X-ray crystallographic investigation of substrate binding to carboxypeptidase A at subzero temperature

(Protein crystallography/x-ray cryoenzymology/enzyme-substrate complex)

DAVID W. CHRISTIANSON AND WILLIAM N. LIPSCOMB

Gibbs Chemical Laboratories, Department of Chemistry, Harvard University, Cambridge, MA 02138

Contributed by William N. Lipscomb, May 22, 1986

ABSTRACT A high-resolution x-ray crystallographic investigation of the complex between carboxypeptidase A (CPA; peptidyl-l-amino-acid hydrolase, EC 3.4.17.1) and the slowly hydrolyzed substrate glycyll-t-tyrosine was done at −9°C. Although this enzyme-substrate complex has been the subject of earlier crystallographic investigation, a higher resolution electron-density map of the complex with greater occupancy of the substrate was desired. All crystal chemistry (i.e., crystal soaking and x-ray data collection) was performed on a diffractometer-mounted flow cell, in which the crystal was immobilized. The x-ray data to 1.6-Å resolution have yielded a well-resolved structure in which the zinc ion of the active site is five-coordinate: three enzyme residues (glutamate-72, histidine-69, and histidine-196) and the carboxyl oxygen and amino terminus of glycyll-t-tyrosine complete the coordination polyhedron of the metal. These results confirm that this substrate may be bound in a nonproductive manner, because the hydrolytically important zinc-bound water has been displaced and excluded from the active site. It is likely that all dipeptide substrates of carboxypeptidase A that carry an unprotected amino terminus are poor substrates because of such favorable bidentate coordination to the metal ion of the active site.

The zinc metalloenzyme carboxypeptidase A (CPA; peptidyl-l-amino-acid hydrolase, EC 3.4.17.1) is an exopeptidase that exhibits preferred specificity toward peptides or esters bearing large hydrophobic carboxyl-terminal residues such as phenylalanine. The proteolytic and esterolytic mechanism(s) of this protease have been the subject of a wealth of chemical and structural studies that has been recently reviewed (1–4). Two possible hydrolytic mechanisms include the attack of water/incipient hydroxide directly at the scissile carbonyl carbon of the substrate, promoted by glutamate-270 and/or zinc, or the attack of the γ-carboxylate of glutamate-270 at the scissile carbonyl carbon, with subsequent formation of a mixed anhydride intermediate, which then undergoes hydrolysis to yield products.

Either the zinc ion of the active site or the positively charged guanidinium moiety of arginine-127, or both, may serve to polarize the substrate carbonyl prior to its attack (regardless of the general mechanism). Mechanistic analyses have been aided in part by structural studies of the native enzyme (5–7) and its complexes with different inhibitors (6–15), including the slowly hydrolyzed (16) substrate glycyll-t-tyrosine (GY; Fig. 1) (6, 7), which is a competitive inhibitor of more rapidly cleaved substrates. Such nonacylated dipeptides are typically cleaved more slowly by a factor of 1000–5000 than their acylated analogues (17). The principal features of the most recent (6) structural study of the CPA–GY complex at 2.0-Å resolution (using x-ray data collected at 4°C) were that its carbonyl oxygen coordinated to zinc at a position different from that of the native zinc-bound water/hydroxide and that the terminal amino group statistically occupied two positions: (i) coordinated to zinc (as the free base), and (ii) hydrogen-bonded to glutamate-270 (possibly as a quaternary ammonium cation). The terminal carboxylate formed a salt link with arginine-145, and tyrosine-248 was in the "down" conformation, presumably donating a hydrogen bond to one of the terminal carboxylate oxygens.

The current high-resolution x-ray crystallographic investigation of the CPA–GY complex was performed at subzero temperature in order to further retard the slow enzyme-catalyzed hydrolysis reaction in the crystal. An aqueous-organic cryosolvent buffer system was used in order to allow for subzero temperatures at the crystal. Because the crystal form CPA used in this study exhibits one-third the activity of the enzyme in solution (18), a diffractometer-mounted flow cell assembly was used to provide a continuous, fresh supply of GY to the enzyme crystal, in case appreciable amounts of GY could be cleaved and released by the enzyme over the time span of crystallographic data collection. X-ray data collected at −9°C to 1.6-Å resolution conclusively reveal the structure of an intact enzyme–substrate complex.

MATERIALS AND METHODS

CPA and GY were purchased from Sigma and used without further purification. CPA was crystallized in space group P<sub>2</sub><sub>1</sub> (a = 51.60 Å, b = 60.27 Å, c = 47.25 Å, β = 97.27 Å), crystal habit elongated along the a axis, by dialysis of the enzyme (solubilized in 1.2 M LiCl/0.02 M Tris-HCl, pH 7.4) against 0.2 M LiCl/0.02 M Tris-HCl, pH 7.4, at 4°C. The CPA crystals were slightly crosslinked (19) with 0.2% (vol/vol) glutaraldehyde for 6 hr following transfer to 0.1 M LiCl/0.02 M barbital-LiOH, pH 7.5. Crystals were mounted in quartz capillaries that, in turn, were fastened to a diffractometer.

FIG. 1. The slowly hydrolyzed CPA substrate glycyll-t-tyrosine. This compound is a competitive inhibitor of more rapidly cleaved substrates; although it is depicted as the neutral zwitterion, both chemical studies and the results of the current crystallographic investigation suggest that it binds to carboxypeptidase A as the anion—i.e., with the terminal amino group in the free base form.

Abbreviations: CPA, carboxypeptidase A<sub>a</sub>; GY, glycyll-t-tyrosine; Me<sub>2</sub>SO, dimethyl sulfoxide.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
mounted yoke adapted from the design of Wyckoff and colleagues (20). The crystal was secured within the capillary by pipe-cleaner fibrils and Sephadex. The use of Sephadex in the vicinity of the crystal facilitated initial optical alignment in the x-ray beam. Polyethylene tubing of the appropriate diameter was attached to each end of the yoke-bound capillary, allowing for a continuous flow of buffer solution past the immobilized enzyme crystal. The flow cell assembly is illustrated in Fig. 2.

A continuous stream of buffer solution (~5 ml/day) passed through the flow cell and was gradually changed, with decreasing temperature, to a cryosolvent mixture consisting of 40% (vol/vol) dimethyl sulfoxide (Me$_2$SO)/0.15 M LiCl/0.02 M cacodylic acid-LiOH, pH 7.4. This particular mixture was chosen for several reasons: (i) of several organic co-solvents tested, Me$_2$SO does not bring about physical or crystallographic disorder to CPA crystals, even at room temperature; (ii) aqueous Me$_2$SO solutions display a dielectric constant very near that of pure water; and (iii) the protonic activity of cacodylic acid buffer solutions is nearly insensitive to temperature variation. Fink and Petsko (21) have recently reviewed general methods in x-ray cryozenyology, and Douzou and colleagues (22) have compiled the physical and chemical constants of various mixed cryosolvent systems and the protonic activities of various buffer reagents as a function of temperature. An FTS Systems (Stone Ridge, NY) XR-85 air jet crystal cooler was adapted to the diffractometer-mounted Syntex (Nicolet; Madison, WI) LT-1 cold air stream delivery system. The temperature inside the flow cell was checked periodically with a Sensortek (Clifton, NJ) IT-23 0.009-inch copper/constantan thermocouple, which easily fits inside the 0.7-mm quartz capillary.

After the system was equilibrated at ~25°C, intensity data were collected from one crystal to a resolution of 2.0 Å and processed as described (12). Data collection was performed on a Syntex (Nicolet) P2i, four-circle automated diffractometer using procedures described elsewhere (12). Initial difference electron-density maps calculated against the native enzyme revealed the presence of a zinc-bound Me$_2$SO molecule that had displaced the zinc-bound water of the native enzyme. No other molecules of Me$_2$SO were observed to bind to the enzyme or displace enzyme-bound solvent water molecules. Building of Me$_2$SO into the maps was done on an Evans and Sutherland PS300 interfaced with a VAX 11/780, with graphics software (FRODO) developed by Jones (23). Reciprocal space least-squares refinement (24) converged smoothly to a final crystallographic R factor of 0.162 at 2.0-Å resolution, where $R = \Sigma |F_o| - |F_c|/\Sigma |F_o|$, and $|F_o|$ and $|F_c|$ are observed and calculated structure factors, respectively.

Another crystal was mounted in the flow cell and equilibrated at ~9°C in a cryosolvent mixture consisting of 20% (vol/vol) Me$_2$SO/0.15 M LiCl/0.02 M cacodylic acid-LiOH, pH 7.4. This solution was then gradually changed to one of similar composition plus 0.1 M GY. This mixture was allowed to flow past the immobilized crystal for 1 day before data collection. It must be stressed that at no time was the crystal exposed to the substrate at higher temperatures; all "crystal soaking" was performed at ~9°C within the diffractometer-mounted flow cell prior to data collection. Thus, the hydrolysis reaction, albeit slow for this particular substrate, was significantly hindered. Additionally, the continuous flow of fresh GY past the crystal throughout data collection helped to ultimately ensure the observation of an intact complex: in case a molecule of GY was cleaved, another one could come along and take its place, assuming that substrate and products would be free to diffuse through the solvent channels in the crystal.

Intensity data were collected from two crystals under these conditions to a resolution of 1.6 Å. These data were processed as described (12) and used in the calculation of difference electron-density maps against the native enzyme. Initial maps revealed an intact CPA-GY complex; the GY molecule was therefore built into the difference density, and the enzyme-substrate model was refined against the observed data. Crystallographic refinement converged to a final R factor of 0.178 at 1.6-Å resolution. A difference electron density map calculated with Fourier coefficients $|F_o| - |F_c|$ and phases calculated from the final model coordinates showed rms residual electron density ($\sigma$) of 0.07 e/Å$^2$; the highest peaks in the vicinity of the active site were just under 3$\sigma$. No peaks corresponding to any residual zinc-bound Me$_2$SO were observed in this map. An rms error in atomic positions was estimated to be about 0.2 Å based on relationships derived by Luzzati (25), which attribute all discrepancies between $|F_o|$ and $|F_c|$ to errors in atomic coordinates. As such, the error of ca. 0.2 Å probably represents an upper limit, since discrepancies between observed and calculated structure factors can also be attributed to a host of other experimental conditions, as well as methods of data collection and reduction.

RESULTS AND DISCUSSION

The structure of CPA in the unique environment afforded by the aqueous/organic cryosolvent and the subzero temperature is remarkably similar to that of the native enzyme. Based upon data collected at ~25°C in a cryosolvent mixture consisting of 40% Me$_2$SO, only one significant difference is observed: a molecule of Me$_2$SO coordinates to the zinc ion of the active site under these conditions, and it displaces the zinc-bound water of the native enzyme in doing so. No Me$_2$SO molecules are observed to displace the water molecules normally residing in the hydrophobic pocket of the native enzyme. Furthermore, no appreciable temperature- or solvent-dependent conformational changes of enzyme residues are observed. A difference electron-density map of the
The CPA–Me₂SO complex is presented in Fig. 3. The prominent feature of this map is a very dense peak corresponding to the zinc-bound Me₂SO molecule, with a zinc–oxygen distance of 2.2 Å. This distance is in agreement with that of the Me₂SO–zinc interaction of horse liver alcohol dehydrogenase in its Me₂SO complex with the NADH holoenzyme (26). It should be noted that the distance of the Me₂SO sulfur from zinc, 3.4 Å, and also the orientation of the pyramidal-like Me₂SO so that the lone electron pair of sulfur is directed away from the active site, preclude a zinc–sulfur bond. Also, no changes are observed in the conformations of enzyme residues serving as zinc ligands, nor does the zinc itself move relative to these residues [the movement of zinc upon the binding of larger inhibitors has been discussed (15)]. The mechanism of CPA inhibition by Me₂SO likely involves Me₂SO coordination to zinc with resultant exclusion of the catalytically important water molecule from access to the reactive carbonyl of the substrate.

When the cryosolvent system is altered to 20% (vol/vol) Me₂SO and 0.1 M GY at -9°C, GY is observed bound to the active site intact. Furthermore, no trace of residual zinc-bound solvent remains. A difference electron-density map of this enzyme–substrate complex calculated at 1.6-Å resolution is presented in Fig. 4. The observed planarity of the difference electron density of the glycine and the amide linkage support this conclusion—if any zinc-bound Me₂SO or water remained at appreciable occupancy, the difference density of this portion of the substrate would be punctuated by “bumps” rather than appearing consistent and flat. The coordination position of Me₂SO is different from that of GY; indeed, the two carbon atoms and the sulfur atom of Me₂SO would lie outside the electron density outlining the substrate in Fig. 4. Additionally, there is no evidence for residual Me₂SO in the final difference electron-density map calculated with Fourier coefficients |Fo| - |Fc| and phases calculated from the final model. No peaks in this map are evident near the coordination position of Me₂SO even at the low level of 1.5σ. Hence, we conclude that it is solely GY that is observed bound to the active site of CPA, and the substrate is bound in an intact, unhydrolyzed condition. The glycine portion of GY spans the zinc ion and chelates in bidentate fashion through both the amide carbonyl oxygen and the terminal amino nitrogen (the O–Zn distance is 2.3 Å, and the N–Zn distance is 1.9 Å). It appears that it is solely the amino nitrogen that is responsible for displacing the formerly zinc-bound solvent molecule; i.e., the coordination position of the carbonyl oxygen would not require the displacement of the zinc-bound water of the native enzyme. The current results differ somewhat from those observed in the lower-resolution study (6), where the amino terminus appeared to be disordered between zinc coordination and hydrogen bonding to glutamate-270; its position is now better defined given the higher resolution and greater occupancy (60% versus 30%) of the bound GY molecule. Additionally, the zinc ion is observed to move about 0.3 Å in the direction toward arginine-127. Such movement of the metal ion upon its interaction with enzyme-bound inhibitors has been discussed (15). A stereoview of the CPA–GY complex is presented in Fig. 5.

It is not surprising to find that the amino terminus of the bound dipeptide is not observed to exist as a quaternary ammonium cation; given all the other strong binding interactions between CPA and GY, the amino terminus finds itself in the very electropositive environment of the zinc ion and nearby arginine-127. Although arginine-127 is not observed to hydrogen-bond to the bound substrate, its close proximity to the bound substrate would certainly favor the free base form of the amino terminus by electrostatic considerations. Additionally, Yanari and Mitz (27) have concluded, on the basis of the pH dependence of GY inhibition kinetics, that GY binds to the enzyme as the anionic species—i.e., with the amino terminus as the free base. On the basis of electron paramagnetic resonance studies of the Co³⁺–substituted enzyme and its complex with GY prepared in the presence of ¹⁸O-enriched water, Kuo and Makinen (28) have concluded that the metal-bound water/hydroxide is displaced upon the binding of GY. This proposition is entirely consistent with the observed mode of binding in the CPA–GY complex. It is likely that all dipeptide substrates of CPA bearing an unprotected amino terminus are poor substrates due to such favorable bidentate chelation to the zinc ion of the active site. This chelation may significantly compete with the catalytic event(s). Since the hydrolytically important zinc-bound water/hydroxide has been excluded from the active site in

**Fig. 3.** Portion of a difference electron-density map calculated with Fourier coefficients |Fø| - |Fc| and phases calculated from the refined CPA–Me₂SO model coordinates less the nonhydrogen atoms of the solvent molecule. Pertinent enzyme residues (glutamate-270, glutamate-72, histidine-69, histidine-196) are indicated by their sequence numbers; the zinc ion appears as a small black sphere. Active site water molecules appear as crosses. The oxygen of the Me₂SO molecule is 2.2 Å from the zinc ion; both the distance of the sulfur from zinc (3.4 Å) and its orientation so that its lone electron pair faces the opposite direction from zinc preclude a zinc–sulfur bond.

**Fig. 4.** Portion of a difference electron-density map calculated with Fourier coefficients |Fø| - |Fc| and phases calculated from the refined CPA–substrate model coordinates less the nonhydrogen atoms of the substrate. Pertinent enzyme residues (glutamate-270, glutamate-72, histidine-69, histidine-196, tyrosine-248) are indicated by their sequence numbers; the zinc ion appears as a small black sphere. Note the continuous and bifurcated electron density spanning the zinc ion. The smoothness and consistency of this density suggest that the intact substrate is bound. Indeed, the density corresponding to the peptide unit of GY is quite planar, indicative of the planar peptide linkage of the intact dipeptide. The amino terminus of GY is 1.9 Å, and its carbonyl oxygen is 2.3 Å from the zinc ion.
the CPA–GY complex, it is likely that the binding mode observed for GY in this enzyme–substrate complex represents a nonproductive—i.e., catalytically inactive—state.

Other binding interactions observed are typical of those found in CPA–inhibitor complexes (8–15): the hydroxybenzyl group of tyrosine resides in the hydrophobic pocket, or specificity pocket, of the S1′ subsite, and the terminal carboxylate is salt linked with the guanidinium moiety of arginine-145. The phenolic residue of tyrosine-248 is in the “down” conformation, and its phenolic oxygen is 2.8 Å away from one of the terminal carboxylate oxygens of GY. This interaction clearly favors the un-ionized state of tyrosine-248. The substrate lies in a slightly different position from that determined in the previous low-resolution study (6); an rms difference in GY atomic positions was calculated to be 0.6 Å between the prior unrefined coordinates and those obtained from the current model refined at high resolution. Arginine-127, although having moved toward the bound substrate relative to its position in the native enzyme, is not within hydrogen-bonding distance to the substrate carbonyl oxygen (the distance to the nearest guanidinium nitrogen is 4.3 Å). This residue is observed to hydrogen-bond to the GY carbonyl in its complex with the apoenzyme (29), as well as hydrogen-bond to the carbonyl of a substrate analogue in a recent x-ray crystallographic study (13). Finally, the γ-carboxylate of glutamate-270 is within hydrogen-bonding distance to the amide nitrogen of the peptide linkage, a feature that supports proposals involving glutamate-270 as a proton donor (12, 30). Selected distances of enzyme–substrate interactions are recorded in Table 1.

A role for the positively charged guanidinium moieties of the three arginine residues lining the active site cleft of CPA has been proposed (31, 32) in view of the precatalytic association mechanism of substrate with enzyme. More recently, the direct involvement of arginine-127 with bound inhibitors prompted speculation as to its possible involvement in actual catalytic steps mediated by the enzyme (1, 12, 14, 15). Such a catalytic role was initially considered to involve the shift of the carboxyl of the putative anhydride intermediate from zinc coordination to hydrogen bonding with arginine-127 in order to facilitate deacetylation by zinc-bound water/hydroxide (1). However, later structural studies that modeled the intermediate of the promoted attack of water directly at the substrate carbonyl prompted speculation as to the role of arginine-127 in the initial binding of the scissile substrate carbonyl before any hydrolytic steps (12) and then to the stabilization of the oxyanion tetrahedral intermediate (14, 15). Such an effect, of course, could be achieved through both a hydrogen bond and/or electrostatic stabilization. Also, the initial step involving the polarization of the substrate carbonyl, previously considered to be the exclusive role of the metal ion at the active site, could be shared with the positively charged side chain of arginine-127. Indeed, both the zinc ion and arginine-127 are oriented in such a way as to provide optimal polarization of the scissile carbonyl. An unidentified arginine residue is observed to be of catalytic significance in the proteolytic mechanism, but not the esterolytic mechanism, based upon chemical modification studies (33). A sequence of binding modes, with different rate-determining steps for peptides and esters involving arginine-127, may account for the observed kinetic anomalies, although it is not clear whether this chemical modification occurs at arginine-71, -127, or -145.

The probably nonproductive mode of binding observed in the CPA–GY complex does not involve a hydrogen-bonding interaction with arginine-127. It is quite possible that the observed bidentate GY may be bound solely because of the favorable chelate interaction provided by the free amino terminus; the involvement of arginine-127 in polarizing the substrate carbonyl is not precluded in a more catalytically favorable productive mode of binding. This mode would involve simply the rotation of the peptide unit of GY by about 30–40° toward arginine-127. Similar roles for potentially electropositive histidine residues in the active sites of the related zinc proteases thermolysin (histidine-231) and peptidase G (histidine-190) are also feasible on the basis of the structure of a thermolysin–inhibitor complex (34) or on the basis of preliminary x-ray crystallographic data on the active-site geometry and proximal residues to the active-site zinc ion of peptidase G (35). Nevertheless, if a metal-bound hydroxide is the catalytically responsible nucleophile, then the coordination of a substrate carbonyl would lessen the nucleophilic capability of such a metal-bound species. That is, in the case of CPA, the zinc- and/or glutamate-270-promoted hydroxide

Table 1. Selected enzyme–substrate distances

<table>
<thead>
<tr>
<th>Protein atom</th>
<th>Substrate atom</th>
<th>Distance, Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg-145 N1</td>
<td>Carboxylate O1</td>
<td>3.3†</td>
</tr>
<tr>
<td>Arg-145 N2</td>
<td>Carboxylate O2</td>
<td>3.0†</td>
</tr>
<tr>
<td>Tyr-248 phenolic O</td>
<td>Carboxylate O1</td>
<td>2.8†</td>
</tr>
<tr>
<td>Glu-270 O1</td>
<td>Amino N</td>
<td>3.6</td>
</tr>
<tr>
<td>Glu-270 O2</td>
<td>Amino N</td>
<td>4.6</td>
</tr>
<tr>
<td>Glu-270 O1</td>
<td>Carboxyl C</td>
<td>3.6</td>
</tr>
<tr>
<td>Glu-270 O2</td>
<td>Amide N</td>
<td>2.6†</td>
</tr>
<tr>
<td>Ser-197 carbonyl O</td>
<td>Amino N</td>
<td>2.8†</td>
</tr>
<tr>
<td>Zinc</td>
<td>Amino N</td>
<td>1.9</td>
</tr>
<tr>
<td>Arg-127 N2</td>
<td>Carboxyl O</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>Carboxyl O</td>
<td>4.3</td>
</tr>
</tbody>
</table>

†Possible hydrogen bond.
would be a more potent nucleophile if the substrate carbonyl initially coordinated to arginine-127, with subsequent shift of the oxyanion tetrahedral intermediate from arginine-127 to zinc. Similar electrophilic roles emerge for histidine-231 of thermolysin and histidine-190 of pepitidase G.

The assignment of the catalytically responsible nucleophile is a task more difficult than that of assigning the identities to active site electrophiles. Glutamate-270 has been proposed (7) to attack the scissile carbonyl of the substrate, with subsequent formation of a covalent mixed anhydride intermediate, which then undergoes hydrolysis to yield products. This mechanistic proposal is supported, in part, by the observation of a five-coordinate metal ion in low-temperature electron paramagnetic resonance spectroscopic experiments with Co–CPA (28). The five metal ligands were presumed to be glutamate-72, histidine-69, histidine-196, a zinc-bound hydroxide, and a carbonyl of an anhydride intermediate. However, such five-coordination is also structurally consistent with intermediates encountered in other mechanistic pathways. If a promoted-water mechanism were followed, a five-coordinate metal complex would result if the tetrahedral intermediate were to become chelated in bidentate fashion to the metal ion. A structural model of such a possibility is illustrated in the complex of CPA with a hydrated aldehyde inhibitor (12), where both oxygens of the tetrahedral gem-diol moiety straddle the zinc ion. Furthermore, a zinc-promoted water would be a much more potent nucleophile than would a carboxylate-promoted water. Glutamate-270, i.e., the carboxylate, may play an important role in this case as the mediator of proton transfer to the leaving group of the collapsing tetrahedral intermediate. As such, glutamate-270 may act in concert with the zinc ion in the promotion of the hydrolytically important zinc-bound water/hydroxide. The resultant tetrahedral intermediate would be exceptionally stabilized if it were coordinated in bidentate fashion to the metal ion, and even more so if the positively charged guanidinium moiety of arginine-127 were to participate in the stabilization mechanism. Low-temperature studies will help to elucidate the structure of the putative five-coordinate intermediate of CPA-catalyzed esterolytic, and quite possibly proteolytic, reactions.

We thank Mr. Peter David for his assistance in preparing the crystals as well as his helpful discussions during the course of this work. The support of W. R. Grace & Co. is gratefully acknowledged; additionally, D.W.C. thanks AT&T Bell Laboratories for a doctoral fellowship. Finally, we thank the National Institutes of Health for Grant GM 06920 in support of this research, and the National Science Foundation for Grant PCM-77-11398 in support of the computational facility.