274 (0.150)	0.277 (0.150)
Average ω (°)	2.3 (2.5)	2.5 (2.5)	2.4 (2.5)
Final r.m.s. coordinate shifts	0.016	0.013	0.012
Final R factor	0.158	0.173	0.157
Resolution (Å) (cutoff on F)	8.0–1.90 (0σ)	8.0–1.8 (0σ)	8.0–1.80 (0σ)
No. of reflections	20084	21798	22954

(*b*) Native CPA and Hg-CPA

Data set	Native-CPA	Native-CPA	Native-CPA	Hg-CPA
Data-collection device	Nicolet P21	R-AXIS IIC (monochr.)	R-AXIS IIC (mirrors)	DIP100
Type of refinement	Restricted LS PROLSQ/RL	Restricted LS PROLSQ/HK	Restricted LS PROLSQ/HK	Restricted LS PROLSQ/HK
No. of cycles	51	36	17	59
No. of protein atoms (excluding metal)	2437	2437	2437	2437
No. of water molecules	192	213	213	214
R.m.s. deviation (σ, target)				
Bond lengths	0.024 (0.017)	0.017 (0.015)	0.016 (0.015)	0.014 (0.015)
Bond angles 1–3	0.041 (0.025)	0.036 (0.030)	0.037 (0.030)	0.035 (0.030)
Dihedral angles 1–4	0.047 (0.033)	0.043 (0.030)	0.039 (0.030)	0.036 (0.030)
Plane restraints	0.038 (0.055)	0.015 (0.015)	0.014 (0.015)	0.014 (0.015)
Chiral volumes	0.263 (0.150)	0.202 (0.150)	0.188 (0.150)	0.138 (0.150)
Average ω (°)	8.5 (6.4)	2.8 (2.5)	2.8 (2.5)	2.6 (3.0)
Final r.m.s. coordinate shifts		0.013	0.012	0.009
Final R factor	0.190	0.148	0.144	0.148
Resolution (Å) (cutoff on F)	1.54	5.0–1.53 (2σ)	5.0–1.64 (2σ)	5.0–1.76 (2σ)
No. of reflections	32080	33101	29528	24188

(*c*) Using *TNT*

Data set	CPA	Apo-CPA	Hg-CPA
Data-collection device	R-AXIS IIC (monochr.)	DIP100	DIP100
Type of refinement	*TNT*	*TNT*	*TNT*
No. of cycles	8	15	7
Resolution (Å, no cutoff)	30–1.53	13–1.88	12.1–1.76
No. of reflections	35152	21331	26751
R factor	0.145	0.143	0.147
R.m.s. deviations			
Bond lengths (Å)	0.020	0.016	0.017
Bond angles (°)	2.9	2.8	2.7
Trigonal planes	0.013	0.013	0.013
General planes	0.019	0.017	0.017
Bad contacts	0.034	0.064	0.043

model of Hg-CPA have been deposited with the Protein Data Bank.†

The starting model for the native structure, for use with the R-AXIS IIC monochromator data was taken from the model of Hg-CPA after 38 cycles of refinement with 206 solvent molecules. No changes were made to the structure except that the zinc was put in place of the mercury, and the Tris molecule and the extra cation were removed. Peak searches were carried out as described above. At the end of refinement, the same method of confirming water molecules was performed as described for Hg-CPA. The final model obtained was then used as the starting model for refinement against the R-AXIS IIC mirror/mirror data. Initial refinement restraints were loosened, to allow the model to relax (seven cycles), followed by a second round of refinement (11 cycles) with tight restraints until the refinement had converged. After completion of the refinement with *PROLSQ*, analysis of the main-chain torsion angles revealed two areas which had very poor φ,ψ values. These regions were around Ser57 and Thr133. Inspection of $2|F_o| - |F_c|$ electron-density maps based on the R-AXIS monochromator data suggested alternate conformations for these two regions, although the electron density was quite poor (Fig. 3). This data set was used because it contained both the highest and lowest resolution ranges of all the data sets. The models built for these two regions were also used in apo-CPA (DIP100) and Hg-CPA. Following the remodeling, a further eight cycles of refinement with the program *TNT* were carried out. Results for both models of native CPA are summarized in Tables 3(*b*) and 3(*c*). Since the improvement in the model was marginal, remodelling was carried out in only three of the models presented here, namely those models which were deposited with the Protein Data Bank. The PDB code for the model of native CPA refined against the monochromator data is 1YME.

4. Results and discussion

4.1. *Comparison of the newest model of CPA with the currently available model*

In order to evaluate the structure of native CPA obtained from the *PROLSQ* refinement of 1.53 Å R-AXIS IIC monochromator data, we have compared it with the currently available structure for native CPA (Rees *et al.*, 1983). It is clear from the results summarized in Table 3(*b*), that there has been a significant improvement in the geometry of the model for CPA. This is presumably because of the high quality of the

data since only one crystal was used for the whole data set instead of three, there was more redundancy in the data set contributing to statistically better data, the beam was stronger, and the crystal was exposed to the X-ray beam for a much shorter time.

An overlap of the two native models gave a value of 0.45 Å root-mean-square (r.m.s.) difference for all atoms, 0.19 Å r.m.s. for all C^α atoms, 0.19 Å r.m.s. for all main-chain N atoms, 0.21 Å r.m.s. for all main-chain O atoms, and 0.18 Å r.m.s. for all main-chain C atoms. These shifts presumably originate from the improved geometry in the newer model. Fig. 4 shows a plot of the r.m.s. difference *versus* sequence number. A plot of mean *B* factor (main chain) per residue *versus* sequence number (Fig. 4) shows that the largest r.m.s. differences correspond to the areas of highest *B* factor, indicating that there are no real differences between the course of the polypeptide chain in the two structures. A plot comparing mean side-chain *B* factors and side-chain r.m.s. values revealed several regions where high r.m.s. differences (>1.0 Å) did not correspond to high temperature factors. An investigation of each of these regions showed that most of the side chains involved are leucine, valine or threonine which had been manually adjusted to have the opposite conformation relative to the original model (see §2). Three side chains had large r.m.s. differences which were not due to this type of manual adjustment. One was Asp16, where the R-AXIS model has the side chain turned 90° about χ_2; no apparent explanation could be found for this difference. The second side chain was that of Ser194, where there was a difference of ~180° about χ_1;
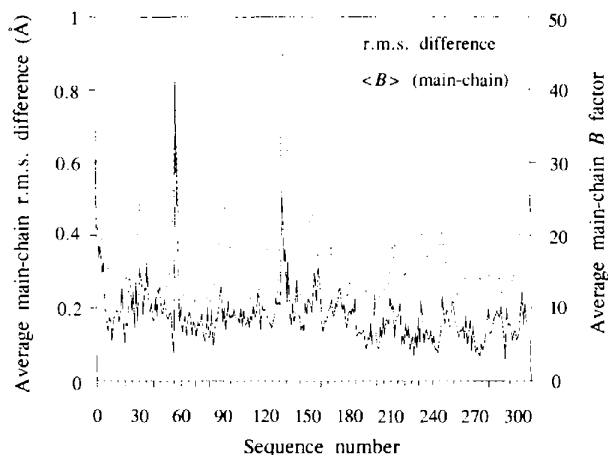

Fig. 4. Plot of average r.m.s. difference per residue between main-chain atoms of the model of CPA derived from the R-AXIS monochromater data and the model available from the Protein Data Bank (PDB), combined with the average main-chain *B* factor per residue for the model derived from the R-AXIS data. Note that the model for native CPA is after *PROLSQ* refinement only, in order that a consistant refinement program be used for this and subsequent comparisons of this type.

movement of the main chain, however, had brought the positions of the hydroxyl groups in each model close to each other. Ser254 showed a similar difference in conformation; these two side chains are discussed below.

Fig. 5 shows a difference B-factor plot, generated by subtracting the mean main-chain B factor for each residue of the original model from the corresponding value from the new model. In order to account for any differences in scaling which would have affected the B factors for each structure, the mean of these differences was then subtracted from each value, essentially zeroing the results. The mean difference was 8.72 Å^2 and the standard deviation was 2.45 Å^2. Only two regions show differences greater than 3σ. The first was the segment of Thr133–Ser134. In both structures these regions have very high B factors (>40 Å^2 in the R-AXIS model, and >50 Å^2 in the PDB model) and so the significance of these differences is questionable. The second region involves Thr304, near the C terminus of the molecule. The B factors in the R-AXIS model are large (~25 Å^2) while the PDB model has B factors around 8 Å^2. Binding experiments using the anti-cancer drug cisplatin (unpublished results) have showed this area to be capable of undergoing some conformational changes without disturbing the crystal structure, thus, there seems to be some inherent flexibility of this area in the crystal.

The gene sequence of bovine pancreatic CPA has changed several residues from uncharged to charged side chains, and two from charged side chains to uncharged. In all, seven residues require correction in the sequence published with the PDB entry (5CPA). Two corrections were in fact made by Pétra et al. (1971), who noted that Glu31 should be amidated and not Gln28, as was assigned in the original sequence

(Bradshaw et al., 1969). It is interesting to note that while Le Huërou et al. (1991) list the correct residues Asp93 in place of Asn93, they do not mention this as one of the changes, and discuss only the remaining four changes.

Examination of the hydrogen bonding around these corrected residues showed that in some instances, the previous assignments gave rise to questionable hydrogen-bond interactions. All the corrected assignments show reasonable interactions, although in most cases a simple rotation of the carboxylate/amide group in the original structure would have given reasonable interactions. These results are summarized in Table 4, and each region is shown in Fig. 6.

One interesting aspect of the original model of CPA was the proposal that the zinc-bound water molecule was not as tightly bound as the other zinc ligands. This was suggested because of the difference between the B factor of zinc (and all of its protein ligands) and the water molecule. Furthermore, a second water site was found further from zinc and closer to Glu270 (see Fig. 7). In the current model based on the R-AXIS data, the B factors of the zinc ligands range from 11 to 14 Å^2. In addition, no density was found to suggest the existence of a second site in the location proposed by Rees et al. (1983). Furthermore, results from other metal derivatives such as Mn-CPA, and Cd-CPA (Greenblatt, Feinberg, Shoham, unpublished work) show no significant difference between the temperature factor of the metal and its water ligand, nor is there any evidence for a second water site.

Another interesting observation is that Ser194 has two conformations, as is apparent from the existence of difference density very close to the C$^\beta$ (<1.5 Å). This is shown in Fig. 8; movement of the O$^\gamma$ of Ser194 to this alternate site brings it within hydrogen-bonding
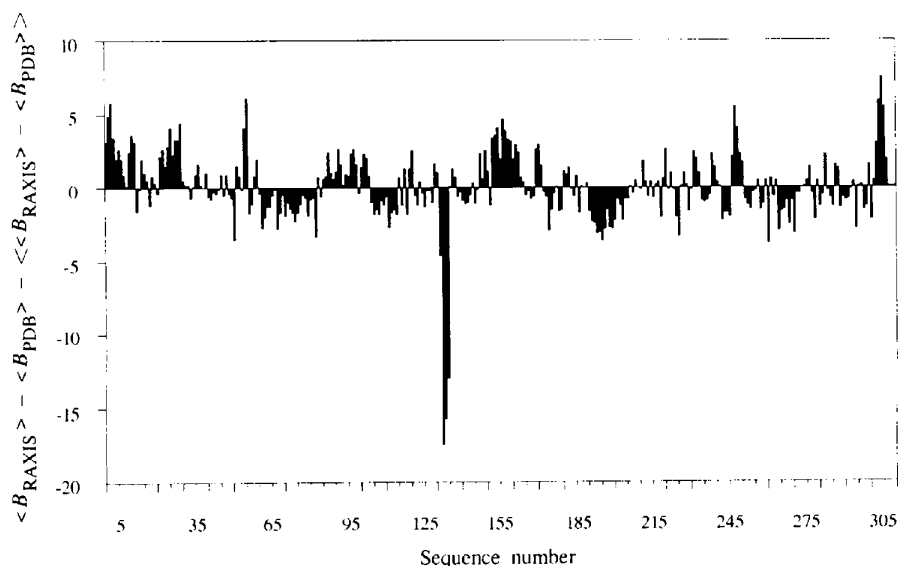

Fig. 5. Plot showing the differences in average main-chain B factors between the PDB model of CPA and the model derived from the R-AXIS data (PROLSQ refinement). In order to account for any differences in B factor which might arise from differences in scaling and the like, the mean of all the differences was subtracted from each value.

Table 4. *Summary of interactions affected by changes to gene sequence in CPA*

Change	Comments on original interactions	New hydrogen-bond interactions (distance, Å)	
Gln28 to Glu28	The former $N^{\epsilon 2}$ of the original model interacts with Thr4 N, itself a hydrogen-bond donor. Rotation of $\chi_2 = 180°$ would remove this unfavourable interaction.	Thr4 N$\cdots$O$^{\epsilon 2}$ (2.8) HOH546$\cdots$O$^{\epsilon 2}$ (2.6) Thr4 O$^{\gamma}\cdots$O$^{\epsilon 1}$ (2.6) HOH548$\cdots$Thr4 O$^{\gamma}$ (2.4) HOH548$\cdots$His29 N$^{\epsilon 2}$ (3.1) Lys84 N$^{\zeta}\cdots$HOH548 (3.2)	Asn5 N$^{\delta 2}$ in the refined model faced Lys84 N$^{\zeta}$; this has been turned around, but not refined. His29 and Lys84 also share water 507.
Glu31 to Gln31	The interaction between Asn307 N$^{\delta 2}$ and His303 O (2.7 Å), dictates that Asn307 O$^{\delta 1}$ must interact with residue 31, and so the presence of Gln31 N$^{\epsilon 2}$ is more reasonable than the former Glu31 O$^{\epsilon 2}$.	Gln31 N$^{\epsilon 2}\cdots$Asn307 O$^{\delta 1}$ (2.6) Gln31 N$^{\epsilon 2}\cdots$Asn306 O (2.8)	Symmetry molecule is generated by $(-x - 1, -1 + y + \frac{1}{2}, -z - 1)$
Asn89 to Asp89		HOH564$\cdots$Asp89 O$^{\delta 2}$ (2.4) Trp294 N$^{\epsilon 1}\cdots$HOH564 (2.8) Lys85 N$^{\zeta}\cdots$Asp89 O$^{\delta 2}$ (2.9) Lys85 N$^{\zeta}\cdots$Asp89 O$^{\delta 1}$ (3.1)	
Asn93 to Asp93	Asn93 N$^{\delta 2}$ in close proximity to Phe96 N (2.9 Å). Could be fixed by simple $\chi_2 = 180°$ rotation.	Phe96 N$\cdots$Asp93 O$^{\delta 2}$ (3.0) HOH564$\cdots$Asp93 O$^{\delta 2}$ (3.5)	HOH564 is shared between Asp93 and Asp89.
Asn114 to Asp114	Original model had interaction between Asn114 N$^{\delta 2}$ and Arg45 N$^{\eta 1}$. Rotation would not have helped in this case, since the N$^{\delta 2}$ would have been brought into close proximity to Glu34 N.	Arg45 N$^{\eta 1}\cdots$Asp114 O$^{\delta 2}$ (2.8) Ser41 O$^{\gamma}\cdots$Asp114 O$^{\delta 2}$ (2.6) Glu43 N$\cdots$Asp114 O$^{\delta 1}$ (2.8) Tyr42 N$\cdots$Asp114 O$^{\delta 1}$ (3.1) HOH402$\cdots$Asp114 O$^{\delta 1}$ (2.8) Arg45 N$^{\eta 1}\cdots$Glu43 O$^{\epsilon 2}$ (2.8)	
Glu122 to Gln122	Interactions between Glu122 O$^{\epsilon 1}$ and symmetry-related Thr245 O and Thr246 O.	Gln122 N$^{\epsilon 2}\cdots$Thr245 O (2.9) Gln122 N$^{\epsilon 2}\cdots$Thr246 O (3.4)	New model was refined with N$^{\epsilon 2}$ and O$^{\epsilon 2}$ reversed. Side chain has been rotated so that the N$^{\epsilon 2}$ faces the carbonyl O atoms of the symmetry-related Thr245 and Thr246 $(-x, -1 + y + \frac{1}{2}, -z)$.
Asn185 to Asp185	Interaction of Asn185 N$^{\delta 2}$ to Asp181 O (3.4 Å) does not occur now.	HOH512$\cdots$Asp185 O$^{\delta 2}$ (2.8) HOH602$\cdots$Asp185 O$^{\delta 2}$ (3.3) Thr209 O$^{\gamma}\cdots$Asp185 O$^{\delta 2}$ (2.6) Thr210 N$\cdots$Asp185 O$^{\delta 1}$ (2.9) Thr210 O$^{\gamma}\cdots$Asp185 O$^{\delta 1}$ (2.9) HOH550$\cdots$Asp185 O$^{\delta 2}$ (3.1)	Contacts with Thr209, Thr210, and HOH550 are to a symmetry-related molecule $(-x, -1 + y + \frac{1}{2}, -z - 1)$.

distance of a water molecule interacting with His196 N$^{\epsilon 2}$. The existence of two conformations for the side chain of Ser194 was also observed by Teplyakov *et al.* (1993). The structures of CPA which have the inhibitor benzyl succinate bound in the active site (Feinberg *et al.*, 1993, 1995), shows Ser194 O$^{\gamma}$ exclusively in this alternate site, interacting with the water molecule. It appears that binding of substrate or inhibitor in the active site causes Ser194 to favour one of the two alternate sites found in the native structure. It is also interesting to note that the R-AXIS native CPA structure has Ser254 in a different conformation from that found in the PDB model. This conformation in the R-

AXIS model brings Ser254 O$^{\gamma}$ closer to Ser194 O$^{\gamma}$ in the position shown in the native enzyme. There is no evidence to suggest that binding of inhibitors or substrate causes any conformation change in Ser254, despite the movement of Ser194 O$^{\gamma}$.

As to the improvement in the model brought about by remodeling the mobile regions around Ser57 and Thr133, the final B factors for Ser57 ranged between 28 and 63 Å^2; all the main-chain atoms for Thr133 had B factors of 100 Å^2 (maximum). It is not clear therefore, if there was any real gain with the new models. They do have the advantage, however, that the φ,ψ angles are more acceptable for these residues, and the new model
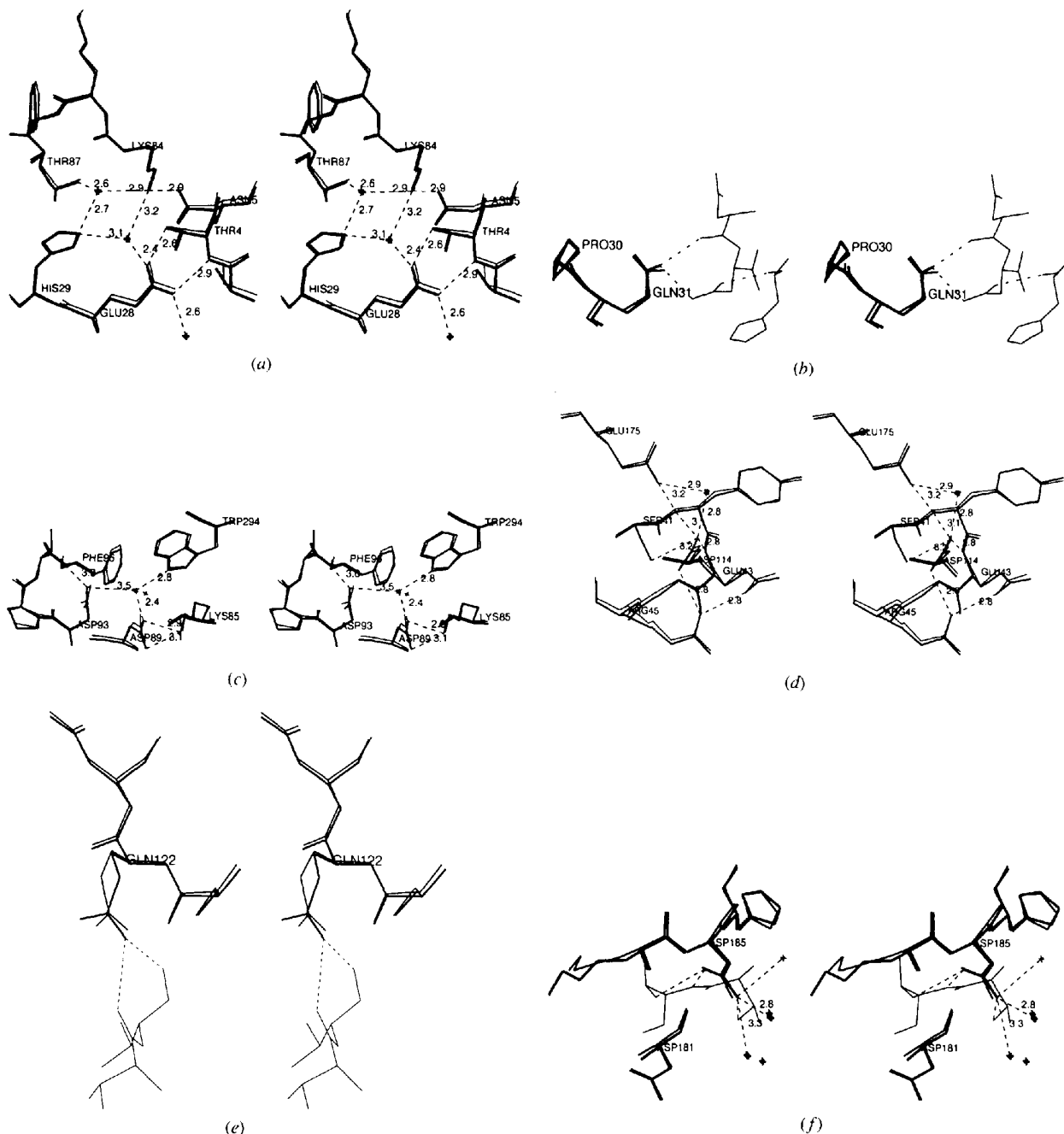
Fig. 6. (a) Stereoview of the region surrounding Glu28, showing the R-AXIS model in thick lines, and the PDB model in thin lines. Dashed lines represent possible hydrogen bonds, with the distances indicated. (b) View of Gln31 (thickest lines) and the possible interactions with symmetry-related Asn306 and Asn307 $(-x - 1, -1 + y + \frac{1}{2}, -z - 1$; thin lines). The PDB structure is drawn in medium lines. Note that the interaction of Asn307 $N^{\delta 2}$ with His303 O forces Asn307 $O^{\delta 1}$ to face the side chain of Gln31. Since this residue was initially identified as Glu31, this would have implied that two hydrogen-bond acceptors face each other. According to the gene sequence this interaction involves Gln31 $N^{\epsilon 2}$, which acts as a hydrogen-bond donor to both Asn307 $O^{\delta 1}$ and Asn306 O. (c) Environment of Asp89 and Asp93, showing the R-AXIS model in thick lines, and the PDB model in thin lines. Dashed lines represent possible hydrogen bonds in the R-AXIS structure, with the distances indicated. (d) Stereoview of the environment of Asp114, with the R-AXIS model drawn in thick lines, and the PDB model in thin lines. Possible hydrogen bonds in the R-AXIS structure and their distances are shown with dashed lines. (e) Possible hydrogen bonds (dashed lines) between Gln122 (thickest lines) and two carbonyl groups from a symmetry-related molecule (Thr245 and Thr246; thin lines). The conformation of Glu122 of the PDB structure is shown in medium thick lines. (f) Environment of Asp185 (thickest lines) showing possible hydrogen bonds (dashed lines) to solvent molecules and a symmetry-related molecule (thin lines). The PDB structure (Asn185) is drawn in medium lines. Water molecules (crosses) from the current model are connected by the hydrogen bonds, while the water molecules from the PDB structure have no hydrogen bonds drawn to them.

for Ser57 seems a marginally better fit for the electron density of this region (Fig. 3).

4.2. Metal-removed and metal replaced-CPA

4.2.1. *Apo-CPA*. The refined structures of native CPA and apo-CPA show clearly that removal of the metal in CPA has no significant effect on the overall structure of the enzyme, since the r.m.s. difference between C^α's of native CPA (R-AXIS model) and apoCPA (DIP100) is only 0.12 Å. A plot of average main-chain r.m.s. difference per residue (Fig. 9) shows

that the regions with high r.m.s. correspond to regions with elevated temperature factors in the apo-CPA structure. The major exception to this is the area centred around Ser212. Examination of this region does not show any obvious reason why this residue has relatively larger r.m.s. deviations between the two structures. In any event, these results demonstrate that zinc plays no structural role by maintaining a certain active conformation of the overall enzyme. It might be thought that the reason why little difference is seen between native and apo-CPA is because the metal was removed from the crystal, and that the molecule is
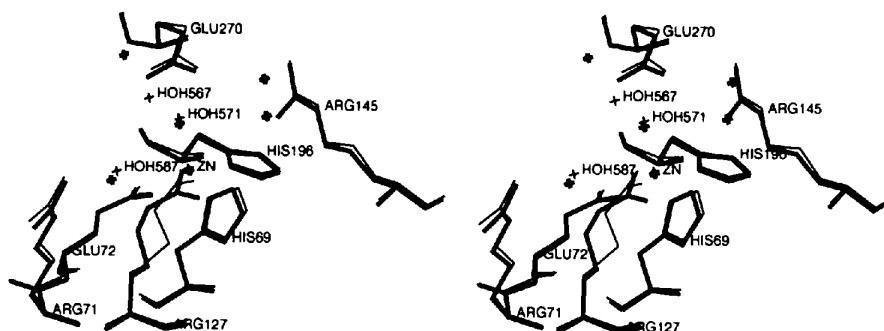

Fig. 7. Active site of native CPA showing the zinc ligands in the R-AXIS structure (thick lines) and the PDB structure (thin lines). Labeled water molecules are from the PDB structure. Not all waters in the active site of the R-AXIS structure are shown; only those in the local vicinity of the zinc ion are shown, to compare with the PDB structure.
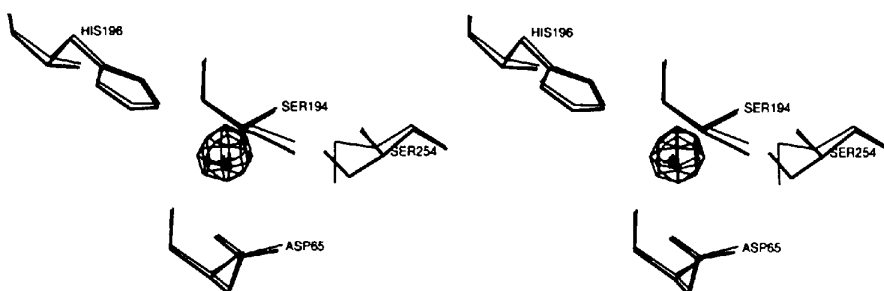

Fig. 8. Stereoview of Ser194, with electron density showing the alternate position. The contour level is 3σ; the R-AXIS model is drawn in thick lines, and the PDB model is shown in thin lines. Note the alternate conformation modeled for Ser254 (see text for discussion).
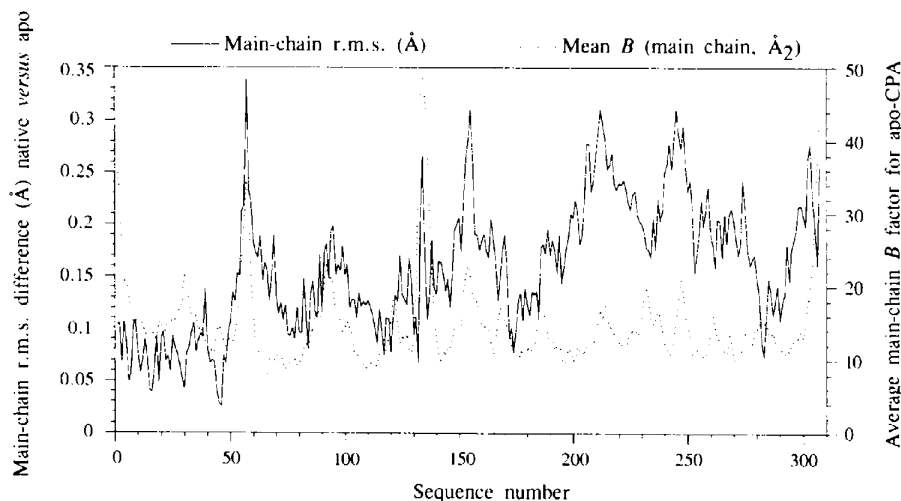

Fig. 9. Plot showing mean r.m.s. difference per residue (solid line) for main-chain atoms of native CPA (R-AXIS IIC monochromator) and apo-CPA (R-AXIS IIC mirrors). Also plotted is the mean-chain B factor (dotted) per residue for apo-CPA (R-AXIS IIC mirrors). Again, the models used here are those refined by *PROLSQ* only.

locked into this conformation. This is unlikely, since apo-CPA and metal replaced CPA can crystallize in a cell isomorphous with native CPA.

Owing to the lack of significant differences between the structures of native CPA and apo-CPA, we shall restrict our discussion to the active-site region. The largest change in the active site of apo-CPA involves the side chain of His196, which undergoes a rotation about χ_2 of ~110° to form an apparent salt bridge with the side chain of Glu270, as shown in Fig. 10. Table 5 lists some relevant interatomic distances found in the active site of apo-CPA. The zinc site does not appear to be occupied by any solvent molecule interacting with His69, despite earlier studies that suggested a water molecule was bound in place of the zinc, and that the rotations about χ_2 of His196 were small (Rees & Lipscomb, 1983). The most likely explanation for these observations was that these were lower resolution, unrefined structures, and so model bias may have affected the resulting maps. While the peak may have come from residual zinc in the crystals, this is unlikely since the studies were carried out in the presence of chelator.

One currently unresolved question centers around the role of the zinc in forming the so-called Michaelis–

Table 5. *Possible hydrogen bonds in the active site of apo-CPA*

Donor	Acceptor	Distance (Å)
His69 $N^{\delta 1}$	Glu72 O^{r1}	3.1
His69 N^{r2}	Asp142 $O^{\delta 1}$	2.9
Arg127 $N^{\eta 1}$	Asp142 $O^{\delta 2}$	2.7
His196 $N^{\delta 1}$	Glu270 O^{r2}	2.9
His196 N^{r2}	HOH440 O	2.8
HOH440 O	Ser254 O^{γ}	3.2
HOH440 O	HOH508 O	3.3

Menten complex. Previous reports have suggested that apo-CPA cannot bind many ester substrates, while peptide substrates do bind in the absence of zinc (Coleman & Vallee, 1962; Steitz et al., 1967; Rees & Lipscomb, 1983). This observation suggested different binding modes for ester and peptide substrates, and was incorporated in a proposal for the mode of action of CPA by Christianson & Lipscomb (1989). In an attempt to characterize the purported apo-CPA substrate complex we soaked crystals of apo-CPA in solutions of dansyl-Gly-Trp, a common peptide substrate used in kinetic studies. Solution of the structure from these crystals did not show any bound peptide substrate. Furthermore, in soaking experiments with apo-CPA
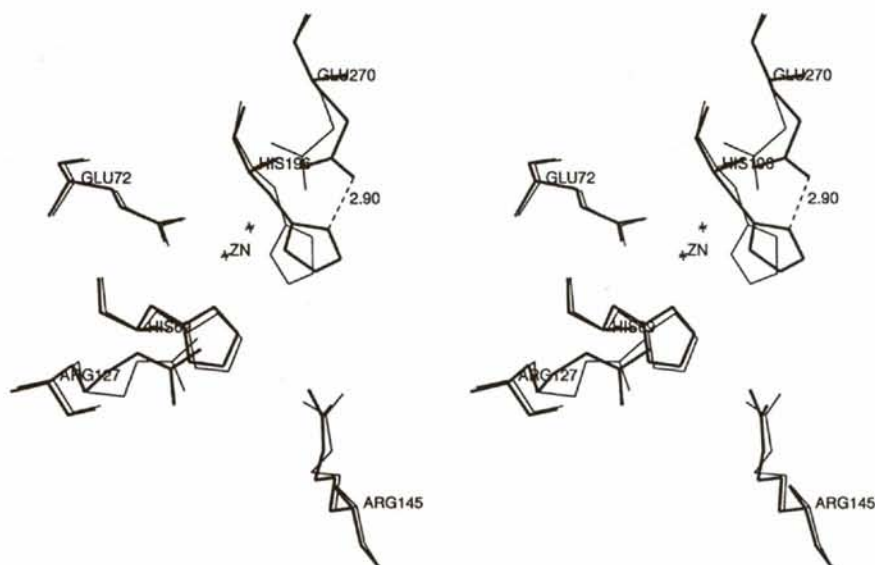

Fig. 10. Comparison between the active sites of apo-CPA (thick lines) and native CPA (thin lines). Note that the largest change occurs as a result of the ion pairing between Glu270 and His196.
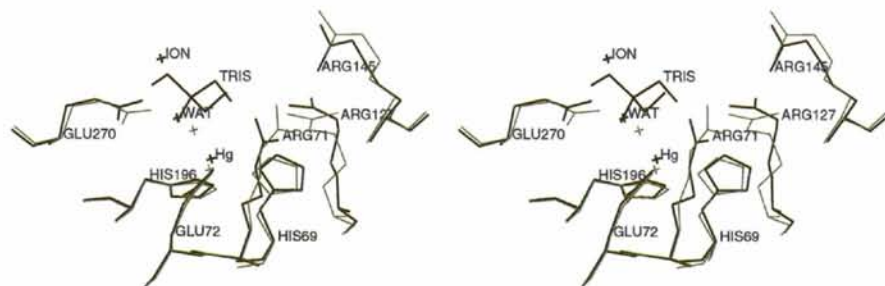
Fig. 11. Stereoview of the overlap of the active site of Hg-CPA (black lines) and native CPA (green lines).

crystals in solutions of benzyl succinate and a keto-methylene inhibitor (BOC-Phe-ψ-Phe), the active site was found to be devoid of any bound inhibitor. Both inhibitors are known to bind to enzyme molecules in crystals of native CPA and other metal derivatives (Shoham et al., 1988; Feinberg et al., 1993). While it might be argued that the conformational changes associated with the side chain of His196 may prevent normal binding of substrates or inhibitors, the implication of Coleman & Vallee (1962) is that bound substrates interfere with the reassociation of zinc and the apo-enzyme. Although the di-peptide Gly-Tyr seems to bind to apo-CPA (Steitz et al., 1967; Rees & Lipscomb, 1983), this molecule also displays anomalous binding to the holo-enzyme (Christianson & Lipscomb, 1986), and so may not be representative of other substrates. One possible reason for the observed lack of binding is that the soaking concentration was too low. Many of the inhibitors and substrates used in these experiments are not easily soluble above concentrations of 10^{-3} M. Since no chemistry is carried out in the active site of apo-CPA, peptide substrates should, at best, bind only as tightly as K_m, the binding constant for the formation of the Michaelis–Menten complex. This would also hold true for those inhibitors which bind tightly by mimicking catalytic intermediates, such as the ketomethylene inhibitors. This was in fact demonstrated by Coleman & Vallee (1962) who showed the inhibition constant for some common substrates which prevented binding of zinc to apo-CPA, to be close to their K_m values.

4.2.2. *Hg-CPA.* There is almost no change in the structure of CPA when Hg occupies the metal binding site, as opposed to zinc. This is shown clearly in the superposition of native CPA on Hg-CPA as shown in Fig. 11. The small shifts that are observed in the ligands of the metal can easily be explained by the increased size of the mercury dication. Table 6 compares the distances between the Hg^{2+} ion and its ligands with those of native CPA. In fact, the largest shifts associated

Table 6. *Comparison of ligand distances (in Å) between Hg-CPA and Native-CPA*

Ligand	Hg	Zn
His69 $N^{\delta 1}$	2.3	2.0
His196 $N^{\epsilon 1}$	2.2	2.1
Glu72 $O^{\epsilon 1}$	2.7	2.9
Glu72 $O^{\epsilon 2}$	2.6	2.7
Water	–	2.0
Tris N	2.4	–

with active-site side chains are seen in Arg71 and Arg127. This is apparently due to the fact that in the native structure, there is a water molecule which interacts with Arg71 (see Fig. 7), which is not present in Hg-CPA, where Arg71 interacts directly with the mercury-bound Tris molecule, drawing it closer to the vicinity of the mercury cation. This movement presumably repels Arg127, and brings about its observed shift.

In addition to the mercury cation bound in the place of zinc, an additional cation was found bound to Glu270. Although this cation was refined as a copper ion, this choice was arbitrary, based on considerations of B factor and occupancy. It is quite possible that the cation is, in fact some other ion with similar scattering characteristics (such as zinc), which would be indistinguishable from copper in these circumstances. While it is possible that this cation is a mercury cation at low occupancy, data collected on different crystals of Hg-CPA, soaked in the same concentration of $HgCl_2$, showed only the Tris molecule bound to the mercury cation in the active site, and not the extra cation seen with these data. Since the data indicating the presence of a cation bound to Glu270 were collected on a crystal soaked in a solution of an inhibitor (benzyl succinate), it appears that some contaminant has been introduced somewhere in the soaking process. Whatever the identity of the cation may be, the result is significant given the observation that excess zinc can inhibit CPA (Vallee et al., 1960). More recent studies have, in fact, proposed
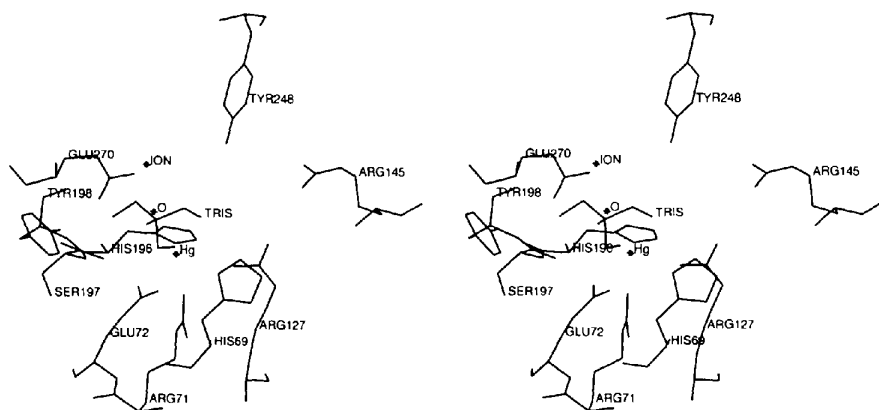

Fig. 12. Active site of Hg-CPA, showing Tris molecule bound to Hg-CPA and another cation (labelled ION) bound to Glu270. The distance between the cation and Glu270 $O^{\epsilon 2}$ is 2.5 Å, and to Glu270 $O^{\epsilon 1}$ is 3.5 Å. The distances of the Tris hydroxyl groups to various side chains in the active site are as follows: to Tyr198 N, 2.5 Å; to Arg71 $N^{\eta 2}$, 2.8 Å; to Tyr248 O^{η}, 3.5 Å.

that the location of binding of inhibitory zinc is Glu270 (see Larsen & Auld, 1991, and references contained therein). This structure then confirms this proposal. Since zinc and lead bind most avidly to this site when compared with other transition metals, the identity of this ion is most probably zinc, although this preference was observed when the catalytic cation was also zinc. It is difficult to assess the effect that mercury has on the cation specificity and binding constant of this secondary site. It may be that Hg-CPA is more susceptible to cation binding at Glu270, relative to native CPA; if this is so, cation contaminants may interfere with the productive binding of substrates to Hg-CPA, and so may contribute to the loss of activity observed upon replacement of the zinc in native CPA with mercury. In addition to the potential inhibition caused by an extra cation bound in the active site of Hg-CPA, the mercury-bound Tris molecule may also interfere with productive substrate binding. This may be due to the mere fact that substrates or inhibitors may not be capable of displacing Tris. In addition, the side chain of Tyr248 appears to favour the 'down' position in Hg-CPA, interacting with an hydroxyl group of the Tris molecule (see Fig. 12). This would effectively block off the active site and limit the substrate's approach. Although the kinetic studies which showed that Hg-CPa had no activity towards peptide substrates were carried out in veronal buffer, and not in Tris, the presence of an extra cation in the active site may have contributed to this observed loss of activity. The kinetics for the esters were in fact carried out in Tris buffer; nevertheless, since according to the hypothesis forwarded by Christianson & Lipscomb (1989), ester substrates interact first with the zinc ion upon binding, the substrate may competitively displace the Tris molecule bound to the mercury. Furthermore, there is no difficulty in the fact that Hg-CPA showed an enhanced ester hydrolysis rate relative to native CPA, since it may be that for this metal derivative without the Tris buffer, the rate of ester hydrolysis would be even faster.

4.3. Comparison of data sets

The structures presented here were all collected on different data-collection devices. Since the proteins are all isomorphous this provides some opportunity to compare data from these devices in a limited way. This is especially true of the apo-CPA structure, since data from crystals of this form were collected on the DIP100 image plate, the FAST area detector, and the R-AXIS IIC imaging plate (mirror/mirror optics). Comparison of the R factor versus resolution plots (Fig. 13) for the data sets shows that all the data sets exhibit the same trends, although the R-AXIS data have the lowest R factors at high resolution, and both the DIP100 and R-AXIS have better R factors than the FAST throughout most of the refined range of data. There are at least

three reasons for the differences. Firstly the trend in R factors is in accord with the exposure times which, per degree of rotation, were in the ratio of 1:2.78:4.50 for the FAST/DIP100/R-AXIS data. Secondly, the data sets from the R-AXIS have reflections with $I/\sigma(I)<1$ omitted whereas only negative intensities were omitted from the FAST and DIP100 data sets. The latter will have a higher proportion of weaker data, especially at high resolution. On the other hand, the R-AXIS mirror/mirror optics are known to reduce the incident intensity of the beam at the crystal (see below). Therefore, it is reassuring that the refinements are all similar (Table 3a). The DIP100 data has marginally the best fit to the target geometry, but the refinement resolution limit is higher and the R factor marginally higher.

In a general comparison of the native (R-AXIS, monochromator), apo-CPA (DIP100), and Hg-CPA (DIP100) refinements, the R factor versus resolution plot (Fig. 14) shows that the R-AXIS and DIP100 seem to give similar results in the range of 2.0–4.0 Å, but show differences at resolutions below 2.0 Å. Here the R-AXIS data is clearly superior, as would be expected for the reasons given above. Even though the higher R factors for the apo-CPA model could be partially due to the fact that no cutoff was used on F during the refinement, the Hg-CPA structure, which did have a 2σ cutoff on F, is still not as good as the R-AXIS data. Note that the superiority of the R-AXIS data is more apparent here than in the comparison among the apo-CPA structures. This is presumably due to the higher quality data obtainable for CPA from the stronger beam of the monochromator relative to the mirror/mirror optics.
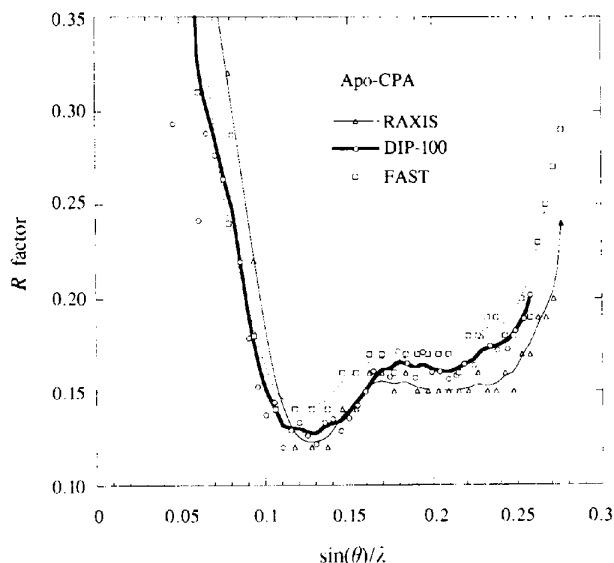

Fig. 13. Plot of R factor versus resolution (PROLSQ refinement) for three data sets of apo-CPA: R-AXIS image plate (mirror/mirror optics), DIP100 image plate, and the FAST area detector.

In order to compare the two optical systems available with the R-AXIS IIC imaging plate (mirror/mirror optics system *versus* graphite monochromator), two data sets of native CPA were collected in Japan at the Rigaku factory. One data set used a standard graphite monochromator, while the other used mirror/mirror optics. A comparison of the signal-to-noise ratios for both data sets as a function of resolution (Fig. 15) clearly shows that the mirror/mirror optics produced weaker data. This is quite understandable given the longer beam path length associated with this optical system, in particular the longer distance in between the source and the first mirror. In this case the advantages gained by having a well focused beam do not compensate for the loss of intensity of the beam which is incident on the crystal. These optics, then, are of no benefit for small or medium unit-cell systems, where spot resolution is not problematic. The main advantage would presumably be gained for large unit cells, where resolving closely spaced spots is critical, or for small crystals where the brightness of the focused beam may still be advantageous. A modification which could improve the beam strength significantly would be to move the double mirror assembly as close to the source as possible, thus giving a well focused beam with minimal loss of beam intensity. The current configuration of the R-AXIS places the mirrors midway between the source and the detector. Although simplistic optics

might suggest this configuration, most practical experience shows that there is no appreciable difference in the focus of the beam when the mirrors are placed very close to the source.

5. Conclusions

Structural changes caused by removal of zinc from CPA are restricted to side chains in the active site, with the largest change being the formation of a salt bridge between His196 and Glu270. No solvent molecule or ion appears to bind in place of the zinc. Attempts to bind substrate and inhibitors to crystallized apo-CPA proved unsuccessful, and so the exact nature of the apparent complex between peptide substrates and apo-CPA remains unknown.

In native CPA there appears to be no difference in the B factor of the zinc-bound water molecule relative to the other zinc ligands. Furthermore, there is no evidence for a secondary location of the zinc-bound water molecule.

Replacement of zinc in CPA by other metals does not appear to alter the conformation of CPA at all, with only minor changes occurring in the side chains of the residues which bind the metals, presumably owing to differences in ion sizes. Thus, changes in activity observed for various metal derivatives are not due to geometric changes in metal coordination. Hg-CPA appears capable of binding a second cation to Glu270,
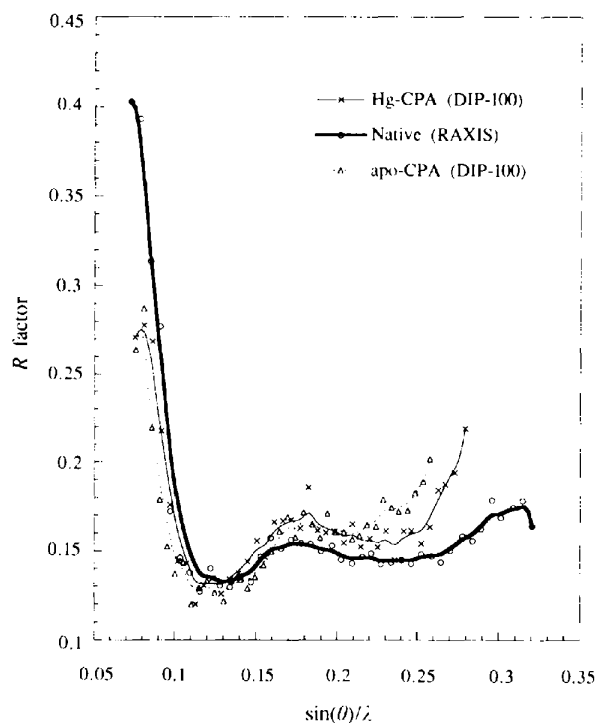

Fig. 14. Plot of R factor *versus* resolution (*PROLSQ* refinement) for two data sets from the DIP100 (apo-CPA and Hg-CPA), and the native CPA data sets from the R-AXIS (monochromator).
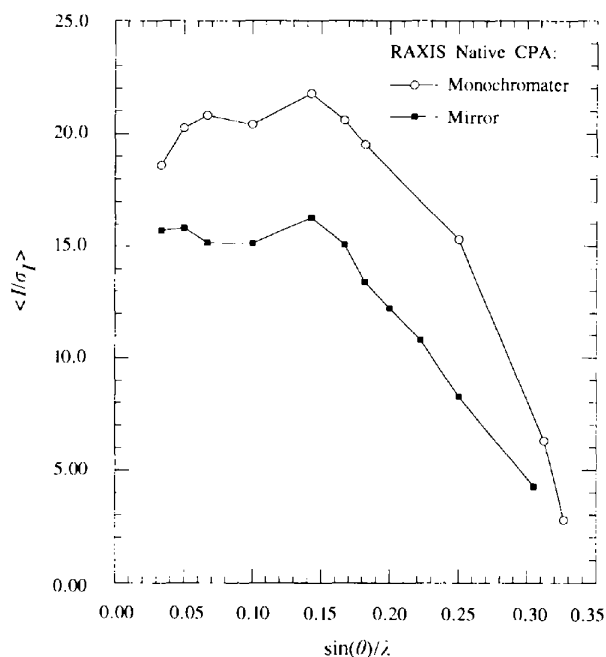

Fig. 15. Comparison of the signal-to-noise ratio for native CPA data sets collected under similar conditions, but one using a monochromatized beam, the other using a beam focused by double mirrors.

and of binding a Tris molecule as a ligand to the mercury cation in place of the apical water molecule seen in native CPA. The location of the extra cation is proposed as the binding site of inhibitory zinc in native CPA. These observations provide a possible explanation for the apparent lack of activity of Hg-CPA towards peptide substrates. Thus, some of the changes in CPA activity observed upon replacement of the zinc with other metals may be artifacts caused by solvent inhibition.

A comparison of several CPA data sets from various data-collection devices shows that the R-AXIS monochromator system produced a refined model to higher resolutions than the other systems. The R-AXIS mirror/mirror optics gives data comparable to the other instruments (DIP100 and FAST).

References

Arndt, U. W. & Thomas, D. J. (1982). *Nucl. Instrum. Methods*, **201**, 21–25.

Arndt, U. W. & in't Veld, D. J. (1988). *Adv. Electron Phys.* **74**, 285–296.

Bradshaw, R. A., Ericsson, L. H., Walsh, K. A. & Neurath, H. (1969). *Proc. Natl Acad. Sci USA*, **63**, 1389–1394.

Christianson, D. W. & Lipscomb, W. N. (1986). *Proc. Natl. Acad. Sci. USA*, **83**, 7568–7572.

Christianson, D. W. & Lipscomb, W. N. (1989). *Acc. Chem. Res.* **22**, 62–69.

Coleman, J. E. & Vallee, B. L. (1961). *J. Biol. Chem.* **236**, 2244–2249.

Coleman, J. E. & Vallee, B. L. (1962). *J. Biol. Chem.* **237**, 3430–3436.

Feinberg, H., Greenblatt, H. M. & Shoham, G. (1993). *J. Chem. Inf. Comput. Sci.* **33**, 501–516.

Feinberg, H., Greenblatt, H. M., Behar, V., Gilon, C., Cohen, S., Bino, A. & Shoham, G. (1995). *Acta Cryst.* **D51**, 428–449.

Fujinaga, M., Delbaere, L. T. J., Brayer, G. D. & James, M. N. G. (1985). *J. Mol. Biol.* **184**, 479–502.

Greenblatt, H. M., Ryan, C. A. & James, M. N. G. (1989). *J. Mol. Biol.* **205**, 201–228.

Hardman, K. D. & Lipscomb, W. N. (1984). *J. Am. Chem. Soc.* **106**, 463–464.

Hendrickson, W. A. & Konnert, J. H. (1980). In *Biomolecular Structure, Function, Conformation and Evolution*, edited by R. Srinivasan, Vol. 1, pp. 43–57. Oxford: Pergamon Press.

Larsen, K. S. & Auld, D. S. (1991). *Biochemistry*, **30**, 2613–2618.

Le Huërou, I., Guilloteau, P., Toullec, R., Puigserver, A. & Wicker, C. (1991). *Biochem. Biophys. Res. Commun.* **175**, 110–116.

Messerschmidt, A. & Pflugrath, J. W. (1987). *J. Appl. Cryst.* **20**, 306–315.

Pétra, P. H., Hermodson, M. A., Walsh, K. A. & Neurath, H. (1971). *Biochemistry*, **10**, 4023–4025.

Rees, D. C., Howard, J. B., Chakrabarti, P., Yeates, T., Hsu, B. T., Hardman, K. D. & Lipscomb, W. N. (1986). *Crystal Structures of Metallosubstituted Carboxypeptidase A*. In *Zinc Enzymes, Progress in Inorganic Biochemistry and Biophysics*, Vol. 1, pp. 155–165. Boston: Birkhauser.

Rees, D. C., Lewis, M. & Lipscomb, W. N. (1983). *J. Mol. Biol.* **168**, 367–387.

Rees, D. C. & Lipscomb, W. N. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 7151–7154.

Sato, M., Yamamoto, M., Imada, K., Katsube, Y., Tanaka, N. & Higashi, T. (1992). *J. Appl. Cryst.* **25**, 348–357.

Shoham, G., Christianson, D. W. & Oren, D. A. (1988). *Proc. Natl Acad. Sci. USA*, **85**, 684–688.

Shoham, G., Nechushtai, R., Steppun, J., Nelson, H. & Nelson, N. (1990). Unpublished results.

Shoham, G., Rees, D. C. & Lipscomb, W. N. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 7767–7771.

Steitz, T. A., Ludwig, M. L., Quiocho, F. A. & Lipscomb, W. N. (1967). *J. Biol. Chem.* **242**, 4662–4668.

Tanaka, I., Yao, M., Suzuki, M., Hikichi, K., Matsumoto, T., Kozasa, M. & Katayama, C. (1990). *J. Appl. Cryst.* **23**, 334–339.

Teplyakov, A., Wilson, K. S., Orioli, P. & Mangani, S. (1993). *Acta Cryst.* **D49**, 534–540.

Tronrud, D. E. (1992). *Acta Cryst.* **A48**, 912–916.

Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987). *Acta Cryst.* **A43**, 489–501.

Tucker, P. A. (1990). Unpublished work.

Vallee, B. L. & Galdes, A. (1984). In *Advances in Enzymology*, edited by A. Meister, Vol. 56, pp. 283–430. New York: John Wiley.

Vallee, B. L., Rupley, J. A., Coombs, T. L. & Neurath, H. (1960). *J. Biol. Chem.* **235**, 64–69.