Cell, Vol. 94, 161–170, July 24, 1998, Copyright ©1998 by Cell Press

Prolyl Oligopeptidase: An Unusual β-Propeller Domain Regulates Proteolysis

Vilmos Fülöp,*†# Zsolt Böcskei,‡§ and László Polgár‖
*Laboratory of Molecular Biophysics
Department of Biochemistry
Oxford Centre for Molecular Sciences
University of Oxford
Oxford OX1 3QU
United Kingdom
†Department of Biological Sciences
University of Warwick
Coventry CV4 7AL
United Kingdom
‡Department of Theoretical Chemistry
L. Eötvös University
H-117 Budapest Pázmány P. s. 2
Hungary
§Chinoin Pharmaceutical and Chemical Works Ltd.
H-1325 Budapest
Hungary
‖Institute of Enzymology
Biological Research Center
Hungarian Academy of Sciences
H-1518 Budapest 112, P.O. Box 7
Hungary

Summary

Prolyl oligopeptidase is a large cytosolic enzyme that belongs to a new class of serine peptidases. The enzyme is involved in the maturation and degradation of peptide hormones and neuropeptides, which relate to the induction of amnesia. The 1.4 Å resolution crystal structure is presented here. The enzyme contains a peptidase domain with an α/β hydrolase fold, and its catalytic triad (Ser554, His680, Asp641) is covered by the central tunnel of an unusual β propeller. This domain makes prolyl oligopeptidase an oligopeptidase by excluding large structured peptides from the active site. In this way, the propeller protects larger peptides and proteins from proteolysis in the cytosol. The structure is also obtained with a transition state inhibitor, which may facilitate drug design to treat memory disorders.

Introduction

There are several regulatory mechanisms that protect proteins in biological systems from incident degradation by peptidases. Thus, peptidases are packaged in lysosomes within the cell or occur in inactive proenzyme (zymogen) forms that can be activated by limited proteolysis under appropriate conditions. Most interestingly, there are active endopeptidases that do not attack proteins. These enzymes must have some mechanism, so far unknown, by which they distinguish between large proteins, native or denatured, and unstructured peptides consisting of no more than about 30 amino acid residues. One such enzyme is prolyl oligopeptidase (Camargo et al., 1979; Moriyama et al., 1988), previously called prolyl endopeptidase or postproline cleaving enzyme. This enzyme is involved in the maturation and degradation of peptide hormones and neuropeptides (Wilk, 1983; Mentlein, 1988). Prolyl oligopeptidase has recently gained pharmaceutical interest, since specific inhibitors reverse scopolamine-induced amnesia in rats (Yoshimoto et al., 1987; Atack et al., 1991; Miura et al., 1995; Portevin et al., 1996). Its activity in plasma correlates with different stages of depression (Maes et al., 1994). The enzyme also has a role in the regulation of blood pressure by participating in the renin–angiotensin system through metabolism of bradykinin and angiotensins I and II (Welches et al., 1993).

Proline, as an imino rather than an amino acid, does not possess the free hydrogen atom that is required to form an S1P1 hydrogen bond with a backbone carbonyl oxygen during catalysis performed by chymotrypsin and subtilisin. Indeed, these serine proteases are unable to hydrolyze peptides at proline residues. On the other hand, prolyl oligopeptidase, a paradigm of the new family of serine peptidases, readily cleaves peptide bonds at the carboxy group of proline residues. It also cleaves at an alanine residue, but the rate of hydrolysis is usually much slower in this case (cf. Polgár, 1994). Several studies have indicated that prolyl oligopeptidase may act directly as an amyloid β protein–generating peptidase in Alzheimer's disease (Ishiura et al., 1990; Fukunari et al., 1994; Shinoda et al., 1997) by processing the C-terminal alanine of the amyloid precursor protein.

Prolyl oligopeptidase is widely distributed and has been isolated and cloned from several sources (Polgár, 1994), including *Flavobacterium meningosepticum* (Yoshimoto et al., 1991; Chevallier et al., 1992; Diefenthal et al., 1993), porcine brain (Rennex et al., 1991), and human lymphocytes (Vanhoof et al., 1994). A method has been developed for isolation of prolyl oligopeptidase from porcine muscle (Polgár, 1991, 1994), and its kinetic behavior was characterized (Polgár, 1991). Several differences have been observed between the extensively studied serine peptidases and prolyl oligopeptidase. The former enzymes exhibit a simple pH-rate profile, controlled by a single ionizing group of pK_a7, and are only slightly, if at all, affected by the ionic strength of the medium (Polgár, 1987, 1989). In contrast, prolyl oligopeptidase catalysis conforms to a doubly sigmoidal curve and is rather sensitive to ionic strength (Polgár, 1991). The doubly sigmoidal pH-rate profile indicates the existence of two pH-dependent enzyme forms, one being predominant at pH 6, and the other at pH 8. Kinetic investigations have also shown that the rate-limiting step for prolyl oligopeptidase catalysis is a conformational change (Polgár, 1991, 1992a), rather than the chemical step characteristic of the reactions of the trypsin family. Unlike other serine peptidases, prolyl oligopeptidase hydrolyzes shorter peptides faster than longer ones, which led to the proposal that the substrate may approach the active site through a tunnel that can exclude large structured peptides (Polgár, 1992b).

No structure from the prolyl oligopeptidase family is available to date. Recently, we have been able to obtain

#To whom correspondence should be addressed.

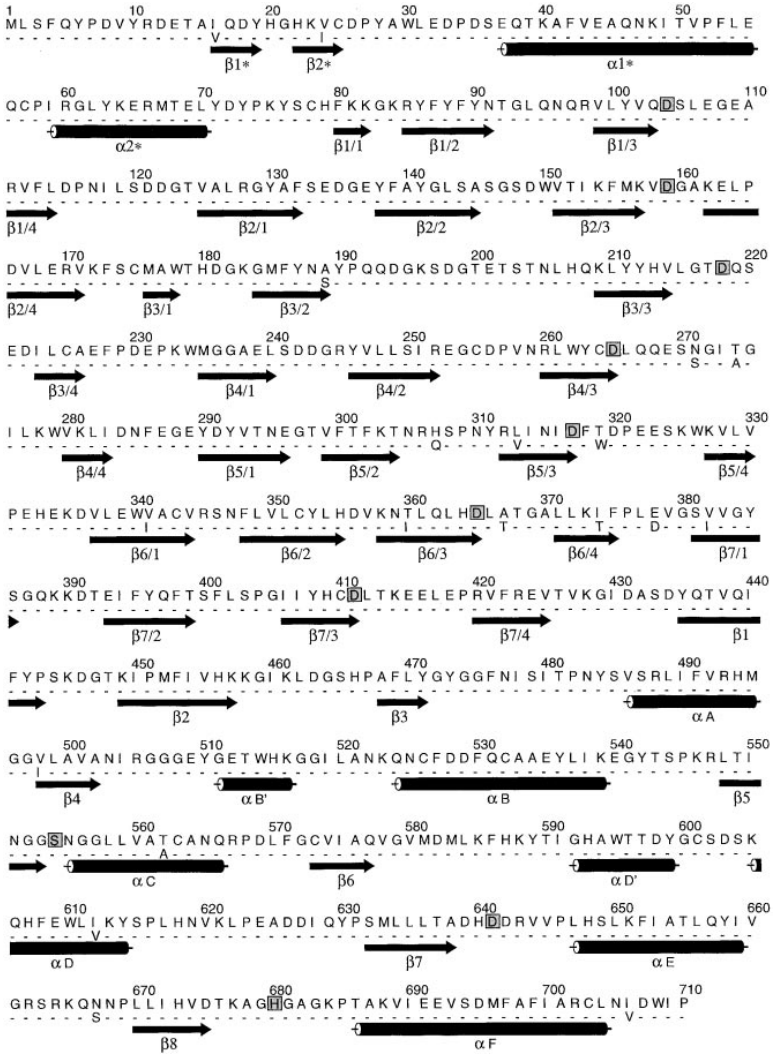

Figure 1. Amino Acid Sequence and Secondary Structure of Prolyl Oligopeptidase

The sequence from porcine brain (Rennex et al., 1991) is shown in the top row. Residues different in the human lymphocyte enzyme (Vanhoof et al., 1994) are shown in the second row. The third strand within the β sheets of the propeller domain always terminates with an aspartate residue (boxed). Members of the catalytic triad (Ser554, His680, and Asp641) are also boxed. The secondary structure assignment is made according to DSSP (Kabsch and Sander, 1983). The figure was produced with ALSCRIPT (Barton, 1993).

well-ordered and strongly diffracting crystals of prolyl oligopeptidase isolated from porcine muscle. This paper describes the structure of the enzyme and reveals the structural basis of how oligopeptides are cleaved and how, at the same time, larger peptides and proteins are protected from proteolysis in the cytosol. We have also determined the structure of prolyl oligopeptidase complexed with its peptide–aldehyde inhibitor, benzyloxycarbonyl-prolyl-prolinal (Z-Pro-prolinal). Prolyl oligopeptidase sequences from different mammalian sources are highly conserved. A comparison (Figure 1) shows that the porcine and human sequences are more than 97% identical. All of the 19 differences are far from the active site and compatible with the present structure; therefore, our Z-Pro-prolinal complex serves as a good model for drug design for treating diseases of the memory system in human.

Results and Discussion

Structure Determination

Crystals of prolyl oligopeptidase belong to the orthorhombic space group $P2_12_12_1$, with unit cell dimensions of a = 70.8 Å, b = 99.7 Å, and c = 111.0 Å (at 100 K). The crystallographic asymmetric unit contains a monomer. The structure of the enzyme was solved by the method of multiple isomorphous replacement, using five heavy atom derivatives (see Table 1). The crystallographic R factor is 19.0% in the resolution range of 30.0–1.4 Å, and the free R value (Brünger, 1992a) is 20.6% (see Experimental Procedures). The model contains all the 710 residues. None of the 16 cysteines are involved in disulfide bridges. The structure of the Z-Pro-prolinal complex was refined in the resolution range of 20.0–2.0 Å and gave an R value of 20.1%, with a free R of 24.3%.

The primary amino acid sequence of the enzyme from porcine muscle is not known. The porcine brain sequence was used for model building and refinement (Rennex et al., 1991). The high-resolution and good quality electron density map did not reveal any difference between the porcine brain and muscle sequences.

Architecture of the Molecule

Prolyl oligopeptidase has a cylindrical shape of an approximate height of 60 Å and diameter of 50 Å. The

Table 1. Data Collection and Phasing Statistics

Compound	Resolution (Å)	Number of Observations	Unique Reflections	Completeness (%)	R_{merge} (%)[a]	R_{iso} (%)[b]	Number of Sites	Phasing Power[c] (Acentric/ Centric)	Cullis R[d] (Acentric/ Centric)
Native (100 K)	30–1.4	366,752	140,350	90.7	5.7				
Z-Pro-prolinal	20–2.0	135,322	51,602	95.7	7.3				
Native (290 K)	20–2.5	70,408	28,064	95.3	7.9				
$UO_2(Ac)_2$	20–2.5	76,601	27,387	93.0	11.2	31.3	1	0.67/0.82	0.94/0.95
$Hg(Ac)_2$	20–2.5	72,464	28,707	97.6	8.2	13.9	1	1.66/1.55	0.72/0.71
EMTS[e]	20–2.5	52,666	24,674	83.7	8.0	22.7	3	1.62/1.81	0.75/0.69
$(Met)_3PbAc$	20–2.5	88,959	28,819	96.1	10.1	24.3	1	0.75/0.74	0.93/0.94
$SmAc_3$	20–3.0	12,300	7253	42.1	11.7	23.5	3	1.81/1.73	0.75/0.73

Derivatives were prepared by soaking native crystals in solutions of heavy atom reagents in the synthetic mother liquor overnight. The concentration was 1 mM for $Hg(Ac)_2$ and 10 mM for all the other derivatives. $Hg(Ac)_2$, EMTS, and $SmAc_3$ shared the same sites. The overall figure of merit was 0.53, and it increased to 0.86 after density modifications in the resolution range of 20–2.8 Å.

[a] $R_{merge} = \Sigma_j\Sigma_h|I_{h,j} - \langle I_h\rangle|/\Sigma_j\Sigma_h\langle I_h\rangle \times 100$, where $I_{h,j}$ is the jth observation of reflection h, and $\langle I_h\rangle$ is the mean intensity of that reflection.

[b] $R_{iso} = \Sigma||F_{PH}|-|F_P||/\Sigma|F_P| \times 100$, where F_P and F_{PH} refer to the native and the derivative structure factor amplitude.

[c] Phasing power = rms ($|F_H|/E$), $|F_H|$ = heavy atom structure factor amplitude, and E = rms lack of closure.

[d] Cullis R = lack of closure/isomorphous difference.

[e] EMTS = ethyl mercury thiosalicylate.

enzyme consists of a peptidase domain with its catalytic triad (Ser554, His680, Asp641) covalently attached to a seven-bladed β propeller (Figure 2). The secondary structure assignment of the molecule is displayed in Figure 1.

Catalytic Domain

The peptidase, or catalytic domain, is built up of residues 1–72 and 428–710. The N terminus contains two short antiparallel β strands and two long helices, which are connected to the much bigger C-terminal region via 23 hydrogen bonds and salt bridges and many hydrophobic interactions. As predicted earlier (Polgár, 1992c), this part of the domain exhibits a characteristic α/β hydrolase fold (Ollis et al., 1992; Smith et al., 1992; Cygler et al., 1993) and contains a central eight-stranded β sheet with all strands except the second one aligned in a parallel manner (Figure 3A). The β sheet is significantly twisted, with the first and last strands arranged at an angle of ∼110° relative to each other. It is flanked by αA and αF on one side, and the six other helices of the domain on the other side. The topology of the β strands

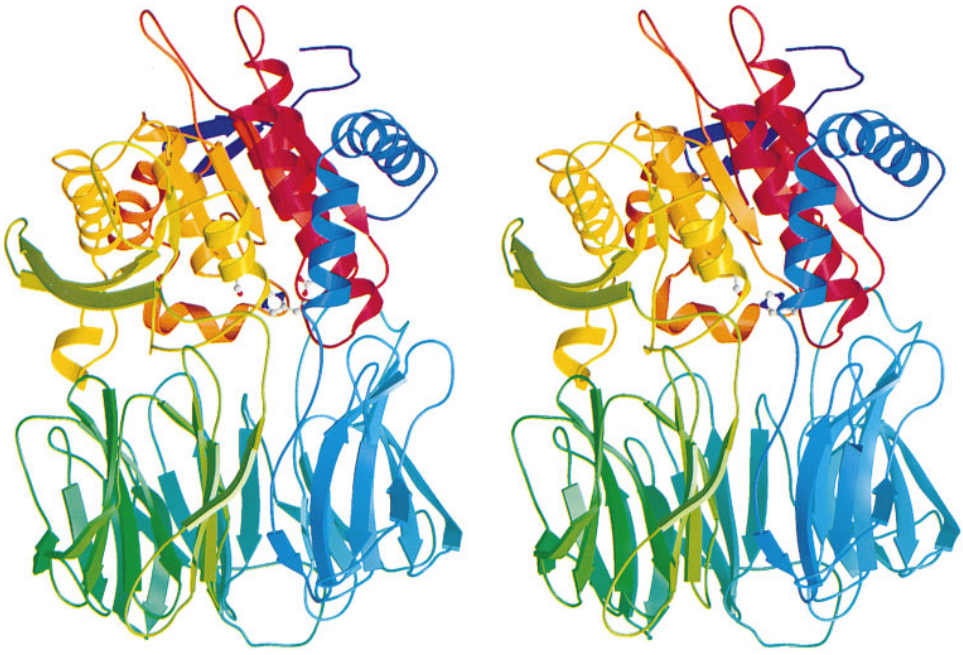

Figure 2. Stereo Representation of the Structure of Prolyl Oligopeptidase

The ribbon diagram is color-ramped blue to red from the N to the C terminus. The catalytic residues are shown in a ball-and-stick representation. The picture was drawn with MolScript (Kraulis, 1991; Esnouf, 1997) and rendered with Raster3D (Merritt and Murphy, 1994).

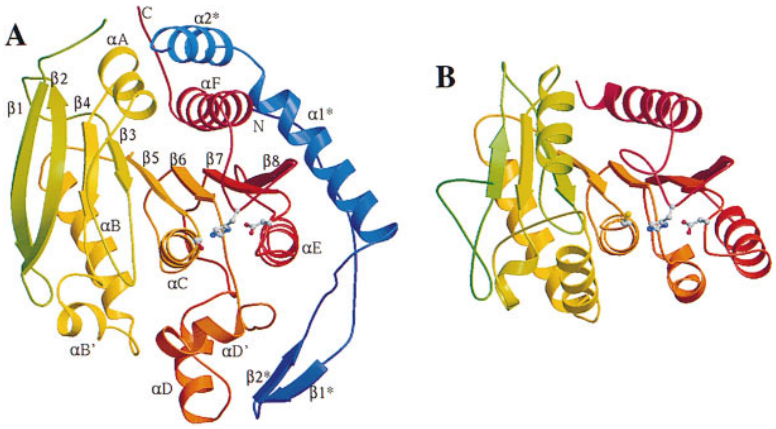

Figure 3. Comparison of the Fold of the Catalytic Domain of Prolyl Oligopeptidase with a Typical α/β Hydrolase Domain Enzyme

(A) The protein chain of the peptidase domain of prolyl oligopeptidase is colored as in Figure 2 and viewed perpendicular to that. The numbering scheme of the secondary structure elements was made according to Ollis et al. (1992). The domain has two extra helices (B′ and D′) and the N-terminal extension (β1*-2*, α1*-2*). The catalytic triad (Ser554, His680, and Asp641) is in a ball-and-stick representation. The N and C termini of the enzyme are labeled with N and C, respectively.

(B) The structure of dienelactone hydrolase (PDB entry 1din). The three-dimensional arrangement and connectivity of the major secondary structure elements, together with the location of its catalytic residues (Cys123, His202, and Asp171) are similar to that of prolyl oligopeptidase. (Drawn with MolScript and rendered with Raster3D.)

is 1, 2, −1x, 2x, and $(1x)_3$ (Richardson, 1981). The secondary structure is close to that predicted by Goossens et al. (1995) and Medrano et al. (1998) except elements of αA, αD′, β1, β2, and β3.

Although the catalytic domain of prolyl oligopeptidase shows low sequence similarity to other α/β hydrolase enzymes, the core of their three-dimensional structures (β strands 2–8 and helices A–F) are closely superimposable. For comparison, Figure 3B displays the structure of the smallest and simplest member of the hydrolase domain enzymes, dienelactone hydrolase (Pathak and Ollis, 1990). The root-mean-square (rms) deviation between the two domains was 1.67 Å for 140 fitted Cα atoms, which gave only 22.1% sequence identity for that region (calculated by STAMP, Russell and Barton, 1992).

The β-Propeller Domain

Residues 73–427 form the noncatalytic domain. This domain belongs to the relatively small but growing number of β propellers. As in the case of G-protein β-subunit structure (Wall et al., 1995; Lambright et al., 1996; Sondek et al., 1996), it is based on a 7-fold repeat of four-stranded antiparallel β sheets (Figure 4). The sheets are twisted and radially arranged around their central tunnel. They pack face-to-face, and the predominantly hydrophobic interaction provides most of the required structural stability. All of the other known propeller proteins have evolved ways to close the circle ("Velcro") between their first and last blades (Neer and Smith, 1996; Baker et al., 1997). In the β subunit of G proteins, for example, six of the seven blades of the propeller domain are built up regularly, with the polypeptide chain progressing outward from the central axis, forming four-stranded antiparallel β sheets. The six blades are joined in succession around the central pseudo-7-fold axis. In blade 7, however, the ring closure is achieved by forming the four antiparallel β strands from both termini of the propeller domain. The N terminus provides the outermost strand, which is connected via main chain hydrogen bonds to the three antiparallel β sheets from the C terminus (Figure 4B). While the six-, seven-, and eight-

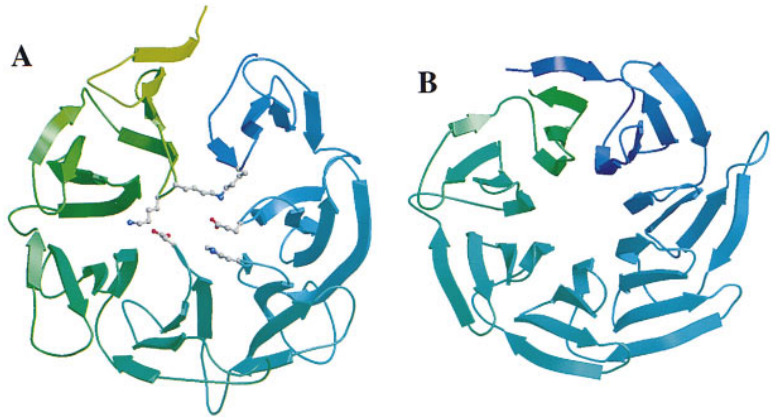

Figure 4. Comparison of the Fold of the Noncatalytic Domain of Prolyl Oligopeptidase with a Typical β-Propeller Structure

(A) The protein chain of the β-propeller domain of prolyl oligopeptidase is colored as in Figure 2 and viewed perpendicular to that, down the pseudo 7-fold axis. The β sheets of the seven blades are joined in succession (β1/1 to β7/4, cf. Figure 1) around the central axis. The "Velcro" is not closed; there are only hydrophobic interactions between the first (blue) and last (green) blades. Residues (Lys82, Glu134, His180, Asp242, Lys389, and Lys390) narrowing the entrance to the tunnel of the propeller are shown in a ball-and-stick representation.

(B) The structure of G-protein β subunit (PDB entry 1tbg). The "Velcro" is closed between the two termini of the polypeptide chain by the main chain hydrogen bonds between the N terminus (blue) and the three antiparallel β strands from the C terminus (green). (Drawn with MolScript and rendered with Raster3D.)

bladed propellers snap the "Velcro" in a similar way (Baker et al., 1997), the smaller four-bladed proteins (hemopexin and collagenase) form a disulfide bond between the first and last blades (Faber et al., 1995; Li et al., 1995). The circular structure is not stabilized in prolyl oligopeptidase by these ways; there are only hydrophobic interactions between the first and last blades (Figure 4A).

β-propeller domains very often show no significant primary sequence similarity, but their three-dimensional structures and individual blades can be very closely superimposed (Neer and Smith, 1996; Baker et al., 1997). The noncatalytic domain of prolyl oligopeptidase looks less regular and less compact than other propellers, e.g., the β subunit of G protein (Figure 4). Consequently, these two domains could not be tightly superimposed; the rms deviation was 2.66 Å for 148 fitted Cα atoms of the propeller (with 17.6% sequence identity for that region). The similarity amongst blades of the propeller is less emphasized in prolyl oligopeptidase—the rms deviation was 1.93 Å for 25 fitted Cα atoms of the seven blades (it is 0.95 Å for the β subunit of G protein). It is worth noting that the third β strand within a blade always terminates with an aspartate (residues 104, 159, 218, 265, 317, 365, and 411; see Figure 1), and they are at the opposite end of the propeller with respect to the catalytic domain. A similar sequence motif was found in the eight-bladed propeller domain of cytochrome cd_1 (Fülöp et al., 1995; Baker et al., 1997). The functional significance of these aspartates is not clear at present. They are predominantly exposed to the solvent, and five of the seven are closely superimposable (the two outliers are Asp159 and 218).

The β propeller is held to the catalytic domain of prolyl oligopeptidase via the two connecting polypeptide main chains (Figure 2) with 27 hydrogen bonds and salt bridges, but mainly with hydrophobic forces.

The Active Site

The active site, containing the catalytic triad (Ser554, Asp641, and His680), is located in a large cavity at the interface of the two domains (Figures 2, 3A, and 5). Ser554 is found at the tip of a very sharp turn between strand 5 and helix C. Such an arrangement, referred to as nucleophile elbow, is characteristic of hydrolases of the α/β type (Ollis et al., 1992). To avoid steric hindrance, several small residues are encountered around the catalytic serine (Gly552, Gly553, Gly558, and Gly559). The main chain conformation of Ser554 is energetically unfavorable, as in other α/β hydrolases (Cygler et al., 1993). Consequently, the serine OH group is well exposed and readily accessible to the catalytic imidazole group on one side and to the substrate on the other. Similar results have most recently been found with proline iminopeptidase (Medrano et al., 1998). His680 is located in the middle of a loop situated between strand 8 and helix F. One of the oxygen atoms of Asp641 is in the plane of the imidazole ring of His680, providing ideal position for hydrogen bond formation, and it is also hydrogen-bonded to a well-ordered water molecule. The other oxygen atom of the carboxylate of Asp641 is coordinated by two main chain NH groups (Arg643 and Val644). The location and geometry of the members of the triad are very similar to that found in other α/β hydrolases providing the same "handedness" (Ollis et al., 1992; see Figure 3).

In the high-resolution structure, a monothioglycerol derivative is covalently bound to the catalytic Ser554 (Figure 5A). Due to the high cysteine content of the enzyme, monothioglycerol was used as a reducing agent for crystallization (see Experimental Procedures). The high-resolution and good quality electron density map permitted the fitting of a 1-thioxy-glycerol molecule to the active site, covalently bound via its secondary carbon atom to Ser554. The precursor of this compound was presumably 1-thioxy-3-hydroxy-acetone, as a trace contaminant in monothioglycerol. At present, there are no high-resolution synchrotron data available from crystals obtained without this compound. Lower-resolution (1.8 Å) in-house data did not reveal any significant changes in the structure, but the Ser554 was not blocked and the geometry of the catalytic triad was identical to that presented here (data are not shown).

Substrate Binding

The substrate binding pocket was mapped by the complex formed between prolyl oligopeptidase and its transition state analog inhibitor, Z-Pro-prolinal. Ser554 nucleophile attacks the aldehyde carbon atom, resulting in a covalent hemiacetal adduct (Figure 5B). The S1 specificity pocket ensures a hydrophobic environment and a snug fit for the proline residue. The pocket is lined by the side chains of Trp595, Phe476, Val644, Val580, and Tyr599 and the side chain carbon atoms of Asn555. The specificity is enhanced by ring stacking between the indole ring of Trp595 and the substrate/inhibitor proline residue (Figure 5B).

The oxyanion-binding site is an essential feature of the serine peptidase catalysis (Polgár, 1987, 1989). The negatively charged oxyanion is generated from the carbonyl oxygen of the scissile bond and stabilized by two hydrogen bonds. In the chymotrypsin-type enzymes, the hydrogen bonds are provided by the main chain NH groups of the catalytic Ser195 and another residue, Gly193. In the subtilisin-type enzymes, the latter residue is substituted by the side chain amide group of an asparagine residue. In common with the α/β hydrolase fold family, one of the groups of the oxyanion-binding site in prolyl oligopeptidase is the main chain NH group adjacent to the catalytic serine (Asn555). This is a direct consequence of handedness (Garavito et al., 1977) of the catalytic triad opposite to that of the classic serine proteases. The second group stabilizing the oxyanion in prolyl oligopeptidase, however, is different from that found in hydrolases, since the hydrogen bond is not provided by another main chain NH group, but the OH group of Tyr473. This hydrogen bonding pattern was found both in the Z-Pro-prolinal complex (Figure 5B) and in the thioxy–glycerol-inhibited enzyme (Figure 5A). Tyrosine residues are better proton donors than main chain NH groups. Consequently, they facilitate the nucleophilic attack of the catalytic serine oxygen on the substrate by the protonation of the ketone oxygen atom, which can explain this unexpected reaction.

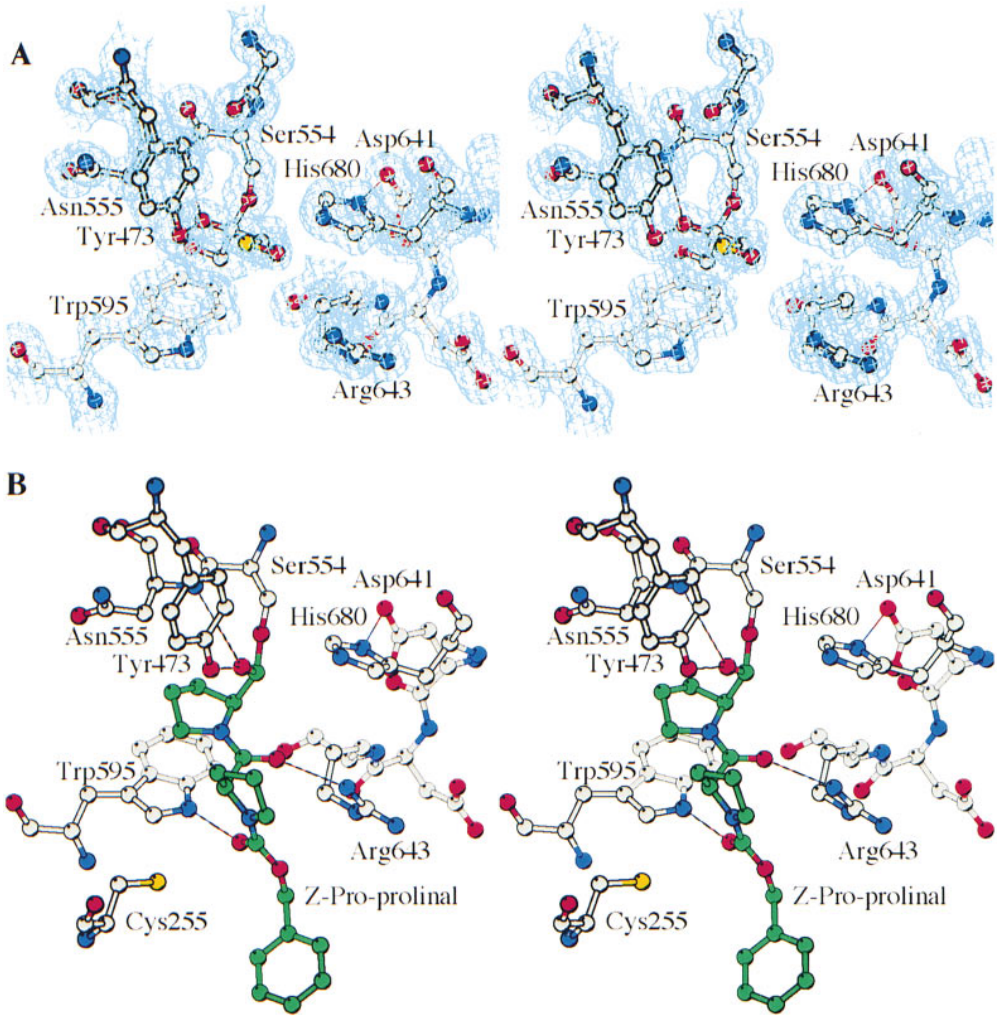

Figure 5. Stereo View of the Active Site of Prolyl Oligopetidase

(A) A 1-thioxy-glycerol molecule is covalently linked to the catalytic Ser554. The SIGMAA (Read, 1986) weighted $2mF_o$-ΔF_c electron density, using phases from the final model, is contoured at 1σ level, where σ represents the rms electron density for the unit cell. Contours more than 1.4 Å from any of the displayed atoms have been removed for clarity. Dashed lines indicate hydrogen bonds.

(B) The carbon atoms of the enzyme and the covalently bound inhibitor, Z-Pro-prolinal, are colored gray and green, respectively. Dashed lines indicate hydrogen bonds. (Drawn with MolScript.)

Most serine proteases have no absolute substrate specificity. They can cleave peptide bonds with a variety of side chains adjacent to the scissile bond. The S1P1 hydrogen bond is a very important factor in the proper substrate orientation and the transition state stabilization. Proline residues do not possess main chain NH groups; therefore, the architecture of the S1-binding site is evolved to be much more specific in prolyl oligopeptidase.

The S2 subsite is less specific for substrate side chains than the S1 pocket. The second proline ring is open toward the cavity containing a host of water molecules. An S2P2 hydrogen bond is formed between the main chain carbonyl group of this proline and the NH1 atom of Arg643 (Figure 5B). The position of the guadinium group is well defined and involved in an extensive hydrogen bonding network.

The S3 site occupied by the benzene ring of the benzyloxycarbonyl group provides a rather nonpolar environment. This pocket is lined by the side chains of several nonpolar residues, including Phe173, Met235, Cys255, Ile591, and Ala594. In agreement with our experimental data (L. P., unpublished data), this subsite prefers hydrophobic residues. The S3P3 hydrogen bond is generated between the carbonyl oxygen of residue P3 and the side chain of Trp595 (Figure 5B).

Figure 5B also shows that Cys255 is situated close to the S1 and S3 subsites, which explains why bulky reagents specific to cysteine residues (N-ethylmaleimide, p-chloromercuribenzoate) inactivate prolyl oligopeptidase almost completely, whereas the small iodoacetamide affords only partial inhibition. Cys255 in the *F. meningosepticum* enzyme is replaced by a threonine residue and is not inhibited by thiol reagents (Yoshimoto et al., 1980). In the case of the homologous acylaminoacyl peptidase, chemical modification of Cys30 and Cys64 inhibited the enzyme, suggesting that one or both of these residues are close to the active site (Scaloni et al., 1994). The corresponding residues in prolyl oligopeptidase are Cys25 and Cys57, which are at the surface

of the protein, far from the active site. Hence, their blocking would not inhibit the enzyme.

Access to the Active Site

The most intriguing problem of oligopeptidase activity concerns the access of substrate to the active site, which is hidden in a large cavity at the interface of the catalytic and β-propeller domains (Figures 2 and 6). There are two contributing factors to be considered: the narrow hole at the bottom of the central tunnel of the propeller, partially covered by flexible side chains (Lys82, Glu134, His180, Asp242, Lys389, and Lys390), and the lack of "Velcro" closing between the first and last blades of the propeller (Figures 2, 4A, and 6). In the resting structure, the hole at the bottom of the propeller is not wide enough (∼4 Å) to permit the diffusion of the substrate into the central cavity. The entrance could be widened by the movements of these flexible side chains, enhanced by opening the propeller between its first and seventh blades. The opening is probably induced by the substrate, and this conformational change may be the rate-limiting step of the catalysis, in agreement with previous results (Polgár, 1992a; Polgár et al., 1993). Such an approach of the substrate is consistent with the tunnel hypothesis put forward earlier to rationalize oligopeptidase activities (Polgár, 1992b). In this manner, the propeller excludes large structured peptides and proteins from the central cavity, thus protecting them from proteolysis in the cytosol.

As in other β-propeller structures (Neer and Smith, 1996; Baker et al., 1997), the central tunnel of the propeller is lined with hydrogen donors and acceptors, and these are water solvated. The cavity is large enough (8562 $Å^3$, ∼8% of the volume of the whole enzyme) to readily accommodate long oligopeptides and is extended from the central tunnel of the β propeller into the catalytic domain (Figure 6). The enzyme is selective for oligopeptides not longer than about the 30 amino acid residues in total (Moriyama et al., 1988). Shorter oligopeptides may not have any secondary structure, or they could easily unfold upon the interaction with the enzyme to be able to diffuse into the active site in the easiest way, as a random coil. The large cavity could, in principle, accommodate oligopeptides longer than 30 residues, but those probably possess some secondary structure elements with stronger interactions. Consequently, they cannot pass the entrance as a random coil.

Conclusions

Prolyl oligopeptidase catalyzes the cleavage of several biologically active peptides, such as angiotensins I and II, bradykinin, oxytocin, and vasopressin (Welches et al., 1993). One of the most unexpected features of the structure of the enzyme is the way in which these oligopeptides are selected and how, at the same time, larger structured peptides and proteins are protected from proteolysis in the cytosol. The C-terminal catalytic domain of prolyl oligopeptidase is covalently linked to an unusual seven-bladed propeller. Unlike other β propellers, the "Velcro" is not closed between the first and last blades of prolyl oligopeptidase. This structural feature permits partial opening of the propeller and provides access via its central tunnel to the active site. Substrates are selected by size exclusion of the propeller and the specificity of the active site.

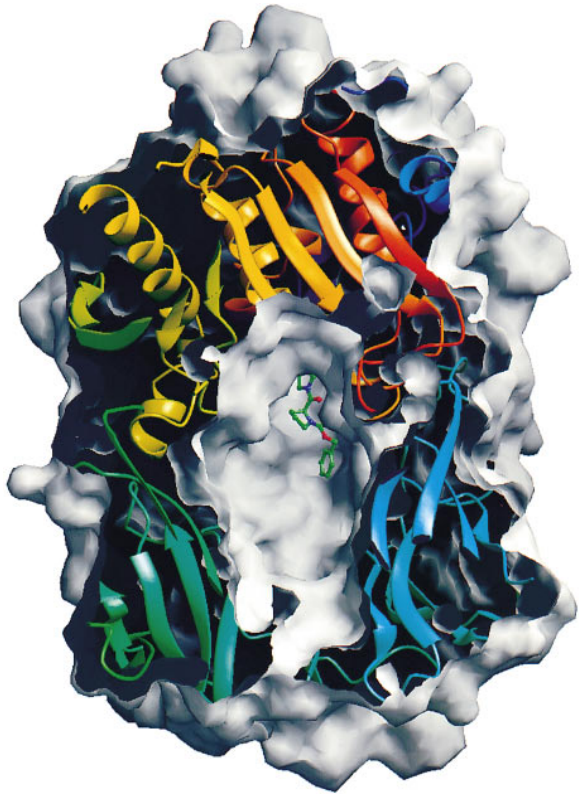

Figure 6. Surface Representation of Prolyl Oligopeptidase

The molecular surface is superimposed on the polypeptide chain. The picture shows a slab of the molecule, hence the cropping of the chain. The large cavity extends from the central tunnel of the β propeller to the catalytic domain and is accessible through the narrow hole at the bottom of the propeller. The covalently bound inhibitor, Z-Pro-prolinal, is shown in a ball-and-stick representation. The molecular surface was calculated by the method published by Connolly (1985), and the figure was prepared using XOBJECTS (M. E. M. Noble, Oxford, unpublished program).

All the other propeller proteins use their central tunnel, or the entrance to that tunnel, to coordinate a ligand or to carry out some catalytic function, which has to be preserved by the structural rigidity of the propeller, provided by the "Velcro" closing (Baker et al., 1997). The functionality of the β-propeller domain in prolyl oligopeptidase requires some structural flexibility. Accordingly, there is no strong interaction between its first and seventh blades. The N-terminal extension, which is structurally part of the catalytic domain, is far from the active site and does not contribute to the characteristic α/β hydrolase fold observed in other enzymes. Its possible role is to provide stability of the circular structure of the propeller by covalently linking it to the peptidase domain.

The catalytic triad (Ser554, His680, and Asp641) is hidden in the central cavity at the interface between the two domains. Specificity is provided by stacking a tryptophan residue against the substrate proline ring. The finding that the S1 and S2 subsites are located in the peptidase domain suggests that the truncated

catalytic domain alone would preserve the enzyme specificity and extend its activity to cleave larger peptides.

Since specific prolyl oligopeptidase inhibitors reverse scopolamine-induced amnesia (Yoshimoto et al., 1987), our Z-Pro-prolinal complex serves as a good model for drug design for treating diseases of the memory system.

Experimental Procedures

Crystallization

Prolyl oligopeptidase was purified from porcine muscle according to the protocols of Polgár (1991, 1994). Crystals were obtained by the hanging-drop vapor diffusion method. The composition of the mother liquor in the reservoir was somewhat different from that reported by Böcskei et al. (1998): 17% (w/v) methoxy-polyethylene glycol (MPEG) 5000, 15% (v/v) glycerol, 1% (v/v) monothioglycerol, 20 mM $Ca(Ac)_2$, and 100 mM TRIS (pH 8.5). The protein solution (10 mg/ml) was mixed with this solution in a 1:1 ratio and left to equilibrate with the reservoir solution at 4°C. Needle-shaped crystal clusters appeared within 4 days. The largest single crystals (0.3 mm × 0.4 mm × 0.5 mm) were obtained at 20°C by seeding preequilibrated drops with microcrystals obtained at 4°C. The Z-Pro-prolinal complex was obtained by cocrystallization without the use of monothioglycerol. Crystals belong to the orthorhombic space group $P2_12_12_1$, with unit cell dimensions of a = 70.8 Å, b = 99.7 Å, and c = 111.0 Å (at 100 K). There is one molecule in the asymmetric unit, with the solvent content of 48% by volume.

X-Ray Data Collection and Structure Determination

Native data to 1.4 Å resolution were collected from frozen crystals at 100 K using 0.87 Å radiation at station 9.6 of SRS, Daresbury. Data from the Z-Pro-prolinal complex were collected at 100 K at the beamline X11, EMBL, Hamburg, using 0.9096 Å radiation. Crystals very often suffered from nonisomorphism upon freezing; therefore, a native data set and all the derivative sets were collected at room temperature, in-house, using Cu Kα radiation. Data were collected either on a 30 cm (synchrotrons) or on an 18 cm (in-house) MAR image plate detector, and they were processed using DENZO and its companion program SCALEPACK (Otwinowski, 1993). Subsequent calculations were performed with the CCP4 program suite (CCP4, 1994). The structure was solved using five heavy atom derivatives (Table 1). Heavy atom sites were located by analysis of isomorphous difference Patterson maps, and their initial parameters were refined by MLPHARE (Otwinowski, 1991). Phases were further improved by SHARP (La Fortelle and Bricogne, 1997) and solvent flipping, using SOLOMON of the CCP4 package (Abrahams and Leslie, 1996). The resulting electron density map allowed the unambiguous tracing and sequence assignment of all but the first five residues.

Model Building and Refinement

Model building and refinement were pursued with alternate cycles of manual refitting using O (Jones et al., 1991) and simulated annealing using X-PLOR (Brünger, 1992b). The resolution was gradually increased from 2.2–1.4 Å, and water molecules were added to the atomic model at the positions of large positive peaks ($>2\sigma$) in the difference electron density at places where the resulting water molecule fell into an appropriate hydrogen bonding environment only. At the final stages, restrained isotropic temperature factor refinements were carried out for each individual atom. Introduction of water molecules in each of the refinement steps resulted in a decrease in both conventional and free R values. A bulk solvent correction allowed all measured data from 30–1.4 Å to be used. The final model contains all 710 residues, the covalently bound 1-thioxy-gycerol and a further noncovalent one, 4 glycerol, and 887 water molecules. Five cysteine residues are oxidized and two others have monothioglycerol covalently attached to them via S–S bridges. The crystallographic R factor is 19.0% in the resolution range of 30.0–1.4 Å (defined as $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}| \times 100$), using 133,425 reflections with $F_{obs} > 2\sigma(F_{obs})$. The free R value (Brünger, 1992a) is 20.6% based on 5631 randomly selected reflections (4% of the total) with $F_{obs} > 2\sigma(F_{obs})$. The final model shows good stereochemistry. Of the residues, 89.4% are in the most favorable regions in the Ramachandran plot with 4 outliers (Tyr311, Ser346, Leu529, and Ser554), all of them in well-defined regions. The structure complexed with Z-Pro-prolinal was refined in the resolution range of 20.0–2.0 Å and gave R factor of 20.1%, using 48,380 reflections, with the free R value of 24.3% for 2,118 reflections. None of the cysteine residues were found to be oxidized. Coordinates have been deposited in the Brookhaven Protein Data Bank.

Acknowledgments

We thank Ms. J. Fejes and Ms. I. Szamosi for technical assistance; the SRS Daresbury for data collection facilities; M. E. M. Noble for help in preparation of Figure 6 and for helpful discussions; M. Groves for his assistance in running SHARP; R. K Bryan and K. Measures for computing; K. Harlos for providing in-house data collection facilities; and S. Lee for figures. We thank the staff of beamlines BW7B and X11 at the DORIS storage ring, and the European Union, for support of the work at EMBL Hamburg through the HCMP Access to Large Installations Project, Contract Number CHGE-CT93-0040. This work was supported by the Royal Society, NATO, and the Wellcome Trust (no. 053147/Z/97/Z/MEP/RC/CRD). V. F. is a Royal Society University Research Fellow.

Received April 27, 1998; revised June 10, 1998.

References

Abrahams, J.P., and Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. *52*, 30–42.

Atack, J.R., Suman-Chauhan, N., Dawson, G., and Kulagowski, J.J. (1991). In vitro and in vivo inhibition of prolyl endopeptidase. Eur. J. Pharmacol. *205*, 157–163.

Baker, S.C., Saunders, N.F.W., Willis, A.C., Ferguson, S.J., Hajdu, J., and Fülöp, V. (1997). Cytochrom cd_1 structure: unusual haem environments in a nitrite reductase and analysis of factors contributing to β-propeller folds. J. Mol. Biol. *269,* 440–455.

Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. Protein Eng. *6*, 37–40.

Böcskei, Zs., Fuxreiter, M., Náray-Szabó, G., Szabó, E., and Polgár, L. (1998). Crystallization and preliminary x-ray analysis of porcine muscle prolyl oligopeptidase. Acta Crystallogr. D., in press.

Brünger, A.T. (1992a). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–474.

Brünger, A.T. (1992b). X-PLOR: Version 3.1; A System for Protein Crystallography and NMR. (New Haven, CT: Yale University Press).

Camargo, A.C.M., Caldo, H., and Reis, M.L. (1979). Susceptibility of a peptide derived from bradykinin to hydrolysis by brain endo-oligopeptidases and pancreatic proteinases. J. Biol. Chem. *254*, 5304–5307.

CCP4 (Collaborative Computational Project, Number 4). (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D. *50*, 760–763.

Chevallier, S., Goeltz, P., Thibault, P., Banville, D., and Gagnon, J. (1992). Characterization of prolyl oligopeptidase from Flavobacterium meningosepticum. J. Biol. Chem. *267*, 8192–8199.

Connolly, M. (1985). Molecular surface triangulation. J. Appl. Crystallogr. *18,* 499–505.

Cygler, M., Schrag, J.D., Sussman, J.L., Harel, M., Silman, I., Gentry, M.K., and Doctor, B.P. (1993). Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases, and related proteins. Protein Sci. *2*, 366–382.

Diefenthal, T., Dargatz, H., Witte, V., Reipen, G., and Svendsen, I. (1993). Cloning of proline-specific endopeptidase gene from Flavobacterium meningosepticum: expression in Escherichia coli and purification of the heterologous protein. Appl. Microbiol. Biotechnol. *40*, 90–97.

Esnouf, R.M. (1997). An extensively modified version of MolScript

that includes greatly enhanced coloring capabilities. J. Mol. Graph. *15*, 133–138.

Faber, H.R., Groom, C.R., Baker, H.M., Morgan, W.T., Smith, A., and Baker, E.N. (1995). 1.8 Å structure of the C-terminal domain of rabbit serum hemopexin. Structure, *3*, 551–559.

Fukunari, A., Kato, A., Sakai, Y., Yoshimoto, T., Ishiura, S., Suzuki, K., and Nakajima, T. (1994). Colocalization of prolyl endopeptidase and amiloid β-peptide in brains of senescence-accelerated mouse. Neurosci. Lett. *176*, 201–204.

Fülöp, V., Moir, J.W.B., Ferguson, S.J., and Hajdu, J. (1995). The anatomy of a bifunctional enzyme: structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome cd_1. Cell *81*, 369–377.

Garavito, R.M., Rossmann, M.G., Argos, P., and Eventoff, W. (1977). Convergence of active center geometries. Biochemistry *16*, 5065–5071.

Goossens, F., De Meester, I., Vanhoof, G., Hendricks, D., Vriend, G., and Scharpé, S. (1995). The purification, characterization and analysis of primary and secondary-structure of prolyl oligopeptidase from human lymphocytes. Evidence that the enzyme belongs to the α/β hydrolase fold family. Eur. J. Biochem. *233*, 432–441.

Ishiura, S., Tsukahara, T., Tabira, T., Shimizu, T., Arahata, K., and Sugita, H. (1990). Identification of a putative amyloid A4-generating enzyme as a prolyl endopeptidase. FEBS Lett. *260*, 131–134.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A. *47*, 110–119.

Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers *22*, 2577–2637.

Kraulis, P.J. (1991). MolScript: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

La Fortelle, E. de, and Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In Meth. Enzymol., Vol. *276*, R.M. Sweet and C.W. Carter, Jr., eds. (New York: Academic Press), pp. 472–494.

Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E., and Sigler, P.B. (1996). The 2.0 Å crystal structure of a heterodimeric G protein. Nature, *379*, 311–319.

Li, J., Brick, P., O'Hare, M.C., Skarzynski, T., Lloyd, L.F., Curry, V.A., Clark, I.M., Bigg, H.F., Hazleman, B.L., Cawston, T.E., et al. (1995). Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed β-propeller. Structure, *3*, 341–349.

Maes, M., Goossens, F., Scharpé, S., Meltzer, H.Y., D'Hondt, P., and Cosyns, P. (1994). Lower serum prolyl endopeptidase enzyme activity in major depression: further evidence that peptidases play a role in the pathophysiology of depression. Biol. Psychiatry *35*, 545–552.

Medrano, F.J., Alonso, J., Garcia, J.L., Romero, A., Bode, W., and Gomis-Ruth, F.X. (1998). Structure of proline iminopeptidase from Xanthomonas campestris pv. citri: a prototype for the prolyl oligopeptidase family. EMBO J. *17*, 1–9.

Mentlein, R. (1988). Proline residues in the maturation and degradation of peptide hormones and neuropeptides. FEBS Lett. *234*, 251–256.

Merritt, E.A., and Murphy, M.E.P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr. D. *50*, 869–873.

Miura, N., Shibata, S., and Watanabe, S. (1995). Increase of the septal vasopressin content by prolyl endopeptidase inhibitors in rats. Neurosci. Lett. *196*, 128–130.

Moriyama, A., Nakanishi, M., and Sasaki, M. (1988). Porcine muscle prolyl endopeptidase and its endogen substrates. J. Biochem. (Tokyo) *104*, 112–117.

Neer, E.J., and Smith, T.F. (1996). G Protein heterodimers: new structures propel new questions. Cell *84*, 175–178.

Ollis, D.I., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., et al. (1992). The α/β hydrolase fold. Protein Eng. *5*, 197–211.

Otwinowski, Z. (1991). Maximum likelihood refinement of heavy atom parameters. In Isomorphous Replacement and Anomalous Scattering, W. Wolf, P.R. Evans, and A.G.W. Leslie, eds. (Warrington, England: Daresbury Laboratory), pp. 80–86.

Otwinowski, Z. (1993). Oscillation data reduction program. In Data Collection and Processing, L. Sawyer, N.W. Isaacs, and S. Bailey, eds. (Warrington, England: Daresbury Laboratory), pp. 55–62.

Pathak D., and Ollis D. (1990). Refined structure of dienelactone hydrolase at 1.8 Å. J. Mol. Biol. *214*, 497–525.

Polgár, L. (1987). Structure and function of serine proteases. New Comp. Biochem. *16*, 159–200.

Polgár, L. (1989). Mechanisms of Protease Action. (Boca Raton, FL: CRC Press), pp. 87–122.

Polgár, L. (1991). pH-Dependent mechanisms in the catalysis of prolyl oligopeptidase from pig muscle. Eur. J. Biochem. *197*, 441–447.

Polgár, L. (1992a). Prolyl endopeptidase catalysis: a physical rather than a chemical step is rate-limiting. Biochem. J. *283*, 647–648.

Polgár, L. (1992b). Unusual secondary specificity of prolyl oligopeptidase and the different reactivities of its two forms towards charged substrates. Biochemistry *31*, 7729–7735.

Polgár, L. (1992c). Structural relationship between lipases and peptidases of the prolyl oligopeptidase family. FEBS Lett. *311*, 281–284.

Polgár, L. (1994). Prolyl oligopeptidases. In Meth. Enzymol., Vol. *244*, A.J. Barrett, ed. (New York: Academic Press), pp. 188–200.

Polgár, L., Kollát, E., and Hollósi, M. (1993). Prolyl oligopeptidase catalysis. Reactions with thiono substrates reveal substrate-induced conformational change to be the rate-limiting step. FEBS Lett. *322*, 227–230.

Portevin, B., Benoist, A., Rémond, G., Hervé, Y., Vincent, M., Lepagnol, J., and De Nanteuil, G. (1996). New prolyl endopeptidase inhibitors: in vitro and in vivo activities of azabicyclo[2.2.2]octane, azabicyclo[2.2.1]heptane, and perhydroindole derivatives. J. Med. Chem. *39*, 2379–2391.

Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr. A. *42*, 140–149.

Rennex, D., Hemmings, B.A., Hofsteenge, J., and Stone, S.R. (1991). cDNA cloning of porcine brain prolyl endopeptidase and identification of the active-site seryl residue. Biochemistry *30*, 2195–2203.

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

Russell, R.B., and Barton, G.J. (1992). Multiple protein-sequence alignment from tertiary structure comparison—assignment of global and residue confidence levels. Proteins *14*, 309–323.

Scaloni, A., Barra, D., Jones, W.M., and Manning, J.M. (1994). Human acylpeptide hydrolase: studies on its thiol groups and mechanism of action. J. Biol. Chem. *269*, 15076–15084.

Shinoda, M., Toide, K., Ohsawa, I., and Kohsaka, S. (1997). Specific inhibitor for prolyl endopeptidase suppresses the generation of amiloid β protein in NG108–15 cells. Biochem. Biophys. Res. Commun. *235*, 641–645.

Smith, L.C., Faustinella, F., and Chan, L. (1992). Lipases: three-dimensional structure and mechanism of action. Curr. Opin. Struct. Biol. *2*, 490–496.

Sondek, J., Bohm, A., Lambright, D.G., Hamm, H.E., and Sigler, P.B. (1996). Crystal structure of a G_A protein βγ dimer at 2.1 Å resolution. Nature *379*, 369–374.

Vanhoof, G., Goossens, F., Hendriks, L., De Meester, I., Hendriks, D., Vriend, G., Van Broeckhoven, C., and Scharpé, S. (1994). Cloning and sequence analysis of the gene encoding human lymphocyte prolyl endopeptidase. Gene (Amst.) *149*, 363–366.

Wall, M.A., Coleman, D.E., Lee, E., Iniquez-Lluhi, J.A., Posner, B.A., Gilman, A.G., and Sprang, S.R. (1995). The structure of the G protein heterotrimer $G_{i\alpha1}\beta_1\gamma_2$. Cell *83*, 1047–1058.

Welches, W.R., Brosnihan, K.B., and Ferrario, C.M. (1993). A comparison of the properties and enzymatic activities of three angiotensin processing enzymes: angiotensin converting enzyme, prolyl endopeptidase and neutral endopeptidase. Life Sci. *52*, 1461–1480.

Wilk, S. (1983). Prolyl endopeptidase. Life Sci. *33*, 2149–2157.

Yoshimoto, T., Walter, R., and Tsuru, D. (1980). Proline-specific endopeptidase from Flavobacterium. Purification and properties. J. Biol. Chem. *255*, 4786–4792.

Yoshimoto, T., Kado, K., Matsubara, F., Koriyama, N., Kaneto, H., and Tsuru, D. (1987). Specific inhibitors for prolyl endopeptidase and their anti-amnesic effect. J. Pharmacobio-Dyn. *10*, 730–735.

Yoshimoto, T., Kanatani, A., Shimoda, T., Inaoka, T., Kobubo, T., and Tsuru, D. (1991). Prolyl endopeptidase from Flavobacterium meningosepticum: cloning and sequencing of the enzyme gene. J. Biochem. (Tokyo) *110*, 873–878.