Caught after the Act: A Human A-Type Metallocarboxypeptidase in a Product Complex with a Cleaved Hexapeptide†

Àlex Bayés,§ Daniel Fernández,§ Maria Solà,‡ Aniebrys Marrero,‡ Sonia García-Piqué,‡ Francesc X. Avilés,§ Josep Vendrell§ and F. Xavier Gomis-Rüth*,‡

Institut de Biotecnologia i de Biomedicina and Departament de Bioquímica i Biologia Molecular, Facultat de Biociències, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain and Department de Biologia Estructural, Institut de Biologia Molecular de Barcelona CSIC, Jordi Girona 18-26, E-08034 Barcelona and Parc Cientific de Barcelona, Josep Samitier 1-5, E-08028 Barcelona, Spain

Received March 9, 2007; Revised Manuscript Received April 2, 2007

ABSTRACT: A/B-type metallocarboxypeptidases (MCPs) are among the most thoroughly studied proteolytic enzymes, and their catalytic mechanisms have been considered as prototypes even for several unrelated metalloprote(in)ase families. It has long been postulated that the nature of the side chains of at least five substrate residues, i.e., $P_4-P_{1'}$, influence K_m and k_{cat} and that once the peptide or protein substrate is cleaved, both products remain in the first instance bound to the active-site cleft of the enzyme in a double-product complex. Structural details of binding of substrate to the nonprimed side of the cleft have largely relied on complexes with protein inhibitors and peptidomimetic small-molecule inhibitors that do not span the entire groove. In the former, the presence of N-terminal globular protein domains participating in large-scale interactions with the surface of the cognate catalytic domain outside the active-site cleft mostly conditions the way their C-terminal tails bind to the cleft. Accordingly, they may not be accurate models for a product complex. We hereby provide the structural details of a true cleaved double-product complex with a hexapeptide of an MCP engaged in prostate cancer, human carboxypeptidase A4, employing diffraction data to 1.6 Å resolution (R_{cryst} and R_{free} = 0.159 and 0.176, respectively). These studies provide detailed information about subsites $S_5-S_{1'}$ and contribute to our knowledge of the cleavage mechanism, which is revisited in light of these new structural insights.

Zinc-dependent metallocarboxypeptidases (MCPs)[1] remove C-terminal amino acids from protein and peptide substrates and can be classified into A/B and N/E forms, previously known as pancreatic and regulatory MCPs, respectively (*1*, *2*). At the onset of biochemistry, A/B-type MCPs (A/B-MCPs) were identified as digestive enzymes produced by the exocrine pancreas of mammals (*3*−*5*). More recently, homologues with functions distinct from the degradation of intake proteins have been found in several other mammalian tissues and fluids and within cells, as well as in every realm of life (*5*). They participate in delicately regulated processes such as insect attack−vegetal defense strategies, parasite invasion, blood coagulation and fibrinolysis, inflammation, local anaphylaxis, fertilization, hormone and neuropeptide processing, resistance to bee and snake poison, and carcinogenesis and tumor invasion, among others (*2*, *6*−*9*).

A/B-MCPs exhibit distinct substrate specificity patterns. While B-type CPs (CPBs) favor substrates with C-terminal basic side chains, A-type CPs (CPAs) are subdivided into A_1 and A_2 forms and prefer smaller aliphatic residues and bulkier aromatic side chains, respectively (*6*, *10*). The archetypal and most thoroughly studied MCP is bovine pancreatic CPA (bCPA), actually an A_1 form in its specificity, discovered in the 1920s (*4*). A more recently identified family member is human carboxypeptidase A4 (hCPA4), detected in human hormone-regulated tissues and thought to play a role in prostate cancer (*11*, *12*). The coding gene may be responsible for prostate-cancer aggressiveness, as it is imprinted and upregulated via the histone hyperacetylation pathway as suggested by the analysis of downstream effects of sodium butyrate treatment of prostate cancer cell lines (*12*). A/B-MCPs are synthesized and secreted as zymogens [procarboxypeptidases (PCPs)] to prevent activity in the

† This study was supported by the following grants and fellowships: BIO2003-00132, GEN2003-20642, BIO2004-20369-E, and BIO2004-05879 from the former Spanish Ministry of Science and Technology, BIO2006-02668, BFU2006-09593, and CONSOLIDER-INGENIO 2010 Project "La Factoría de Cristalización" (CSD2006-00015) from the Spanish Ministry of Education and Science, ON03-7-0 from the "Fundació La Caixa", EU FP6 Integrated Project LSHC-CT-2003-503297 "CANCERDEGRADOME", "AVON-Project" 2005X0648 from the Scientific Foundation of the Spanish Association Against Cancer, and Grants 2005SGR00280 and 2005SGR01037 from the Generalitat of Catalunya. M.S. is a beneficiary of the "Ramón y Cajal" Program of the Spanish Ministry of Science and Education. D.F. is a recipient of a Ph.D. fellowship from the EU Al*β*an Programme (Contract E04D035150AR). Funding for X-ray diffraction data collection at ESRF synchrotron was provided by ESRF and the EU.

* To whom correspondence should be addressed: Departament de Biologia Estructural, Institut de Biologia Molecular de Barcelona (CSIC), Jordi Girona 18-26, E-08034 Barcelona, Spain. Telephone: +34 934 006 144. Fax: +34 932 045 904. E-mail: xgrcri@ibmb.csic.es.

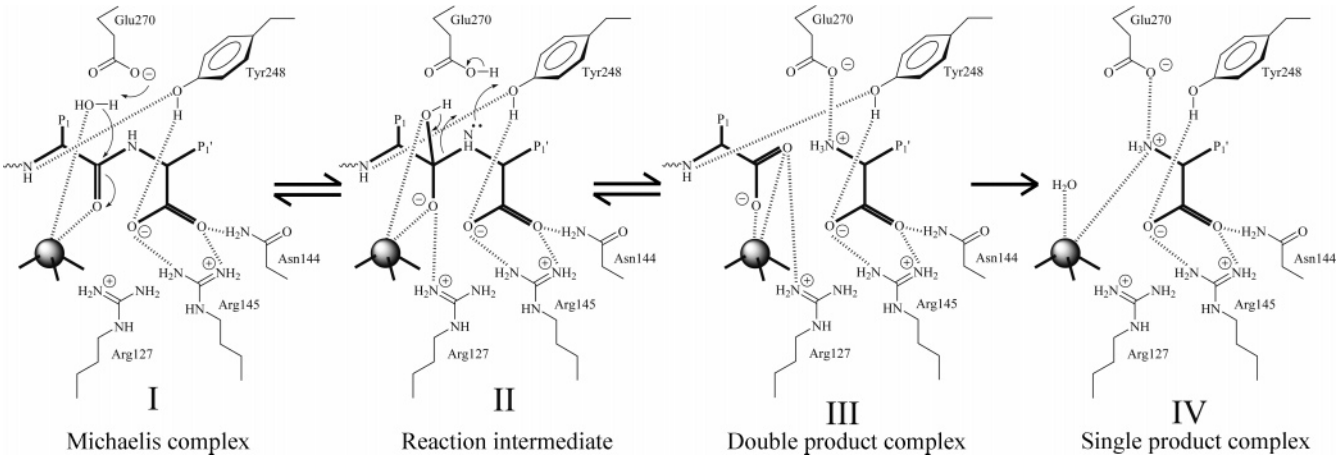
§ Universitat Autònoma de Barcelona.

‡ Institute de Biologia Molecular de Barcelona, CSIC.

[1] Abbreviations: Z, benzyloxycarbonyl; MCP, metallocarboxypeptidase; MP, metalloprote(in)ase; MS, mass spectrometry; (P)CPA, (P)-CPA2, (P)CPA4, (P)CPB, etc., (pro)carboxypeptidase A, A2, A4, B, etc.; PDB, Protein Data Bank.

10.1021/bi700480b CCC: $37.00
Published on Web 05/17/2007

Scheme 1: Catalytic Mechanism of MCP-Mediated Peptide Cleavage[a]



[a] This scheme is an upgrade that includes both the well-established features of catalysis and those novel details proposed on the basis of this structure and our previous studies (*17*, *34*, *39*, *71*−*73*). The catalytic zinc ion is shown as a sphere, and the three protein ligands are represented by three sticks. Hydrogen bonds are depicted with dashed lines. Residues key for catalysis are shown and labeled.

absence of the appropriate environment and stimuli. The structural effector for latency is an N-terminal 90−95-residue independently folding prodomain that prevents substrate access to the catalytic CP domain, which is already in an active conformation in the zymogen. Activation occurs through limited proteolysis in a connecting segment at the end of the prodomain through endopeptidases such as trypsin (*13*, *14*). A number of endogenous and exogenous protein inhibitors of A/B-MCPs have been reported: the archetypal potato carboxypeptidase inhibitor (PCI) and its orthologue from tomato (MCPI), inhibitors from the intestinal parasite *Ascaris suum* and the medical leech *Hirudo medicinalis* (LCI), the inhibitor from the tick *Rhipicephalus bursa* (TCI), and rodent and human latexin alias ECI (*2*, *15*−*17*). These protein inhibitors display equilibrium dissociation constants in the nanomolar range against vertebrate A/B-MCPs but are inactive against N/E-type MCPs (*15*−*19*).

The enzymatic activity of A/B-MCPs has been the subject of a number of studies, chiefly on bCPA, that have led to the following consensus mechanism of catalysis (*5*, *20*−*26*) (Scheme 1). In the Michaelis complex, the substrate side chains (P_1−P_3, etc.; the nonprimed side) preceding the scissile bond interact with substrate-binding subsites S_1−S_3, etc., of the enzyme (nomenclature from refs *27* and *28*). The C-terminal residue to be removed (side chain $P_{1'}$) interacts with cognate enzyme substrate-binding subsite $S_{1'}$, the specificity pocket, on the primed side of the groove (stage I) (*27*). This is followed by the nucleophilic attack of a catalytic solvent molecule, further bound to and thus polarized by Glu270, on the scissile carbonyl carbon atom. This generates a negatively charged tetrahedral reaction intermediate (stage II). The scissile bond is cleaved, and the cleavage fragments are subsequently removed so that the enzyme is poised for a new catalytic round. This mechanism entails two product complexes, a double-product complex (stage III), in which both the upstream and downstream cleavage products are still bound to the enzyme, and a single-product complex (stage IV), in which only the excised C-terminal residue remains anchored to the specificity pocket (see Scheme 1 and Figure 68.3 in ref *24*).

This mechanistic view is supported by studies with peptidomimetic compounds with a phosphinato group [-P(O)-(O^-)- (alias -PO_2-)] replacing the sp^3-hybridized diolate group [-C(OH)(O^-); Scheme 1, stage II] of the reaction intermediate, as observed in phosphonamidates (-PO_2-NH-), phosphonates (-PO_2-O-), and phosphinates (-PO_2-CH_2-). To date, phosphinate complexes with zinc-dependent proteases have only been reported for astacin and matrix metalloproteinase 11, but not for MCPs (*29*). In contrast, bCPA has been the subject of studies with tripeptide-mimicking phosphonates with the generic formula φ-CH_2-OCO-NH-CH(R_1)-CO-NH-CH(R_2)-PO_2-O-CH(CH_2-φ)-COO^-, where R_1 = CH_2-φ and R_2 = $CH(CH_3)_2$ for *O*-{[(1*R*)-{[*N*-(phenylmethoxycarbonyl)-L-phenylalanyl]amino}isobutyl]-hydroxyphosphinyl}-L-3-phenyllactate [alias $ZFV^P(O)F$]; R_1 = R_2 = CH_3 for *O*-{[(1*R*)-{[*N*-(phenylmethoxycarbonyl)-L-alanyl]amino}ethyl]hydroxyphosphinyl}-L-3-phenyllactate [alias $ZAA^P(O)F$]; and R_1 = CH_3 and R_2 = H for *O*-{[(1*R*)-{[*N*-(phenylmethoxycarbonyl)-L-alanyl]amino}-methyl]hydroxyphosphinyl}-L-3-phenyllactate [alias ZAG^P-(O)F] (*26*, *30*). These and similar phosphorus-containing compounds were initially identified as high-affinity inhibitors of CPA with dissociation constants in the pico- or femtomolar range (*26*, *31*). Other transition-state analogues based on sulfur-containing compounds have also been described and their structures determined (*32*, *33*). In addition, structures representing the single-product complex of stage IV are available for porcine PCPA [Protein Data Bank (PDB) entry 1PCA (*34*)] and hCPA4 in complex with latexin [PDB entry 2BO9 (*17*)]. Here, leftover valine residues are found occupying the $S_{1'}$ pocket. In contrast, structural information about stage III of the mechanism is restricted to complexes with the inhibitors PCI, LCI, and TCI (*35*−*37*). These protein inhibitors recognize the surface of the target molecule via large contact areas with prefixed structural cores, and their C-terminal tails are mostly disordered in the unbound structures (*36*, *38*). These tails become rigid upon complex formation, and reminiscent of a stopperlike mechanism, they penetrate the active-site cleft in a substratelike manner that allows for an occasional hydrolysis of the C-terminal residues (*35*−*37*). However, superimposition of the complex structures of PCI, LCI, and TCI reveals that the conformation of these tails is conditioned by globular protein scaffolds, which accounts for the variability in the polypeptide chain traces

in the positions preceding S_2. Early studies of bCPA had revealed that the nature of the side chains of at least five residues, $P_4-P_{1'}$, influences K_m and k_{cat} (*21*, *28*). Accordingly, a more accurate model for a bound and cleaved substrate, i.e., of a double-product complex, would be better provided by a structurally unrestrained oligopeptide.

We have determined the high-resolution three-dimensional structure of the first complex of an A/B-MCP with a free-standing oligopeptide in the form of hCPA4 bound to a hexapeptide emanating from the prodomain processed during the activation process. The peptide is tightly bound to the enzyme, and its C-terminal residue is severed and occupies the specificity pocket. Given the lack of restraints due to preceding protein domains, this complex can be envisaged as a true double-product complex. The structure enables the structural determinants of the subsites shaping the nonprimed region of the active-site cleft to be assessed in detail and contributes to the unraveling of novel details of the catalytic mechanism of A/B-MCPs that had not been addressed previously.

EXPERIMENTAL PROCEDURES

Protein Production and Purification. hCPA4 was produced as a zymogen (hPCPA4) through recombinant heterologous overexpression using vector pPIC9 and the methylotrophic yeast *Pichia pastoris* as an expression host and purified as described elsewhere (*17*, *39*). The active enzyme was obtained through tryptic activation (at a 200:1 w/w ratio) for 90 min at room temperature, and the resulting protein−product complex was subsequently purified by anion-exchange chromatography (TSK-DEAE 5PW) using a FPLC-Äkta system with a linear salt gradient from 0 to 60% of 0.8 M ammonium acetate in 20 mM Tris-HCl (pH 9.5). Eluted fractions were analyzed by SDS−PAGE, and the purest samples containing the complex were pooled, desalted, and concentrated to ~15 mg/mL by Amicon centrifugal filter devices.

Activity and Mass Spectrometry Measurements. hCPA4 proteolytic activity was analyzed using peptides Phe-Asn-Arg-Pro-Val-Asp, Phe-Asn-Arg-Pro-Val-Val, and Phe-Asn-Arg-Pro-Val (purchased from EZBiolabs). The reaction took place in a 20 μL solution of 150 mM NaCl and 10 mM Tris-HCl (pH 8.0) and with peptide and enzyme concentrations of 1 μmol and 1 nmol, respectively. The reaction products were analyzed by mass spectrometry (MS) following the α-cyano-4-hydroxycinnamic acid affinity sample preparation protocol with a 600 μm AnchorChipTM MALDI sample support in a Bruker Daltonics Ultraflex MALDI-TOF spectrometer. Inhibitory studies of hCPA4 were performed with the chromogenic substrate *N*-(4-methoxyphenyl-azoformyl)-L-Phe-OH in 50 mM Tris-HCl and 0.45 M NaCl (pH 7.5) with a peptide concentration of 100 μM and an enzyme concentration of 28 nM. Assayed peptides bore the following sequences: Phe-Asn-Arg-Pro-Val-Asp, Phe-Asn-Arg-Ala-Val-Asp, Phe-Asn-Arg-Ala-Val-Val, Phe-Asn-Arg-Pro-Val-Val, Phe-Asn-Arg-Pro-Val, Gly-Asn-Arg-Pro-Val-Thr, and Gly-Asn-Arg-Pro-Thr.

Crystallization of the Protein−Product Complex. Crystals were obtained from sitting drops subjected to vapor diffusion and containing 100 nL of hCPA4−hexapeptide complex [15 mg/mL in 5 mM Tris-HCl (pH 7.5)] and 100 nL of reservoir solution (0.2 M potassium thiocyanate and 20% PEG 3350). Crystallization drops were dispensed on 96 × 3-well Greiner plates by a Tecan robot and a Cartesian nanodrop robot (Genomic Solutions) at the joint IBMB-CSIC/Barcelona Science Park Automated Crystallization Platform (PAC). Crystals appeared after incubation for 10−15 days in a Bruker steady-temperature crystal farm at 4 °C. The cryo-protection protocol included soaking crystals in a mixture containing reservoir solution and 20% glycerol. A complete diffraction data set was collected at 100 K from a single N_2 flash-cryocooled (Oxford Cryosystems) crystal on an ADSC Q210 2D detector at beamline ID29 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) within the Block Allocation Group "BAG Barcelona". Crystals were orthorhombic, harbored one complex per asymmetric unit, and diffracted beyond 1.6 Å resolution. Diffraction data were integrated, scaled, merged, and reduced with MOSFLM (*40*) and SCALA (*41*) within the CCP4 suite (*42*) (see Table 1).

Table 1: Crystallographic Statistics on Data Collection and Refinement

space group/cell constants [*a*, *b*, *c* (Å)]	$P2_12_12_1$/50.3, 72.5, 81.6
wavelength (Å)	1.036
no. of measurements/no. of unique reflections	240598/39536
resolution range (Å)	72.6−1.60 (1.69−1.60)[a]
completeness (%)	98.7 (95.1)[a]
R_{rim}[b]	0.102 (0.411)[a]
R_{pim}[b]	0.041 (0.218)[a]
average intensity ($\langle[\langle I\rangle/\sigma(\langle I\rangle)]\rangle$)	13.0 (3.3)[a]
B-factor (Wilson) (Å^2)	21.5
average multiplicity	6.1 (3.2)[a]
resolution range used for refinement (Å)	50.0−1.60
no. of reflections used (test set)	38869 (609)[a]
crystallographic R_{factor} (R_{free})[c]	0.159 (0.176)[a]
estimated overall coordinate error based on R_{free} (Å)	0.076
no. of protein atoms,[d] peptide atoms,[d] solvent molecules, ions, other molecules	2439/59/267/1 (Zn^{2+})/ 1 thiocyanate, 1 glycerol
rmsd from target values	
bonds (Å)	0.017
angles (deg)	1.63
bonded *B*-factors (main chain/side chain) (Å^2)	1.09/2.67
average *B*-factors for protein/peptide atoms (Å^2)	25.4/20.3

[a] Values in parentheses refer to the outermost resolution shell if not otherwise indicated. [b] $R_{rim} = \sum_{hkl}[n_{hkl}/(n_{hkl}-1)^{1/2}]\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, and $R_{pim} = \sum_{hkl}[1/(n_{hkl}-1)^{1/2}]\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the *i*th intensity measurement, n_{hkl} the number of observations of reflection *hkl*, including symmetry-related reflections, and $\langle I(hkl)\rangle$ is its average intensity. R_{rim} (alias R_{meas}) and R_{pim} are improved multiplicity-weighted indicators of the quality of the data, the redundancy-independent merging *R*-factor and the precision-indicating merging *R*-factor, the latter computed after averaging the multiple measurements (*69*, *70*). [c] Crystallographic $R_{factor} = \sum_{hkl}||F_{obs}| - k|F_{calc}||/\sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. R_{free}, same for a test set of reflections (>500) not used during refinement. [d] Including atoms in an alternate conformation.

Structure Solution and Refinement. The structure was determined by Patterson search methods with AMoRe (*43*) within ccp4i (*44*), employing all diffraction data up to 4 Å resolution. The coordinates of hCPA4, as in its complex with latexin (PDB entry 2BO9), were used as a searching model. A single solution was encountered at 50.5, 83.8, and 302.0

(α, *β*, and *γ*, respectively, in Eulerian angles) and 0.030, 0.437, and 0.136 (*x*, *y*, and *z*, respectively, as fractional unit-cell coordinates) after rigid-body fitting with the routine "fiting" (*43*). This solution gave a correlation coefficient in structure factor amplitudes (CC_F) of 36.3% and a crystallographic R_{factor} of 40.8% (for definitions, see Table 1 and ref *43*), with a second-highest unrelated peak, $CC_F = 10.8\%$, $R_{factor} = 57.4\%$). Subsequently, manual model building on a SiliconGraphics graphic workstation using TURBO-Frodo (*45*) alternated with crystallographic refinement with REFMAC5 (*46*) within CCP4, until the final model was obtained (see Table 1). This model contained protein residues Asn5A−Leu308A of the mature protease moiety (for the numbering and nomenclature of hCPA4, see refs *17* and *39*) and hexapeptide residues Phe1B-Asn2B-Arg3B-Pro4B-Val5B-Asp6B. All residues were in most favored and additionally favored regions of the Ramachandran plot except for Ser199A, which was in a generously allowed region, as seen in other h(P)CPA4 structures and other (P)CPs of the A/B subfamily like hPCPA2 (*17*, *39*, *47*). An N-glycosylation site consisting of a single *N*-acetylglucosamine module (Nag901A) was modeled, attached to Asn148A Nδ2 based on weak electron density, in particular for the distal part of the sugar ring. Three peptide bonds were in the cis conformation (Ser197A−Tyr198A, Pro205A−Tyr206A, and Arg272−Asp273A), and a disulfide bond linked Cys138A and Cys161A. Residues Arg35A, Met47A, Val62A, Ser75A, Cys244A, His303A, and Asp6B were modeled with two conformations. The 50 loop (Gly55A−Arg58A) was flexible, and the electron density map did not allow the side chains to be placed unambiguously. Two hundred sixty-seven solvent molecules (Hoh503−Hoh769), one (tentatively assigned) glycerol molecule (Gol501), and one thiocyanate molecule (Scn502) were also identified. As there is only one polypeptide chain and one oligopeptide in the asymmetric unit, the chain identifiers (A and B, respectively) were omitted in the following.

Miscellaneous. Figures were prepared with TURBO-Frodo and SETOR (*48*); superimpositions were performed with TURBO-Frodo. The final coordinates of the hCPA4−hexapeptide complex were deposited with the Protein Data Bank at the RCSB (PDB entry 2PCU).

RESULTS AND DISCUSSION

Complex Preparation. Human CPA4 was produced as a zymogen by heterologous recombinant overexpression in a *P. pastoris* system, purified, and activated with trypsin to produce crystals suitable for inhibitory drug design, given its potential role in human pancreatic cancer. To our surprise, the initial unbiased electron density map calculated after structure solution by Patterson search unambiguously revealed the presence of a hexapeptide bound to the active-site cleft, with its C-terminal residue severed and occupying the $S_{1'}$ pocket, i.e., featuring a double-product complex (stage III in Scheme 1; see Figure 1A). This finding is reminiscent of studies performed on *γ*-chymotrypsin, for which a residual tetrapeptide was encountered bound to the catalytic serine, i.e., featuring the acylenzyme intermediate (*49*, *50*). This would be a precedent for the observation of a substrate in a protease active site following activation.

The excellent quality and high resolution of the experimental diffraction data of our complex with an unexpected ligand bound to the active site enabled us to identify the chemical sequence of the first five residues of the peptide as [Phe,His]-[Asp,Asn]-Arg-Pro-[Val,Thr]. The electron density map corresponding to the C-terminal residue clearly evinced four atoms in a plane adopting a "Y" structure adjacent to a sp^3-hydrized atom (see Figure 1A). The shape of the map clearly indicated the presence of a planar group; i.e., it excluded the presence of a molecule only composed of tetrahedral carbon atoms as found in, for example, glycerol, a molecule that was considered during the initial stages of structure solution and refinement. The density map, taken together with the surrounding protein residues (in particular Arg145; see below and Figure 1E), suggested a carboxymethylene moiety attached (1.51 Å distant) to a carbon atom in sp^3 hybridization. After several attempts to interpret this density, it became clear that no further atom was attached to the carbon next to the carboxylate group. This ruled out the possibility that it was an α-carbon atom and restricted the identity of the residue to an aspartate that interacted with Arg145 through its side chain carboxylate and displayed alternate occupancy for its α-amino and α-carboxylate groups (see Figure 1A). This train of thoughts was confirmed by the refined *B*-factor values of the intervening atoms (overall *B*-factor for the whole residue, 23.2 $Å^2$; whole peptide, 20.3 $Å^2$): lowest for the carboxylate oxygens (13.3 and 14.6 $Å^2$), as they constitute the main anchor points to the protein; average for the adjacent carboxylate and methylene carbon atoms (18.2 and 23.3 $Å^2$, respectively); and highest for the remaining distal atoms in alternate occupancy (ranging from 23.8 to 28.5 $Å^2$). Although acidic residues are not among the favorite residues of A-type CPs in $P_{1'}$, it has been reported that human CPA2 cleaves the C-terminal glutamate residue of LCI (*36*). Furthermore, the oxygen atoms of the α-carboxylate group evince rotational disorder and are undefined by electron density (Figure 1A). This finding correlates well with the different orientations found for isolated residues in $S_{1'}$ in the complexes of MCPs with PCI, LCI, and ECI and in porcine PCPA1 (*17*, *34*, *36*, *51*) and accounts for a certain degree of liberty of the severed C-terminal residues to reallocate within the specificity pocket after cleavage, in particular if the preference of the enzyme for a hydrophobic side chain in the specificity pocket is not fulfilled. Sequence-pattern searches based on the previous sequence information performed on proteins of the expression host and the full-length hPCPA4 protein found only one hit, PheA48-AsnA49-ArgA50-ProA51-ValA52-AspA53 from the prodomain of hPCPA4 [for the complete sequence and numbering of h(P)CPA4, see ref *39*)]. Accordingly, this hexapeptide, whose sequence and numbering are Phe1-Asn2-Arg3-Pro4-Val5-Asp6, is likely to be a collateral product generated during the enzymatic activation of the metalloenzyme that was trapped and cleaved in the active-site cleft of hCPA4, where it withstood the final anion-exchange chromatography purification step and the crystallization conditions to give a true double-product complex, as depicted in Scheme 1. All atoms of the hexapeptide except the α-amino and α-carboxylate group of Asp6 and the α-carbon of Phe1 are unambiguously defined by excellent electron density. Their average temperature factor is lower than the average of the protease atoms (Table 1), thus accounting for the marked stability and rigidity of the complex. Indeed, the hexapeptide participates in building up

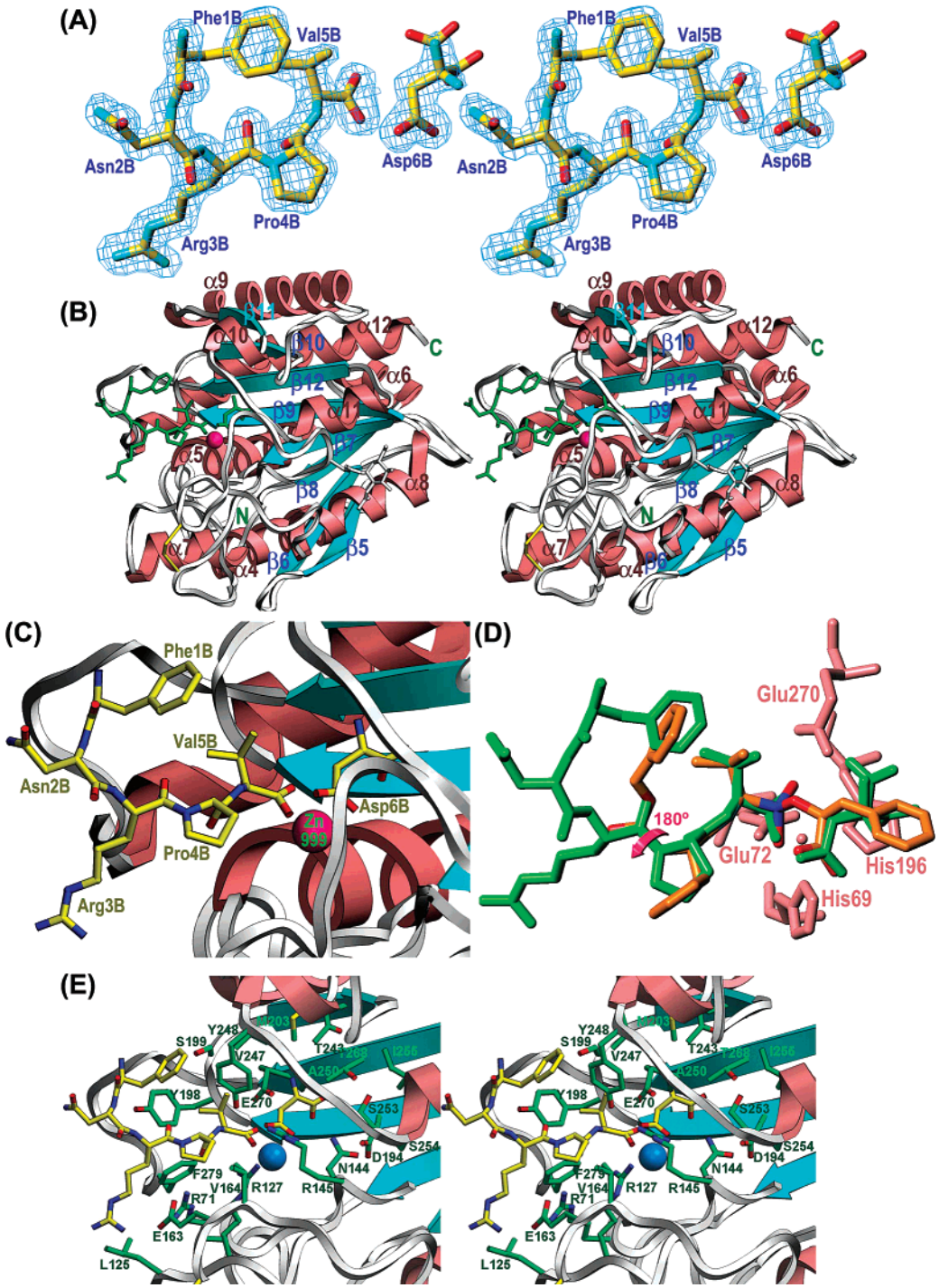
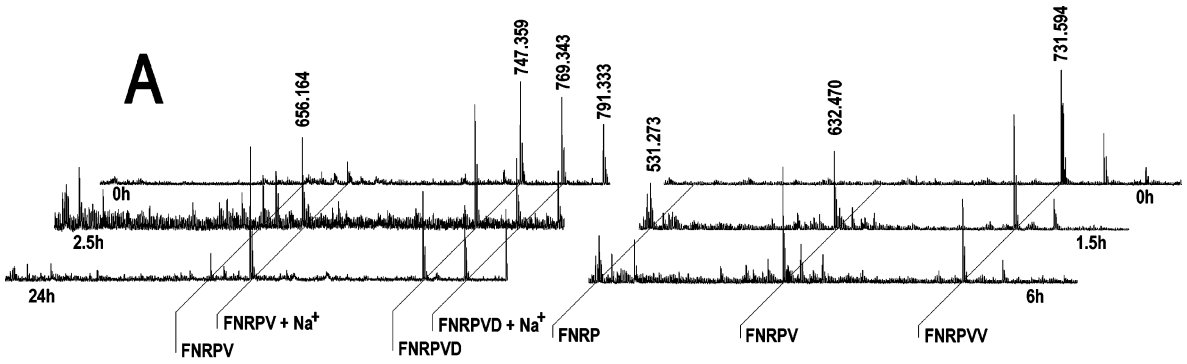


FIGURE 1: Human carboxypeptidase A4 in complex with a double product. (A) Segment of the initial σA-weighted ($2mF_{obs} - dF_{calc}$)-type electron density map contoured at 0.8σ above background superimposed with the bound and cleaved hexapeptide. The six residues are labeled, and the atom color coding is as follows: yellow for C, cyan for N, and red for O. (B) Richardson plot in stereo of hCPA4 in complex with the hexapeptide (green stick model) in standard orientation. Regular secondary structure elements are displayed as cyan arrows for β-strands (β5−β12) and chestnut ribbons for α-helices (α4−α12) and labeled. The N- and C-termini, the catalytic zinc ion (magenta sphere), and the only disulfide bond present in the structure (yellow sticks) are also indicated. (C) Close-up view of panel B in the same orientation zooming in on the active-site cleft with the bound hexapeptide. Color coding as in panel A except for N atoms, which are colored blue. (D) Superimposition of the hexapeptide (green sticks) onto the reaction-intermediate-mimicking phosphonate inhibitor $ZFV^{P}(O)F$ (color coding, orange for C, red for O, and blue for P) after optimal least-squares fitting of the respective hCPA4 and bCPA moieties. The catalytic zinc of hCPA4 (small sphere) and its protein ligands are further shown as salmon sticks. The structures deviate in the positions upstream of P_2 due to an ~180° rotation in $ZFV^{P}(O)F$ around the bond that is equivalent to the P_3−P_2 peptide bond of the hexapeptide. (E) Stereo cartoon showing the Richardson plot of hCPA, as well as the stick models of the bound hexapeptide (color coding as in panel C) and the protein residues (green one-letter-code labels; same atom color coding as for the hexapeptide except for the C atoms, which are colored green) participating in subsite shaping and substrate−product binding. The protein ligands of the catalytic zinc ion (blue sphere) have been omitted for clarity.

the crystal through a total of six crystal contacts with a neighboring hCPA4 molecule.

Cleavage and Inhibitory Assays in Vitro. The sequence of the active-site-bound hexapeptide corresponds to a region



B

	IC_{50} values (μM)	
Peptide	bCPA1	hCPA4
FNRPVD	530	910
FNRAVD	1700	720
FNRAVV	530	630
FNRPVV	560	710
FNRPV	NM	4000-6000
GNRPVT	830	NM
GNRPT	NM	2500

FIGURE 2: hCPA4 proteolytic activity and inhibition by peptides. (A) MALDI-TOF analysis of the time course of hCPA4-promoted cleavage of peptides Phe-Asn-Arg-Pro-Val-Asp (left) and Phe-Asn-Arg-Pro-Val-Val (right). The molecular masses of some peaks are indicated, and they are labeled according to the sequence that corresponds to their molecular mass. Digestion times are shown at the left of the spectra. Na^+ indicates that the molecular mass corresponds to the peptide−Na^+ ion pair. (B) IC_{50} values for a series of peptides assayed with bCPA1 and hCPA4. Activity measurements were taken with the chromogenic substrate *N*-(4-methoxyphenyl-azoformyl)-L-Phe-OH in 50 mM Tris-HCl and 0.1 M NaCl (pH 7.5). NM stands for not measurable.

where the activation domain folds in a loop and the beginning of strand β3 (*39*). Mapping hCPA4 for tryptic digestion indicates that there are two putative cleavage sites for trypsin, 4 and 12 residues upstream of the position of the hexapeptide sequence, respectively. Thus, the hexapeptide detected is likely to be a mature product of a series of cleavage events of undetermined sequence within the prodomain, or, alternatively, a byproduct due to a minor chymotryptic activity present in the trypsin used for activation. To test the hypothesis of peptide uptake and cleavage, the ability of hCPA4 to cleave the C-terminal residue of the molecule bound at the enzyme active site was examined with a synthetic hexapeptide identical to the one found in the crystal (Phe-Asn-Arg-Pro-Val-Asp). In addition, a second hexapeptide with a C-terminal valine instead of an aspartate was used to fit the preference of A-type MCPs for aliphatic residues. Figure 2A depicts some points of the time course of peptide degradation as followed by MS analysis. The enzyme is able to excise the C-terminal residues of both peptides, though, as expected, with a much higher efficiency in the case of the Phe-Asn-Arg-Pro-Val-Val substrate; while approximately half of this peptide was cleaved after 90 min and degradation is seen to proceed further to generate a tetrapeptide, much longer digestion times (i.e., 24 h) are needed to clearly observe excision of the C-terminal aspartate of Phe-Asn-Arg-Pro-Val-Asp. Interestingly, the main peaks for both the uncleaved and cleaved forms of the latter peptide correspond to their ion pairs with sodium cations in the solution. Preservation of the analyte solution charge state and detection of peptide−sodium ion pairs have been described in MS analysis (*52*). A detailed analysis of the time course degradation of the peptide Phe-Asn-Arg-Pro-Val-Val shows that, although the MALDI-TOF analysis lacks quantitative precision, a decrease in the amount of the resulting pentapeptide occurred with time, suggesting some kind of enzyme inhibition.

Activity measurements with chromogenic substrates of the enzyme incubated with increasing quantities of both hexapeptides gave the IC_{50} values that are shown in Figure 2B. Both peptides behave as inefficient inhibitors of hCPA4 in the high micromolar range, with their common pentapeptide product, Phe-Asn-Arg-Pro-Val, being the least inhibitory species. The observable, albeit low, inhibitory activity of the Phe-Asn-Arg-Pro-Val-Asp peptide may explain both the persistence of the double-product complex and its stability during purification and crystallization procedures. These studies were complemented with the inhibitory analysis of further pentapeptides and hexapeptides (see Figure 2B) that confirmed (i) the significantly higher inhibitory activity of hexapeptides over pentapeptides, (ii) the requirement of a bulky aromatic residue in P_5, and (iii) the fact that proline is not an absolute requirement in P_2.

Overall Structure of hCPA4. The mature hCPA4 enzyme has a compact globular shape and shows the classic α/β-hydrolase fold of A/B- and N/E-type MCPs. It has a central eight-stranded β-sheet made up of strands $\beta5-\beta12$ (numbering of the regular secondary structure elements according to the zymogen; see ref *39*) of strand connectivity $+1,+2,-1x,-2x,-2,+1x,-2$ according to ref *53*. Metalloprote(in)-ases (MPs) are best shown in a canonical standard orientation

(*54*), i.e., with the view onto the active-site groove and the nonprimed side on the left of the zinc ion (Figure 1B). From this perspective, the sheet vertically spans the molecule from the bottom to the top accumulating a vertical clockwise twist of ~135°. This gives rise to a concave front side of the sheet that accommodates helical segments α8, α10, and α11, as well as the active-site cleft (Figure 1B). At the rear, the convex side harbors sheet-adjacent helical segments α4−α7, α9, and α12 and the superficial N- and C-termini of the molecule. The access to the active site is reminiscent of a funnel (*17*, *39*), whose rim is shaped by a series of irregular loop segments of varying length. These segments, i.e., the loop connecting strand *β*12 with helix α12 (L*β*12α12), Lα10α11, L*β*7α5, L*β*10α9, and the 53-residue loop Lα7α8, are required for interactions of the CP moiety with the prodomain and cognate protein inhibitors (*17*, *39*). Lα7α8 is engaged in closing the front and bottom of the active site and the specificity pocket that form the characteristic cul-de-sac of exopeptidases (Figure 1B,C). This segment contains an N-glycosylation site at Asn148 (*17*), and its conformation is stabilized by a disulfide bond (Cys138−Cys161). The catalytic zinc ion (Zn999) resides at the bottom of the funnel-like cleft and is coordinated by His69 N*δ*1 (2.07 Å distant) and Glu72 in a slightly asymmetric bidentate manner (Zn999−Glu72 O*ϵ*2, 2.19 Å; Zn999−Glu72 O*ϵ*1, 2.63 Å). These two residues are provided by L*β*7α5 and are imbedded in a consensus sequence, HXXE [amino acid one-letter code; X for any residue (*55*)], characteristic of A/B- and N/E-type MCPs. The third hCPA4 zinc ligand is His196 N*δ*1 (2.09 Å).

As the hCPA4 structure closely resembles the bCPA structure (55% identical sequence for the respective zymogens), they can be considered equivalent in mechanistic terms. The numbering adopted for the human enzyme is equivalent to that of bCPA (see ref *39*) and has only a one-residue insertion between Gly55 and Gly56, termed Lys55A. Comparison of the hCPA4−hexapeptide complex with the complex between bCPA and the phosphonate inhibitor ZFV^{P}-(O)F [PDB entry 7CPA (*26*)], hereafter taken as a model for the reaction intermediate structure (see below and stage II in Scheme 1), shows that the 304 common Cα atoms deviating less than 3 Å show a rmsd of 0.65 Å upon superimposition, indicative of a close structural relationship. Merely the N-terminus, loop L*β*6*β*7, and the Pro134−Ser137 and Gly153−Gly155 segments (Lys153−Gly155 in bCPA) of Lα7α8 show somewhat larger deviations of up to 2.3 Å, although none of them affects the active-site groove.

A Bound Hexapeptidic Product. The pentapeptidic part of the hexapeptide bound to the nonprimed site of the groove adopts an approximately extended conformation for residues in P_1 (Val5), P_2 (Pro4), and P_3 (Arg3), reaching the molecular surface with the third residue (Figure 1B,C,E). Unlike other MPs, anchoring is not accomplished through inter-main chain *β*-ribbonlike interactions with a *β*-strand of the enzyme (*56*). Here, binding occurs through recognition of the peptide main chain by enzyme side chains (see below). The peptide chain trace undergoes a sharp turn in position P_4 (Asn2) due to a rotation of ~170° around the Ψ angle of the latter residue (Figure 1B,C,E). Thus, the peptide folds back to creep along the molecular surface of the enzyme. The C-terminus of the pentapeptide is bound to the catalytic zinc ion in a bidentate manner (Val5 O−Zn999, 2.32 Å; Val5 OT−Zn999, 2.21 Å), giving rise, together with the protein ligands, to the cation being coordinated in a distorted trigonal-bipyramidal fashion by four oxygen and two nitrogen atoms. Val5 O is also within binding distance of atoms O*ϵ*1 (2.96 Å) and O*ϵ*2 (2.90 Å) from the general base/acid Glu270, suggesting that either the latter residue or Val5 O is protonated; i.e., the second proton transfer to the α-amino group of Asp6 in $S_{1'}$ has not yet occurred (see Scheme 1). Val5 OT, in turn, establishes a salt bridge with Arg127 N*η*2 (2.85 Å), a residue that stabilizes the negatively charged reaction intermediate as seen in the ZFV^{P}(O)F complex (see below). The P_1 side chain of Val5 is placed in the S_1 subsite shaped in hCPA4 and in bCPA by the side chains of Phe279 and Tyr198. Val5 N is bound by Ty248 O*η* (2.77 Å). The preceding residue in P_2 (Pro4) shows its main chain torsion angles in the *β*-region of a Ramachandran plot (Φ = −66°; Ψ = 157°), and its carbonyl oxygen is bound by Arg71 N*η*2 (3.06 Å) and, more weakly, by Arg127 N*η*2 (3.28 Å). The P_2 side chain is accommodated in shallow subsite S_2, shaped by the apolar side chain parts of Arg127, Glu163, and Val164 (Thr164 in bCPA). Arg3, the peptide residue in P_3 (Φ = −130°; Ψ = 154°, also in the *β*-region), is bound via its carbonyl oxygen and its α-amino group with Tyr198 O*η* through a solvent-mediated (Hoh697) hydrogen bond. The P_3 side chain protrudes away, running down along the enzyme surface, and is held in place by Glu163 (Arg3 N*η*1−Glu163 O*ϵ*1, 3.09 Å), which adopts a novel conformation (see below). Interestingly, the position equivalent to 163 is always acidic in the reported CP sequences, except for human CPO, where a glutamine is found. The side chain of Arg71, Leu125, and, to a lesser extent, Phe279 further contribute to the shallow S_3 site. P_4 residue Asn2 points to the bulk solvent, while Phe1, featuring P_5, nestles snugly into a mainly hydrophobic pocket formed by the aromatic part of Tyr198, Tyr248, Val247 (Ile247 in bCPA), and Ser199, as well as the side chain of the product P_1 residue Val5. The participation of Ser199 and Tyr198 in subsite shaping may provide an explanation for (i) the main chain angles of the former residue systematically corresponding to disallowed or generously allowed regions of a Ramachandran plot and (ii) Ser197 and Tyr198 being linked by a conserved cis peptide bond.

On the opposite side of the groove, Asp6 fits into the $S_{1'}$ or specificity pocket, which is conceived to accommodate hydrophobic side chains. This residue binds the primary substrate-anchoring residue in MCPs, Arg145, through a bidentate salt bridge with its *β*-carboxylate group instead of its α-carboxylate group (Asp6 O*δ*2−Arg145 N*η*2, 2.77 Å; Asp6 O*δ*1−Arg145 N*η*1, 2.89 Å). Asp6 O*δ*2 is further anchored to Tyr248 O*η* (2.58 Å) and Arg127 N*η*1 (3.11 Å), while Asp6 O*δ*1 additionally contacts Asn144 N*δ*2 (2.92 Å). These interactions are usually performed by the α-carboxylate group in non-aspartate C-terminal residues of substrates (see refs *17*, *26*, and *34* for examples). As Asp6 is present with multiple conformations for its α-amino and -carboxylate groups, it can be stated in merely tentative terms that interactions of these groups may occur with Glu270 O*ϵ*2, Asn144 N*δ*2, and Thr268 O*γ*1. In accordance with previous descriptions for bCPA, the $S_{1'}$ pocket is delimited further in hCPA4 by the side chains of Asp194 (Ser194 in bCPA), Met203 (Leu203), Thr243 (Ile243), Val247 (Ile247), Ala250, Ser253 (Gly253), Ser254, Ile255, and Thr268. These residues

render a dead-end pocket characteristic of exopeptidases.

Among these residues, Asp194 may be hallmarked as being unique for hCPA4, since its position is always occupied by a serine or a threonine in mammalian A/B-MCPs (see ref *39*). Here, it is simultaneously strongly bound to one of the zinc-binding protein ligands (Asp194 Oδ2−His196 Nϵ2, 2.60 Å) and to Ser254, at the bottom of the cleft (Asp194 Oδ2−Ser254 Oγ, 2.73 Å). This arrangement is reminiscent of the catalytic triad of serine proteases, with the relative positions of the aspartate and the histidine residues swapped. The strong interaction between the latter two residues may contribute to the enhancement of the nucleophilicity of the zinc ligand, which, in turn, affects the catalytic efficiency of hCPA4, when compared with bCPA. One of the few instances in which position 194 is occupied by an aspartate is in insect CPA from *Helicoverpa armigera*. Interestingly, this enzyme exhibits broad substrate specificity and excises C-terminal acidic residues (*57*) with an efficiency comparable to that of a human CPB specifically engineered to display such specificity (*58*).

Comparison with a Stage III Model. As reported, structural information about the reaction intermediate of MCP catalysis relies on complexes with phosphinyl-derived peptidomimetic compounds. A search through the Cambridge Structural Database for small-molecule structures (*59*) with the motif C-PO_2-[O,NH,CH_2]-C found that phosphinyl P−O bonds span 1.49 ± 0.04 Å (n = 50), while the bonds mimicking the scissile peptide bond in the compounds found, P−O, P−N, and P−C, span 1.60 ± 0.03 (n = 25), 1.69 (n = 1), and 1.81 ± 0.01 Å (n = 2), respectively. In contrast, the only available structure mimicking a tetrahedral reaction intermediate carbon, 1,1,1,3,3,3-hexafluoropropanediolate {[$(CF_3)_2$-C(O)-OH]$^-$}, evinces C−C bonds that are 1.54 ± 0.01 Å long (n = 2) and C−O bonds that are 1.38 ± 0.02 Å long (n = 2) (*60*). Despite these differences in bond lengths and the pentavalence of phosphorus, phosphinyl-based compounds are widely accepted as valid models for a reaction intermediate and provide an unrivaled aid to the understanding of enzyme catalysis and specificity (*61*) due to the similarities in hybridization and geometry (*29*, *62*). Furthermore, they constitute a validated scaffold for the development of potent inhibitors of zinc proteases (*63*). Among the three compound classes, phosphonamidates are less well suited due to the intrinsic instability of the P−N bond (*61*, *63*). Moreover, phosphinates should be preferred over phosphonates due to the reported capacity of the phosphinyl-adjacent methylene to participate in substratelike hydrogen bonds (*29*), thus representing a closer model for a substrate.

Analysis of the ZFV^p(O)F complex reveals that the phosphinyl group binds the catalytic zinc ion in a bidentate manner as suggested for the transition-state diolate (Figure 1D) (*26*). The adjacent phenyllactate group properly mimics a substrate comprising a C-terminal phenylalanine, and the C-terminus is anchored to the side chains of Arg145 and Tyr248. On the opposite side of the phosphinyl group, structural parts mimicking a P_1 and a P_2 side chain allow proper mapping of the cognate S_1 and S_2 sites of the enzyme and coincide with this study (Figure 1D). However, the distal benzyloxycarbonyl (Z) moiety theoretically mimicking P_3 invades the space occupied by the side chain in P_5 in our hexapeptide complex (see Figure 1D). While this accounts for the hydrophobic nature of S_5, comparison of the complexes strongly suggests that the absence of residues and/or atoms mimicking the upstream part of a substrate leads the Z moiety to rotate ∼180° around the bond equivalent to the Arg3−Pro4 peptide bond (Figure 1D). This rotation entails the aromatic Z group occupying S_5.

A further noteworthy point is the novel conformation of the Glu163 side chain due to, we believe, a hitherto unseen interaction with the side chain of the residue nestling in S_3. This interaction is particularly important, as it involves the acidic side chain rotating anticlockwise ∼80° around χ_2, in comparison with the reaction-intermediate-mimicking ZFV^p-(O)F complex of bCPA. This conformation of Glu163 is buttressed by a bidentate salt bridge with Arg71 (Glu163 Oϵ1−Arg71 Nη1, 3.06 Å; Glu163 Oϵ2−Arg71 Nη2, 2.89 Å) and can be envisaged as being pivotal for P_3 recognition.

An Updated Reaction Mechanism for MCPs. The structure described here, along with the number of structural studies performed on MCPs, enables us to add some details to the consensus mechanism, in particular for secondary players, and to provide an update (see Scheme 1). As reported, the zinc ion coordinates the scissile carbonyl oxygen atom and the catalytic solvent molecule polarized by Glu270. Substrate binding further entails that Tyr248 makes a large movement from a superficial "up" position in the unbound state to a "down" position (*34*, *47*, *64*, *65*), although there is some controversy about this point (*66*). In the down position, the tyrosine hydroxyl group hydrogen bonds the amide nitrogen atom of the peptide bond between P_1 and P_2 and the C-terminal carboxylate group (*26*). This group is further anchored to the active-site cleft by a bidentate salt bridge with Arg145 and by a hydrogen bond with Asn144 Nδ2. This process can be described in four steps. (i) In the first step, Glu270 acts as a general base and abstracts a proton from the catalytic solvent molecule, which conducts a nucleophilic attack on the scissile carbonyl carbon atom. (ii) This leads to a negatively charged tetrahedral reaction intermediate that is stabilized by the side chain of Arg127 (*67*) and possibly interacts in a bidentate manner with the zinc ion that becomes pentacoordinate (*26*). This basic residue replaces structural elements forming what is known as the "oxyanion hole", which stabilizes similar types of intermediates during cleavage reactions performed by other proteases (*68*). Glu270 subsequently acts as a general acid catalyst delivering the captured proton to the scissile amide nitrogen atom. A second proton is transferred to the amide nitrogen, and (iii) this results in the metal-bound tetrahedral intermediate collapsing to products that initially remain bound to the enzyme in the form of a double-product complex, described here for the first time. At this stage, the enzyme can perform the reverse reaction, i.e., the synthesis of a peptide bond (*24*). The new C-terminus of the upstream part of the substrate interacts with the catalytic zinc in a bidentate manner, and the amide nitrogen of the preceding peptide bond remains hydrogen-bonded to Tyr248 Oη. In turn, the novel N-terminus probably interacts with Glu270 Oϵ2 and the catalytic zinc ion, but the isolated amino acid in $S_{1'}$ could be free to reallocate within the cleft if it does not fit the preference of the enzyme for a hydrophobic side chain. (iv) The N-terminal product leaves, and the catalytic solvent molecule is replenished; the C-terminal product is still bound to the zinc ion, Glu270, and Arg145 (and

putatively to Asn144 and Tyr248) (*34*, *39*) featuring the single-product complex. Eventually, the amino acid in $S_{1'}$ leaves and the enzyme is poised for a new catalytic round.

ACKNOWLEDGMENT

We are grateful to Robin Rycroft for helpful discussions on the manuscript and acknowledge the help provided by EMBL and ESRF synchrotron local contacts during data collection.

REFERENCES

1. Willstätter, R., Waldschmidt-Leitz, E., Harden, A., Keilin, D., Haldane, J. B. S., Quastel, J. H., Meldrum, N. U., and Roughton, F. J. W. (1932) Discussion on recent advances in the study of enzymes and their action, *Proc. R. Soc. London, Ser. B 111*, 280−297.
2. Vendrell, J., Querol, E., and Avilés, F. X. (2000) Metallocarboxypeptidases and their protein inhibitors. Structure, function and biomedical properties, *Biochim. Biophys. Acta 1477*, 284−298.
3. Waldschmidt-Leitz, E., Ziegler, F., Schäffner, A., and Weil, L. (1931) Über die Struktur der Protamine. I. Protaminase und die Produkte ihrer Einwirkung auf Clupein und Salmin, *Hoppe-Seyler's Z. Physiol. Chem. 197*, 219−236.
4. Waldschmidt-Leitz, E., and Purr, A. (1929) Über Proteinase und Carboxy-Polypeptidase aus Pankreas. (XVII. Mitteilung zur Spezifizität tierischer Proteasen.), *Ber. Dtsch. Chem. Ges. 62*, 2217−2226.
5. Auld, D. S. (2004) 240. Carboxypeptidase A, in *Handbook of proteolytic enzymes* (Barrett, A. J., Rawlings, N. D., and Woessner, J. F., Jr., Eds.) 2nd ed., pp 812−821, Elsevier Academic Press, London.
6. Reznik, S. E., and Fricker, L. D. (2001) Carboxypeptidases from A to Z: Implications in embryonic development and Wnt binding, *Cell. Mol. Life Sci. 58*, 1790−1804.
7. Normant, E., Gros, C., and Schwartz, J. C. (1995) Carboxypeptidase A isoforms produced by distinct genes or alternative splicing in brain and other extrapancreatic tissues, *J. Biol. Chem. 270*, 20543−20549.
8. Miller, L. A., Cochrane, D. E., Feldberg, R. S., and Carraway, R. E. (1998) Inhibition of neurotensin-stimulated mast cell secretion and carboxypeptidase A activity by the peptide inhibitor of carboxypeptidase A and neurotensin-receptor antagonist SR 48692, *Int. Arch. Allergy Immunol. 116*, 147−153.
9. Metz, M., Piliponsky, A. M., Chen, C. C., Lammel, V., Abrink, M., Pejler, G., Tsai, M., and Galli, S. J. (2006) Mast cells can enhance resistance to snake and honeybee venoms, *Science 313*, 526−530.
10. Neurath, H. (1960) Carboxypeptidases A and B, in *Enzymes* (Boyer, P. D., Lardy, H. A., and Myrback, K., Eds.) pp 11−36, Academic Press, New York.
11. Huang, H., Reed, C. P., Zhang, J. S., Shridhar, V., Wang, L., and Smith, D. I. (1999) Carboxypeptidase A3 (CPA3): A novel gene highly induced by histone deacetylase inhibitors during differentiation of prostate epithelial cancer cells, *Cancer Res. 59*, 2981−2988.
12. Kayashima, T., Yamasaki, K., Yamada, T., Sakai, H., Miwa, N., Ohta, T., Yoshiura, K., Matsumoto, N., Nakane, Y., Kanetake, H., Ishino, F., Niikawa, N., and Kishino, T. (2003) The novel imprinted carboxypeptidase A4 gene (CPA4) in the 7q32 imprinting domain, *Hum. Genet. 112*, 220−226.
13. Uren, J. R., and Neurath, H. (1972) Mechanism of activation of bovine procarboxypeptidase AS5. Alterations in primary and quaternary structure, *Biochemistry 11*, 4483−4492.
14. Anson, M. L. (1937) Carboxypeptidase. II. Partial purification of pro-carboxypeptidase, *J. Gen. Physiol. 20*, 777−780.
15. Reverter, D., Vendrell, J., Canals, F., Horstmann, J., Avilés, F. X., Fritz, H., and Sommerhoff, C. P. (1998) A carboxypeptidase inhibitor from the medical leech *Hirudo medicinalis*. Isolation, sequence analysis, cDNA cloning, recombinant expression, and characterization, *J. Biol. Chem. 273*, 32927−32933.
16. Arolas, J. L., Lorenzo, J., Rovira, A., Castella, J., Avilés, F. X., and Sommerhoff, C. P. (2004) A carboxypeptidase inhibitor from the tick *Rhipicephalus bursa*. Isolation, cDNA cloning, recombinant expression, and characterization, *J. Biol. Chem. 280*, 3441−3448.
17. Pallarès, I., Bonet, R., García-Castellanos, R., Ventura, S., Avilés, F. X., Vendrell, J., and Gomis-Rüth, F. X. (2005) Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin, *Proc. Natl. Acad. Sci. U.S.A. 102*, 3978−3983.
18. Mao, S. S., Colussi, D., Bailey, C. M., Bosserman, M., Burlein, C., Gardell, S. J., and Carroll, S. S. (2003) Electrochemiluminescence assay for basic carboxypeptidases: Inhibition of basic carboxypeptidases and activation of thrombin-activatable fibrinolysis inhibitor, *Anal. Biochem. 319*, 159−170.
19. Normant, E., Martres, M. P., Schwartz, J. C., and Gros, C. (1995) Purification, cDNA cloning, functional expression, and characterization of a 26-kDa endogenous mammalian carboxypeptidase inhibitor, *Proc. Natl. Acad. Sci. U.S.A. 92*, 12225−12229.
20. Polgár, L. (1999) Basic kinetic mechanisms of proteolytic enzymes, in *Proteolytic enzymes: Tools and targets* (Sterchi, E. E., and Stöcker, W., Eds.) pp 148−166, Springer-Verlag, Berlin.
21. Hartsuck, J. A., and Lipscomb, W. N. (1971) Carboxypeptidase A, in *The Enzymes. Hydrolysis: Peptide bonds* (Boyer, P. D., Ed.) 3rd ed., pp 1−56, Academic Press, New York.
22. Auld, D. S. (1987) Acyl group transfer: Metalloproteinases, in *Enzyme Mechanisms* (Page, M. I., and Williams, A., Eds.) pp 241−258, Royal Society of Chemistry, London.
23. Matthews, B. W. (1988) Structural basis of the action of thermolysin and related zinc peptidases, *Acc. Chem. Res. 21*, 333−340.
24. Auld, D. S. (2004) 68. Catalytic mechanisms for metallopeptidases, in *Handbook of proteolytic enzymes* (Barrett, A. J., Rawlings, N. D., and Woessner, J. F., Jr., Eds.) 2nd ed., pp 268−289, Elsevier Academic Press, London.
25. Christianson, D. W., and Lipscomb, W. N. (1989) Carboxypeptidase A, *Acc. Chem. Res. 22*, 62−69.
26. Kim, H., and Lipscomb, W. N. (1991) Comparison of the structures of three carboxypeptidase A−phosphonate complexes determined by X-ray crystallography, *Biochemistry 30*, 8171−8180.
27. Schechter, I., and Berger, A. (1967) On the size of active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun. 27*, 157−162.
28. Abramowitz, N., Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. II. Carboxypeptidase-A, *Biochem. Biophys. Res. Commun. 29*, 862−867.
29. Dive, V., Georgiadis, D., Matziari, M., Makaritis, A., Beau, F., Cuniasse, P., and Yiotakis, A. (2004) Phosphinic peptides as zinc metalloproteinase inhibitors, *Cell. Mol. Life Sci. 61*, 2010−2019.
30. Kim, H., and Lipscomb, W. N. (1990) Crystal structure of the complex of carboxypeptidase A with a strongly bound phosphonate in a new crystalline form: Comparison with structures of other complexes, *Biochemistry 29*, 5546−5555.
31. Kaplan, A. P., and Bartlett, P. A. (1991) Synthesis and evaluation of an inhibitor of carboxypeptidase A with a K_i value in the femtomolar range, *Biochemistry 30*, 8165−8170.
32. Cappalonga, A. M., Alexander, R. S., and Christianson, D. W. (1992) Structural comparison of sulfodiimine and sulfonamide inhibitors in their complexes with zinc enzymes, *J. Biol. Chem. 267*, 19192−19197.
33. Park, J. D., Kim, D. H., Kim, S. J., Woo, J. R., and Ryu, S. E. (2002) Sulfamide-based inhibitors for carboxypeptidase A. Novel type transition state analogue inhibitors for zinc proteases, *J. Med. Chem. 45*, 5295−5302.
34. Guasch, A., Coll, M., Avilés, F. X., and Huber, R. (1992) Three-dimensional structure of porcine pancreatic procarboxypeptidase A. A comparison of the A and B zymogens and their determinants for inhibition and activation, *J. Mol. Biol. 224*, 141−157.
35. Rees, D. C., and Lipscomb, W. N. (1980) Structure of the potato inhibitor complex of carboxypeptidase A at 2.5-Å resolution, *Proc. Natl. Acad. Sci. U.S.A. 77*, 4633−4637.
36. Reverter, D., Fernandez-Catalan, C., Baumgartner, R., Pfander, R., Huber, R., Bode, W., Vendrell, J., Holak, T. A., and Aviles, F. X. (2000) Structure of a novel leech carboxypeptidase inhibitor determined free in solution and in complex with human carboxypeptidase A2, *Nat. Struct. Biol. 7*, 322−328.
37. Arolas, J. L., Popowicz, G. M., Lorenzo, J., Sommerhoff, C. P., Huber, R., Avilés, F. X., and Holak, T. A. (2005) The three-dimensional structures of tick carboxypeptidase inhibitor in complex with A/B carboxypeptidases reveal a novel double-headed binding mode, *J. Mol. Biol. 350*, 489−498.
38. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) Three-dimensional structure of potato carboxypeptidase inhibitor in solution. A study using nuclear magnetic resonance,

distance geometry, and restrained molecular dynamics, *Biochemistry 26*, 8012−8023.
39. García-Castellanos, R., Bonet-Figueredo, R., Pallarès, I., Ventura, S., Aviles, F. X., Vendrell, J., and Gomis-Rüth, F. X. (2005) Detailed molecular comparison between the inhibition mode of A/B-type carboxypeptidases in the zymogen state and by the endogenous inhibitor latexin, *Cell. Mol. Life Sci. 62*, 1996−2014.
40. Leslie, A. G. W. (1999) Integration of macromolecular diffraction data, *Acta Crystallogr. D55*, 1696−1702.
41. Evans, P. (1993) Data reduction, in *Data collection and processing: Proceedings of the CCP4 Study Weekend 29−30 January 1993* (Sawyer, L., Isaacs, N., and Bailey, S., Eds.) pp 114−122, SERC Daresbury Laboratory, Warrington, U.K.
42. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
43. Navaza, J. (1994) AMoRe: An automated package for molecular replacement, *Acta Crystallogr. A50*, 157−163.
44. Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite, *Acta Crystallogr. D59*, 1131−1137.
45. Carranza, C., Inisan, A.-G., Mouthuy-Knoops, E., Cambillau, C., and Roussel, A. (1999) Turbo-Frodo, in *AFMB Activity Report 1996−1999*, pp 89−90, CNRS-UPR 9039, Marseille, France.
46. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240−255.
47. García-Sáez, I., Reverter, D., Vendrell, J., Avilés, F. X., and Coll, M. (1997) The three-dimensional structure of human procarboxypeptidase A2. Deciphering the basis of the inhibition, activation and intrinsic activity of the zymogen, *EMBO J. 16*, 6906−6913.
48. Evans, S. V. (1993) SETOR: Hardware lighted three-dimensional solid model representations of macromolecules, *J. Mol. Graphics 11*, 134−138.
49. Dixon, M. M., and Matthews, B. W. (1989) Is γ-chymotrypsin a tetrapeptide acyl-enzyme adduct of α-chymotrypsin? *Biochemistry 28*, 7033−7038.
50. Harel, M., Su, C. T., Frolow, F., Silman, I., and Sussman, J. L. (1991) γ-Chymotrypsin is a complex of α-chymotrypsin with its own autolysis products, *Biochemistry 30*, 5217−5225.
51. Rees, D. C., and Lipscomb, W. N. (1982) Refined crystal structure of the potato inhibitor complex of carboxypeptidase A at 2.5 Å resolution, *J. Mol. Biol. 160*, 475−498.
52. Wang, P., Wesdemiotis, C., Kapota, C., and Ohanessian, G. (2007) The sodium ion affinities of simple di-, tri, and tetrapeptides, *J. Am. Soc. Mass Spectrom. 18* (in press).
53. Richardson, J. S. (1981) The anatomy and taxonomy of protein structure, *Adv. Protein Chem. 34*, 167−339.
54. Gomis-Rüth, F. X., Stöcker, W., Huber, R., Zwilling, R., and Bode, W. (1993) Refined 1.8 Å X-ray crystal structure of astacin, a zinc-endopeptidase from the crayfish *Astacus astacus* L. Structure determination, refinement, molecular structure and comparison with thermolysin, *J. Mol. Biol. 229*, 945−968.
55. Hooper, N. M. (1994) Families of zinc metalloproteases, *FEBS Lett. 354*, 1−6.
56. Gomis-Rüth, F. X. (2003) Structural aspects of the *metzincin* clan of metalloendopeptidases, *Mol. Biotechnol. 24*, 157−202.
57. Bayes, A., Sonnenschein, A., Daura, X., Vendrell, J., and Avilés, F. X. (2003) Procarboxypeptidase A from the insect pest *Helicoverpa armigera* and its derived enzyme. Two forms with new functional properties, *Eur. J. Biochem. 270*, 3026−3035.
58. Edge, M., Forder, C., Hennam, J., Lee, I., Tonge, D., Hardern, I., Fitton, J., Eckersley, K., East, S., Shufflebotham, A., Blakey, D., and Slater, A. (1998) Engineered human carboxypeptidase B enzymes that hydrolyse hippuryl-L-glutamic acid: Reversed-polarity mutants, *Protein Eng. 11*, 1229−1234.
59. Allen, F. H. (2002) The Cambridge Structural Database: A quarter of a million crystal structures and rising, *Acta Crystallogr. B58*, 380−388.
60. Roesky, H. W., Lucas, J., Keller, K., Dhathathreyan, K. D., Noltemeyer, M., and Sheldrick, G. M. (1985) Reaktionen von Hexafluoraceton mit Alkalicyanaten, *Chem. Ber. 118*, 2659−2670.
61. Hanson, J. E., Kaplan, A. P., and Bartlett, P. A. (1989) Phosphonate analogues of carboxypeptidase A substrates are potent transition-state analogue inhibitors, *Biochemistry 28*, 6294−6305.
62. Bartlett, P. A., and Marlowe, C. K. (1987) Evaluation of intrinsic binding energy from a hydrogen bonding group in an enzyme inhibitor, *Science 235*, 569−571.
63. Yiotakis, A., Lecoq, A., Nicolaou, A., Labadie, J., and Dive, V. (1994) Phosphinic peptide analogues as potent inhibitors of *Corynebacterium rathayii* bacterial collagenase, *Biochem. J. 303* (Part 1), 323−327.
64. Reeke, G. N., Hartsuck, J. A., Ludwig, M. L., Quiocho, F. A., Steitz, T. A., and Lipscomb, W. N. (1967) The Structure of Carboxypeptidase A. VI. Some results at 2.0-Å resolution, and the complex with glycyl-tyrosine at 2.8-Å resolution, *Proc. Natl. Acad. Sci. U.S.A. 58*, 2220−2226.
65. Rees, D. C., Lewis, M., and Lipscomb, W. N. (1983) Refined crystal structure of carboxypeptidase A at 1.54 Å resolution, *J. Mol. Biol. 168*, 367−387.
66. Bukrinsky, J. T., Bjerrum, M. J., and Kadziola, A. (1998) Native carboxypeptidase A in a new crystal environment reveals a different conformation of the important tyrosine 248, *Biochemistry 37*, 16555−16564.
67. Phillips, M. A., Fletterick, R., and Rutter, W. J. (1990) Arginine 127 stabilizes the transition state in carboxypeptidase, *J. Biol. Chem. 265*, 20692−20698.
68. Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H. Y., and Poulos, T. (1986) Site-directed mutagenesis and the role of the oxyanion hole in subtilisin, *Proc. Natl. Acad. Sci. U.S.A. 83*, 3743−3745.
69. Weiss, M. S. (2001) Global indicators of X-ray quality, *J. Appl. Crystallogr. 34*, 130−135.
70. Evans, P. (2006) Scaling and assessment of data quality, *Acta Crystallogr. D62*, 72−82.
71. Gomis-Rüth, F. X., Companys, V., Qian, Y., Fricker, L. D., Vendrell, J., Avilés, F. X., and Coll, M. (1999) Crystal structure of avian carboxypeptidase D domain II: A prototype for the regulatory metallocarboxypeptidase subfamily, *EMBO J. 18*, 5817−5826.
72. Gomis-Rüth, F. X., Gómez, M., Bode, W., Huber, R., and Avilés, F. X. (1995) The three-dimensional structure of the native ternary complex of bovine pancreatic procarboxypeptidase A with pro-proteinase E and chymotrypsinogen C, *EMBO J. 14*, 4387−4394.
73. Pereira, P. J. B., Segura-Martín, S., Oliva, B., Ferrer-Orta, C., Avilés, F. X., Coll, M., Gomis-Rüth, F. X., and Vendrell, J. (2002) Human procarboxypeptidase B: Three-dimensional structure and implications for thrombin-activatable fibrinolysis inhibitor (TAFI), *J. Mol. Biol. 321*, 537−547.

BI700480B