Critical Reviews in Biochemistry and Molecular Biology, 43:319–345, 2008

ISSN: 1040-9238 print / 1549-7798 online
DOI: 10.1080/10409230802376375

informa
healthcare

Structure and Mechanism of Metallocarboxypeptidases

F. Xavier Gomis-Rüth
Molecular Biology Institute of Barcelona, CSIC, Barcelona, Spain

Metallocarboxpeptidases cleave C-terminal residues from peptide substrates and participate in a wide range of physiological processes, but they also contribute to human pathology. On the basis of structural information, we can distinguish between two groups of such metallopeptidases: *cowrins* and *funnelins*. Cowrins comprise protozoan, prokaryotic, and mammalian enzymes related to both neurolysin and angiotensin-converting enzyme and their catalytic domains contain 500–700 residues. They are ellipsoidal and traversed horizontally by a long, deep, narrow active-site cleft, in which the C-terminal residues are cut from oligopeptides and unstructured protein tails. The consensus cowrin structure contains a common core of 17 helices and a three-stranded β-sheet, which participates in substrate binding. This protease family is characterized by a set of spatially conserved amino acids involved in catalysis, HEXXH+EXXS/G+H+Y/R+Y. Funnelins comprise structural relatives of the archetypal bovine carboxypeptidase A1 and feature mammalian, insect and bacterial proteins with strict carboxypeptidase activity. Their ~300-residue catalytic domains evince a consensus central eight-stranded β-sheet flanked on either side by a total of eight helices. They also contain a characteristic set of conserved residues, HXXE+R+NR+H+Y+E, and their active-site clefts are rather shallow and lie at the bottom of a funnel-like cavity. Therefore, these enzymes act on a large variety of well-folded proteins. In both cowrins and funnelins, substrate hydrolysis follows a common general base/acid mechanism. A metal-bound solvent molecule ultimately performs the attack on the scissile peptide bond with the assistance of a strictly conserved glutamate residue.

Keywords peptide bond hydrolysis, three-dimensional X-ray crystal structure, zinc enzyme, catalysis

INTRODUCTION AND SCOPE

Proteolytic enzymes are essential for life. Catalyzing the essentially irreversible step of peptide bond hydrolysis, they participate in broad spectrum degradation of peptides and proteins, as observed during digestion, remodeling and resorption of tissue, and intracellular scavenging of mis-folded proteins and peptides. However, these enzymes also carry out highly selective cleavage of specific substrates, as observed during shedding and (in)activation of signal molecules, factors, peptide hormones, pro-proteins, and (pro-)enzymes. In this dual function, peptidases govern processes such as DNA replication, embryonic development and morphogenesis, angiogenesis and alimentary uptake, aging and apoptosis, cell-cycle regulation, bone formation, neuronal outgrowth, immune and inflammatory responses, and coagulation and fibrinolysis, among several others (Neurath and Walsh, 1976; López-Otín and Overall, 2002; Barrett *et al*., 2004). Peptidases are exquisitely modulated and deregulation due to either excessive or diminished activity may induce embryonic lethality, neurodegenerative and other diseases of the central nervous system, thrombotic and cardiovascular disorders, inflammation and autoimmune diseases, arthritis and osteoporosis, microbial infection and virulence, and cancer (Nagase, 2001; Barrett *et al*., 2004). Therefore, a better understanding of the architecture, interacting partners, and working mechanisms of these players, both *in vivo* and at the molecular level, is a requisite for the development of therapeutic strategies aimed at enhancing, palliating or reversing their destructive potential. Detailed three-dimensional structural information can contribute much to this understanding (Mittl and Grütter, 2006).

Peptidases are divided into endopeptidases and exopeptidases, which cut in the middle or at the ends of a protein or peptide substrate, respectively (Barrett *et al*., 2004). The latter bifurcate into aminopeptidases and carboxypeptidases, which cleave off N-terminal and C-terminal amino acid residues, respectively. Enzymes targeting C-terminal dipeptides, tripeptides, etc., of substrates are termed peptidyl dipeptidases, peptidyl tripeptidases, etc. In addition, peptidases are classified, according to the protein residues that are essential for catalysis, into serine, cysteine, aspartic acid, threonine, and glutamic acid proteases. A sixth class is provided by the metallopeptidases (MPs), which

Address correspondence to F. Xavier Gomis-Rüth, Prostelysis Lab Molecular Biology Institute of Barcelona, CSIC, Department of Structural Biology, Barcelona Science Park, Helix Building, c/Baldiri Reixac, 15–21, E-08028 Barcelona, Spain. E-mail: xgrcri@ibmb.csic.es

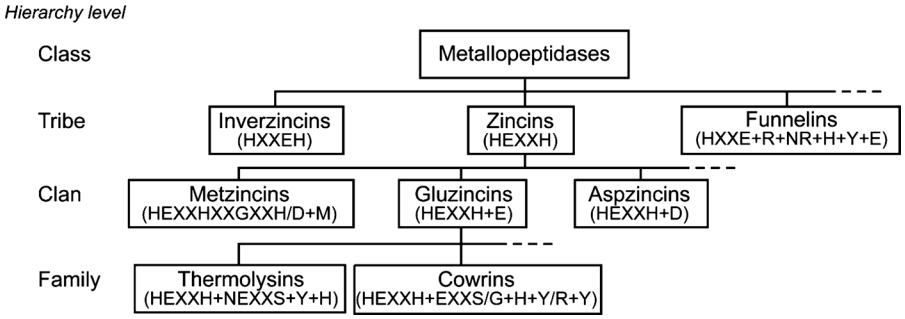


FIG. 1. Classification of MPs. Scheme depicting a possible subdivision of the MP class of enzymes based on features essential for metal-binding and catalysis. Only groupings and hierarchy levels relevant for the present discussion are displayed.

exert catalysis through one or two bound divalent metal ions, mostly zinc but also sometimes cobalt, nickel, or manganese (Hooper, 1994; Auld, 2004a). These enzymes elicit peptide-bond hydrolysis through a general base/acid mechanism, which positions a metal-bound solvent molecule for attack on the scissile peptide bond (Matthews, 1988; Auld, 2004a). For this to happen, the substrate must first be inserted into the active-site cleft with the side chain of the residues downstream of the scissile bond, termed P'_1, P'_2, P'_3, etc., fitting into the cognate enzyme sub-sites S'_1, S'_2, S'_3, etc. Analogously, the residue side chains upstream of the scissile bond (P_1, P_2, P_3, etc.) are inserted into the enzyme sub-sites S_1, S_2, S_3, etc. (Abramowitz *et al*., 1967; Hooper, 2002). In the case of aminopeptidases, the non-primed side is restricted to P_1/S_1 and in carboxypeptidases the primed side is limited to P'_1/S'_1. The latter enzymes recognize the C-termini of protein and peptide substrates through dedicated S'_1 or specificity pockets and through basic residues that trap the terminal carboxylate group. Substrate binding, as well as structural details of the catalytic metal-binding site and the active-site cleft, are best visualized by adopting a standard orientation for structural representations of MPs. This consists of a frontal view into the cleft, with the substrate binding horizontally from left (N-terminal non-primed side) to right (C-terminal primed side) of the catalytic metal ion (Gomis-Rüth *et al*., 1993).

Initial structural studies on MPs were performed on bovine pancreatic carboxypeptidase A1 (BCPA1) (Lipscomb *et al*., 1968) and *Bacillus thermoproteolyticus* thermolysin (Matthews *et al*., 1972). Since then, MPs have been the object of a number of structural studies. In 1993, the term *metzincin* was coined to refer to a clan of MPs composed of several families, which shared the presence of an extended zinc-binding consensus sequence, HEXXHXXGXXH/D (amino acid one-letter code; X for any residue). This motif comprises three metal-bonding protein residues, three histidines or two histidines plus an aspartate, and a general base/acid glutamate. In addition, a strictly conserved methionine was observed embedded in a *Met-turn* forming a hydrophobic pillow below the catalytic zinc (Bode *et al*., 1993; Stöcker *et al*., 1995; Gomis-Rüth, 2003). The name metzincin led to the creation of the higher-hierarchy terms *zincins* and *inverzincins* to refer to those tribes of MPs—each potentially made up of several clans—that harbored the short consensus sequences HEXXH and HXXEH, respectively, containing two zinc-binding histidines and a general base/acid glutamate (Jongeneel *et al*., 1989; Bode *et al*., 1993; Hooper, 1994) (Figure 1). Zincins were split, in addition to metzincins, into the *gluzincin* and *aspzincin* clans, which have an additional glutamate and aspartate, respectively, as the third zinc-binding residue downstream in the sequence (Hooper, 1994; Fushimi *et al*., 1999). Examples of families of the gluzincin clan are the *thermolysins* and the relatives of neurolysin/angiotensin-converting enzyme, hereafter referred to as *cowrins*. On the same hierarchy level as zincins and inverzincins are the relatives of BCPA1, hereafter *funnelins*, and those of DD-carboxypeptidase (Jiang and Bond, 1992; Hooper, 1994) (Figure 1).

Funnelins and cowrins contain MPs with carboxypeptidase activity, i.e. metallocarboxypeptidases (MCPs) (Coleman *et al*., 1994; Skidgel, 1996; Vendrell *et al*., 2000; Reznik and Fricker, 2001; Arolas *et al*., 2007), and they are the focus of the present study. Microbial peptidoglycan-processing enzymes such as DD-carboxypeptidases are not considered. In addition, relatives of γ-D-glutamyl-L-diamino acid endopeptidase 1, a remote structural relative of funnelins that participates in sporulation, have also been excluded. Finally, glutamate carboxypeptidases *alias* carboxypeptidase G enzymes, such as *Pseudomonas* sp. carboxypeptidase G2 and human prostate-specific membrane antigen, remove C-terminal glutamate residues, but act on non-peptidic substrates such as folic acid and its analog methotrexate. They share the overall folding pattern with funnelins, but they display a two-zinc catalytic site as found in some aminopeptidases (Rowsell *et al*., 1997; Davis *et al*., 2005; Mesters *et al*., 2006). Glutamate carboxypeptidases have also been omitted from the present study.

COWRINS

Carboxypeptidases, Peptidyl Dipeptidases, and Oligoendopeptidases

Cowrins agglutinate around the three most emblematic mammalian members, angiotensin-I-converting enzyme (alias peptidyl dipeptidase A and kininase II), neurolysin (alias

endopeptidase 24.16, mitochondrial oligopeptidase, neurotensin-degrading enzyme, and oligopeptidase M), and thimet oligopeptidase (alias endopeptidase 24.15, Pz-peptidase, and endo-oligopeptidase A). These three prototypes are currently ascribed to families M2 and M3 in the MEROPS protease database (http://merops.sanger.ac.uk; Rawlings *et al*., 2008). Angiotensin-I-converting enzyme (ACE) is present in humans as 140-kDa somatic and 77-kDa testicular variants. Somatic ACE is a type-I protein bound to the plasma membrane and it is heavily glycosylated. It consists of two similar 66-kDa catalytic domains with slight differences in activity and specificity originated by gene-duplication, an N-terminal (HACEN) and a C-terminal (HACEC) domain, followed by a transmembrane anchor and a short cytosolic tail (Soubrier *et al*., 1988). Testicular ACE is identical to HACEC except for an additional short glycosylated N-terminal extension (Ehlers *et al*., 1989). The somatic and testicular ACE variants arise from alternative tissue-specific transcription initiation (Soubrier *et al*., 1988) and both are shed from the cell surface via sheddase-mediated cleavage of their ectodomains (Woodman *et al*., 2006). ACE is a rather unspecific peptidyl dipeptidase that governs the renin-angiotensin-aldosterone system by processing the decapeptide angiotensin I to the potent vasopressor angiotensin II through removal of a C-terminal histidine-leucine dipeptide. ACE also processes the vasodilator bradykinin by removing its C-terminal proline-phenylalanine tail. Via these functions, the enzyme contributes to blood-pressure regulation and electrolyte homeostasis and it is a validated drug target for cardiovascular conditions and diabetic nephropathy (Sturrock *et al*., 2004). In addition, ACE also acts on neurotensin, gonadotropin-releasing hormone, luteinizing hormone-releasing hormone, substance P, and β-neoendorphin$_{1-9}$ (Natesh *et al*., 2003; Sturrock *et al*., 2004). A human ACE paralog is 90-kDa ACE2 (HACE2), which is also a glycosylated type-I integral membrane protein. Like ACE, it undergoes ectodomain shedding from the membrane, giving rise to a soluble form (Guy *et al*., 2008). In addition to a vasopeptidase, HACE2 is an essential regulator for heart function and it is the functional receptor for cell attachment and membrane fusion of the human coronavirus that causes severe acute respiratory syndrome. HACE2 has, in addition to an ACE-like catalytic domain with carboxypeptidase instead of peptidyl dipeptidase activity, a further C-terminal domain reminiscent of human collectrin. Identified substrates of HACE2 include apelin-13 and -36, ΔArg9-bradykinin, angiotensin II, and dynorphin A (Towler *et al*., 2004; Turner *et al*., 2004). AnCE is a single-domain homolog of ACE with a role in spermiogenesis in *Drosophila melanogaster* (DMACE). It hydrolyzes several ACE substrates and is inhibited by specific ACE inhibitors such as captopril and lisinopril (Cornell *et al*., 1995; Hurst *et al*., 2003; Kim *et al*., 2003).

Neurolysin (NEURO) is a soluble 78-kDa oligopeptidase that modulates the activity of peptide neurotransmitters and neurohormones in mammals. Like ACE, it is not sequence specific and it recognizes a certain number of sequences within oligopeptides such as dynorphin A, substance P, angiotensin II, bradykinin, and luteinizing hormone-releasing hormone *in vitro*. Its main substrate *in vivo* is the tridecapeptide neurotensin, a vasodilator implicated in the release of luteinizing hormone and prolactin and in the regulation of the dopaminergic system. Processing through NEURO leads to inactivation of neurotensin through cleavage between positions 10 and 11 of the peptide (Brown *et al*., 2001). The enzyme shares high sequence identity with thimet oligopeptidase (HSTOP), a broadly-distributed 77-kDa processor of bioactive oligopeptides such as neurotensin, bradykinin, somatostatin, opioids, and angiotensin I. It also degrades proteasome-released peptides (Saric *et al*., 2001) and amyloid β-peptide, a component of amyloid plaques in Alzheimer's disease (Yamin *et al*., 1999).

In addition to these mammalian and insect enzymes, cowrins comprise more distant relatives from unicellular organisms. These are a 78-kDa peptidyl dipeptidase from *Escherichia coli* (Dcp *alias* Dcp II; hereafter ECDCP; MEROPS family M3), which is susceptible to captopril and may be involved in intracellular degradation of peptides (Comellas-Bigler *et al*., 2005). Oligoendopeptidase F from *Bacillus stearothermophilus* (OPEPF) is a 62-kDa peptidase, which hydrolyzes leucine-glycine peptide bonds (Gerdts *et al*., 2006). Two hyperthermophilic relatives have been reported from the archaeon *Pyrococcus furiosus* (59 kDa; hereafter PFUCP) and the bacterial thermophile *Thermus thermophilus* (58 kDa; hereafter TTHCP). PFUCP has proven carboxypeptidase activity against a broad range of small fluorigenic substrates and natural peptides such as renin (Arndt *et al*., 2002; Cheng *et al*., 1999). Furthermore, two closely-related 58-kDa MCPs (64% sequence identity), TcMCP-1 (TCMCP) and -2 (64% sequence identity), are present in the protozoan responsible for Chagas' disease, *Trypanosoma cruzi*. They display carboxypeptidase activity towards fluorigenic substrates comprising a basic and an aromatic residue at P'_1, respectively. However, while TCMCP appears in all life cycles of *T. cruzi*, its paralog is mainly expressed in the stages that proceed within the insect vector (Niemirowicz and Cazzulo, 2007). PFUCP, TTHCP and TCMCP have all been assigned to MEROPS family M32 (Arndt *et al*., 2002; Niemirowicz *et al*., 2007).

Evolutionary and Structural Relationship among Cowrins

Several three-dimensional structures are available for cowrins, either unbound or in complex with another protein, a catalytic product or an inhibitor (Brown *et al*., 2001; Arndt *et al*., 2002; Kim *et al*., 2003; Natesh *et al*., 2003; Ray *et al*., 2004; Towler *et al*., 2004; Comellas-Bigler *et al*., 2005; Li *et al*., 2005; Corradi *et al*., 2006; Gerdts *et al*., 2006; Niemirowicz *et al*., 2008). In addition, the coordinates of the structures of oligopeptidase F from *Enterococcus faecium* and of a putative MP from *Chlamydophila abortus* have been deposited by a structural genomics consortium but not published (Table 1). Analysis of a subset of 11 representative proteins, HACEC, HACEN, HACE2, DMACE, HSTOP, NEURO, ECDCP, OPEPF, PFUCP, TCMCP, and TTHCP, reveals that cowrin catalytic domains are related by

TABLE 1
Proteins and associated PDB access codes relevant for the present study

Cowrins			
Neurolysin	*Rattus norvegicus*	NEURO	**1I1I**, 2O3E
Thimet oligopeptidase	*Homo sapiens*	HSTOP	**1S4B**, 2O36
Somatic ACE C-dom.	*H. sapiens*	HACEC	**1O8A**, 1O86, 2C6N, 2OC2, 2IUL, 2IUX, 1UZE, 1UZF
Somatic ACE N-dom.	*H. sapiens*	HACEN	**2C6F**
ACE2	*H. sapiens*	HACE2	**1R42**, 1R4L, 2AJF (complex with SARS CoV spike protein)
ANCE	*Drosophila melanogaster*	DMACE	1J36, 1J37, **1J38**
TcMCP-1	*Trypanosoma cruzi*	TCMCP	**3DWC**
Dcp	*Escherichia coli*	ECDCP	**1Y79**
Oligopeptidase F	*Bacillus stearothermophilus*	OPEPF	**2H1N**, 2H1J
Oligoendopeptidase F	*Enterococcus faecium*		2QR4
Putative peptidase	*Chlamydophila abortus*		3CE2
MCP	*Pyrococcus furiosus*	PFUCP	**1K9X**, 1KA2, 1KA4
MCP	Thermus thermophilus	***TTHCP***	**1WGZ**
Funnelins – A/B type			
CPA2*	*H. sapiens*	HCPA2	1AYE (zymogen), **1DTD** (complex with LCI), 1O6X (pro-domain)
CPA4(=CPA3)	*H. sapiens*	HCPA4	**2PCU** (with cleaved substrate), 2BOA (zymogen), 2BO9 (complex with latexin)
CPA1	*Bos taurus*	BCPA1	1BAV, 1ELL, 1ELM, 1EE3, 1HDU, 1CPX, 1CBX, 1HDQ,1ARL, 1ARM, 1CPS, 1F57, 1HDQ, 1HDU, 1HEE, 1IY7, 1M4L, 1YME, 1ZLH (complex with TCI), 2ABZ (complex with mutant LCI), 2CTB, 2CTC, 3CPA, 4CPA (complex with PCI), **5CPA**, 6CPA, 7CPA, 8CPA, 1PYT (zymogen in ternary complex)
TAFIa	*B. taurus*	BTAFI	**3D4U** (complex with TCI)
CPB1	*H. sapiens*	HCPB1	1KWM (zymogen), **1ZLI** (complex with TCI)
	Sus scrofa	*SSCPB*	**1Z5R**, 1CPB, 1ZG7, 1ZG8, 1ZG9, 1NSA (zymogen),
CPB	*Heliothis zea* (corn earworm)	HZCPB	**2C1C**
CPT	*Thermoactinomyces vulgaris*	TVCPT	**1OBR**
CPA	*Helicoverpa armigera* (cotton bollworm)	HACPA	**1JQG**
Funnelins – N/E type			
CPM	*H. sapiens*	HSCPM	**1UWY**
CPN	*H. sapiens*	HSCPN	**2NSM**
CPD (repeat II)	*Lophonetta specularioides* (crested duck)	DCPD2	1QMU, **1H8L**

Unless otherwise indicated, reported PDB entries contain the unbound wild-type active molecule, point mutants, metal-depleted or –substituted forms or small-molecule inhibitor complexes. The PDB entries selected for the present structural comparison of catalytic domains are displayed in **bold**. Their corresponding shorthand names used throughout the text are indicated.

*The structure of CPA2 from *R. norvegicus* has been reported, but the PDB access code indicated in the primary reference (Faming *et al*., 1991) currently corresponds to another protein.

genetic distance and sequence similarity, and also by structural similarity, especially the latter.

A phylogenetic analysis based on a structure-based multiple-sequence alignment (see legend to Figure 2) clusters ACE-like enzymes, which are separated by a genetic distance ranging from 605 to 1428 single-residue substitutions. Another cluster is formed by NEURO and HSTOP, which display the closest genetic distance with 397 substitutions. A third cluster contains the two thermophilic (PFUCP and TTHCP) and the protozoan (TCMCP) sequences (1318–1395 substitutions). The two most widely separated sequences correspond to the two remaining bacterial proteins, OPEPF and ECDCP, which are separated from their closest sequences by 3722 and 2274 substitutions, respectively (Figure 2b). Pairwise scores of the multiple structure-based sequence alignment confirm several of these findings (Figure 2a). Closest sequence similarity is found between HSTOP and NEURO (64% sequence identity; the root mean square deviation (rmsd) among topologically equivalent Cα atoms equals 1.4Å), among HACEC, HACEN and DMACE (52–38%; 1.3–1.7Å), and between PFUCP and TTHCP (36%; 1.9Å). A second, less similar group is made up by the pairs TCMCP and PFUCP (29%; 2.2Å) and TCMCP and TTHMP (31%; 2.3Å), as well as by HACE2 with either HACEC, HACEN or DMACE (29–25%; 3.5–3.7Å). The latter rmsd values are already beyond the threshold (3Å) at which two Cα positions can be considered structurally similar (Domingues *et al*., 2000). A third group, with sequence identities clearly below twilight values (20–35%; (Rost, 1999), is made up by the pairs between ECDCP and either NEURO or HSTOP (19–18%; 3.7–3.9Å). All the remaining pairs display sequence identity values below 10% and rmsd values above 3.6Å, i.e. they are not identified as structural relatives by purely sequence-based methods. This again confirms that structural relatedness is more strongly conserved than sequence identity (Chothia and Lesk, 1986).

Catalytic Domain of Cowrins

Viewed in standard orientation, cowrin catalytic domains are prolate spheroids of ~80 and ~60Å diameter. They resemble a conch shell or cowry, with the active-site cleft mimicking the aperture (Figures 3a, b). The cleft lies between two halves of similar size: an upper moiety (UM) above and a lower moiety (LM) below. They are connected by the cleft floor and conform the proteolytic sub-domain. In addition, an N-terminal sub-domain to the left of the UM participates in dimerization and cleft delimitation (dimerization sub-domain (DS); see also the two next sections, and Figures 3c, d). Structural superimposition of the 11 representatives studied renders a classical pairwise rmsd and a maximum-likelihood rmsd of 4.8Å and 1.1Å, respectively. They all share a common core comprising 17 α-helices and three β-strands (β1–β), and each has additional secondary-structure elements and segments that account for differences in polypeptide chain length, which ranges from ~500 (PFUCP) to ~700 residues (ECDCP).

The DS includes the first three helices (α1–α3), which are arranged as a roughly parallel α-hairpin (α1α2) and a subsequent helix (α3) running below the hairpin. Beside these consensus features, all cowrins except the ACE relatives display a further helix inserted between α2 and α3 ($\alpha2'$ in PFUCP, see Figure 3d). This extra helix, together with α3, gives rise to a second α-hairpin within DS, $\alpha2'\alpha$3, which is nearly coplanar with α1α2. In addition, OPEPF, ECDCP, HSTOP, and NEURO have a characteristic N-terminal helix preceding α1, which roughly parallels α2. Like the DS, the downstream LM is fully helical, and it forms the lower half of the molecule (Figures 3c, d). It starts with α-hairpin α4α5, which shapes the lower front of the molecule. The simple loop connecting the two helices at the tip of the hairpin found in PFUCP, TTHMP, and TCMCP (Figure 3d) is replaced in other cowrins by an extra β-hairpin and two (in ACE-like cowrins and OPEPF), three (in HSTOP and NEURO), or even four extra helices (in ECDCP). After α5, a loop connects helix α5 with helix α6 (Lα5α6), whose axis is approximately perpendicular to α-hairpin α4α5. This helix ends at the bottom of the active-site cleft on its non-primed side. Thereafter, Lα6α7, which is particularly long in HACE2, joins an extended helix, α7, which runs horizontally across the back surface reaching the right flank of the molecule. After α7, the polypeptide reaches the first crossing point between LM and UM with a varying segment in extended conformation, which runs along the right front. This connecting segment, Lα7α8, interacts with the likewise variable structure created by Lα4α5. It shows the largest deviation among cowrins and may include one (OPEPF) or two extra helices (NEURO, HSTOP, HACEC, HACEN, ECDCP, and DMACE). These variations give rise to unique active-site floors and walls on the primed side of the active-site clefts.

In contrast to the DS and the LM, the UM has $\alpha\beta$ topology and its first regular secondary structure element is α8, which crests the molecule from front to back (Figure 3d). In the minimal consensus structure, the ensuing loop (Lα8β1) leads to the uppermost strand (β1) of a three-stranded mixed β-sheet of connectivity +2, –1 (according to Richardson, 1981) with the strands linked by same-end connections. This causes the lowermost strand, β2, to run antiparallel to β1 and β3 and to the active-site cleft (Figures 3c–e). In ACE relatives, a short helical segment precedes β1 and, in ECDCP, HSTOP, and NEURO, the sheet is expanded on its cleft-distal side by a β-ribbon inserted between α8 and β1 to give rise to a five-stranded β-sheet of connectivity +1, +1, +2, –1. In all structures, strand β3 is connected to α9, the *active-site helix*, which traverses the molecule almost horizontally. Thereafter, the polypeptide chain enters α10, the *glutamate helix*, which runs below α9 from front to back. A subsequent rotation in the direction of the polypeptide chain entails that the next helix, α11, is nearly vertical and ends at the right top. Thereafter, the polypeptide chain leaves the UM to enter the second half of the LM with α-hairpin α12α13, whose tip creates the bottom of the molecule (Figure 3d). All cowrins display distinct additional helices after α11 and/or α13. After α13,

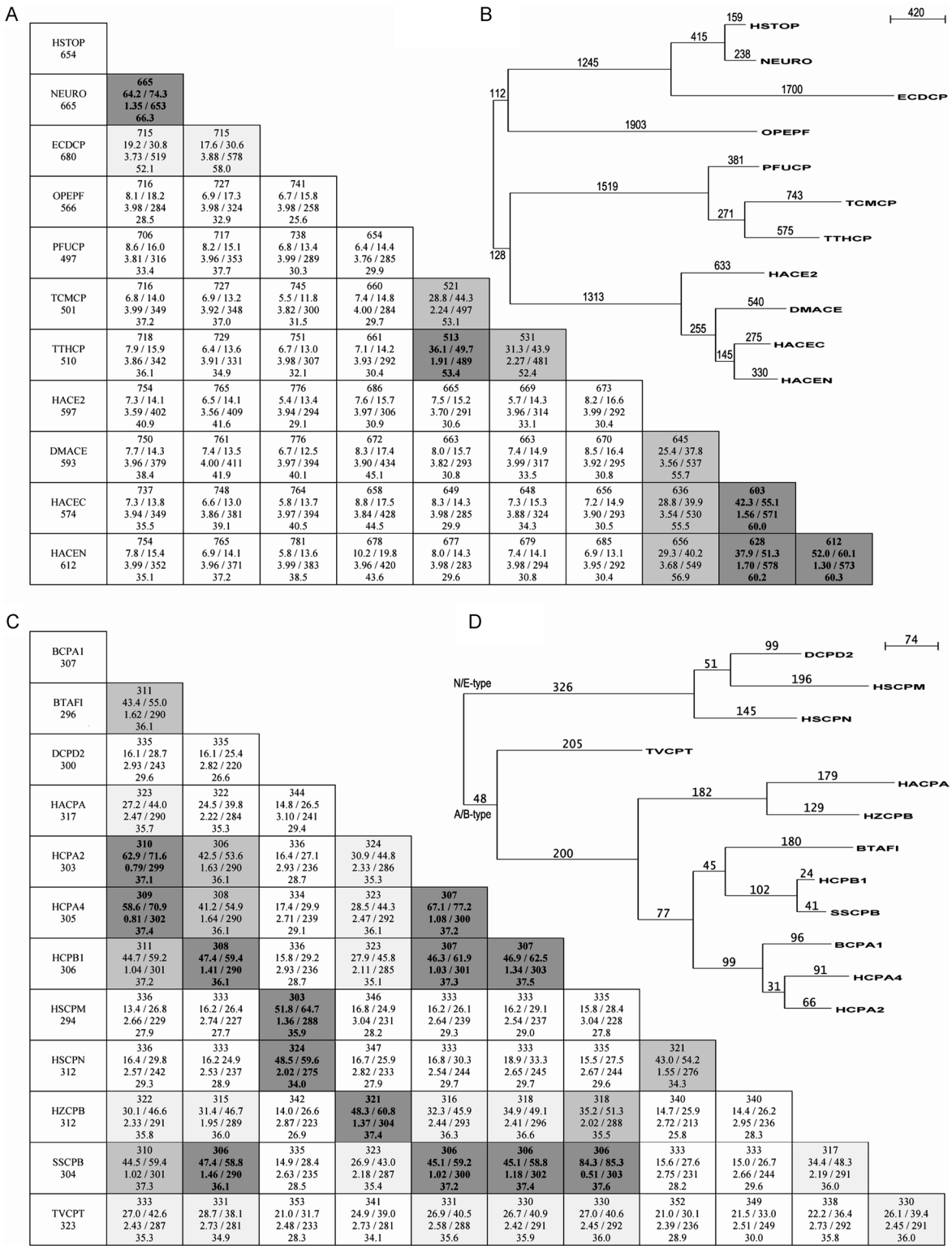


FIG. 2. Sequence and genetic relationships. (A) Figure depicting the pairwise relationship within cowrins. For each of the 11 proteins studied (leftmost column), the total number of residues of the catalytic domain used for structural superimposition is indicated (see Table 1 for the corresponding PDB entries). The following parameters are displayed at each intersection: total number of residues of the common sequence stretch of the pairwise alignment; sequence identity/similarity in %; rmsd of Cα superimposition/number of topologically equivalent residues; and Z-score of the structural fit. Closest sequences (identity > 35%)

are shown in bold over dark gray background, moderately similar sequences (25% < identity ≤ 35%) over gray, more distant sequences over light gray (15% < identity ≤ 25%), and dissimilar sequences over white (identity ≤ 15%). (B) Rootless phylogenetic tree reflecting genetic distances between cowrins. The value indicated in each branch path indicates the number of single-residue substitutions. The bar in the upper right shown for reference represents 420 substitutions. (C) Same as (A) but for the 12 funnelins analyzed. Due to the much closer relationship among these proteins, the threshold values for grouping based on identity were reset to 45%, 35% and 25%. (D) Same as (B) for funnelins, with the bar representing 74 substitutions. The sequence of each of the selected catalytic domains (11 cowrins and 12 funnelins) was extracted from the appropriate PDB entry (see Table 1) with program THESEUS (using the –F option) (Theobald and Wuttke, 2006) and dumped to files in FASTA format. The coordinates of the Cα atoms of the residues comprising each catalytic domain were fed into program MAMMOTH-MULT (Lupyan *et al*., 2005; Ortiz *et al*., 2002) in a single file through the dedicated server at http://ub.cbm.uam.es/mammoth/mult for a sequence-independent multiple structural alignment. This program provides an acceptable balance of accuracy and speed. These calculations rendered the transformation matrices for each structure and a multiple structure-based sequence alignment in CLUSTALW format. In addition, the number of topologically equivalent residues and the rmsd for each pairwise Cα superimposition, as well as a Z-score of the structural fit (according to Ortiz *et al*., 2002), were obtained. After format-exchange to Newick phylogenetic tree format at www.genopole-lille.fr/spip/spip.php?article39&incsuite=clustalW_convert_simple.html, the sequence alignment was used in a phylogenetic analysis based on maximum-likelihood criteria with program PHYML (Guindon and Gascuel, 2003). The resulting rootless phylogenetic trees were represented with program NJPLOT (Perriere and Gouy, 1996). The values obtained for each branch path are the absolute number of substitutions, not single amino acid substitutions, as each position in a sequence may undergo a number of changes during evolution. Therefore, this genetic distance is a much more accurate reflection of relatedness than mere sequence identity or similarity. To obtain the pairwise scores of sequence identity and similarity, the structure-based multiple sequence alignment in CLUSTALW format was format-exchanged to FASTA format and analyzed with program IDENT&SIM (Stothard, 2000) at server http://bioinformatics.org/sms2/ident_sim.html. The following groups of amino acids were considered to be similar: ILV, FWY, KRH, DE, GAS, P, C, and TNQM. In addition, the multiple sequence alignment was used together with the single-sequence files in FASTA format (see above) to perform a final sequence-independent structural superimposition based on maximum-likelihood criteria with THESEUS (Theobald and Wuttke, 2006). This procedure requires a prior structure-based sequence alignment and assumes that the structures are distributed according to a matrix Gaussian distribution and that the eigenvalues of the atomic covariance matrix are hierarchically distributed according to an inverse γ-distribution. This algorithm renders much more accurate results by down-weighting variable regions and by correcting for correlations among atoms. This superimposition rendered a classical pairwise rmsd and a maximum-likelihood rmsd, respectively, of 4.8Å and 1.1Å for cowrins and of 3.4Å and 0.3Å for funnelins.

an inversion of the direction of the polypeptide eventually leads to the second moiety cross-point to rejoin the UM at α14, the *tyrosine helix*, which traverses the molecule from front to back. The active-site helix, the glutamate helix and the tyrosine helix constitute the main structural core of cowrins. The three helices are tightly packed, with their axes spaced ~8–10Å and intersecting on a vertical projection at ~50° (see Figure 3f). This is allowed for by the presence of small side chains at the two cross-point interfaces. Cowrins share the presence of this three-helical bundle with thermolysin-related enzymes. In PFUCP, TTHCP, and TCMCP, a single helix is found after α14 (Figure 3d) in a segment that approaches the upper right of the molecule in the consensus structure and enters a final triple-helical segment (α15–α17). The first two helices are arranged as an α-hairpin, which, through the last helix α17, leads to the C-terminus in the right back, roughly at half the height of the molecule. Exceptionally, the four ACE-relatives comprise hallmark extensions at their C-termini, accounting for longer polypeptide chains, which include an additional α-helix and a long segment in extended conformation, so that their C-termini are located in the bottom at the front, interacting with Lα13α14.

Active-Site Cleft

The active-site cleft completely traverses the cowrin catalytic domains from left to right, forming a deep narrow canyon with a height of 10–20Å and a width of ~35Å (Figure 3a). Such narrow clefts have been implicated in cleavage activities of cowrins against oligopeptides or partially unfolded substrates only (Brown *et al*., 2001). The cleft is delimited by six elements: the DS, which creates a protrusion above the cleft on its non-primed side; strand β2 of the three-stranded β-sheet in the UM, which gives rise to an upper rim above the cleft; the beginning of the tyrosine helix; the active-site helix; the glutamate helix; and mainly helix α4 of the LM (Figures 3d, e). The active site lies midway along the cleft rather than at one of its ends even in those cowrins with strict carboxypeptidase activity, i.e. the cleft is open beyond S_1' (Figures 3a, d). This is more reminiscent of endopeptidases than exopeptidases. The active site is centered on a metal ion. A zinc has been reported in all cowrin structures except PFUCP and TCMCP, which are (most likely) cobalt-dependent (Cheng *et al*., 1999; Niemirowicz *et al*., 2008). This assumption is supported by biochemical studies on a related enzyme from *Thermus aquaticus*., which was reported to be cobalt-dependent

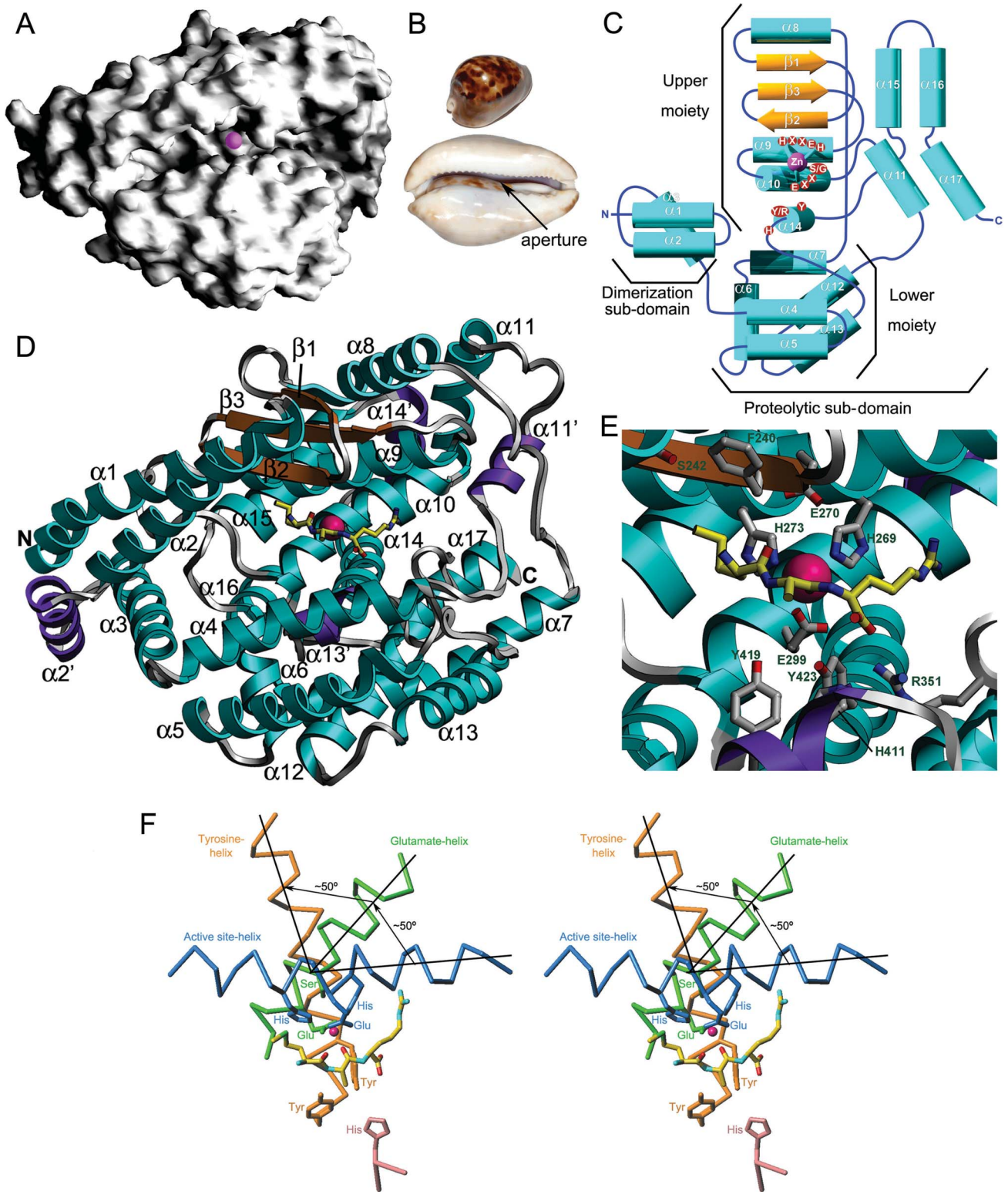


FIG. 3. Cowrin structural features. (A) Molecular surface representation of TCMCP in standard orientation for cowrins with the catalytic metal ion as a magenta sphere. Picture generated with program GRASP (Nicholls *et al*., 1993). (B) Top view (up) and bottom view (down) into the characteristic aperture of a cowry of the marine gastropod mollusc, *Zoila venusta episema* (picture courtesy of Frédéric Weber). (C) Topology scheme illustrating the consensus secondary-structure elements found in all cowrins studied (α-helices in cyan, $\alpha 1$–$\alpha 17$, and β-strands in orange, $\beta 1$–$\beta 3$). The catalytic domains can be divided into a dimerization

(Lee *et al*., 1992; Lee *et al*., 1994). Also, PFUCP apo-enzyme was shown to be reconstituted by Co^{2+} but not by Zn^{2+} (Cheng *et al*., 1999), although in subsequent structural studies, a Mg^{2+}-cation was unexpectedly placed at the active site (Arndt *et al*., 2002). Substitution of divalent cobalt for zinc was reported to give rise to functional forms of the metalloendopeptidases thermolysin (Hashida and Inouye, 2007) and astacin (Gomis-Rüth *et al*., 1994b), thus indicating that cobalt is compatible with catalytic activity in MPs. Given the close relationship between PFUCP, TCMCP, and TTHCP (see above), it is likely that the latter, structurally solved by a structural genomics consortium, but not analyzed to date beyond a preliminary crystallization note (Nagata *et al*., 2004), depends on cobalt rather than zinc, as found in the structure deposited (Protein Data Bank (PDB) access code 1WGZ).

The catalytic metal ion is coordinated in cowrins by the Nε2 atoms of the two histidine residues imbedded in the HEXXH sequence, which is contained in the active-site helix, and by a downstream glutamate comprised within a short sequence, EXXS/G, within the glutamate helix (Figures 3e, f). In the consensus cowrin structure, a single solvent molecule most likely completes a distorted tetrahedral geometry, as shown for HACE2 and found in other MPs (Matthews *et al*., 1972; Rees *et al*., 1983; Gomis-Rüth *et al*., 1994a; Towler *et al*., 2004). Cobalt-containing forms such as TCMCP may present differences in the coordination sphere, with additional solvent molecules (Niemirowicz *et al*., in preparation). The number of residues between the second metal-binding histidine and the downstream glutamate range from 24 (HACEC, HACEN, HACE2, and DMACE) and 25 (ECDCP, NEURO, and HSTOP) to 26 (TTHCP and PFUCP), 27 (TCMCP), and 28 (OPEPF), thus accounting for varying lengths of Lα9α10. A substrate is bound at several anchor points. In particular, upper-rim strand β2 exerts a similar function as the homonymous strand in metzincins and thermolysins in binding a substrate in an antiparallel manner (Matthews, 1988; Gomis-Rüth, 2003; see also Figure 3e and the Mechanisms of Hydrolysis section below). Further residues that have been implicated in substrate binding and catalysis in cowrins include a conserved histidine within Lα13α14 and two tyrosine residues or a tyrosine and an arginine within the beginning of the tyrosine helix. Binding of the C-terminal carboxylate group of substrates is carried out by an arginine in cowrins with strict carboxypeptidase activity, such as HACE2, PFUCP and TCMCP, and the peptidyl dipeptidase ECDCP. In the other peptidyl dipeptidases, DMACE, HACEC, and HACEN, binding is by a lysine and a tyrosine, whereas in the endopeptidases (HSTOP, NEURO, and OPEPF) this feature is missing (see the Mechanisms of Hydrolysis section).

The structures reported show varying degrees of aperture of the cleft (Figures 4a, b). Only HACE2 has been unambiguously analyzed in both unbound and bound states and these studies suggested that the cleft opens and closes like a clam shell (Figure 5a) (Towler *et al*., 2004). Additional evidence for such a motion was provided by ECDCP, the only family member trapped in a product complex (Comellas-Bigler *et al*., 2005). No crystals could be obtained with the unbound enzyme, suggesting major rearrangement of the protease upon binding. Similar motions have previously been described for other MPs such as astacin, matrix metalloprotease 3, anthrax-toxin lethal factor, and thermolysin-family members (Holland *et al*., 1992; Gerstein *et al*., 1994; Grams *et al*., 1996; Gerstein and Krebs, 1998; see also http://www.molmovdb.org). Accordingly, an open cleft may facilitate the initial accommodation of the substrate in cowrins through provision of extra space. In turn, the closed cleft may account for the subsequent tight substrate binding required for a productive Michaelis complex.

A conspicuous structural element found in cowrins is a serine placed immediately underneath the metal site, reminiscent of the position of a conserved methionine imbedded in the Met-turn of metzincins (Gomis-Rüth, 2003). This residue is contained within the EXXS/G motif and is part of the second turn of the glutamate helix, two positions downstream of the metal-binding

←

and a proteolytic sub-domain, the latter consists of an upper and a lower moiety with respect to the central active-site cleft. The approximate position of the amino acids comprised in the hallmark set of conserved residues, HEXXH+EXXS/G+H+Y/R+Y, is further pinpointed. The metal ion is shown as a magenta sphere bonded by residues from helices α9 and α10. (D) Richardson-diagram (Richardson, 1985) of PFUCP, the shortest cowrin studied, in standard orientation. The consensus secondary-structure elements and the metal site are colored and labeled as in (C). Additional elements not present in other cowrins are shown in purple and labeled α2', α11', α13', and α14'. A potential substrate of sequence Met-Ala–Arg was modeled based on the coordinates of the product complex of ECDCP (Comellas-Bigler *et al*., 2005) superimposed on the PFUCP structure and subsequently energy-minimized with the Dundee PRODRG2-server at http://davapc1.bioch.dundee.ac.uk/prodrg. To delineate the active-site cleft, the substrate is shown as a stick model with atom color code: yellow for C, blue for N, yellow for S, and red for O. The N- and C-termini are labeled for reference. Picture made with program SETOR (Evans, 1993). (E) Close-up view of (D) around the active site further displaying protein side chains essential for metal and substrate binding, and catalysis. Picture made with program SETOR. (F) Stereo-cartoon showing the Cα-traces of the active-site helix (blue), the glutamate helix (green), and the tyrosine helix (orange), as seen after a horizontal $\sim$90°-rotation of the orientation shown in (D). The active-site metal (magenta sphere) and residues from the cowrin set of spatially conserved residues, HEXXH+EXXS/G+H+Y/R+Y, are shown as sticks colored according to the helix they belong to. The histidine in pink is within the loop preceding the tyrosine helix. Picture created with program TURBO-FRODO (Carranza *et al*., 1999).

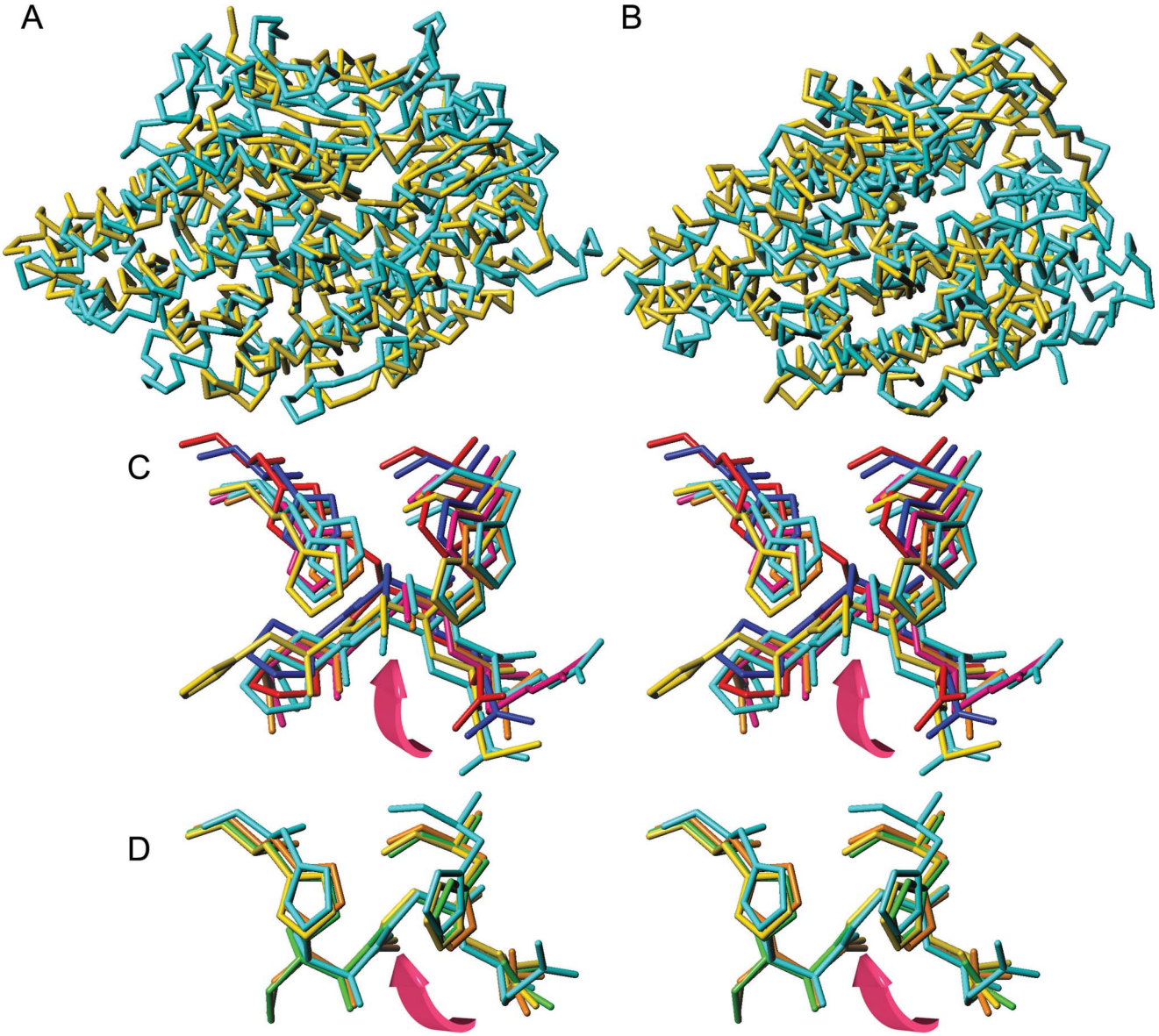


FIG. 4. Cowrin structure superposition. (A) Cα-traces after optimal superimposition of OPEPF (yellow) and ECDCP (cyan) in standard orientation. (B) Same as (A) but showing TCMCP (yellow) and HACEC (cyan). (C) Detail of those cowrin structures containing a serine within the Ser/Gly-turn under the catalytic site, approximately in standard orientation. ECDCP is displayed in cyan, OPEPF in yellow, PFUMCP in magenta, NEURO in dark blue, HSTOP in red, TCMCP in orange, and TTHCP in white. In each case, the two metal-binding histidines have been depicted for orientation (above). The magenta curved arrow points to the serine side chain. (D) Same as (C) but showing the cowrins with a glycine within the Ser/Gly-turn. DMACE is shown in orange, HACEC in yellow, HACEN in green, and HACE2 in cyan. The arrow points to the glycine Cα atom. All figure parts were made with program TURBO-FRODO.

glutamate. All structures have similar main-chain angles, within the core α-region of a Ramachandran plot ($-77^\circ \leq \Phi \leq -66^\circ$; $-49^\circ \leq \Psi \leq -35^\circ$), and χ_1-angles ($-75^\circ \leq \chi_1 \leq -62^\circ$), except HSTOP ($\chi_1 = 163^\circ$; see Figure 4c). Interestingly, thermolysin also displays a serine at this position (Matthews *et al*., 1972). This residue might stabilize the metal site or even modulate metal nucleophilicity through its side chain. The only exceptions to the strict conservation of this serine are the four ACE relatives, which all have a glycine instead, despite showing similar main-chain angles to the serine-containing structures ($-78^\circ \leq \Phi \leq -59^\circ$; $-45^\circ \leq \Psi \leq -30^\circ$; see Figure 4d). Here, the putative function of the serine may be replaced by a charge-relay system absent from other cowrins, which involves three elements: the second HEXXH-metal-binding histidine (HisA), an additional

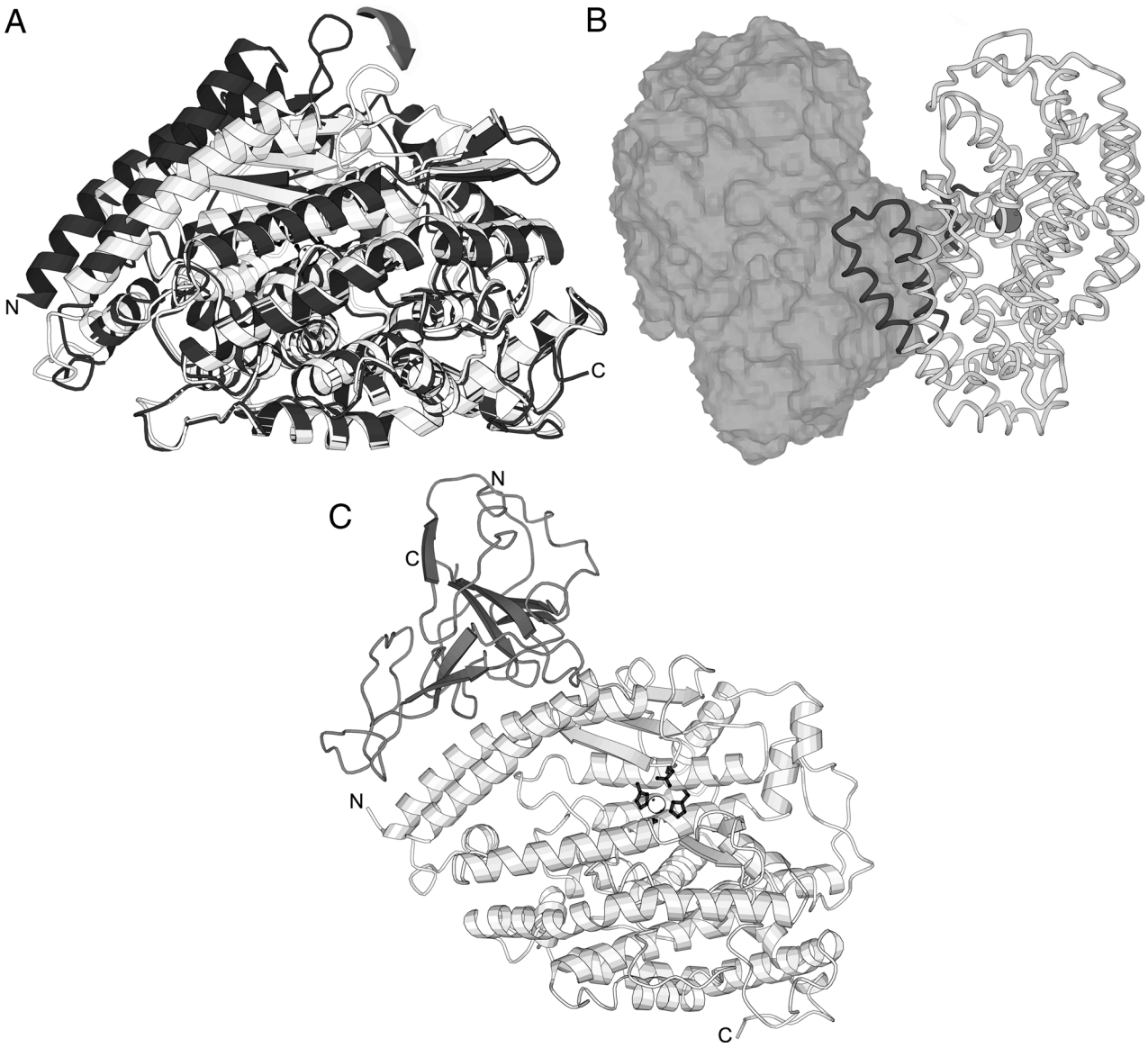


FIG. 5. Cowrin flexibility, quaternary structures, and complexes. (A) Richardson plot of unbound (dark gray; PDB 1R42) and bound (light gray; PDB 1R4L) HACE2 structures. Upon superposition of the LMs, it appears that most of the UM undergoes major rigid-body rearrangement (pinpointed by the arrow). The metal ion of the bound structure is shown as a white sphere and the N- and C-termini of the unbound structure are labeled for orientation. (B) TCMCP homodimer. One monomer is shown superimposed with its Connolly surface and the other as a ribbon. The proteolytic sub-domain is shown as a white rope and the dimerization sub-domain as a dark gray rope. The catalytic metal ion of the right monomer is presented as a gray sphere. (C) Richardson plot of the complex between HACE2 (white ribbon) and the spike receptor-binding domain of severe acute respiratory syndrome coronavirus (dark gray ribbon) (PDB 2AJF). The three protein ligands of the catalytic zinc, displayed as a white sphere, are shown as sticks and the N- and C-termini of each polypeptide chain are labeled. All figure parts were generated with programs VOLUMES and BOBSCRIPT (Esnouf, 1997).

central histidine (HisB) preceding the metal-binding glutamate, and a tyrosine at the end of the active-site helix. In all four ACE-relatives, the side chains implicated rotate to establish two hydrogen bonds, HisA Nδ1-HisB Nδ1 and HisB Nε-Tyr Oη, with interatomic distances of 2.6–2.9Å and 2.9–3.2Å, respectively. So, analogously to the Met-turn in metzincins, this serine or glycine-containing helix turn may be referred to as a *Ser/Gly-turn*.

Summarizing, the conserved residues accounting for the specific metal-binding and catalytic features of cowrins are the two histidine residues and the glutamate comprised in the HEXXH

consensus sequence, as well as both a glutamate engaged in metal binding and a serine or a glycine contained in the downstream motif EXXS/G. These elements, together with the relative disposition of the active site, glutamate, and tyrosine helices and with a β-sheet containing an upper-rim strand above the active-site cleft are shared with thermolysin-related enzymes and can thus be considered as hallmark elements for the gluzincin clan of MPs (Bode *et al*., 1993; Hooper, 1994). In addition, cowrins display consensus residues engaged in substrate binding and/or stabilization of the reaction intermediate (see also the Mechanisms of Hydrolysis section below), i.e. an additional histidine, a tyrosine or an arginine, and a tyrosine. All these elements give rise to a set of spatially conserved residues involved in catalysis in cowrins, HEXXH+EXXS/G+H+Y/R+Y.

Quaternary Structure of Cowrins and Reported Complexes with other Proteins

P. furiosus carboxypeptidase (PFUCP) was shown by biophysical techniques to be a homodimer, which is consistent with its crystal structure and that of its close relative TTHCP (Cheng *et al*., 1999; Arndt *et al*., 2002; Nagata *et al*., 2004). Furthermore, TCMCP was shown to be a dimer both in solution and in the crystal structure, and the same quaternary structure was reported for its paralog, TcMCP-2, in solution (Niemirowicz *et al*., 2007; in preparation). All three structurally-characterized proteins dimerize similarly through a vertical non-crystallographic two-fold axis (Figure 5b). The superstructure shows dimensions $\sim$95 (width) $\times$ $\sim$60 (depth) $\times$ $\sim$75Å (height) and dimerization occurs through the upper left front surface of each monomer. This association resembles the mythical two-headed god Janus, with the fully accessible active-site clefts mimicking a mouth on each of the opposite faces of the dimer. The axes of the clefts intersect at an angle of $\sim$55° and dimerization is caused by the respective DSs via α1, Lα1α2 and α2, as well as by the concave face of the β-sheet of each UM. Participating segments include Lα8β1, the beginning of β1, Lβ2β3 and the beginning of β3. All remaining cowrins have been reported to be active as monomers. Only HSTOP was reported to be regulated in its activity through covalent cysteine-mediated oligomerization (Ray *et al*., 2004). However, its functional oligomerization state is monomeric.

With respect to interactions with other proteins, HACE2 has an additional function unrelated to catalysis: it interacts at the cell surface with the spike receptor-binding domain of the coronavirus causing severe acute respiratory syndrome in humans. It does so through the first α-hairpin of the structure, an element engaged in dimerization in dimeric cowrins, in addition to loop Lβ2β3 of the UM β-sheet above the cleft and a short helical segment between α8 and β1, which is unique to this enzyme (Li *et al*., 2005). Accordingly, complex formation does not affect the active-site cleft and should thus not hamper the catalytic activity of HACE2 (Figure 5c). In any case, it remains to be ascertained whether the hinge motion observed for HACE2 upon binding has implications in the binding of the viral protein.

FUNNELINS

Funnelins Divide into A/B- and N/E-type Enzymes

Funnelins comprise two groups, which show strict carboxypeptidase activity and share a common fold and catalytic mechanism. These are the A/B-type and N/E-type MCPs, previously known as pancreatic and regulatory carboxypeptidases, respectively, and they have been ascribed to MEROPS families M14A and M14B, respectively (Vendrell *et al*., 2000; Arolas *et al*., 2007). A/B-MCPs were among the first peptidases studied as digestive enzymes produced by the pancreas of mammals (Waldschmidt-Leitz and Purr, 1929; Willstätter *et al*., 1932). These enzymes are biosynthesized with a signal peptide and secreted as inactive zymogens. They can be classified, according to their specificity for C-terminal hydrophobic and basic amino acids, into A- and B-type enzymes, respectively, whose archetypes are BCPA1 and pig CPB (SSCPB). A-type carboxypeptidases are further classified into A1- and A2-forms, which prefer small aliphatic residues and bulky aromatic side chains, respectively (Neurath, 1960; Reznik and Fricker, 2001). In humans, the digestive counterparts include CPA1, CPA2 (HCPA2) and CPB1 (HCPB1). More recently, A/B-MCPs have been found in archaea and bacteria, such as CPT from the thermophile *Thermoactinomyces vulgaris* (TVCPT), with a mixed A+B-type specificity (Teplyakov *et al*., 1992). Orthologs have also been found in protozoa, fungi, nematodes, insects and other invertebrates, such as the A-type carboxypeptidase from the cotton bollworm *Helicoverpa armigera* (HACPA) and the B-type enzyme from the corn earworm *Heliothis zea* (HZCPB). Further relatives have been discovered in plants, amphibians, birds, and mammals (Arolas *et al*., 2007; Auld, 2004b). In recent years, functional and local ascription of A/B-MCPs has moved away from the mere breakdown of intake proteins in the digestive tract. They have been localized in mammalian brain, heart, stomach, colon, testis, and lung (Normant *et al*., 1995a). They participate in peptide hormone homeostasis, and tissue growth and differentiation. In addition, a major role in fibrinolysis inhibition and bradykinin activation in blood serum has been described for a funnelin called tissue-activatable fibrinolysis inhibitor (TAFI), which has been studied from bovine (BTAFI), human and rodents (Hillmayer *et al*., 2006; Valnickova *et al*., 2007; Sanglas *et al*., 2008). Moreover, A/B-MCPs may act during cellular response and complementing the serine endopeptidase chymase in mast cells (Springman, 2004). Another example is a new gene product, human CPA4 (HCPA4), detected in human hormone-regulated tissues and thought to play a role in prostate cancer (Huang *et al*., 1999; Kayashima *et al*., 2003).

N/E-type funnelins are engaged in a variety of functions in mammals including neuropeptide and peptide hormone activation, processing of vasoactive peptides and complement proteins, and regulation and modulation of protein–protein and protein–cell interactions (Skidgel, 1988; Arolas *et al*., 2007). They include five active members, CPD, CPE, CPM, CPN and

CPZ, which remove basic C-terminal residues from their substrates. In addition, three catalytically inert forms with potential functions as binding proteins have been described: CPX1, CPX2 and aortic carboxypeptidase-like protein. The gene of the latter may give rise to adipocyte enhancer binding protein-1 through alternative splicing. In contrast to A/B-type enzymes, N/E funnelins are not secreted as pro-enzymes and rely entirely on their subcellular localization to regulate their activity (Arolas *et al*., 2007). Downstream of the catalytic domain, they all show a ~80-residue domain resembling the plasma protein prealbumin alias transthyretin, which may contribute to protein folding (Arolas *et al*., 2007; Gomis-Rüth *et al*., 1999). In addition, CPZ displays an N-terminal domain reminiscent of frizzled receptors and other Wnt-binding proteins (Song and Fricker, 1997). CPD is a 180-kDa polypeptide containing three ~390-residue tandem repeats, each comprising a funnelin moiety followed by a C-terminal transthyretin-like module, a transmembrane domain and a ~60-residue cytosolic tail. The first two repeats are catalytically active and contain all the characteristic funnelin residues required for substrate binding and activity (Tan *et al*., 1997; Xin *et al*., 1997). The third repeat is inactive and has been shown to be the receptor for preS envelope protein of hepatitis B virus particles in duck (Kuroki *et al*., 1995), a function which is reminiscent of that of the cowrin HACE2 (see the section above on Quaternary Structure of Cowrins and Reported Complexes with other Proteins). Like CPD, CPM is membrane-associated and has a 25-residue C-terminal extension, to which a glycosylphosphatidylinositol-anchor is attached. The latter enables CPM to interact with the membrane and approach other membrane-bound substrates or peptides (Skidgel *et al*., 1989). In contrast, functional CPN is found in plasma as a soluble 280-kDa heterotetramer made up by two catalytically active CPN1 moieties (48–55 kDa) and two leucine-repeat-rich 83-kDa CPN2 subunits (Levin *et al*., 1982).

Funnelins are More Closely Related than Cowrins

A number of three-dimensional structures are available for A/B-MCPs, either in their active or zymogenic state, unbound or as inhibitor or product complexes (Lipscomb *et al*., 1968; Schmid and Herriott, 1976; Rees and Lipscomb, 1980; Kim and Lipscomb, 1990; 1991; Coll *et al*., 1991; Faming *et al*., 1991; Guasch *et al*., 1992; Teplyakov *et al*., 1992; Gomis-Rüth *et al*., 1995; 1997a; García-Sáez *et al*., 1997; Reverter *et al*., 2000; Vendrell *et al*., 2000; Pereira *et al*., 2002; Arolas *et al*., 2005b; Pallarès *et al*., 2005; Bayés *et al*., 2007; see also Table 1). In contrast, N/E proteins have received less attention. Structures analyzed include human CPM (HSCPM; Reverter *et al*., 2004), human CPN (HSCPN; Keil *et al*., 2007), and crested-duck CPD repeat II (DCPD2; Gomis-Rüth *et al*., 1999). The study of a set of representative funnelins, which are all monomers, namely BCPA1, BTAFI, DCPD2, HACPA, HCPA2, HCPA4, HCPB1, HSCPM, HSCPN, HZCPB, SSCPB, and TVCPT, reveals that their catalytic domains are more closely related to each other than cowrins, in terms of both genetic distance and sequence and structural similarity.

A phylogenetic analysis based on a structure-based multiple-sequence alignment shows that the most dissimilar sequences, HSCPM and HACPA, are separated by a genetic distance corresponding to 1182 single-residue substitutions. This is about a fifth of the greatest distance found between cowrins if we take into account the differences in chain length between cowrins and funnelins (see Evolutionary and Structural Relationship among Cowrins). The shortest genetic distances between funnelins are found among pancreatic B-type (65 substitutions) and mammalian A-type proteins (at most 193 substitutions). BTAFI, which is a B-type protein, is only slightly closer to a pancreatic B-type (306 substitutions) than to an A-type counterpart (420 substitutions). These genetic similarities are comparable with those observed between A- and B-type enzymes (at least 363 substitutions), the two insect forms (308 substitutions), and among N/E funnelins (at most 392 substitutions). The shortest genetic distance between an N/E- and an A/B-representative is roughly three times greater (967 substitutions). Finally, the only bacterial representative, TVCPT, is approximately equidistant from A/B- (at least 653 substitutions) and N/E-type relatives (at least 724 substitutions). TVCPT could constitute a separate subgroup (Figure 2d), although it is formally ascribed to A/B-MCPs.

Assessment of pairwise sequence alignments essentially supports these features (Figure 2c). Closest sequence similarity is found among pancreatic B-type enzymes (84% sequence identity; rmsd of 0.5Å), followed by mammalian A-type enzymes (59–67%; 0.8–1.1Å). BTAFI is only slightly closer to pancreatic B funnelins (47%; 1.4–1.5Å) than to mammalian A funnelins (41–43%; 1.6Å), with values similar to those found for mixed mammalian A- and B-type pairs (44–47%; 1.0–1.3Å). Within A/B funnelins (excluding TVCPT), values are 25–45% and 1.0–2.2Å. N/E funnelins show higher sequence similarities (43–52%) with similar rmsd values (1.4–2.0Å). Comparisons of N/E- and A/B funnelins reveal sequence identities of 13–19% and rmsd values of 2.5–3.1Å. These rmsd values are close to the limit for structurally similar positions and the sequence identity values do not reveal any similarities. Finally, as suggested by the phylogenetic analyses, TVCPT differs almost equally from both A/B- and N/E-type funnelins (21–29%; 2.4–2.7Å).

Catalytic Domain and Active Site of Funnelins

Funnelin catalytic domains have maximal dimensions of ~50 × ~40 × ~40Å (Hartsuck and Lipscomb, 1971) and a multiple-structure superimposition renders a classical pairwise rmsd and maximum-likelihood rmsd of 3.4Å and 0.3Å, respectively. They have a compact globular shape, which resembles the volume obtained when a cone is extracted from a sphere (Figures 6a, b, d). There is a funnel-like opening at the top, which accounts for the name funnelins; the active-site cleft lies inside, at the base of the opening (García-Castellanos *et al*., 2005; Pallarès *et al*., 2005). The cleft is rather shallow, which explains the capacity of funnelins to cleave a large variety of well-folded proteins, as only few contacts between the enzyme and the C-terminal stretch of a

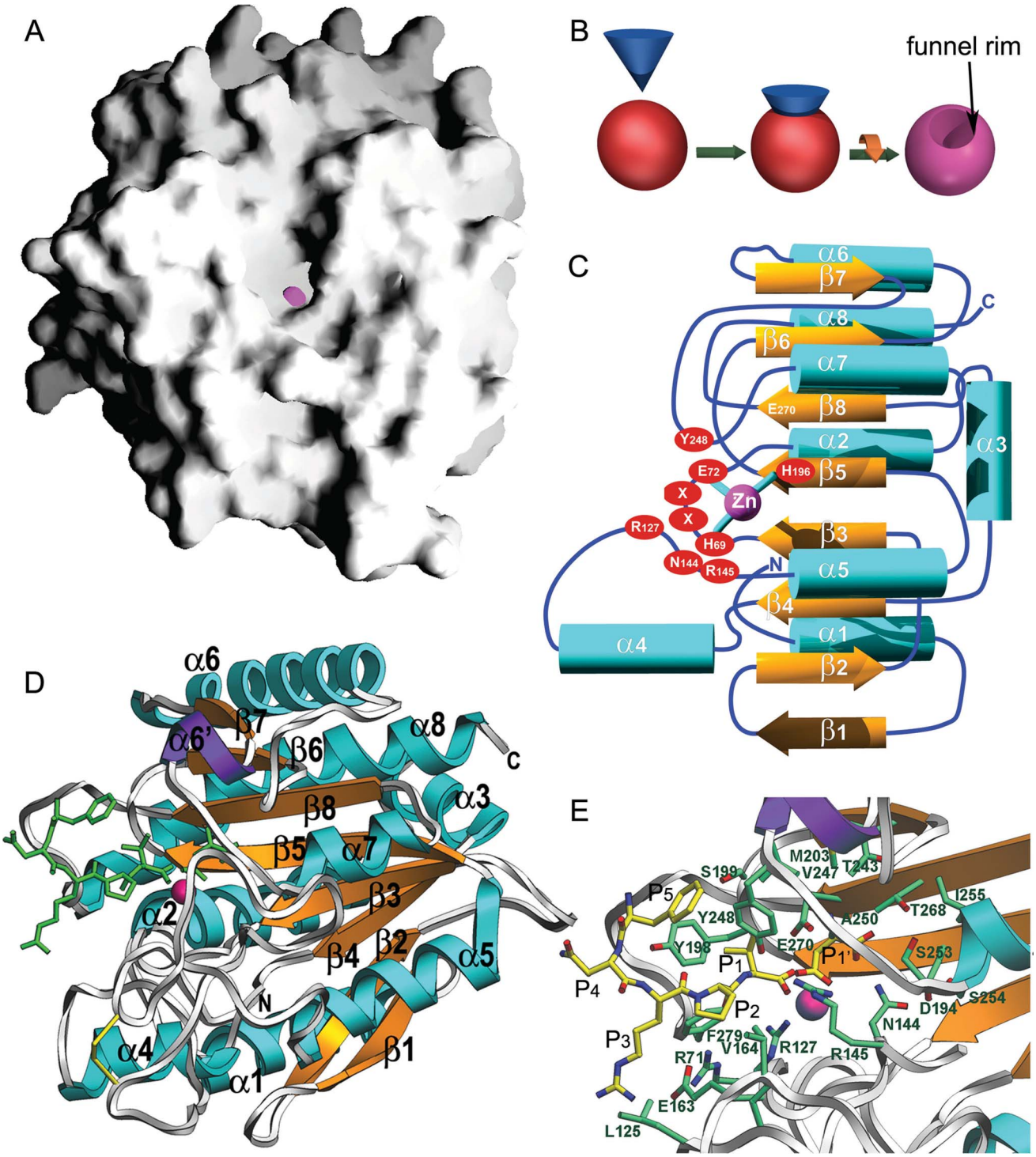


FIG. 6. Funnelin structural features. (A) Molecular surface representation of HCPA4 calculated with GRASP, with the view along the funnel-like aperture that leads to the catalytic zinc cation at the bottom (magenta sphere). The view corresponds to the standard orientation for funnelins subjected to a vertical ~90°-rotation. (B) Scheme illustrating the generation of a major (or inverted) spherical cone (magenta) through the intersection of a regular cone (blue) and a sphere (red). The resulting border is like a funnel rim, see also Figure 7c. (C) Topology scheme illustrating the consensus secondary-structure elements (α-helices in cyan and β-strands in orange) found in all funnelins analyzed. Here, a central β-sheet ($\beta1$–$\beta8$) is flanked on either side by a total of eight α-helices ($\alpha1$–$\alpha8$). The highlighted amino acids are contained in the characteristic set of conserved residues,

HXXE+R+NR+H+Y+E, and are numbered according to the BCPA1 consensus. The metal ion is shown as a magenta sphere bound by residues from β-strand $\beta5$ and loop L$\beta3\alpha2$. (D) Richardson diagram of HCPA4 shown in standard orientation as a representative member of the funnelin catalytic domains (Bayés *et al*., 2007). The consensus secondary-structure elements and the metal site are colored and labeled as in (C). An additional element to the consensus α_8/β_8-structure is helix $\alpha6$' (purple ribbon), integrated into loop L$\beta7\alpha7$. A bound and subsequently cleaved hexapeptide (green stick model) is also displayed. The N- and C-termini are labeled for reference. Picture made with program SETOR. (E) Close-up view of (D). The cleaved product is shown with atom color code: yellow for C, blue for N and red for O. Enzyme residues important for substrate binding and catalysis are depicted for their side chains (atom color code: green for C, blue for N, and red for O) and labeled. The protein ligands of the catalytic ion have been omitted for clarity. The C-terminus of the pentapeptide is bound to the catalytic zinc ion in a bidentate manner, giving rise to a distorted trigonal-bipyramidal coordination of the metal through four oxygen and two nitrogen atoms. In this complex, an isolated aspartate residue occupies the pocket of the A-type enzyme HCPA4, which prefers an aliphatic residue at this position. Therefore, the amino acid is rotated with respect to a canonical orientation and it is the β-carboxylate group which is bound to the substrate-fixing residue Arg145. See Bayés *et al*. (2007) for details. The substrate positions from P_5 to P_1' are labeled. Picture made with program SETOR.

substrate are required to fix the latter to the active-site. Funnelin catalytic domains are made up by ~300 residues and show a consensus α_8/β_8-topology conforming to an α/β-hydrolase or PLEES fold (Ollis *et al*., 1992; Puente and López-Otín, 1997) (Figures 6c, 6d, 7a, 7b). They contain a central doubly-wound eight-stranded β-sheet ($\beta1$–$\beta8$) of mixed parallel/antiparallel topology with strand connectivity +1, +2, −1x, −2x, −2, +1x, −2 (according to Richardson, 1981). When viewed from the standard orientation (Figures 6c, d), the sheet spans the molecule vertically from the bottom to the top, with a twist of ~130°. The core of the sheet consists of four parallel coplanar central strands ($\beta3$–$\beta5$ and $\beta8$) and the catalytic site is located at the C-terminal end of these strands, as observed in other α/β-enzymes (Brändén, 1980). These four strands are flanked by two parallel strands on the top ($\beta6$–$\beta7$) and a β-ribbon ($\beta1$–$\beta2$) at the bottom (Figures 6c, d), which confer the overall twist on the sheet. This leads to a concave front side of the sheet, which accommodates helices $\alpha5$ and $\alpha7$, as well as the active-site cleft. All funnelins studied except DCPD2, HSCPM, and TVCPT contain an additional short helix inserted after the upper strand $\beta7$ ($\alpha6'$ in Figure 6d). At the back of the molecule, the convex side of the sheet embraces helices $\alpha1$–$\alpha4$, $\alpha6$ and $\alpha8$ and the surface N- and C-termini of the molecule. The funnel-like access to the active site is shaped by a series of irregular segments of varying length, namely L$\beta8\alpha8$, L$\beta5\beta6$, L$\beta7\alpha7$, L$\beta3\alpha2$, and, in particular, the ~50-residue loop L$\alpha4\alpha5$ in the lower left of the molecule (Figures 6d, 7c). These segments are required for interactions of the catalytic moiety with the pro-domain in the zymogen and with cognate protein inhibitors (García-Castellanos *et al*., 2005; Pallarès *et al*., 2005). L$\alpha4\alpha5$ delimits the active-site cleft (Figures 6d, e) and its conformation is stabilized by a disulfide bond (Cys138-Cys161; the residue numbering of funnelins conforms to BCPA1; each of the distinct tribe members displays insertions/deletions with respect to this convention; see e.g. Coll *et al*., 1991; Pereira *et al*., 2002). Most departures from the canonical BCPA1 fold are found in N/E funnelins, and they affect the five loop segments mentioned above (Figure 7). In particular, L$\alpha4\alpha5$ deviates most in HSCPN and it adopts the conformation of a fold-back flap, which includes a single extra α-helix, $\alpha4'$ (Keil *et al*., 2007).

Most funnelins catalyze the digestion of proteins, which imposes a requirement for resistance against extreme pH values and (auto)proteolysis. One exception is TAFI, which undergoes proteolysis consisting of removal of its pro-domain in the bloodstream by thrombin/thrombomodulin (Latt *et al*., 1972; Hendriks *et al*., 1989; Eaton *et al*., 1991; Bajzar *et al*., 1995; Arolas *et al*., 2007; Willemse and Hendriks, 2007). The result is TAFIa, which has a half-life of less than ten minutes *in vivo*. So, the structure of bovine TAFIa (BTAFI) could only be obtained after stabilization of the molecule in a complex with a protein inhibitor (see the Funnelin Inhibitors section ; and also Sanglas *et al*., 2008). Two surface loops are flexible, which correlates with proteolytic susceptibility and thus degradation (Zappacosta *et al*., 1996), especially in a protease-rich medium like blood. The equivalent regions of all other A/B-type funnelins so far analyzed are well ordered and rigid. Accordingly, this flexible region could be conceived of as a "fibrinolysis switch" to trigger rapid clearance of TAFIa from the bloodstream (Sanglas *et al*., 2008).

The catalytic zinc ion resides at the bottom of the active-site cleft and is coordinated by His69 N$\delta1$ and by Glu72 atoms O$\varepsilon2$ and O$\varepsilon1$ in a (mostly asymmetric) bidentate manner with *syn* stereochemistry. These two residues are provided by L$\beta7\alpha5$ and are embedded in a consensus sequence characteristic of funnelins, HXXE (Hooper, 1994). A third protein zinc ligand is His196 N$\delta1$ at the end of strand $\beta5$ (Figure 6c), which together with a metal-bound solvent molecule (see the Mechanisms of Hydrolysis section below) gives rise to an overall distorted tetrahedral coordination sphere, in which one of the tetrahedron corners bifurcates into the two carboxylate oxygen atoms of Glu72. This metal-binding constellation is reminiscent of cowrins. In contrast, no conserved structure or sequence element similar to the Ser/Gly-turn in cowrins (see the Active-Site Cleft section) or the Met-turn in metzincins (Gomis-Rüth, 2003) is found below the zinc-binding site in funnelins.

The length, shape, and nature of the active-site cleft have been investigated by a series of biochemical and structural studies

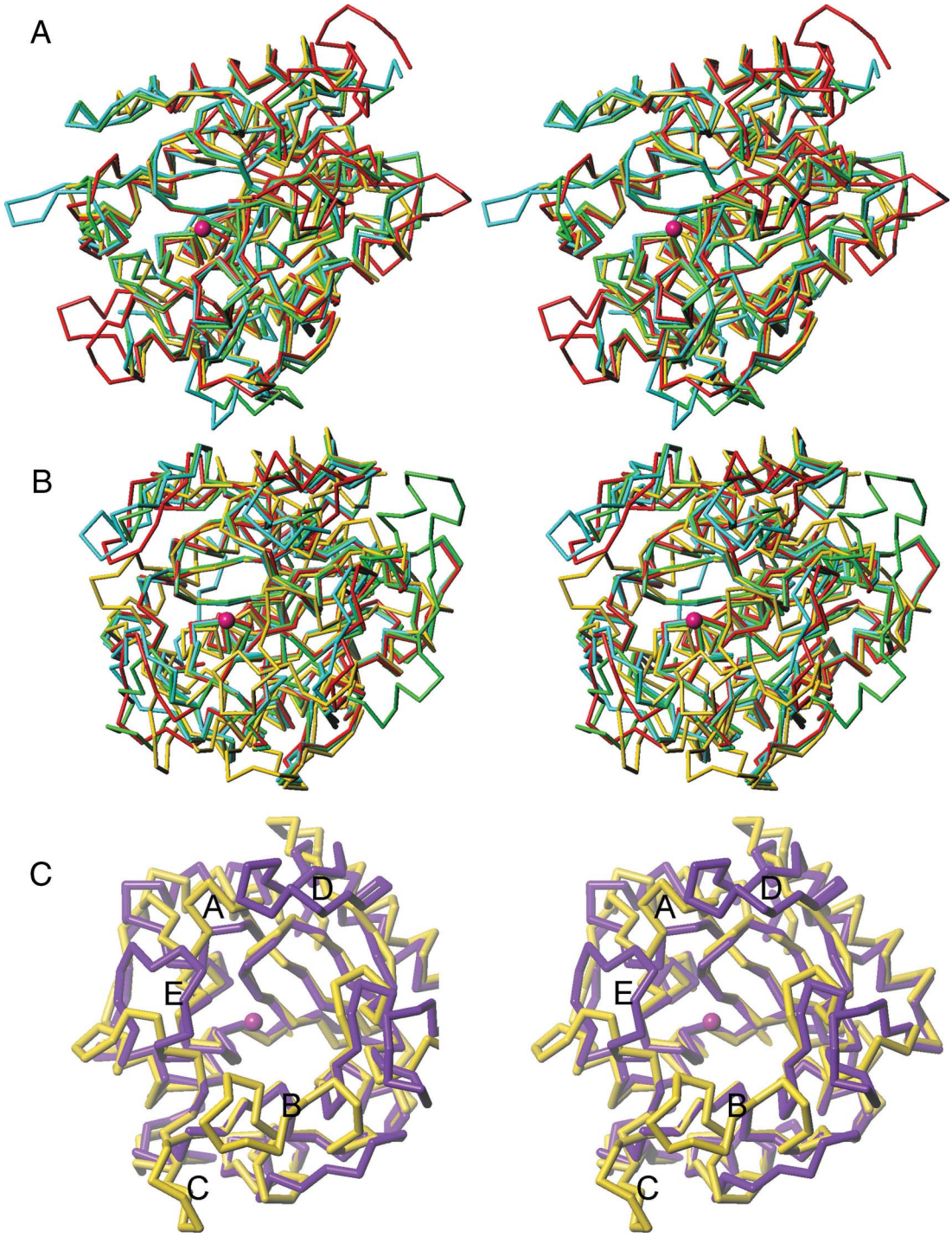


FIG. 7. Superimposition of funnelins. (A) Cα-traces in stereo after optimal superimposition of BCPA1 (yellow), HACPA (cyan), HZCPB (green) and TVCPT (red) in standard orientation. The catalytic zinc of BCPA1 is shown as a magenta sphere. (B) Same as (A) but showing BCPA1 (yellow), DCPD2 (cyan), HSCPN (green) and HSCPM (red). (C) Stereo-superimposition of HCPA4 (yellow) and DCPD2 (violet) as representatives of A/B- and N/E-type funnelins, respectively, to illustrate the regions shaping the rim of the funnel that leads to the active site of the enzymes, pinpointed by the catalytic zinc ion of HCPA4 (magenta sphere). The five main insertions/deletions accounting for lack of interactions (A to C) or potential steric clashes (D to E) with latexin as a representative of funnelin protein inhibitors on going from HCPA4 to DCPD2 are labeled (see García-Castellanos *et al*., 2005; Pallarès *et al*., 2005). The orientation displayed was obtained after rotating the molecules as shown in (B), first vertically ~90° and then in the plane by ~45°. All figure parts were made with program TURBO-FRODO.

with transition-state analogs, reaction products and inhibitors (Rees and Lipscomb, 1980; Clore *et al*., 1987; Kaplan and Bartlett, 1991; Kim and Lipscomb, 1990; 1991; Cappalonga *et al*., 1992; Guasch *et al*., 1992; Reverter *et al*., 2000; Park *et al*., 2002; Arolas *et al*., 2005b; Pallarès *et al*., 2005; Bayés *et al*., 2007). In contrast to what happens in cowrins and other MPs, a funnelin substrate is not anchored through β-ribbon-like interactions with a β-strand of the enzyme but rather is bound by enzyme side chains. The peptide chain, in a nearly extended conformation at P_1', P_1, P_2 and P_3, is bent at P_4 and folds back to creep along the molecular surface (Figures 6d, e). The scissile residue side chain is accommodated in the S_1' site—the specificity pocket of funnelins—which is delimited by the protein side chains in positions 194, 203, 207, 243, 247, 250, 253–255 and 268. They give rise to the characteristic cul-de-sac of exopeptidases and give the distinct enzyme types their specificity. In most cases, an aspartate at position 255 accounts for B-type specificity in A/B-type funnelins, while an isoleucine is found in A-type enzymes. TVCPT, with mixed A+B specificity, has a threonine at this position and N/E-type funnelins have a glutamine. The B-type preference of the N/E type is provided by an aspartate in the position equivalent to 207 of BCPA1 (a glycine in A/B-type funnelins). The scissile residue C-terminus is bound by Arg145, Tyr248 and Asn144 (Kim and Lipscomb, 1991; Guasch *et al*., 1992; Pallarès *et al*., 2005). Other highly-conserved fixing and catalytic residues are Arg127, which stabilizes the negatively-charged reaction intermediate and, together with Arg71, binds the carbonyl oxygen in P_2; Phe279 and Tyr198, which shape sub-site S_1; and Tyr248, which undergoes major rearrangement upon substrate binding and binds the main-chain amido nitrogen of the residue in P_1 (Figure 6e) (see the Mechanisms of Hydrolysis section below).

Summarizing, the residues that are conserved and account for the specific zinc-binding and catalytic features of funnelins are His69, Glu72, Arg127, Asn144, Arg145, His196, Tyr248, and Glu270. They constitute a set of spatially conserved residues involved in catalysis in funnelins, HXXE+R+NR+H+Y+E.

Other Domains of Funnelins

A/B-type funnelins are synthesized and secreted as zymogens with an N-terminal ~95-residue pro-domain. This pro-domain contributes to folding during synthesis of the pro-enzyme (Vendrell *et al*., 2000) and to the maintenance of latency. It blocks access to the active-site cleft of the enzyme, which is already in an active conformation in the zymogen (Figure 8a). The pro-domain comprises a major ~80-residue globular part and a C-terminal α-helical connecting segment and it contacts the mature enzyme moiety through the funnel-rim loops of the latter (see the previous section). Limited proteolysis at the end of this segment activates the zymogen through release of the pro-domain, which is degraded *in vivo*. *In vitro*, the isolated pro-domain acts as an autologous inhibitor *in trans* (Vendrell *et al*., 2000). The globular part of the pro-domain follows an open-sandwich topology with a twisted four-stranded antiparallel β-sheet packed through its face distal to the enzyme against two nearly parallel α-helices. This fold is also found in the pro-domain of the otherwise unrelated serine proteinase subtilisin (Gomis-Rüth *et al*., 1997a). Overall, the pro-domain/mature enzyme complexes of funnelins involve large interfaces (e.g. 1122Å^2 in human procarboxypeptidase A4; García-Castellanos *et al*., 2005) but few direct contacts. A key role in inhibition is played by Asp36A and Trp38A (pro-domain residues contain an "A"-suffix; see e.g. Coll *et al*., 1991; Pereira *et al*., 2002), which interact with the active-site proximal residues Arg71 and Phe279/Tyr198, respectively, of the mature enzyme moiety in A-type zymogens. The human and porcine B-type zymogens have an additional salt bridge made up by Asp41A and Arg145. The lack of this interaction accounts for the capacity of A-type zymogens to cleave small fluorigenic substrates, while B-type zymogens are inactive (Arolas *et al*., 2007). All these interactions are missing in the TAFI zymogen, which may explain why it also catalyzes larger peptides (Valnickova *et al*., 2007).

N/E funnelins contain a characteristic ~80-residue domain with a rod-like shape of approximate dimensions ~25 × ~25 × ~40Å downstream of the catalytic domain (Figure 8b). Its N- and C-termini are on opposite sides of the rod, which folds into an all-β seven-stranded flat β-sandwich. The β-sandwich consists of two layers of three mixed and four antiparallel strands, respectively, which are held together by a hydrophobic core. These strands are arranged as two subsequent Greek-key-like elements related by a two-fold axis that is perpendicular to the sandwich surface (Gomis-Rüth *et al*., 1999; Reverter *et al*., 2004; Keil *et al*., 2007). This domain has topological similarity with transthyretin, which contains an additional C-terminal β-strand that is missing in funnelins (Gomis-Rüth *et al*., 1999). Interactions between the transthyretin-like and the catalytic domains are established by hydrophobic residues and a salt bridge formed by an aspartate at BCPA1 position 216 at the beginning of helix α6, which is conserved among N/E funnelins (Asp206 in DCPD2, Asp227 in HSCPN and Asp201 in HSCPM; numbering according to the corresponding PDB entry, see Table 1), with arginine residues within the transthyretin-like domains. Although possible roles for these domains have been proposed, such as folding of the funnelin domain, regulation of enzyme activity or mediation of protein–protein interactions, their function remains to be established (Reznik and Fricker, 2001).

Funnelin Inhibitors

Heterologous funnelin inhibitors have been reported from potato (potato carboxypeptidase inhibitor; PCI), tomato, the intestinal parasitic nematode *Ascaris suum*, the medical leech *Hirudo medicinalis* (LCI), and the ixodid ticks *Haemaphysalis longicornis* and *Rhipicephalus bursa* (TCI).The only mammalian inhibitor described is latexin *alias* ECI, characterized from human and rodents (Vendrell *et al*., 2000; Díez-Díaz *et al*., 2004; Arolas *et al*., 2005a; 2007; Gong *et al*., 2007). All these molecules are potent inhibitors of A/B-type funnelins, but are ineffective against N/E-type members, against which no protein

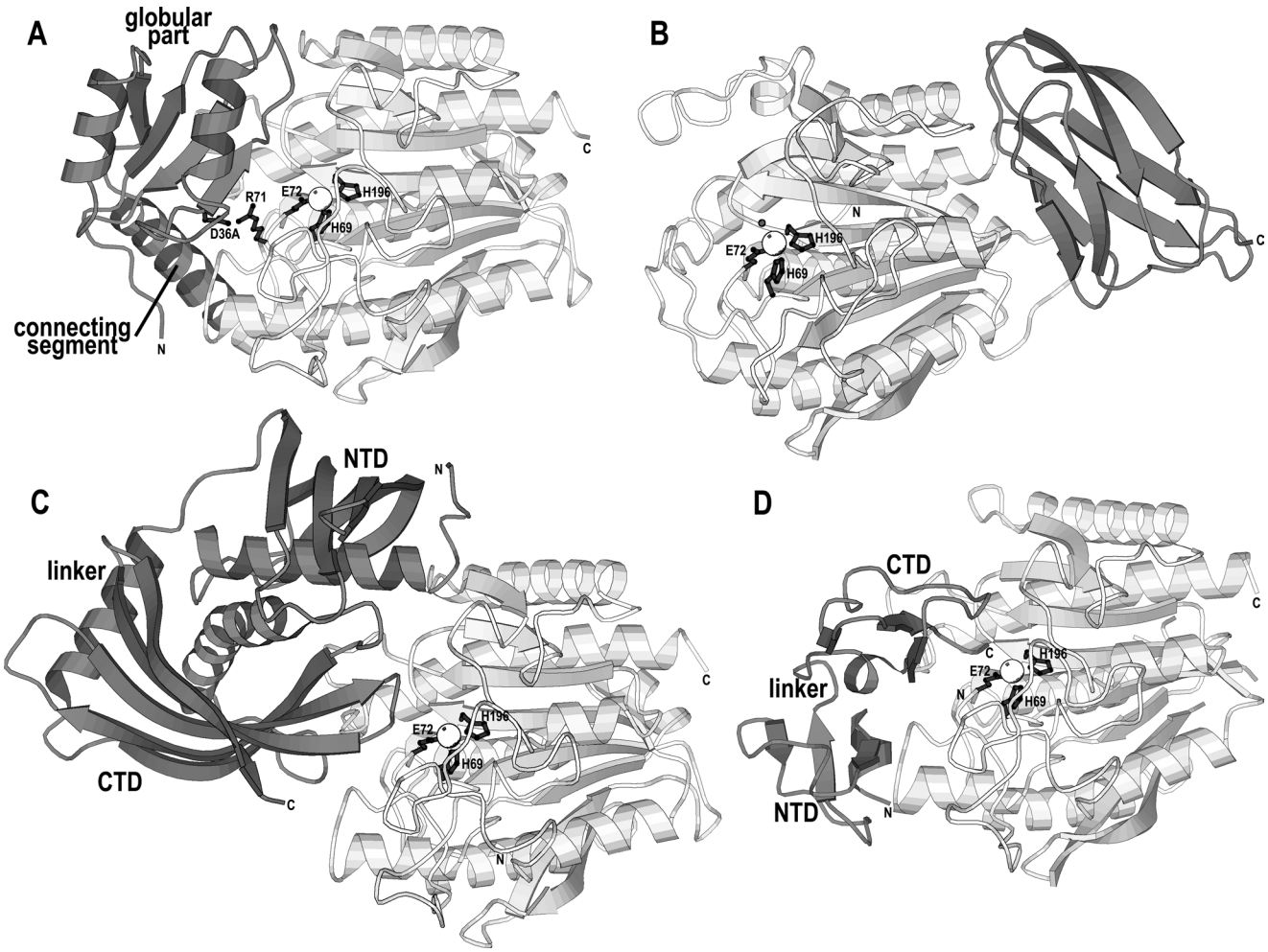


FIG. 8. Additional domains in funnelins and complexes with protein inhibitors. (A) Richardson diagram of the zymogenic form of HCPA4 in standard orientation as a paradigm of A/B-type pro-carboxypeptidases. The pro-domain, shown in dark gray, consists of a globular part and a helical connecting segment that leads to the mature enzyme moiety, in white. The three protein ligands of the catalytic zinc are shown as sticks and labeled, as are residues Asp36A of the pro-domain and Arg71 of the catalytic domain. (B) Diagram of DCPD2 showing the catalytic domain in white and the C-terminal transthyretin-like domain in dark gray. (C) Ribbon plot of the complex between HCPA4 (white) and the double-headed endogenous protein inhibitor latexin, which consists of an N-terminal (NTD) and a C-terminal domain (CTD) connected by a helical linker (Pallarès *et al*., 2005). (D) Complex structure of BTAFI with a double-headed inhibitor from *R. bursa* (Sanglas *et al*., 2008). This inhibitor also consists of an N-terminal (NTD) and a C-terminal domain (CTD) connected by a linker. All figure parts were fabricated with program BOBSCRIPT and in all cases, the N- and C-termini of each polypeptide are labeled.

inhibitors have been reported (Normant *et al*., 1995b; Reverter *et al*., 1998; Mao *et al*., 2003; Arolas *et al*., 2005a; Pallarès *et al*., 2005). The reason for this disparity is that these inhibitors interact with their target enzymes through the funnel-rim loops, which display major differences between A/B- and N/E-type funnelins (see Catalytic Domain and Active Site of Funnelins and Figure 7c). These discrepancies lead to steric clashes or prevent the interactions required for complex formation in N/E-type funnelins.

One inhibitory mechanism is exemplified by human latexin. This 222-residue molecule comprises two topologically equivalent domains linked by a connecting segment encompassing an α-helix. These domains are reminiscent of cysteine protease inhibitors (cystatins) and consist of an α-helix surrounded by a curved β-sheet. Both domains are packed against each other through their respective helices (Aagaard *et al*., 2005; García-Castellanos *et al*., 2005; Pallarès *et al*., 2005; see also Figure 8c). The enzyme/inhibitor complex of latexin with HCPA4 has a large interface (1,170Å^2; see García-Castellanos *et al*., 2005; Pallarès *et al*., 2005) but few inter-molecular contacts. These complex features lead to high inhibitory potency but low specificity, which explains why latexin inhibits

all vertebrate A/B-MCPs tested, even across species barriers (García-Castellanos *et al*., 2005; Pallarès *et al*., 2005). The enzyme is bound at the funnel border of its active site through the interface of the two inhibitor domains and neither of the latexin termini participates in complex formation. Main contacts are observed between an inhibitory loop in the central part of the β-sheet of the C-terminal latexin domain. Important for inhibition are the interactions of His185I and Gln190I (inhibitor residues contain an "I"-suffix) of this loop with Arg71 and Phe279/Tyr198 of the protease moiety, respectively. These protease residues also contribute to latency maintenance in A/B-type zymogens (see Other Domains of Funnelins), but are absent in N/E-forms, which are not secreted as pro-enzymes. Overall, the contact modus of latexin with the protease moiety is reminiscent of the way the pro-domain blocks the mature enzyme in A/B funnelin zymogens.

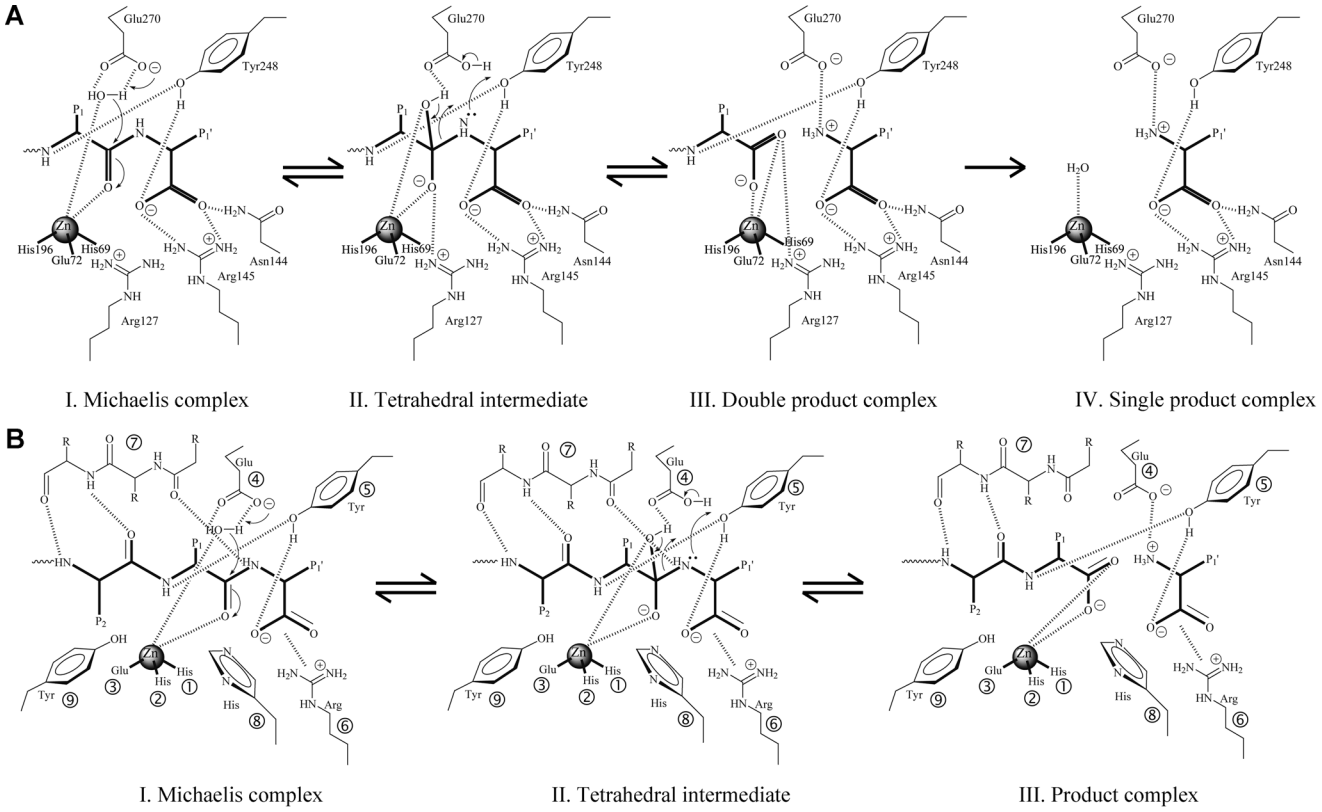
A second mechanism of inhibition is exerted through the exogenous inhibitors PCI (39 residues), LCI (66 residues), and *R. bursa* TCI (75 residues). These are small multi-cysteine proteins, whose structure is basically maintained by intra-molecular disulfide bonds (three in PCI, four in LCI, and six in TCI). Only LCI contains substantial regular secondary-structure elements in the form of a twisted antiparallel four-stranded β-sheet, which accommodates an α-helix on its convex side. In turn, TCI consists of two small ~35-residue domains connected by a short link. While the C-terminal sub-domain exerts the inhibitory function at the active site, the N-terminal domain approaches an exosite as shown for the BTAFI/TCI complex (Figure 8d; Sanglas *et al*., 2008). In all three cases, a conserved seven-residue segment at the C-terminus binds in a substrate-like manner to the non-primed side of the cleft and thus contributes to inhibition (Rees and Lipscomb, 1980; Reverter *et al*., 2000; Arolas *et al*., 2005b; Sanglas *et al*., 2008). The C-terminal residues are cleaved by the enzymes, but remain bound to the specificity pocket. The newly generated C-terminus coordinates the catalytic zinc and the preceding two residues interact with S_1 and S_2. In addition, strong ancillary contacts with funnel-rim loops stabilize the complexes. Such substrate-like binding has been described for other MP inhibitors, such as the tissue inhibitors of metalloproteinases, which target matrix metalloproteinases, and endogenous serralysin inhibitors (Baumann *et al*., 1995; Gomis-Rüth *et al*., 1997b). In these cases, however, binding is exerted through the N-terminal chain ends, which occupy the primed side of the cleft.

Mechanisms of Hydrolysis

Extensive studies on the mechanisms of funnelins and other MCPs have been based on complexes with substrates, products and transition-state analogs of BCPA1 and related enzymes such as HCPA4. In addition, many conclusions have been derived from the study of the endopeptidases thermolysin and astacin, which are postulated to share the bulk of their catalytic mechanism with funnelins, as a result of convergent evolution (Christianson, 1991; Grams *et al*., 1996; Gomis-Rüth, 2003). In chemical terms, peptide bond hydrolysis is the substitution of a hydroxide for the amide nitrogen and the attached downstream (poly)peptide chain. The current view of catalysis assumes a water-induced pathway following a general base/acid mechanism. A polarized solvent molecule or hydroxide performs a nucleophilic attack on the scissile peptide bond, mostly at pH values close to 7 (Hartsuck and Lipscomb, 1971; Kester and Matthews, 1977; Auld, 1987; Matthews, 1988; Christianson and Lipscomb, 1989; Christianson, 1991; Kim and Lipscomb, 1991; Grams *et al*., 1996; Polgár, 1999; Auld, 2004a; 2004b; Bayés *et al*., 2007). This catalytic mechanism is reminiscent of that of the type $B_{AC}2$ described for secondary amide hydrolysis (de Wolfe and Newcomb, 1971; March, 1985).

In the unbound enzyme, the catalytic zinc ion coordinates to His69, Glu72, His196, and a solvent molecule (see Catalytic Domain and Active Site of Funnelins), which in turn hydrogen-bond to the side chain of Glu270. This interaction fixes the position of the solvent and limits the conformational disorder about the zinc-solvent oxygen bond, thus reducing the entropic barrier to catalysis (Merz, 1990; Christianson, 1991). Another function of Glu270 is to prevent small anions from displacing the metal-bound solvent (Auld, 1997). The cation enhances solvent nucleophilicity and reduces its pKa value from 15.7 to ~7, as shown for a small-molecule metalloamide model complex (Groves and Olson, 1985; Christianson, 1991). In contrast, the presence of an acidic group, Glu72, in the metal coordination sphere diminishes the effective charge of the latter and its capacity to act as a Lewis acid. It also makes solvent polarization more difficult, which reduces the activity (Christianson and Cox, 1999) in comparison to cations coordinated to protein side chains, which are uncharged in physiological conditions. This is the case for metzincins, in which the metal is bonded by three histidines (first pKa value of 6–6.5) (Christianson, 1991; Gomis-Rüth, 2003; Auld, 2004a). Therefore, the presence of a second polarizing element in the form the glutamate carboxylate of Glu270 may provide the degree of solvent nucleophilicity required for the enzymatic reaction. Protease mutants, in which this glutamate was replaced by non-acidic residues, were catalytically inactive or showed only residual activity (Devault *et al*., 1988; Toma *et al*., 1989; Cha and Auld, 1997).

Before catalysis, the substrate is accommodated within the active-site cleft, forming a Michaelis complex (Figure 9a). Here, the P_1' side chain of the scissile residue interacts with the specificity pocket. The shallow active-site cleft of funnelins entails that substrate binding on the non-primed side is rather weak and unspecific, which contributes to high catalytic efficiency and broad substrate specificity (Laethem *et al*., 1996). It is generally accepted that substrate binding leads to the rearrangement of the side chain of Tyr248, which moves up to 12Å from a surface "up" position in the unbound enzyme to a "down" position to participate in substrate fixation (Reeke *et al*., 1967; Rees *et al*., 1983; Gardell *et al*., 1985; Hilvert *et al*., 1986; Guasch *et al*., 1992; García-Sáez *et al*., 1997), although there is some controversy on this point (Bukrinsky *et al*., 1998). The tyrosine



C

	①	②	③	④	⑤	⑥	⑦	⑧	⑨
HSTOP	473	477	502	474	612	EP	425-427	600	605
NEURO	474	478	503	475	613	EP	426-428	601	606
ECDCP	469	473	498	470	614	593	424-426	601	607
OPEPF	356	360	384	357	490	EP	327-329	479	486
PFUCP	269	273	299	270	423	351	239-241	411	419
TCMCP	265	269	296	266	421	348	236-238	409	417
TTHCP	276	280	306	277	430	358	246-248	418	426
HACE2	374	378	402	375	515	273	346-348	505	Arg514
DMACE	367	371	395	368	507	Lys495+Tyr504	338-340	497	Arg506
HACEC	383	387	411	384	523	Lys511+Tyr520	354-356	513	Arg522
HACEN	361	365	389	362	501	Lys489+Tyr498	332-334	491	Arg500

FIG. 9. MCP-mediated catalytic mechanism. (A) Scheme for the consensus mechanism established for funnelins with the residues key for catalysis depicted and labeled (BCPA1 numbering). The catalytic zinc ion is shown as a sphere and hydrogen bonds as dashed lines. The three protein ligands are represented by sticks. A possible alternative is that in III and IV the proton of the newly created N-terminus is bound to the general base/acid Glu270. (B) Mechanism proposed for cowrins. In the absence of a uniform numbering for this family, the different (potentially) intervening residues are denoted with a circled number (1–9).1–3 are the zinc-binding protein residues contained in the active-site and glutamate helices, 4 is the general base/acid glutamate, 5 is a tyrosine residue from the tyrosine helix responsible for substrate fixation, 6 is an arginine that may bind the C-terminal carboxylate in MCP-cowrins, 7 denotes the upper-rim strand, 8 is a (putative) binding/stabilizing histidine, and 9 is a potential substrate-binding tyrosine. EP stands for endopeptidase, i.e. in these cowrins no residue is expected to be responsible for C-terminus binding. In ACE-like peptidyl dipeptidases, the function of 6 might be performed by a lysine and a tyrosine. (C) Residues hypothetically occupying the positions depicted in (B) in each structure according to the corresponding PDB entry (see Table 1).

hydroxyl receives a hydrogen bond from the P_1 amide nitrogen and hydrogen bonds to the C-terminal carboxylate of the substrate (Kim and Lipscomb, 1991). This group is further anchored to the cleft by a bidentate salt bridge with Arg145 and by a hydrogen bond to Asn144 Nδ2. In addition, there is evidence for a shift from a bidentate to a monodentate zinc coordination mode of Glu72 (Christianson, 1991).

Once the substrate is bound, the scissile carbonyl oxygen enters the metal coordination sphere, which becomes pentacoordinate. This is facilitated by the flexible requirements for coordination geometry of and by the speed a ligand is exchanged in divalent zinc, which evinces *borderline* hardness (Pearson, 1963; Fraústo da Silva and Williams, 2001) but no ligand-field stabilization (Christianson and Cox, 1999). The solvent molecule remains bound to the cation, but is displaced slightly towards the carboxylate of Glu270, which acts as a general base and abstracts a proton from the water (Figure 9a, step I). Some authors postulate that the solvent is a ground-state hydroxide (Christianson and Cox, 1999). Whatever the case, the resulting (incipient) hydroxide performs a nucleophilic attack on the scissile amide carbonyl carbon following a tetrahedral mechanism of type $B_{AC}2$ rather than a conventional S_N2 mechanism (March, 1985). It is unclear whether these two steps are concerted or whether the reaction proceeds via a hydroxide and a protonated Glu270 carboxylate. In either case, this would lead to a negatively-charged tetrahedral reaction intermediate centered on an sp^3-hybridized *gem*-diolate group. The latter interacts in a bidentate manner with, and is thus stabilized by, the zinc ion, which remains pentacoordinate (Kim and Lipscomb, 1991). Further stabilization of the diolate is provided by the side chain of Arg127 (Phillips *et al*., 1990). This residue plays the role of an *oxyanion hole*, which stabilizes reaction intermediates in other classes of proteases (Bryan *et al*., 1986). Glu270 subsequently acts as a general acid catalyst, delivering the proton captured from the solvent to the scissile amide nitrogen, which becomes a secondary ammonium (Monzingo and Matthews, 1984). The remaining *gem*-diolate proton is either captured by Glu270, which again acts as a base, and then transferred to the secondary ammonium, or is conveyed to it directly (Figure 9a, step II) (Hangauer *et al*., 1984). In support of the first hypothesis is the finding that an oxyanionic *gem*-diolate can lose its proton, giving rise to an oxydianionic form, as observed during the alkaline hydrolysis of anilides (Biechler and Taft, 1957). In any case, this second proton-shuffling results in the reaction intermediate collapsing to two products containing a carboxylate and an α-ammonium group, respectively, i.e. the peptide bond does not break until the nitrogen has received a second proton (Eriksson and Holst, 1966) (Figure 9a, step III). An alternative to this order of events entails that transfer of the first proton to the amide nitrogen by Glu270 is simultaneous with the capture of the remaining *gem*-diolate proton by Glu270 and with peptide bond rupture (see Figure 31 of Christianson, 1991). In this case, Glu270 would be protonated after the cleavage event and the newly generated N-terminus of the downstream product would be an amino group. This hypothesis would entail a subsequent transfer of the proton from Glu270 to this amino group while still bound to the active-site cleft to render an α-ammonium group. However, this alternative is less favored as $R\text{-}NH_2$ is a worse leaving group for the resolution of the tetrahedral intermediate than $R\text{-}NH_3^+$ (March, 1985). The two cleavage products remain initially bound to the enzyme in the form of a double-product complex (Bayés *et al*., 2007). At this stage, the enzyme could perform the reverse reaction, i.e. the synthesis of a peptide bond (Auld, 2004a). The newly formed C-terminus interacts with the catalytic zinc in a bidentate manner and the upstream amide nitrogen remains hydrogen-bonded to Tyr248 Oη. In turn, the new N-terminus hydrogen bonds to Glu270 Oε2 and/or the catalytic zinc. Subsequently, the N-terminal product leaves and the catalytic solvent molecule is replenished, while the C-terminal product is still bound to the zinc ion, Glu270 and Arg145 (and, putatively, to Asn144 and Tyr248) as a single product complex (Guasch *et al*., 1992; García-Castellanos *et al*., 2005) (Figure 9a, step IV). Eventually, the amino acid in S_1' leaves and the enzyme is ready for subsequent catalysis.

In contrast to funnelins, there are very few studies on the mechanism of activity of cowrins. Furthermore, only ECDCP has been reported in complex with a substrate or a reaction product (Comellas-Bigler *et al*., 2005). An additional difficulty stems from the fact that cowrins encompass carboxypeptidases but also peptidyl dipeptidases and oligo-endopeptidases. In any case, it has been suggested that these enzymes that have mostly zinc but also occasionally cobalt in their active sites catalyze their substrates in a similar way to funnelins and thermolysin (Matthews, 1988; Arndt *et al*., 2002; Towler *et al*., 2004; Comellas-Bigler *et al*., 2005). Likewise, the substrate must be anchored to the active-site cleft on the enzyme to form the Michaelis complex. However, main-chain atoms of the upper-rim strand of the enzyme are involved here in the formation of an antiparallel β-ribbon with the substrate in extended conformation (7 in Figures 9b, 9c). In addition, cowrins close their entrance to bring the protein sub-domains closer and to narrow the active-site cleft once a substrate is trapped (see the Active-Site Cleft section). Solvent polarization and general base/acid function is likewise carried out by a glutamate residue, viz the one imbedded in the short metal-binding consensus sequence, HEXXH, which further provides the two histidine metal ligands (2–4 in Figures 9b, 9c). The third metal ligand is a downstream glutamate, as found in funnelins, so the cation may be similarly polarized in funnelins and cowrins. The C-terminal carboxylate group of a substrate is bound by an arginine in those cowrins with carboxypeptidase activity and in the peptidyl dipeptidase ECDCP (6 in Figures 9b, 9c). This binding is carried out by a conserved lysine-tyrosine tandem in ACE-like cowrins. Furthermore, a tyrosine recalling Tyr248 of funnelins (5 in Figures 9b, 9c), together with a second conserved tyrosine (a conserved arginine in ACE-like cowrins; 9 in Figures 9b, 9c), fixes the substrate main chain on the non-primed side of the cleft. The Tyr248-equivalent may account for scissile carbonyl binding and stabilization of the tetrahedral reaction intermediate (Comellas-Bigler *et al*., 2005). Finally,

substrate binding on the primed side implicates a conserved histidine residue (8 in Figures 9b, 9c), which may be topologically equivalent to His231 in thermolysin. This residue hydrogen-bonds to the incoming substrate and stabilizes the tetrahedral reaction intermediate in the bacterial enzyme (Matthews, 1988). This could also happen in cowrins, thus complementing the function of the Tyr248-equivalent. Except for these differences, the working mechanism of cowrins is expected to resemble that of funnelins (Figure 9).

ACKNOWLEDGEMENTS

This study was supported by the following grants: BIO2006-02668, PSE-010000-2007-1, and CONSOLIDER-INGENIO 2010 Project "La Factoría de Cristalización" (CSD2006-00015) from Spanish ministries; EU FP6 Strep Project LSHG-2006-018830 "CAMP"; and EU FP7 Collaborative Project 223101 AntiPathoGN and 2005SGR00280 from the National Catalan Government. Robin Rycroft is thanked for helpful contributions to the manuscript and Frédéric Weber for assessment in conchology (http://pagesperso-orange.fr/frederic.weber/coquillages%20accueuil.htm).

REFERENCES

Aagaard, A., Listwan, P., Cowieson, N., Huber, T., Ravasi, T., Wells, C., Flanagan, J.U., Hume, D.A., Kobe, B., and Martin, J.L. 2005. An inflammatory role for the mammalian carboxypeptidase inhibitor latexin: relationship to cystatins and the tumor suppressor TIG1. *Structure* 13:309–317.

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.

Arndt, J.W., Hao, B., Ramakrishnan, V., Cheng, T., Chan, S.I., and Chan, M.K. 2002. Crystal structure of a novel carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Structure* 10:215–224.

Arolas, J.L., Lorenzo, J., Rovira, A., Castella, J., Aviles, F.X., and Sommerhoff, C.P. 2005a. A carboxypeptidase inhibitor from the tick Rhipicephalus bursa: isolation, cDNA cloning, recombinant expression, and characterization. *J Biol Chem* 280:3441–3448.

Arolas, J.L., Popowicz, G.M., Lorenzo, J., Sommerhoff, C.P., Huber, R., Avilés, F.X., and Holak, T.A. 2005b. 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.

Arolas, J.L., Vendrell, J., Avilés, F.X., and Fricker, L.D. 2007. Metallocarboxypeptidases: emerging drug targets in biomedicine. *Curr Pharm Des* 13:349–366.

Auld, D.S. 1987. Acyl group transfer–metalloproteinases, pp. 241–258. In: Page, M.I., and Williams, A., eds, *Enzyme Mechanisms*, London: Royal Society of Chemistry.

Auld, D.S. 1997. Zinc catalysis in metalloproteases. *Struct Bond* 89:29–50.

Auld, D.S. 2004a. 68. Catalytic mechanisms for metallopeptidases, pp. 268–289. In: Barrett, A.J., Rawlings, N.D., and Woessner Jr., J.F., eds, *Handbook of Proteolytic Enzymes*, Vol. 1. London: Elsevier Academic Press.

Auld, D.S. 2004b. 240. Carboxypeptidase A, pp. 812–821. In: Barrett, A.J., Rawlings, N.D., and Woessner Jr., J.F., eds, *Handbook of Proteolytic Enzymes*, Vol. 1. London: Elsevier Academic Press.

Bajzar, L., Manuel, R., and Nesheim, M.E. 1995. Purification and characterization of TAFI, a thrombin-activable fibrinolysis inhibitor. *J Biol Chem* 270:14477–14484.

Barrett, A.J., Rawlings, N.D., and Woessner, J.F. 2004. *Handbook of Proteolytic Enzymes*. Amsterdam: Elsevier.

Baumann, U., Bauer, M., Letoffe, S., Delepelaire, P., and Wandersman, C. 1995. Crystal structure of a complex between *Serratia marcescens* metallo-protease and an inhibitor from Erwinia chrysanthemi. *J Mol Biol* 248:653–661.

Bayés, A., Fernández, D., Solà, M., Marrero, A., García-Piqué, S., Avilés, F.X., Vendrell, J., and Gomis-Rüth, F.X. 2007. Caught after the Act: A human A-type metallocarboxypeptidase in a product complex with a cleaved hexapeptide. *Biochemistry* 46:6921–6930.

Biechler, S.S. and Taft, R.W.J. 1957. The effect of structure on kinetics and mechanism of the alkaline hydrolysis of anilides. *J Am Chem Soc* 79:4927–4935.

Bode, W., Gomis-Rüth, F.X., and Stöcker, W. 1993. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. *FEBS Lett* 331:134–140.

Brändén, C.-I. 1980. Relation between structure and function of α/β proteins. *Q Rev Biophys* 13:317–338.

Brown, C.K., Madauss, K., Lian, W., Beck, M.R., Tolbert, W.D., and Rodgers, D.W. 2001. Structure of neurolysin reveals a deep channel that limits substrate access. *Proc Natl Acad Sci USA* 98:3127–3132.

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 USA* 83:3743–3745.

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.

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.

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.

Cha, J., and Auld, D.S. 1997. Site-directed mutagenesis of the active site glutamate in human matrilysin: investigation of its role in catalysis. *Biochemistry* 36:16019–16024.

Cheng, T.C., Ramakrishnan, V., and Chan, S.I. 1999. Purification and characterization of a cobalt-activated carboxypeptidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Protein Sci* 8:2474–2486.

Chothia, C. and Lesk, A.M. 1986. The relation between the divergence of sequence and structure in proteins. *EMBO J* 5:823–826.

Christianson, D.W. 1991. Structural biology of zinc. *Adv Protein Chem* 42:281–355.

Christianson, D.W. and Cox, J.D. 1999. Catalysis by metal-activated hydroxide in zinc and manganese metalloenzymes. *Annu Rev Biochem* 68:33–57.

Christianson, D.W. and Lipscomb, W.N. 1989. Carboxypeptidase A. *Acc Chem Res* 22:62–69.

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.

Coleman, D.E., Berghuis, A.M., Lee, E., Linder, M.E., Gilman, A.G., and Sprang, S.R. 1994. Structures of active conformations of Gia1 and the mechanism of GTP hydrolysis. *Science* 265:1405–1412.

Coll, M., Guasch, A., Aviles, F.X., and Huber, R. 1991. Three-dimensional structure of porcine procarboxypeptidase B: a structural basis of its inactivity. *EMBO J* 10:1–9.

Comellas-Bigler, M., Lang, R., Bode, W., and Maskos, K. 2005. Crystal structure of the *E. coli* dipeptidyl carboxypeptidase Dcp: further indication of a ligand-dependent hinge movement mechanism. *J Mol Biol* 349:99–112.

Cornell, M.J., Williams, T.A., Lamango, N.S., Coates, D., Corvol, P., Soubrier, F., Hoheisel, J., Lehrach, H., and Isaac, R.E. 1995. Cloning and expression of an evolutionary conserved single-domain angiotensin converting enzyme from *Drosophila melanogaster*. *J Biol Chem* 270:13613–13619.

Corradi, H.R., Schwager, S.L., Nchinda, A.T., Sturrock, E.D., and Acharya, K.R. 2006. Crystal structure of the N domain of human somatic angiotensin I-converting enzyme provides a structural basis for domain-specific inhibitor design. *J Mol Biol* 357:964–974.

Davis, M.I., Bennett, M.J., Thomas, L.M., and Bjorkman, P.J. 2005. Crystal structure of prostate-specific membrane antigen, a tumor marker and peptidase. *Proc Natl Acad Sci USA* 102:5981–5986.

de Wolfe, R.H., and Newcomb, R.C. 1971. Hydrolysis of formanilides in alkaline solutions. *J Org Chem* 36:3870–3878.

Devault, A., Nault, C., Zollinger, M., Fournie-Zaluski, M.C., Roques, B.P., Crine, P., and Boileau, G. 1988. Expression of neutral endopeptidase (enkephalinase) in heterologous COS-1 cells. Characterization of the recombinant enzyme and evidence for a glutamic acid residue at the active site. *J Biol Chem* 263:4033–4040.

Díez-Díaz, M., Conejero, V., Rodrigo, I., Pearce, G., and Ryan, C.A. 2004. Isolation and characterization of wound-inducible carboxypeptidase inhibitor from tomato leaves. *Phytochemistry* 65:1919–1924.

Domingues, F.S., Lackner, P., Andreeva, A., and Sippl, M.J. 2000. Structure-based evaluation of sequence comparison and fold recognition alignment accuracy. *J Mol Biol* 297:1003–1013.

Eaton, D.L., Malloy, B.E., Tsai, S.P., Henzel, W., and Drayna, D. 1991. Isolation, molecular cloning, and partial characterization of a novel carboxypeptidase B from human plasma. *J Biol Chem* 266:21833–21838.

Ehlers, M.R., Fox, E.A., Strydom, D.J., and Riordan, J.F. 1989. Molecular cloning of human testicular angiotensin-converting enzyme: the testis isozyme is identical to the C-terminal half of endothelial angiotensin-converting enzyme. *Proc Natl Acad Sci USA* 86:7741–7745.

Eriksson, S.O., and Holst, C. 1966. Hydrolysis of anilides. II. Hydrolysis of trifluoro- and trichloroacetanilide by hydroxyl ions and by some other catalysts. *Acta Chem Scan* 20:1892–1906.

Esnouf, R.M. 1997. An extensively modified version of Molscript that includes greatly enhanced coloring capabilities. *J Mol Graph* 15:133–138.

Evans, S.V. 1993. SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J Mol Graphics* 11:134–138.

Faming, Z., Kobe, B., Stewart, C.B., Rutter, W.J., and Goldsmith, E.J. 1991. Structural evolution of an enzyme specificity. The structure of rat carboxypeptidase A2 at 1.9-A resolution. *J Biol Chem* 266:24606–24612.

Fraústo da Silva, J.J.R. and Williams, R.J.P. 2001. *The Biological Chemistry of the Elements: the inorganic chemistry of life*. New York: Oxford University Press.

Fushimi, N., Ee, C.E., Nakajima, T., and Ichishima, E. 1999. Aspzincin, a family of metalloendopeptidases with a new zinc-binding motif. Identification of new zinc-binding sites (His(128), His(132), and Asp(164)) and three catalytically crucial residues (Glu(129), Asp(143), and Tyr(106)) of deuterolysin from *Aspergillus oryzae* by site-directed mutagenesis. *J Biol Chem* 274:24195–24201.

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.

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.

Gardell, S.J., Craik, C.S., Hilvert, D., Urdea, M.S., and Rutter, W.J. 1985. Site-directed mutagenesis shows that tyrosine 248 of carboxypeptidase A does not play a crucial role in catalysis. *Nature* 317:551–555.

Gerdts, C.J., Tereshko, V., Yadav, M.K., Dementieva, I., Collart, F., Joachimiak, A., Stevens, R.C., Kuhn, P., Kossiakoff, A., and Ismagilov, R.F. 2006. Time-controlled microfluidic seeding in nL-volume droplets to separate nucleation and growth stages of protein crystallization. *Angew Chem Int Ed Engl* 45:8156–8160.

Gerstein, M., and Krebs, W. 1998. A database of macromolecular motions. *Nucl Acids Res* 26:4280–4290.

Gerstein, M., Lesk, A.M., and Chothia, C. 1994. Structural mechanisms for domain movements in proteins. *Biochemistry* 33:6739–6749.

Gomis-Rüth, F.X. 2003. Structural aspects of the *metzincin* clan of metalloendopeptidases. *Mol Biotech* 24:157–202.

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.

Gomis-Rüth, F.X., Kress, L.F., Kellermann, J., Mayr, I., Lee, X., Huber, R., and Bode, W. 1994a. Refined 2.0Å X-ray crystal structure of the zinc-endopeptidase adamalysin II. Primary and tertiary structure determination, refinement, molecular structure and comparison with astacin, collagenase and thermolysin. *J Mol Biol* 239:513–544.

Gomis-Rüth, F.X., Nar, H., Grams, F., Yallouros, I., Küsthardt, U., Zwilling, R., Bode, W., and Stöcker, W. 1994b. Crystal structures, spectroscopic features and catalytic properties of cobalt(II)-, copper(II)-, nickel(II)-, and mercury(II)-derivatives of the zinc-endopeptidase astacin. A correlation of structure and proteolytic activity. *J Biol Chem* 269:17111–17117.

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 proproteinase E and chymotrypsinogen C. *EMBO J* 14:4387–4394.

Gomis-Rüth, F.X., Gomez-Ortiz, M., Vendrell, J., Ventura, S., Bode, W., Huber, R., and Aviles, F.X. 1997a. Crystal structure of an oligomer of proteolytic zymogens: detailed conformational analysis of the bovine ternary complex and implications for their activation. *J Mol Biol* 269:861–880.

Gomis-Rüth, F.X., Maskos, K., Betz, M., Bergner, A., Huber, R., Suzuki, K., Yoshida, N., Nagase, H., Brew, K., Bourenkov, G.P., Bartunik, H., and Bode, W. 1997b. Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1. *Nature* 389:77–81.

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.

Gong, H., Zhou, J., Liao, M., Hatta, T., Harnnoi, T., Umemiya, R., Inoue, N., Xuan, X., and Fujisaki, K. 2007. Characterization of a carboxypeptidase inhibitor from the tick *Haemaphysalis longicornis*. *J Insect Physiol* 53:1079–1087.

Grams, F., Dive, V., Yiotakis, A., Yiallouros, I., Vassiliou, S., Zwilling, R., Bode, W., and Stöcker, W. 1996. Structure of astacin with a transition-state analogue inhibitor. *Nature Struc Biol* 3:671–675.

Groves, J.T., and Olson, J.R. 1985. Models of zinc-containing proteases – Rapid amide hydrolysis by an unusually acidic Zn^{2+}-OH_2 complex. *Inorg Chem* 24:2715–2717.

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.

Guindon, S., and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704.

Guy, J.L., Lambert, D.W., Turner, A.J., and Porter, K.E. 2008. Functional angiotensin converting enzyme-2 (ACE2) is expressed in human cardiac myofibroblasts. *Exp Physiol* 93:421:551–554.

Hangauer, D.G., Monzingo, A.F., and Matthews, B.W. 1984. An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxymethyl dipeptides. *Biochemistry* 23:5730–5741.

Hartsuck, J.A., and Lipscomb, W.N. 1971. Carboxypeptidase A, pp. 1–56. In: Boyer, P.D., ed. *The Enzymes. Hydrolysis: peptide bonds*, Vol. III. New York: Academic Press.

Hashida, Y., and Inouye, K. 2007. Kinetic analysis of the activation-and-inhibition dual effects of cobalt ion on thermolysin activity. *J Biochem* 141:843–853.

Hendriks, D., Scharpé, S., van Sande, M., and Lommaert, M.P. 1989. Characterisation of a carboxypeptidase in human serum distinct from carboxypeptidase N. *J Clin Chem Clin Biochem* 27:277–285.

Hillmayer, K., Macovei, A., Pauwels, D., Compernolle, G., Declerck, P.J., and Gils, A. 2006. Characterization of rat thrombin-activatable fibrinolysis inhibitor (TAFI) a comparative study assessing the biological equivalence of rat, murine and human TAFI. *J Thromb Haemost* 4:2470–2477.

Hilvert, D., Gardell, S.J., Rutter, W.J., and Kaiser, E.T. 1986. Evidence against a crucial role for the phenolic hydroxyl of Tyr-248 in peptide and ester hydrolyses catalyzed by carboxypeptidase A: Comparative studies of the pH dependencies of the native and Phe-248 mutant forms. *J Am Chem Soc* 108:5298–5304.

Holland, D.R., Tronrud, D.E., Pley, H.W., Flaherty, K.M., Stark, W., Jansonius, J.N., McKay, D.B., and Matthews, B.W. 1992. Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis. *Biochemistry* 31:11310–11316.

Hooper, N.M. 1994. Families of zinc metalloproteases. *FEBS Lett* 354:1–6.

Hooper, N.M. 2002. Proteases: a primer, pp. 1–8. In: Hooper, N.M., ed.*Essays in Biochemistry—Proteases in biology and medicine*. London: Portland Press.

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.

Hurst, D., Rylett, C.M., Isaac, R.E., and Shirras, A.D. 2003. The *Drosophila* angiotensin-converting enzyme homologue AnCE is required for spermiogenesis. *Dev Biol* 254:238–247.

Jiang, W. and Bond, J.S. 1992. Families of metalloendopeptidases and their relationships. *FEBS Lett* 312:110–114.

Jongeneel, C.V., Bouvier, J., and Bairoch, A. 1989. A unique signature identifies a family of zinc-dependent metallopeptidases. *FEBS Lett* 242:211–214.

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.

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.

Keil, C., Maskos, K., Than, M., Hoopes, J.T., Huber, R., Tan, F., Deddish, P.A., Erdos, E.G., Skidgel, R.A., and Bode, W. 2007. Crystal structure of the human carboxypeptidase N (kininase I) catalytic domain. *J Mol Biol* 366:504–516.

Kester, W.R. and Matthews, B.W. 1977. Crystallographic study of dipeptide inhibitors to thermolysin: implications for the mechanism of catalysis. *Biochemistry* 16:2506–2516.

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.

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.

Kim, H.M., Shin, D.R., Yoo, O.J., Lee, H., and Lee, J.O. 2003. Crystal structure of *Drosophila* angiotensin I-converting enzyme bound to captopril and lisinopril. *FEBS Lett* 538:65–70.

Kuroki, K., Eng, F., Ishikawa, T., Turck, C., Harada, F., and Ganem, D. 1995. gp180, a host cell glycoprotein that binds duck hepatitis B virus particles, is encoded by a member of the carboxypeptidase gene family. *J Biol Chem* 270:15022–15028.

Laethem, R.M., Blumenkopf, T.A., Cory, M., Elwell, L., Moxham, C.P., Ray, P.H., Walton, L.M., and Smith, G.K. 1996. Expression and characterization of human pancreatic preprocarboxypeptidase A1 and preprocarboxypeptidase A2. *Archives of Biochemistry & Biophysics* 332:8–18.

Latt, S.A., Auld, D.S., and Vallee, B.L. 1972. Distance measurements at the active site of carboxypeptidase A during catalysis. *Biochemistry* 11:3015–3022.

Lee, S.H., Minagawa, E., Taguchi, H., Matsuzawa, H., Ohta, T., Kaminogawa, S., and Yamauchi, K. 1992. Purification and characterization of a thermostable carboxypeptidase (carboxypeptidase Taq) from *Thermus aquaticus* YT-1. *Biosci Biotechnol Biochem* 56:1839–1844.

Lee, S.H., Taguchi, H., Yoshimura, E., Minagawa, E., Kaminogawa, S., Ohta, T., and Matsuzawa, H. 1994. Carboxypeptidase Taq, a thermostable zinc enzyme, from *Thermus aquaticus* YT-1: molecular cloning, sequencing, and expression of the encoding gene in *Escherichia coli*. *Biosci Biotechnol Biochem* 58:1490–1495.

Levin, Y., Skidgel, R.A., and Erdos, E.G. 1982. Isolation and characterization of the subunits of human plasma carboxypeptidase N (kininase i). *Proc Natl Acad Sci USA* 79:4618–4622.

Li, F., Li, W., Farzan, M., and Harrison, S.C. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309:1864–1868.

Lipscomb, W.N., Hartsuck, J.A., Reeke Jr., G.N., Quiocho, F.A., Bethge, P.H., Ludwig, M.L., Steitz, T.A., Muirhead, H., and Coppola, J.C. 1968. The structure of carboxypeptidase A. VII. The 2.0-Å resolution studies of the enzyme and of its complex with glycyltyrosine, and mechanistic deductions. *Brookhaven Symp Biol* 21:24–90.

López-Otín, C. and Overall, C.M. 2002. Protease degradomics: a new challenge for proteomics. *Nat Rev Mol Cell Biol* 3:509–519.

Lupyan, D., Leo-Macias, A., and Ortiz, A.R. 2005. A new progressive-iterative algorithm for multiple structure alignment. *Bioinformatics* 21:3255–3263.

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.

March, J. 1985. *Advanced Organic Chemistry—Reactions, mechanisms and structure*. New York: John Wiley & Sons.

Matthews, B.W. 1988. Structural basis of the action of thermolysin and related zinc peptidases. *Acc Chem Res* 21:333–340.

Matthews, B.W., Jansonius, J.N., Colman, P.M., Schoenborn, B.P., and Dupourque, D. 1972. Three-dimensional structure of thermolysin. *Nature* 238:37–41.

Merz, K.M., Jr. 1990. Insights into the function of the zinc hydroxide-Thr199-Glu106 hydrogen bonding network in carbonic anhydrases. *J Mol Biol* 214:799–802.

Mesters, J.R., Barinka, C., Li, W., Tsukamoto, T., Majer, P., Slusher, B.S., Konvalinka, J., and Hilgenfeld, R. 2006. Structure of glutamate carboxypeptidase II, a drug target in neuronal damage and prostate cancer. *EMBO J* 25:1375–1384.

Mittl, P.R. and Grütter, M.G. 2006. Opportunities for structure-based design of protease-directed drugs. *Curr Opin Struct Biol* 16:769–775.

Monzingo, A.F. and Matthews, B.W. 1984. Binding of N-carboxymethyl dipeptide inhibitors to thermolysin determined by X-ray crystallography: a novel class of transition-state analogues for zinc peptidases. *Biochemistry* 23:5724–5729.

Nagase, H. 2001. Metalloproteases. *Curr Protoc Protein Sci* Suppl. 24:21.24.21-21.24.13.

Nagata, K., Tsutsui, S., Lee, W.C., Ito, K., Kamo, M., Inoue, Y., and Tanokura, M. 2004. Crystallization and preliminary X-ray analysis of carboxypeptidase 1 from *Thermus thermophilus*. *Acta Crystallogr sect D* 60:1445–1446.

Natesh, R., Schwager, S.L.U., Sturrock, E.D., and Acharya, K.R. 2003. Crystal structure of the human angiotensin-converting enzyme-lysinopril complex. *Nature*:in press.

Neurath, H. 1960. Carboxypeptidases A and B, pp. 11–36. In: Boyer, P.D., Lardy, H.A., and Myrback, K., eds. *Enzymes*, Vol. 4. New York: Academic Press.

Neurath, H. and Walsh, K.A. 1976. Role of proteolytic enzymes in biological regulation. *Proc Natl Acad Sci USA* 73:3825–3832.

Nicholls, A., Bharadwaj, R., and Honig, B. 1993. GRASP: graphical representation and analysis of surface properties. *Biophys J* 64(2, part 2):A166–A166.

Niemirowicz, G. and Cazzulo, J.J. 2007, Changing the substrate specificities of M32 peptidases of *T. cruzi* by site-directed mutagenesis. *XLIII. Annual Congress of the Argentinian Society for Research in Biochemical and Molecular Biology (SAIB)*, 17–20 November 2007, Mar del Plata (Argentina).

Niemirowicz, G., Fernandiz, D., Sola, M., et al. 2008. The molecular analysis of trypanosoma cruzi metallocarboxypeptidase 1 provides insight into fold and substrate specificity. *Mol. Microbiol*., in press.

Niemirowicz, G., Parussini, F., Agüero, F., and Cazzulo, J.J. 2007. Two metallocarboxypeptidases from the protozoan *Trypanosoma cruzi* belong to the M32 family, found so far only in prokaryotes. *Biochem J* 401:399–410.

Normant, E., Gros, C., and Schwartz, J.C. 1995a. Carboxypeptidase A isoforms produced by distinct genes or alternative splicing in brain and other extrapancreatic tissues. *J Biol Chem* 270:20543–20549.

Normant, E., Martres, M.P., Schwartz, J.C., and Gros, C. 1995b. Purification, cDNA cloning, functional expression, and characterization of a 26-kDa endogenous mammalian carboxypeptidase inhibitor. *Proc Natl Acad Sci USA* 92:12225–12229.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G., and Goldman, A. 1992. The α/β hydrolase fold. *Prot Engng* 5:197–211.

Ortiz, A.R., Strauss, C.E., and Olmea, O. 2002. MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison. *Prot Sci* 11:2606–2621.

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 USA* 102:3978–3983.

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.

Pearson, R.G. 1963. Hard and soft acids and bases. *J Am Chem Soc* 85:3533–3539.

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.

Perriere, G., and Gouy, M. 1996. WWW-query: an on-line retrieval system for biological sequence banks. *Biochimie* 78:36436–36439.

Phillips, M.A., Fletterick, R., and Rutter, W.J. 1990. Arginine 127 stabilizes the transition state in carboxypeptidase. *J Biol Chem* 265:20692–20698.

Polgár, L. 1999. Basic kinetic mechanisms of proteolytic enzymes, pp. 148–166. In: Sterchi, E.E., and Stöcker, W., eds.*Proteolytic enzymes–Tools and targets*. Berlin/Heidelberg: Springer Verlag.

Puente, X.S. and López-Otín, C. 1997. The PLEES-proteins–A family of structurally related enzymes widely distributed from bacteria to humans. *Biochem J* 332:947–949.

Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J., and Barrett, A.J. 2008. MEROPS: the peptidase database. *Nucleic Acids Res* 36 (Database issue):D320–D325.

Ray, K., Hines, C.S., Coll-Rodríguez, J., and Rodgers, D.W. 2004. Crystal structure of human thimet oligopeptidase provides insight into substrate recognition, regulation, and localization. *J Biol Chem* 279:20480–20489.

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 glycyltyrosine at 2.8-Å resolution. *Proc Natl Acad Sci USA* 58:2220–2226.

Rees, D.C. and Lipscomb, W.N. 1980. Structure of the potato inhibitor complex of carboxypeptidase A at 2.5-A resolution. *Proceedings of the National Academy of Sciences of the United States of America* 77:4633–4637.

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.

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.

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.

Reverter, D., Maskos, K., Tan, F., Skidgel, R.A., and Bode, W. 2004. Crystal structure of human carboxypeptidase M, a membrane-bound enzyme that regulates peptide hormone activity. *J Mol Biol* 338:257–269.

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.

Richardson, J.S. 1981. The anatomy and taxonomy of protein structure. *Adv Prot Chem* 34:167–339.

Richardson, J.S. 1985. Schematic drawings of protein structures. *Meth Enzymol* 115:359–380.

Rost, B. 1999. Twilight zone of protein sequence alignments. *Protein Eng* 12:85–94.

Rowsell, S., Pauptit, R.A., Tucker, A.D., Melton, R.G., Blow, D.M., and Brick, P. 1997. Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure* 5:337–347.

Sanglas, L., Valnickova, Z., Arolas, J.L., Pallarés, I., Guevara, T., Solà, M., Kristensen, T., Enghild, J.J., Avilés, F.X., and Gomis-Rüth, F.X. 2008. Structure of activated thrombin-activatable fibrinolysis inhibitor, a molecular link between coagulation and fibrinolysis. *Mol Cell* 31:598–606.

Saric, T., Beninga, J., Graef, C.I., Akopian, T.N., Rock, K.L., and Goldberg, A.L. 2001. Major histocompatibility complex class I-presented antigenic peptides are degraded in cytosolic extracts primarily by thimet oligopeptidase. *J Biol Chem* 276:36474–36481.

Schmid, M.F. and Herriott, J.R. 1976. Structure of carboxypeptidase B at 2.8 Å resolution. *J Mol Biol* 103:175–190.

Skidgel, R.A. 1988. Basic carboxypeptidases: regulators of peptide hormone activity. *Trends Pharmacol Sci* 9:299–304.

Skidgel, R.A. 1996. Structure and function of mammalian zinc carboxypeptidases, pp. 241–283. In: Hooper, N.M., ed. *Zinc Metalloproteases in Health and Disease*. London: Taylor and Francis.

Skidgel, R.A., Davis, R.M., and Tan, F. 1989. Human carboxypeptidase M. Purification and characterization of a membrane-bound carboxypeptidase that cleaves peptide hormones. *J Biol Chem* 264:2236–2241.

Song, L. and Fricker, L.D. 1997. Cloning and expression of human carboxypeptidase Z, a novel metallocarboxypeptidase. *Journal of Biological Chemistry* 272:10543–10550.

Soubrier, F., Alhenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G., and Corvol, P. 1988. Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proc Natl Acad Sci USA* 85:9386–9390.

Springman, E.B. 2004. Mast cell carboxypeptidase, pp. 828–830. In: Barrett, A.J., Rawlings, N.D., and Woessner Jr., J.F., eds, *Handbook of Proteolytic Enzymes,* Vol. 1. London: Elsevier.

Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.X., McKay, D.B., and Bode, W. 1995. The metzincins – Topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Prot Sci* 4:823–840.

Stothard, P. 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 28:1102–1104.

Sturrock, E.D., Natesh, R., van Rooyen, J.M., and Acharya, K.R. 2004. Structure of angiotensin I-converting enzyme. *Cell Mol Life Sci* 61:2677–2686.

Tan, F., Rehli, M., Krause, S.W., and Skidgel, R.A. 1997. Sequence of human carboxypeptidase D reveals it to be a member of the regulatory carboxypeptidase family with three tandem active site domains. *Biochem J* 327:81–87.

Teplyakov, A., Polyakov, K., Obmolova, G., Strokopytov, B., Kuranova, I., Osterman, A., Grishin, N., Smulevitch, S., Zagnitko, O., Galperina, O., Matz, M., and Stepanov, V. 1992. Crystal structure of carboxypeptidase T from *Thermoactinomyces vulgaris*. *Eur J Biochem* 208:281–288.

Theobald, D.L. and Wuttke, D.S. 2006. THESEUS: maximum likelihood superpositioning and analysis of macromolecular structures. *Bioinformatics* 22:2171–2172.

Toma, S., Campagnoli, S., De Gregoriis, E., Gianna, R., Margarit, I., Zamai, M., and Grandi, G. 1989. Effect of Glu-143 and His-231 substitutions on the catalytic activity and secretion of *Bacillus subtilis* neutral protease. *Protein Eng* 2:359–364.

Towler, P., Staker, B., Prasad, S.G., Menon, S., Tang, J., Parsons, T., Ryan, D., Fisher, M., Williams, D., Dales, N.A., Patane, M.A., and Pantoliano, M.W. 2004. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem* 279:17996–18007.

Turner, A.J., Hiscox, J.A., and Hooper, N.M. 2004. ACE2: from vasopeptidase to SARS virus receptor. *Trends Pharmacol Sci* 25:291–294.

Valnickova, Z., Thogersen, I.B., Potempa, J., and Enghild, J.J. 2007. Thrombin-activable fibrinolysis inhibitor (TAFI) zymogen is an active carboxypeptidase. *J Biol Chem* 282:3066–3076.

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.

Waldschmidt-Leitz, E. and Purr, A. 1929. Über Proteinase und Carboxy-Polypeptidase aus Pankreas. (XVII. Mitteilung zur Spezifizität tierischer Proteasen.). *Ber Dt Chem Ges* 62:2217–2226.

Willemse, J.L. and Hendriks, D.F. 2007. A role for procarboxypepidase U (TAFI) in thrombosis. *Front Biosci* 12:1973–1987.

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 Royal Soc London (Series B)* 111:280–297.

Woodman, Z.L., Schwager, S.L., Redelinghuys, P., Chubb, A.J., van der Merwe, E.L., Ehlers, M.R., and Sturrock, E.D. 2006. Homologous substitution of ACE C-domain regions with N-domain sequences: effect on processing, shedding, and catalytic properties. *Biol Chem* 387:1043–1051.

Xin, X., Varlamov, O., Day, R., Dong, W., Bridgett, M.M., Leiter, E.H., and Fricker, L.D. 1997. Cloning and sequence analysis of cDNA encoding rat carboxypeptidase D. *DNA Cell Biol* 16:897–905.

Yamin, R., Malgeri, E.G., Sloane, J.A., McGraw, W.T., and Abraham, C.R. 1999. Metalloendopeptidase EC 3.4.24.15 is necessary for Alzheimer's amyloid-β peptide degradation. *J Biol Chem* 274:18777–18784.

Zappacosta, F., Pessi, A., Bianchi, E., Venturini, S., Sollazzo, M., Tramontano, A., Marino, G., and Pucci, P. 1996. Probing the tertiary structure of proteins by limited proteolysis and mass spectrometry: the case of Minibody. *Prot Sci* 5:802–813.

Editor: Michael M. Cox