Carboxypeptidase A

DAVID W. CHRISTIANSON*

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323

WILLIAM N. LIPSCOMB

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

Scientific and technological advances brought about in the recent part of this decade have provided great insight regarding the catalytic mechanism of the prototypical zinc-requiring protease carboxypeptidase A (CPA). Importantly, some of our ideas regarding the general theory of enzyme catalysis have been inspired and developed through studies of this enzyme. Just as importantly, CPA serves as a model for many zinc proteases of unknown structure which serve as pharmaceutical targets, such as angiotensin-converting enzyme, enkephalinase, and collagenase. Although the CPA mechanism has been the subject of much debate spanning several decades, recent success in X-ray crystallography, site-directed mutagenesis, and rapid spectrokinetics is leading to an unparalleled understanding of the catalytic mechanism. This Account is set within the context of our high resolution X-ray crystallographic studies of enzyme–substrate and enzyme–inhibitor complexes. We shall review the structural features of the enzyme active site and aspects concerning the chemistry of zinc. Importantly, we propose a catalytic role for the active site zinc ion which is different from that historically established for CPA. Finally, we set forth a model for the mechanism of peptide hydrolysis which reflects our interpretation of the most recent structural, genetic, and chemical information.

The Enzyme: Subsite Structure and Specificity

Bovine carboxypeptidase A (peptidyl-L-amino-acid hydrolase, EC 3.4.17.1) is a metalloexopeptidase of molecular weight 34 472 containing one zinc ion bound to a single polypeptide chain of 307 amino acids.[1,2] Its biological function is the hydrolysis of C-terminal amino acids from polypeptide substrates, and it exhibits a preference toward those substrates possessing large, hydrophobic C-terminal side chains such as phenylalanine (Figure 1). Important residues for catalysis and binding are Glu-270, Arg-71, Arg-127, Asn-144, Arg-145, Tyr-248, Zn^{2+}, and the zinc-bound water molecule (Figure 2). The substrate binding region has been observed kinetically to extend over five amino acids of

the substrate, and the hydrolysis of longer substrates is not subject to large deviations from Michaelis–Menten kinetics.[3] The most rapidly cleaved substrates display values of $k_{cat.}$ near 10^2 s^{-1} for peptides and 10^3 s^{-1} for esters. The structure of the native enzyme has been determined by X-ray crystallographic methods to a resolution of 1.54 Å,[4] and additional X-ray experiments have identified four out of five subsites on the enzyme in the product complex with the 39 amino acid inhibitor from the potato.[5] Recent high resolution crystallographic experiments have shown that typical substrate analogues,[6] transition-state analogues,[7,8] reaction-coordinate analogues,[9–12] slowly hydrolyzed substrates,[13] or a substrate–product pair also representing hydrolysis products[14] each bind to the active site of CPA and reveal common features of enzyme–ligand specificity (Figure 3). Significantly, the enzyme endures a large conformational change of Tyr-248 and its associated polypeptide chain (particularly within a loop region consisting of about four to five residues) upon the binding of ligands to the active site (Figure 4). The phenolic hydroxyl of this residue can donate a hydrogen bond to the terminal carboxylate of a bound substrate, and it can accept an additional hydrogen bond from the amide proton of the penultimate peptide linkage of longer peptide substrates. These interactions, since they require substantial movement on the part of the enzyme in order to fully accommodate the binding of inhibitors or substrates, provided the first structural perspective of Koshland's "induced fit" hypothesis.[15]

David W. Christianson did his undergraduate and graduate work at Harvard, receiving his A.B. in 1983 and his Ph.D. in 1987. He is currently Assistant Professor of Chemistry at the University of Pennsylvania, and he is sometimes found off the northeastern coastal states as captain of the sloop PEPTIDE.

William N. Lipscomb did undergraduate work at the University of Kentucky and in 1946 took his Ph.D. from the California Institute of Technology. He was Professor of Chemistry at the University of Minnesota before moving to Harvard in 1959. He is currently Abbot and James Lawrence Professor of Chemistry at Harvard, and in 1976 he received the Nobel Prize in Chemistry. Prof. Lipscomb is also a Kentucky Colonel.
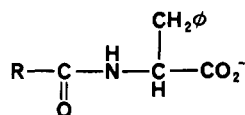
(1) Quiocho, F. A.; Lipscomb, W. N. Adv. Protein Chem. 1971, 25, 1–78.
(2) Hartsuck, J. A.; Lipscomb, W. N. In The Enzymes, 3rd ed.; Boyer, P.; Ed.; Academic: New York, 1971; Vol. 3, pp 1–56.
(3) Abramowitz, N.; Schechter, I.; Berger, A. Biochem. Biophys. Res. Commun. 1967, 29, 862–867.
(4) Rees, D. C.; Lewis, M.; Lipscomb, W. N. J. Mol. Biol. 1983, 168, 367–387.
(5) Rees, D. C.; Lipscomb, W. N. J. Mol. Biol. 1982, 160, 475–498.
(6) Christianson, D. W.; Kuo, L. C.; Lipscomb, W. N. J. Am. Chem. Soc. 1985, 107, 8281–8283.
(7) Christianson, D. W.; Lipscomb, W. N. J. Am. Chem. Soc. 1986, 108, 545–546.
(8) Christianson, D. W.; Lipscomb, W. N. J. Am. Chem. Soc. 1988, 110, 5560–5565.
(9) Christianson, D. W.; David, P. R.; Lipscomb, W. N. Proc. Natl. Acad. Sci. U.S.A. 1987, 84, 1512–1515.
(10) Christianson, D. W.; Lipscomb, W. N. Proc. Natl. Acad. Sci. U.S.A. 1985, 82, 6840–6844.
(11) Christianson, D. W.; Lipscomb, W. N. J. Am. Chem. Soc. 1986, 108, 4998–5003.
(12) Shoham, G.; Christianson, D. W.; Oren, D. A. Proc. Natl. Acad. Sci. U.S.A. 1988, 85, 684–688.
(13) Christianson, D. W.; Lipscomb, W. N. Proc. Natl. Acad. Sci. U.S.A. 1986, 83, 7568–7572.
(14) Christianson, D. W.; Lipscomb, W. N. J. Am. Chem. Soc. 1987, 109, 5536–5538.

$$R—\underset{\underset{O}{\|}}{C}—\underset{\underset{}{H}}{N}—\underset{\underset{H}{}}{\overset{\overset{CH_2\phi}{|}}{C}}—CO_2^-$$

Figure 1. The typical C-terminal portion (with terminal phenylalanine) of a peptide CPA substrate (ϕ = phenyl, R = remainder of the peptide).

The S_1' subsite of the enzyme is a dead-end hydrophobic pocket, and Tyr-248 essentially "caps" this pocket once substrate or inhibitor is bound.[16] Specificity toward the C-terminal carboxylate of substrates is provided by hydrogen-bond interactions with a triad of residues: a salt link with the positively charged guanidinium moiety of Arg-145, a hydrogen bond donated from Tyr-248, and a hydrogen bond donated from the amide amino group of Asn-144. The latter two interactions display anti stereochemistry with respect to each carboxylate oxygen.[17]

The S_1 subsite of the enzyme is the locus of catalysis, and the catalytically obligatory zinc ion is liganded by enzyme residues Glu-72 (bidentate in the native enzyme), His-69, and His-196. Catalytically important residues Glu-270 and Arg-127 are on opposite sides of the zinc–water couple, and each of these four species lies roughly in a plane bisecting the active-site groove. The region that binds the P_1 side chain of substrates is a hydrophobic cleft adjacent to, and not to be confused with, the S_1' hydrophobic pocket. Although P_1 valine was observed to occupy the S_1 hydrophobic cleft in the complex with the cleaved 39 amino acid potato inhibitor,[5] the enzyme displays preference toward substrates with P_1 aromatic side chains.[18,19] Recent X-ray crystallographic studies show that benzyl side chains in this location exploit favorable, weakly polar interactions with aromatic enzyme residues, particularly Tyr-198.[12,14] Such "edge-to-face" interactions have been discussed recently by Burley and Petsko,[20,21] and these interactions may account, in part, for the observed specificity phenomena in the S_1 subsite. Enzyme subsites beyond S_1 do not comprise a well-defined active-site groove, and no hydrogen-bond contacts for P_2 and P_3 residues are observed in the CPA–potato inhibitor complex.[5] An S_4 subsite has been detected in substrate specificity experiments,[3] but its occupancy has not yet been observed in X-ray crystallographic studies.

Catalytic Mechanism(s)

The hydrolytic mechanism of CPA has been the topic of much debate, particularly since its structure was determined by X-ray crystallography. There are currently two basic hydrolytic pathways under consideration: (1) a *nucleophilic pathway* involving a covalent

acyl enzyme intermediate with active-site base Glu-270 (hence, an anhydride), subsequently labile to hydrolysis; or (2) a *promoted-water pathway* involving the attack of a water molecule, promoted by zinc and assisted by Glu-270, directly at the scissile peptide linkage of the substrate. Historically, a precatalytic zinc–carbonyl interaction with substrate has been invoked in each pathway.

Chemical and kinetic support for the nuclephilic pathway has been relatively recent, yet sometimes ambiguous. Makinen's investigations of certain ester substrates resulted in the detection of an accumulating intermediate, presumably the acyl enzyme with Glu-270, at subzero temperature.[22,23] However, this result was questioned in Kaiser's laboratory by the results of a parallel resonance Raman cryospectroscopic study.[24] Later, Suh and colleagues demonstrated the accumulation of an intermediate, presumed to be the acyl enzyme, at -2 °C in the hydrolysis of different ester substrates.[25] The designation of acyl enzyme was made in each study without confirmation by chemical trapping experiments. In the absence of such experiments, then, the observed intermediate is not necessarily distinguishable from an enzyme–products or enzyme–substrate product complex where facile product (or substrate) diffusion is hindered by substrate concentration, cryosolvent viscosity, and/or low temperature.[14] Sander and Witzel have provided the only direct, but marginal, chemical evidence for a mixed anhydride intermediate—quantitative recovery of the adduct with the trapping reagent was not achieved.[26] Breslow and Wernick demonstrated in a landmark ^{18}O isotope labeling experiment that an acyl enzyme intermediate is not involved in the synthesis or hydrolysis of typical peptide substrates, but these investigators did not rule out an intermediate in esterolysis.[27,28] Although different mechanisms may be favored for proteolysis and esterolysis, it has been suggested that proteolysis and esterolysis proceed via similar mechanisms but with different rate-determining steps.[29]

Low-temperature studies from Auld's laboratory are said to be in accord with the promoted-water pathway for both proteolysis and esterolysis. Auld and colleagues have achieved short-term isolation of metastable enzyme-bound intermediates that do not react with trapping reagents.[30] These investigators found that peptides and esters are hydrolyzed via a simplified reaction scheme $E + S \rightleftharpoons ES_1 \rightleftharpoons ES_2 \rightleftharpoons E + P$ through subzero radiationless energy transfer experiments on Zn^{2+}–CPA and Co^{2+}–CPA with N-dansylated substrates. Results from subsequent cryospectroscopic and cryokinetic studies suggest that peptides and esters

(15) Koshland, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.

(16) The amino acid residues of the proteolytic substrate (or product) are conventionally designated as $P_n(')$; the scissile peptide bond connects the P_1 and P_1' residues of the substrate. C-Terminal residues of the substrate, adjacent to the scissile peptide linkage, are designated P_1' through P_n', and the N-terminal residues are designated P_1 through P_n. Enzyme subsites complementary to these substrate residues are similarly designated as $S_n(')$, where the scissile peptide linkage spans enzyme subsites S_1 and S_1' in the enzyme–substrate complex; see: Schecter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.

(17) Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169–176.

(18) Smith, E. L. *J. Biol. Chem.* **1948**, *175*, 39–47.

(19) Schechter, I.; Berger, A. *Biochemistry* **1966**, *5*, 3371–3375.

(20) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23–28.

(21) Burley, S. K.; Petsko, G. A. *J. Am. Chem. Soc.* **1986**, *108*, 7995–8001.

(22) Makinen, M. W.; Kuo, L. C.; Dymowski, J. J.; Jaffer, S. *J. Biol. Chem.* **1979**, *254*, 356–366.

(23) Kuo, L. C.; Makinen, M. W. *J. Biol. Chem.* **1982**, *257*, 24–27.

(24) Hoffman, S. J.; Chu, S. S.-T.; Lee, H.-H.; Kaiser, E. T.; Carey, P. R. *J. Am. Chem. Soc.* **1983**, *105*, 6971–6973.

(25) Suh, J.; Cho, W.; Chung, S. *J. Am. Chem. Soc.* **1985**, *107*, 4530–4535.

(26) Sander, M. E.; Witzel, H. *Biochem. Biophys. Res. Commun.* **1985**, *132*, 681–687.

(27) Breslow, R.; Wernick, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 1303–1307.

(28) Breslow, R.; Wernick, D. L. *J. Am. Chem. Soc.* **1976**, *98*, 259–261.

(29) Cleland, W. W. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1977**, *45*, 273–387.

(30) Auld, D. S.; Galdes, A.; Geoghegan, K. F.; Holmquist, B.; Martinelli, R. A.; Vallee, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 5041–5045.
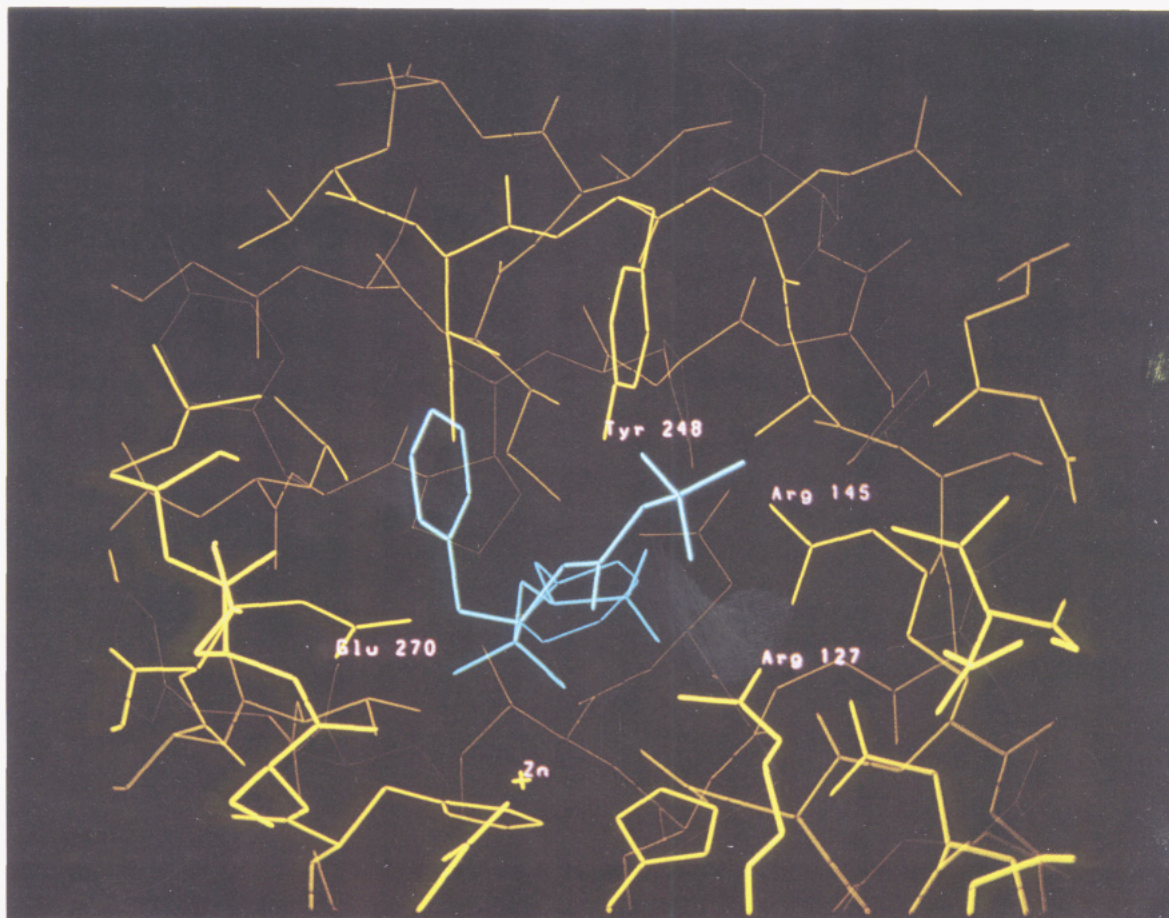
Figure 2. The active site of CPA with a bound inhibitor as depicted by conventional stick-figure representation. Important residues for binding and catalysis are labeled by their sequence numbers: Glu-270, Arg-127, Arg-145, Tyr-248, Zn^{2+}. The inhibitor, colored light blue, is the hydrated ketonic analogue (Figure 7D) of the peptide substrate *tert*-butoxycarbonyl-L-phenylalanyl-L-phenylalanine.
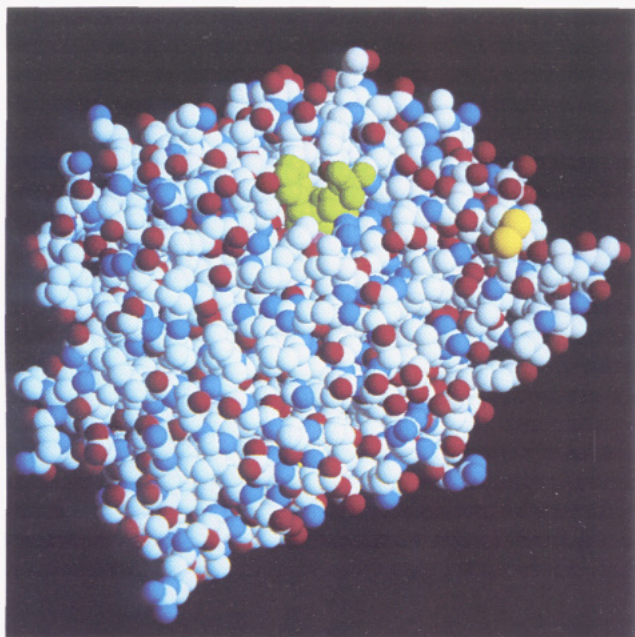

Figure 3. Space-filling representation of the complex found in Figure 2. Carbon, nitrogen, oxygen, and sulfur atoms of the enzyme are colored white, blue, red, and yellow, respectively; the zinc ion (barely observable) is colored magenta. Inhibitor atoms are colored uniformly green.

coordinate to the metal ion prior to the rate-determining step, and that respective chemical intermediates are structurally distinct.[31–33] Hydrolysis of the ester occurs prior to the rate-determining step so that the esterolytic intermediate ES_2 is actually a product complex EP_1P_2.[33] In contrast, peptide hydrolysis occurs during (or after) the rate-determining step. There are proposals for remarkably different binding modes for peptides and esters,[34,35] but we suggest that these binding modes may be distinguished simply by different roles for zinc and Arg-127.

In the promoted-water pathway, the zinc ion of CPA is a classical electrophilic catalyst: it provides electrostatic stabilization for negatively charged intermediates formed during the course of a hydrolytic reaction. Zinc will likewise stabilize the fractional negative charges formed in the transition states flanking such intermediates.[36] Additionally, a precatalytic zinc–carbonyl interaction (which would polarize the peptide bond and make it more susceptible to nucleophilic attack) has been invoked historically. Support for this catalytic ion–dipole interaction has been provided by studies of metal ion catalysis in the hydrolysis of model com-

(31) Geoghegan, K. F.; Galdes, A.; Marinelli, R. A.; Holmquist, B.; Auld, D. S.; Vallee, B. L. *Biochemistry* **1983**, *22*, 2255–2262.

(32) Geoghegan, K. F.; Galdes, A.; Hanson, G.; Holmquist, B.; Auld, D. S.; Vallee, B. L. *Biochemistry* **1986**, *25*, 4669–4674.

(33) Galdes, A.; Auld, D. S.; Vallee, B. L. *Biochemistry* **1986**, *25*, 646–651.

(34) Vallee, B. L.; Galdes, A.; Auld, D. S.; Riordan, J. F. In *Metal Ions in Biology*; Spiro, T. G., Ed.; Wiley: New York, 1983; Vol. 5, pp 25–75.

(35) Vallee, B. L.; Galdes, A. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1984**, *56*, 283–430.

(36) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334–338.
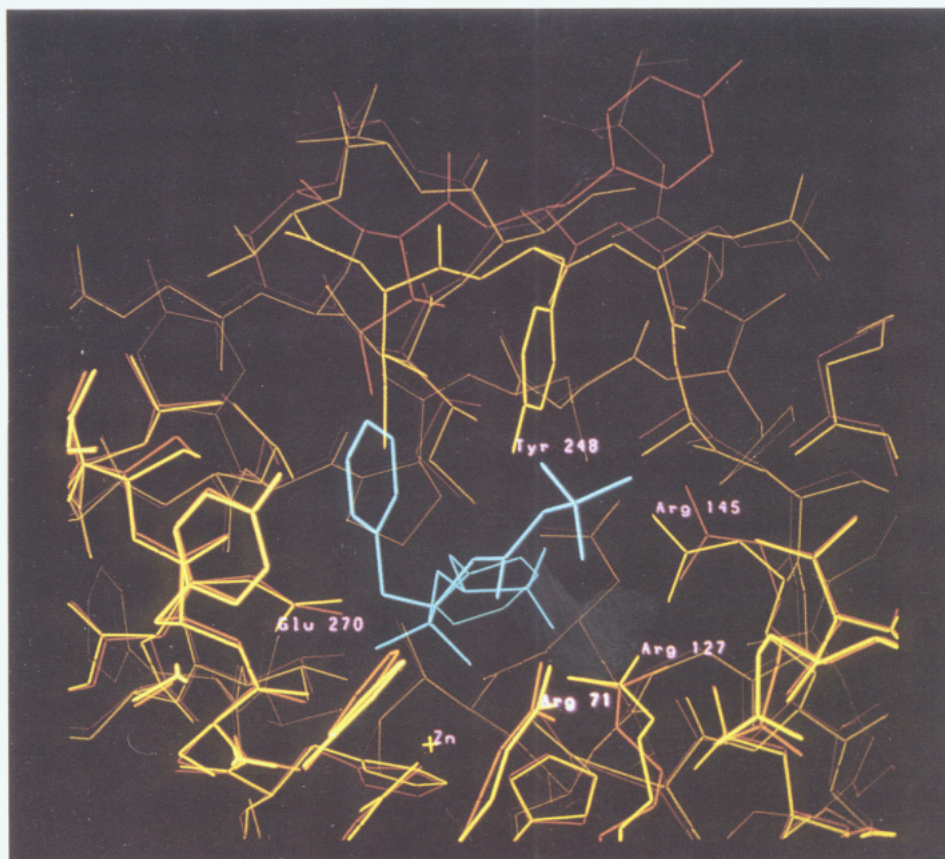
Figure 4. Stick-figure representation of the complex found in Figure 2, superimposed on the coordinates of the native enzyme. The inhibitor is shown in light blue and the complexed enzyme in yellow; the native enzyme is shown in red. Note the significant conformational change of Tyr-248, plus the backbone movement in the loop region including Tyr-248.

pounds.[37,38] However, in these studies, the ion–dipole interaction is facilitated by—indeed, perhaps is only a consequence of—a chelate interaction involving the labile carbonyl and some other suitably located donor atom of the model compound. For example, Fife and Przystas recently found that such a chelate effect is required for the hydrolysis of *N*-acylimidazoles in solution.[38] The scissile carbonyl of typical peptide substrates may not tend toward zinc coordination with CPA if a chelate interaction is inaccessible.

The coordination of the scissile carbonyl exclusively to zinc would raise the pK_a of the zinc-bound water molecule, perhaps sufficiently to hinder attack at a stubborn peptide carbonyl. The results of Fife and Przystas[38] do indicate that zinc promotes a water molecule as a potent nucleophile in both proteolytic and esterolytic reactions. Groves and Olson[39] have demonstrated that a zinc-coordinated water can display a pK_a as low as 7, i.e., be hydroxide-like or nucleophilic at neutral pH. In CPA, the nucleophilicity of metal-bound water is further enhanced by a hydrogen bond with general base/general acid Glu-270. The zinc ion may be only partially involved, if at all, in the polarization of the substrate carbonyl prior to catalysis. The mechanistic role of zinc may instead be to promote a water molecule, with the assistance of Glu-270, to attack a peptide bond polarized by *Arg-127*.

Enzyme–Ligand Structures: Implications toward Mechanism

The pseudosubstrate glycyl-L-tyrosine binds to CPA at –9 °C as the unequivocally intact peptide (contrary to the results of solid-state NMR experiments performed under different conditions[40]), but both the peptide carbonyl oxygen and the amino terminus (as the free base) coordinate to the active-site zinc ion in a chelate interaction (Figure 5).[13,41] This nonproductive binding mode is inaccessible for typical substrates with blocked P_1 amino groups, so the zinc–carbonyl interaction may be simply a consequence of the chelate effect. Hence, the observed zinc–carbonyl interaction cannot model a productive interaction in the mechanism of peptide hydrolysis. Glycyl-L-tyrosine is hydrolyzed nearly 5000 times more slowly than N-protected dipeptides,[1,42] and chemical studies have confirmed the competitive chelate interaction.[43] It might be said that the zinc–carbonyl interaction, as part of the chelate, actually hinders the prevailing hydrolyic mechanism.

A more productive representation of enzyme–substrate interaction may be found in the complex of CPA with the substrate analogue (–)-2-benzyl-3-(*p*-methoxybenzoyl)propanoic acid.[6,44,45] The intact carbonyl

(37) Woolley, P. *Nature* **1975**, *258*, 677–682.
(38) Fife, T. H.; Przystas, T. J. *J. Am. Chem. Soc.* **1986**, *108*, 4631–4636.
(39) Groves, J. T.; Olson, J. R. *Inorg. Chem.* **1985**, *24*, 2715–2717.
(40) Mackenzie, N. E.; Fagerness, P. E.; Scott, A. I. *J. Chem. Soc., Chem. Commun.* **1985**, 635–637.
(41) Rees, D. C.; Lewis, M.; Honzatko, R. B.; Lipscomb, W. N.; Hardman, K. D. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 3408–3412.
(42) Izumaya, N.; Uchio, H. *J. Biochem. (Tokyo)* **1959**, *46*, 235–245.
(43) Yanari, S.; Mitz, M. A. *J. Am. Chem. Soc.* **1956**, *79*, 1154–1158.
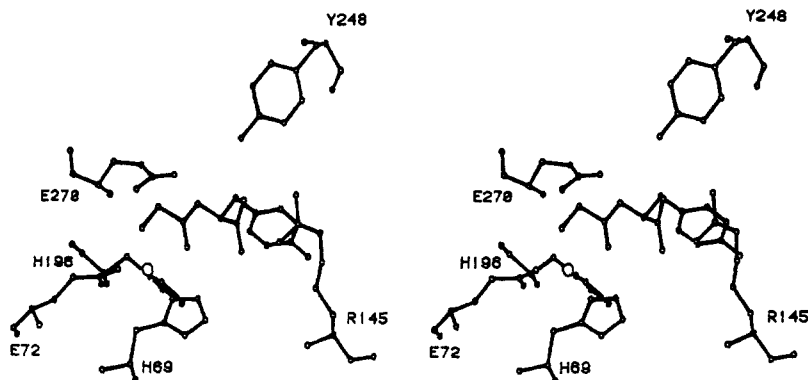(44) Suigimoto, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1978**, *100*, 7750–7751.

Figure 5. Stereoview of the complex between CPA and the slowly hydrolyzed substrate glycyl-L-tyrosine; note the chelate interaction between the scissile carbonyl, the terminal amino group, and zinc. Important active-site amino acids are indicated by one-letter abbreviations and sequence numbers.
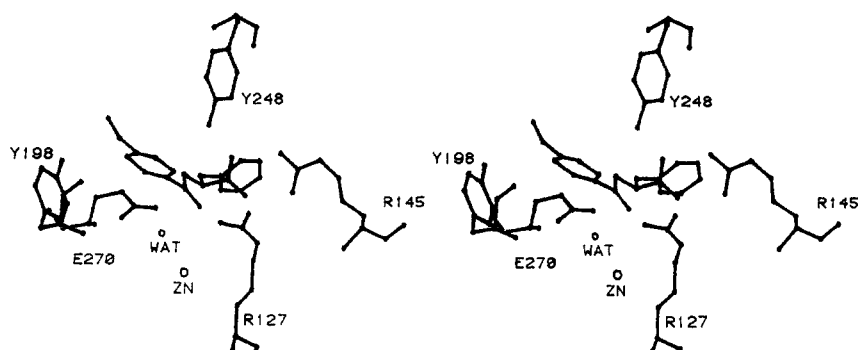

Figure 6. Stereoview of the complex between CPA and (-)-2-benzyl-3-(p-methoxybenzoyl)propanoic acid. The inhibitor binds with its carbonyl oxygen hydrogen bonded to Arg-127, and this binding mode may provide a model for the precatalytic binding of peptide substrates. Contrast this binding mode with that found for glycyl-L-tyrosine in Figure 5. Important active-site residues are indicated by one-letter abbreviations and sequence numbers.

of this ketone is hydrogen bonded to the guanidinium moiety of Arg-127, and the coordination polyhedron of the native zinc ion is not perturbed (Figure 6). The zinc-bound water molecule in this complex appears to be in good orientation for attack at the π^* orbital of the carbonyl system. We believe that this enzyme–inhibitor complex provides a model for the Michaelis complex of peptide substrates with CPA, and this binding mode may resolve questions arising from several chemical and kinetic experiments: (1) K_M values for peptides with various metallosubstituted enzymes do not vary, whereas those for ester substrates show considerable variation;[46] (2) peptides (but not esters) have been shown to bind to the apoenzyme through kinetic methods;[46] (3) X-ray crystallographic studies with the apoenzyme show that peptides bind to Arg-127;[47] (4) peptides, but not esters, have been shown to bind to Co³⁺-substituted CPA—since Co³⁺ is substitution-inert, these results indicate that the intact peptide carbonyl favors a nonmetal binding location in the precatalytic enzyme–substrate complex;[48] and (5) a yet-unidentified arginine residue is required for proteolysis but not esterolysis.[49] Thus, we propose that Arg-127 is the nonmetal location of the intact peptide carbonyl just prior to catalysis. Differences between proteolysis and esterolysis may be ascribed to different roles for zinc
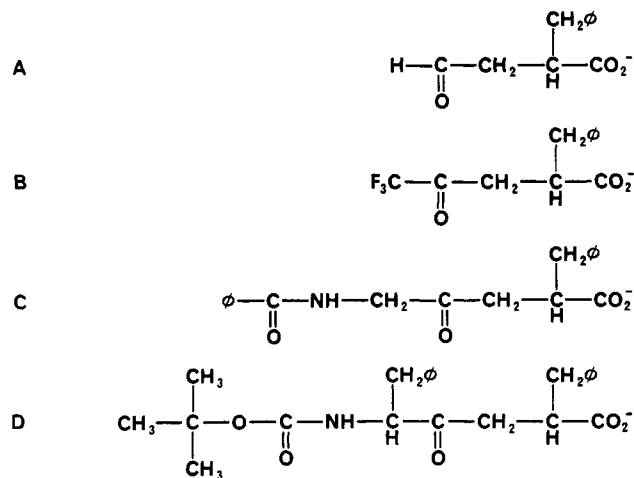

Figure 7. Reaction-coordinate analogues of the typical CPA substrate found in Figure 1: (A) 2-benzyl-3-formylpropanoic acid, $K_i = 4.8 \times 10^{-7}$ M (enantiomeric mixture);[51] (B) 2-benzyl-4-oxo-5,5,5-trifluoropentanoic acid, $K_i = 2/\times 10^{-7}$ M (enantiomeric mixture);[52] (C) 5-benzamido-2-benzyl-4-oxopentanoic acid, $K_i = 4.8 \times 10^{-5}$ M (enantiomeric mixture);[53] (D) N-(tert-butoxycarbonyl)-5-amino-2-benzyl-4-oxo-6-phenylhexanoic acid, $K_i = 6.7 \times 10^{-7}$ M (diastereomeric mixture).[54]

and Arg-127, and not to drastically different substrate binding modes.[34,35] Interestingly, a nonmetal location for the peptide carbonyl has been proposed on the basis of detailed molecular orbital calculations.[50] Site-directed mutagenesis studies on Arg-127 are necessary in

(45) Suigimoto, T.; Kaiser, E. T. *J. Am. Chem. Soc.* **1979**, *101*, 3946–3951.
(46) Auld, D. S.; Holmquist, B. *Biochemistry* **1974**, *13*, 4355–4361.
(47) Rees, D. C.; Lipscomb, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 7151–7154.
(48) Van Wart, H. E.; Vallee, B. L. *Biochemistry* **1978**, *17*, 3385–3394.
(49) Riordan, J. F. *Biochemistry* **1973**, *12*, 3915–3923.

(50) Nakagawa, S.; Umeyama, H.; Kitaura, K.; Morokuma, K. *Chem. Pharm. Bull.* **1981**, *29*, 1–6.

Carboxypeptidase A

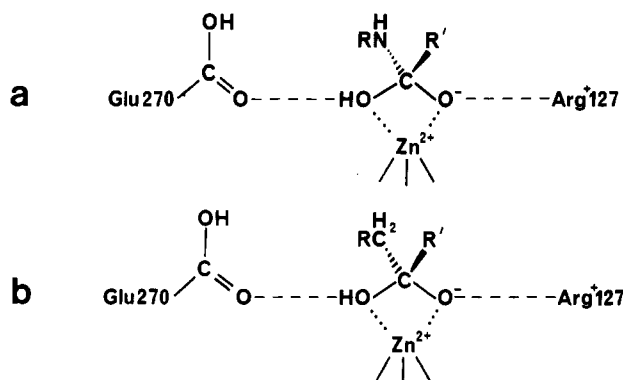*Acc. Chem. Res., Vol. 22, No. 2, 1989* **67**

Figure 8. The first step of the CPA-catalyzed proteolytic mechanism involves the direct attack of water, promoted by zinc and assisted by Glu-270, at the scissile peptide carbonyl to result in the tetrahedral intermediate (a). Carbonyl-containing reaction-coordinate analogues (Figure 7) bind to CPA as covalent hydrate adducts (b). The chemistry of hydration may be performed by the enzyme; both zinc and Arg-127 stabilize the bound hydrates as oxyanionic *gem*-diolates.

order to clarify its particular role in proteolytic and esterolytic reactions.

The binding of the scissile peptide carbonyl to Arg-127 allows for subsequent attack by zinc-bound water, and the nucleophilic promotion of this water molecule may be assisted by Glu-270. Several examples of the resulting tetrahedral intermediate are provided by the complexes of CPA with four reaction-coordinate analogues (Figure 7). The aldehyde 2-benzyl-3-formyl-propanoic acid[10,51] binds to CPA as the hydrate, as does the highly electrophilic α-trifluoro ketone 2-benzyl-4-oxo-5,5,5-trifluoropentanoic acid.[11,52] These results are not surprising in view of the enhanced electrophilicity of these compounds. It is surprising, however, to find that the nonactivated ketones 5-benzamido-2-benzyl-4-pentanoic acid[9,53] and N-(*tert*-butoxycarbonyl)-5-amino-2-benzyl-4-oxo-6-phenylhexanoic acid[12,54] bind to the enzyme as tetrahedral hydrates. The latter two compounds are not very electrophilic at their ketonic carbonyls, and in solution they must exist in hydrated form to less than 0.2% if their degree of hydration is comparable with that of acetone;[55] this percentage must certainly represent an upper limit for sterically hindered ketonic centers. The enzyme may well have performed a hydration reaction on the intact carbonyls of the ketones to result in the *gem*-diolate analogues of the proteolytic tetrahedral intermedate—the polarization of one hydrate oxygen by zinc and Arg-127 stabilizes that oxygen as the oxyanion.[56] Therefore, the observed X-ray structures bear both a steric and electronic resemblance to the actual proteolytic tetrahedral intermediate (Figure 8). Color representations of the complex between CPA and a bound reaction-coordinate analogue are found in Figures 2 and 3. Whether or not CPA hydrates these reaction-coordinate analogues, the enzyme displays elegant selectivity in the binding of

each tetrahedral *gem*-diolate instead of the planar carbonyl species. We consider it mechanistically significant that neither of these nonactivated ketones accepts a lone pair of electrons from Glu-270.

A preformed tetrahedral center contained within a phosphonamidate inhibitor comprises a possible transition-state analogue of the proteolytic reaction catalyzed by CPA. The intact phosphonamidate N-[[[(benzyloxycarbonyl)amino]methyl]hydroxy-phosphinyl]-L-phenylalanine (an analogue of the peptide substrate Cbz-Gly-Phe)[57] binds to CPA at pH 8.5 with the anionic phosphonamidate moiety asymmetrically straddling the zinc ion; the P_2–P_1 portion of the inhibitor binds anomalously in the S_1 hydrophobic cleft.[8] Surprisingly, though, the compound is observed to bind as the hydrolysis products at pH 7.5, probably due to the inherent instability of such linkages to acidic hydrolysis and the time span usually required for X-ray data acquisition.[7] It is undetermined at this time whether the enzyme actually participated in the hydrolysis. At any rate, the intact phosphonamidate and the hydrated *gem*-diolates represent the proteolytic tetrahedral intermediate bound to the active site of CPA. This tetrahedral intermediate requires a proton donor in order to facilitate the departure of the amino leaving group; Rees and Lipscomb once considered Glu-270, yet favored Tyr-248, for this role.[5] Monzingo and Matthews elaborated upon this role for Glu-270 as an intermediate proton donor in CPA and a similar role for Glu-143 in thermolysin.[58,59] This proposal was corroborated by the near-normal activity observed for the Phe-248 mutant of CPA, generated by site-directed mutagenesis experiments with the rat enzyme.[60,61] Additionally, on the basis of model studies, Breslow has proposed that Glu-270 mediates two proton transfers: one as intermediate proton donor, and a second one to generate the salt pair between the P_1 product carboxylate and the P_1' product amino group.[62] Mock considered Glu-270 as a proton donor to the leaving amino group in proteolysis, but his mechanistic proposal did not concern a Glu-270 and/or zinc-promoted water nucleophilic step.[63]

The site-directed mutagenesis experiment referred to in the preceding paragraph, in which Tyr-248 of the wild-type rat enzyme was replaced by a phenylalanine residue, represents perhaps one of the most important contributions of frontier genetic methods toward the elucidation of enzymatic reaction mechanism.[60,61] A proton donor is required for an amino leaving group in the collapse of a tetrahedral intermediate. Historically, the role of intermediate proton donor in the CPA mechanism had been ascribed to Tyr-248,[1,34] although water or Glu-270 had been considered for the role.[5] The near-normal activity of the mutant Phe-248 enzyme[60,61] strongly suggests that residue 248 is not obligatory for

(51) Galardy, R. E.; Kortylewicz, Z. P. *Biochemistry* **1984**, *23*, 2083–2087.
(52) Gelb, M. H.; Svaren, J. P.; Abeles, R. H. *Biochemistry* **1985**, *24*, 1813–1817.
(53) Grobelny, D.; Goli, U. B.; Galardy, R. E. *Biochemistry* **1985**, *24*, 7612–7617.
(54) Oren, D.; Ewenson, A.; Gilon, C.; Shoham, G. *Biochem. Biophys. Res. Commun.*, in press.
(55) Lewis, C. A.; Wolfenden, R. *Biochemistry* **1977**, *16*, 4886–4890.
(56) Teater, C.; Grobelny, D.; Galardy, R. E. *Biochem. Biophys. Res. Commun.* **1988**, *153*, 773–778.

(57) Jacobsen, N. E.; Bartlett, P. A. *J. Am. Chem. Soc.* **1981**, *103*, 654–657.
(58) Monzingo, A. F.; Matthews, B. W. *Biochemistry* **1984**, *23*, 5724–5729.
(59) Hangauer, D. G.; Monzingo, A. F.; Matthews, B. W. *Biochemistry* **1984**, *23*, 5730–5741.
(60) Gardell, S. J.; Craik, C. S.; Hilvert, D.; Urdea, M. S.; Rutter, W. J. *Nature (London)* **1985**, *317*, 551–555.
(61) Hilvert, D.; Gardell, S. J.; Rutter, W. J.; Kaiser, E. T. *J. Am. Chem. Soc.* **1986**, *108*, 5298–5304.
(62) Schepartz, A.; Breslow, R. *J. Am. Chem. Soc.* **1987**, *109*, 1814–1826.
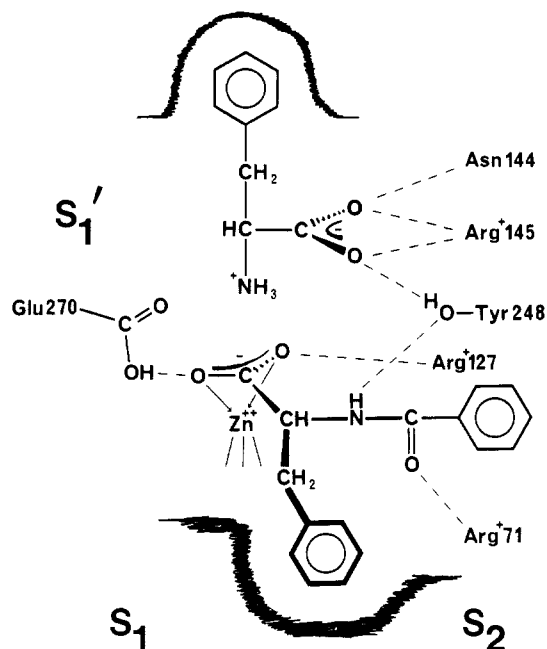(63) Mock, W. L. *Bioorg. Chem.* **1975**, *4*, 270–278.

Figure 9. CPA crystals soaked in a buffer solution containing 0.1 M benzoyl-L-phenylalanine show enzymatic activity during the time frame of the X-ray diffraction experiment. A schematic view of the resulting enzyme–substrate–product complex is presented; note that this complex also represents an enzyme–products complex for the hydrolysis of N-benzoyl-L-phenyl-alanyl-L-phenylalanine.

catalysis. However, it is found that substrate affinity is noticeably altered in the Phe-248 mutant. Hence, Tyr-248 is confirmed as a residue that is important for substrate binding, probably by hydrogen bonding via its phenolic hydroxyl moiety. We note that the role of "capping" the S_1' hydrophobic pocket can be performed equally well by a phenylalanine at position 248, and we await an X-ray crystallographic study of substrate binding on a Phe-248 mutant of CPA. As for Tyr-248 of the wild-type enzyme, we would expect Phe-248 of the mutant enzyme to be in the "down" conformation upon the binding of substrates or inhibitors.

The collapse of the proteolytic tetrahedral intermediate results in the formation of product carboxylate and amino groups. A ternary enzyme–substrate–product complex with S_2–S_1 N-benzoyl-L-phenylalanine and S_1' phenylalanine also would be expected for the hydrolysis products of the peptide substrate N-benzoyl-L-phenylalanyl-L-phenylalanine (Figure 9).[14]   The distinction between a stable product complex and a metastable intermediate may have been overlooked by other investigators.[22,23,25] We find in the enzyme–substrate–product (or enzyme–products) complex that the zinc-bound carboxylate makes a hydrogen-bond contact with Arg-127. Hence, Arg-127 is now implicated in the binding of proteolytic products as well as precatalytic substrate and the oxyanion of the tetrahedral intermediate.

Summary and Conclusions

A proteolytic mechanism is presented in Figure 10 which encompasses all structural and chemical aspects discussed in this Account. Figure 10 is thus our schematic conclusion regarding the current interpretation of chemical and structural data pertaining to the CPA mechanism. We emphasize that our interpretation favors the *promoted-water pathway* in the hydrolysis of
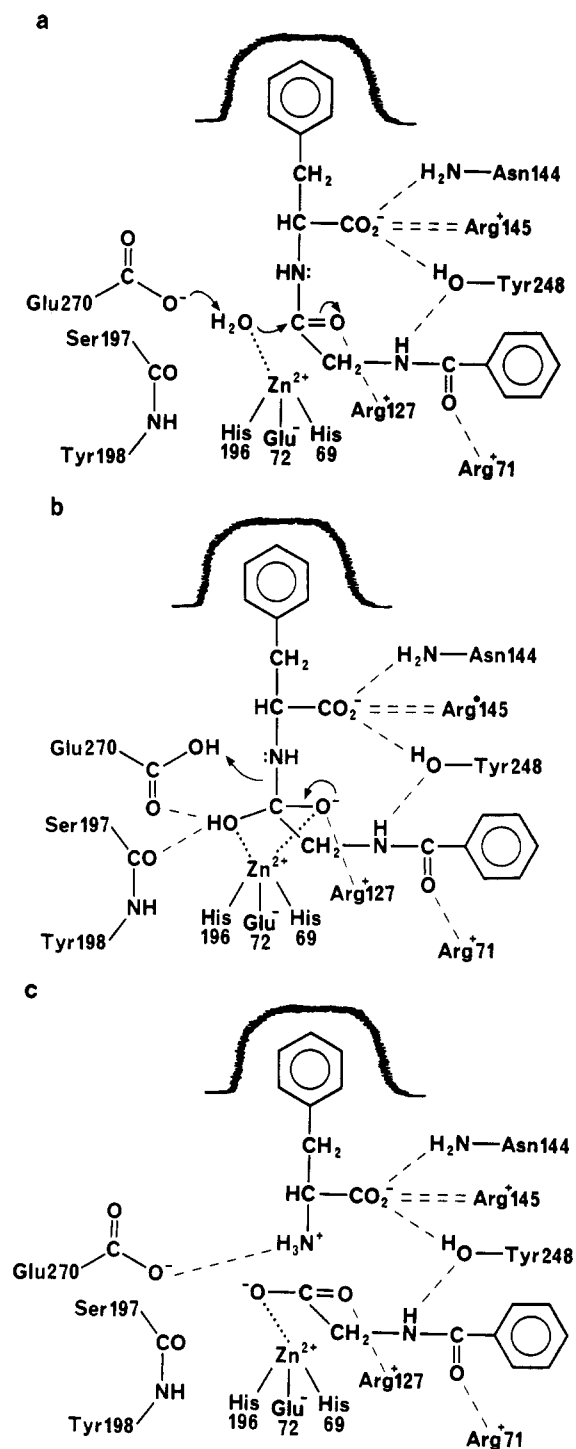

Figure 10. Proposed mechanism for CPA-catalyzed proteolysis: (a) the precatalytic Michaelis complex with the substrate carbonyl hydrogen bonded to Arg-127 allows for nucleophilic attack by a water molecule promoted by zinc and assisted by Glu-270; (b) the stabilized tetrahedral intermediate collapses with required proton donation by Glu-270; (c) the final product complex, after a second and final proton transfer mediated by Glu-270—unfavorable electrostatic interactions between Glu-270 and the ionized product carboxylate may facilitate product release.

peptide substrates. We propose that the mechanistic sequence of Figure 10 may be generalizable to many other zinc proteases (in addition to thermolysin, which has been recently reviewed by Matthews[64]) which possess an additional electrophile, e.g., the side chains of protonated arginine, lysine, or histidine residues, in

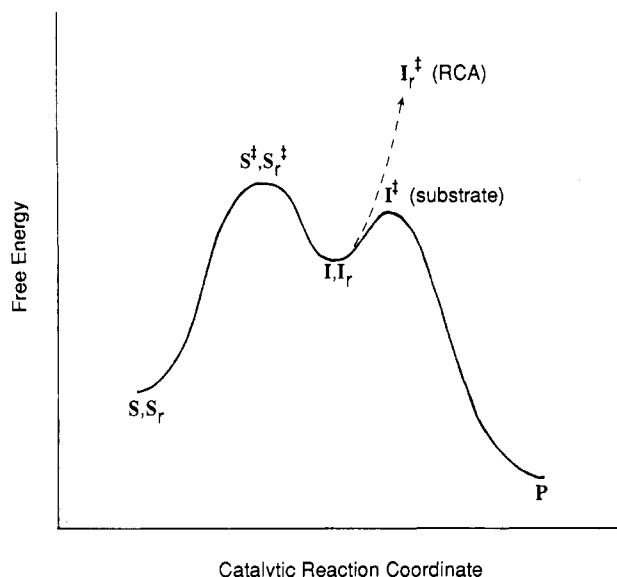(64) Matthews, B. W. *Acc. Chem. Res.* **1988,** *21,* 333–340.

Catalytic Reaction Coordinate

Figure 11. An inert transition-state analogue mimics the structure of only one point along a given reaction coordinate (optimally $S^\ddagger$) in the conversion of substrate S into product P. A reaction-coordinate analogue (RCA), however, is a reactive substrate analogue that undergoes a reversible chemical transformation parallel to that of an actual substrate in the first step of catalysis. Hence, the RCA undergoes a transformation from S_r (subscript r distinguishes the RCA from substrate), *through* $S_r^\ddagger$, to I_r. The RCA cannot complete the chemistry of the catalytic cycle, but it must stop, in dead-end fashion, at the intermediate point I_r due to an inherent and insurmountable barrier to subsequent reactivity ($I_r^\ddagger$). For example, this barrier might arise from an inability of the tetrahedral *gem*-diolate of Figure 8b to expel a carbanion. Hence, the RCA cannot form a catalytic product. The X-ray crystallographic observation of I_r, the selectivity of the enzyme toward the binding of I_r versus S_r, and the implication of the chemistry leading from S_r to I_r provide powerful evidence for the parallel chemistry of the S → I conversion with actual substrates.

their active sites.[11,13,65] Determinations of the three-dimensional structures of other zinc proteases may ultimately corroborate the ubiquitous presence of this additional electrophile.

Inspired by the course of our recent studies of CPA, we have recently suggested the reclassification of a large family of proteolytic enzyme inhibitors as *reaction coordinate analogues*.[8] Consider the carbonyl-containing inhibitors of Figure 7. Each of these compounds is noteworthy since each undergoes a chemical transformation prior to the observation of the enzyme–inhibitor

complex; i.e., each ketone, or the aldehyde, binds to CPA as the covalent nucleophilic adduct. For the zinc protease, this adduct is a *gem*-diolate. In other proteolytic systems such as the serine or sulfhydryl proteases, the related enzyme–analogue adduct is instead a covalent hemiketal with the active-serine or cysteine,[66-68] or even a tetrahedral boronic acid anhydride if the parent inhibitor is a boronic or alkylboronic acid.[69] Aspartyl proteases are known to be strongly inhibited by the hydrate form of ketonic substrate analogues.[70]

The term "reaction-coordinate analogue" describes reversibly reactive substrate analogues that bind to enzyme active sites as analogues of catalytic intermediates. A reaction-coordinate analogue can undergo a reversible chemical reaction identical with the first elementary step(s) of catalysis, yet it cannot complete the chemistry of the entire catalytic cycle due to a subsequently insurmountable barrier (e.g., expulsion of an unfavorable carbanion or hydride ion; see Figure 11). Reaction-coordinate analogues differ from transition-state analogues in that the former are chemically reactive, whereas the latter are typically inert. Moreover, since the reversible reactivity of the reaction-coordinate analogue parallels the reversible reactivity of substrates, it cannot be classified as an irreversible mechanism based or suicide enzyme inhibitor. Reaction-coordinate analogues are perhaps the best candidates for X-ray crystallographic study because their binding invokes the same reversible chemistry that occurs during catalysis. This chemistry is long-lived and largely amenable to the time scale of typical X-ray diffraction data collection. Importantly, the K_i measured in solution for a reaction-coordinate analogue may not correspond to the dissociation constant of the actual inhibiting species detected by X-ray crystallography, and this discrepancy may span several orders of magnitude.

(65) Christianson, D. W.; Lipscomb, W. N. In *Mechanistic Principles of Enzyme Activity*; Liebman, J. F., Greenberg, A., Eds.; VCH Publishers, Inc.: New York, 1989; pp 1–25.

(66) Liang, T. C.; Abeles, R. H. *Biochemistry* 1987, 26, 7603–7608.
(67) Brayer, G. D.; Delbaere, L. T. J.; James, M. N. G.; Bauer, C.-A.; Thompson, R. C. *Proc. Natl. Acad. Sci. U.S.A.* 1979, 76, 96–100.
(68) Westerick, J. O.; Wolfenden, R. *J. Biol. Chem.* 1972, 247, 8195–8197.
(69) Bone, R.; Shenvi, A. B.; Kettner, C. A.; Agard, D. A. *Biochemistry* 1987, 26, 7609–7614.
(70) Rich, D. H. *J. Med. Chem.* 1985, 28, 263–273.