research papers

Structure of aminopeptidase N from *Escherichia coli* complexed with the transition-state analogue aminophosphinic inhibitor PL250

Marie-Claude Fournié-Zaluski,[a] Hervé Poras,[a] Bernard P. Roques,[a]* Yoshitaka Nakajima,[b] Kiyoshi Ito[b] and Tadashi Yoshimoto[b]

[a]Pharmaleads, France, and [b]Pharmaceutical Sciences, Graduate School of Biomedical Sciences, Nagasaki University, Japan

Correspondence e-mail:
bernard.roques@pharmaleads.com

Aminopeptidase N (APN; EC 3.4.11.2) purified from *Escherichia coli* has been crystallized with the optically pure aminophosphinic inhibitor PL250, H_3N^+-CH(CH_3)-P(O)(OH)-CH_2-CH(CH_2Ph)-CONH-CH(CH_2Ph)CO_2^-$, which mimics the transition state of the hydrolysis reaction. PL250 inhibits APN with a K_i of 1.5–2.2 nM and its three-dimensional structure in complex with *E. coli* APN showed its interaction with the S_1, S_1' and S_2' subsites of the catalytic site. In this structure, the Zn ion was shown to be pentacoordinated by His297, His301 and Glu320 of APN and the two O atoms of the phosphinic moiety of PL250. One of these O atoms is also involved in a hydrogen bond to Tyr381, supporting the proposed role of this amino acid in the stabilization of the transition state of the enzymatic process. The strength of the phosphinic zinc binding and the occupancy of the S_2' subsite account for the 100-fold increase in affinity of PL250 compared with the dipeptide-derived inhibitor bestatin ($K_i = 4.1 \times 10^{-6} M$). Accordingly, the removal of the C-terminal phenylalanine of PL250 resulted in a large decrease in affinity ($K_i = 2.17 \times 10^{-7} M$). Furthermore, it was observed that the C-terminal carboxyl group of the inhibitor makes no direct interactions with the amino acids of the APN active site. Interestingly, PL250 exhibits the same inhibitory potency for *E. coli* APN and for mammalian enzymes, suggesting that the structure of the complex could be used as a template for the rational design of various human APN inhibitors needed to study the role of this aminopeptidase in various pathologies.

1. Introduction

Aminopeptidase N (APN; EC 3.4.11.2) is a well conserved exopeptidase (Maroux *et al.*, 1973) that is expressed in a broad spectrum of species (bacteria, plants, insects and mammals) in which it plays an important role in the final digestion of peptides. This type II membrane-bound zinc peptidase is also widely distributed in mammalian tissues, including the central nervous system (Hersh *et al.*, 1987; Noble *et al.*, 2001; Larrinaga *et al.*, 2005), and is particularly abundant in the kidney, the intestine and the lung (Maroux *et al.*, 1973; Hersh *et al.*, 1987). *In vivo*, this enzyme is involved in the catabolism of biologically active peptides such as angiotensin III (Zini *et al.*, 1996), nociceptin (Montiel *et al.*, 1997) and enkephalins in association with neprilysin (Roques *et al.*, 1993) in the brain and in peripheral organs. Furthermore, APN is identical to the human lymphocyte surface cluster differentiation antigen CD13 (Look *et al.*, 1989) and behaves as a receptor for coronaviruses TGEV and 229E in pigs and humans (Delmas *et al.*, 1992; Yeager *et al.*, 1992). Moreover, APN has recently

been shown to play an important role in the invasion of metastatic tumours (Saiki *et al.*, 1993; Hashida *et al.*, 2002) and in angiogenesis (Sato, 2004).

Owing to the importance of APN in many physiological and pathological processes, there is a challenging interest in the development of potent and selective inhibitors of this enzyme for potential therapeutic applications. Various families of inhibitors have been described (Fournié-Zaluski & Roques, 2002; Bauvois & Dauzonne, 2006), most of which have inhibitory potencies in the micromolar range and low selectivity. Consequently, to aid in the design of new potent and specific inhibitors of this enzyme, it seems important to determine the structural parameters which allow optimized recognition of the aminopeptidase N active site.

APN belongs to the M1 family of zinc metallopeptidases (clan MA; Rawlings & Barrett, 1995), which also contains leukotriene A4 hydrolase (LTA4H), aminopeptidase A (APA) and aminopeptidase B (APB) and is characterized by two consensus sequences: (i) the active-site sequence HEX-XH(X_{18})E, where the two histidine residues and the second glutamate are zinc ligands, with the first glutamate being involved in the catalytic process, and (ii) the sequence GXMEN containing the glutamate residue which ensures the aminopeptidase selectivity of this class of enzymes (Luciani *et al.*, 1998; Vazeux *et al.*, 1998).

Only three members of this family have been crystallized and their three-dimensional structures established by X-ray crystallography: human LTA4H (Thunnissen *et al.*, 2001), *Escherichia coli* APN (Ito *et al.*, 2006; Addlagatta *et al.*, 2006; Nocek *et al.*, 2008) and the archaeal tricorn interacting factor F3 (Kyrieleis *et al.*, 2005). In spite of low primary sequence identities, these three peptidases show great similarities in their three-dimensional structure, particularly at the active-site level, and in their mechanism of action. Thus, bestatin (Fig. 1), a nonspecific aminopeptidase inhibitor, has been cocrystallized with both LTA4H (Thunnissen *et al.*, 2001) and with *E. coli* APN (Ito *et al.*, 2006; Addlagatta *et al.*, 2006) revealing very similar structural organizations of the complexes.

A large homology in primary sequence is observed between aminopeptidases N from various species (80% identity between different mammalian enzymes and 44.1% similarity between human and *E. coli* APN), suggesting that the structure of the bacterial enzyme could be used as a model of its human counterpart.

In this paper, we describe the three-dimensional structure of PL250, the most potent aminopeptidase N inhibitor reported to date (Chen *et al.*, 1999), in the active site of *E. coli* APN. This phosphinic derivative, (2S)-2-((2S)-3-{[(1R)-1-ammonioethyl](hydroxy)phosphoryl}-2-benzylpropanamido)-
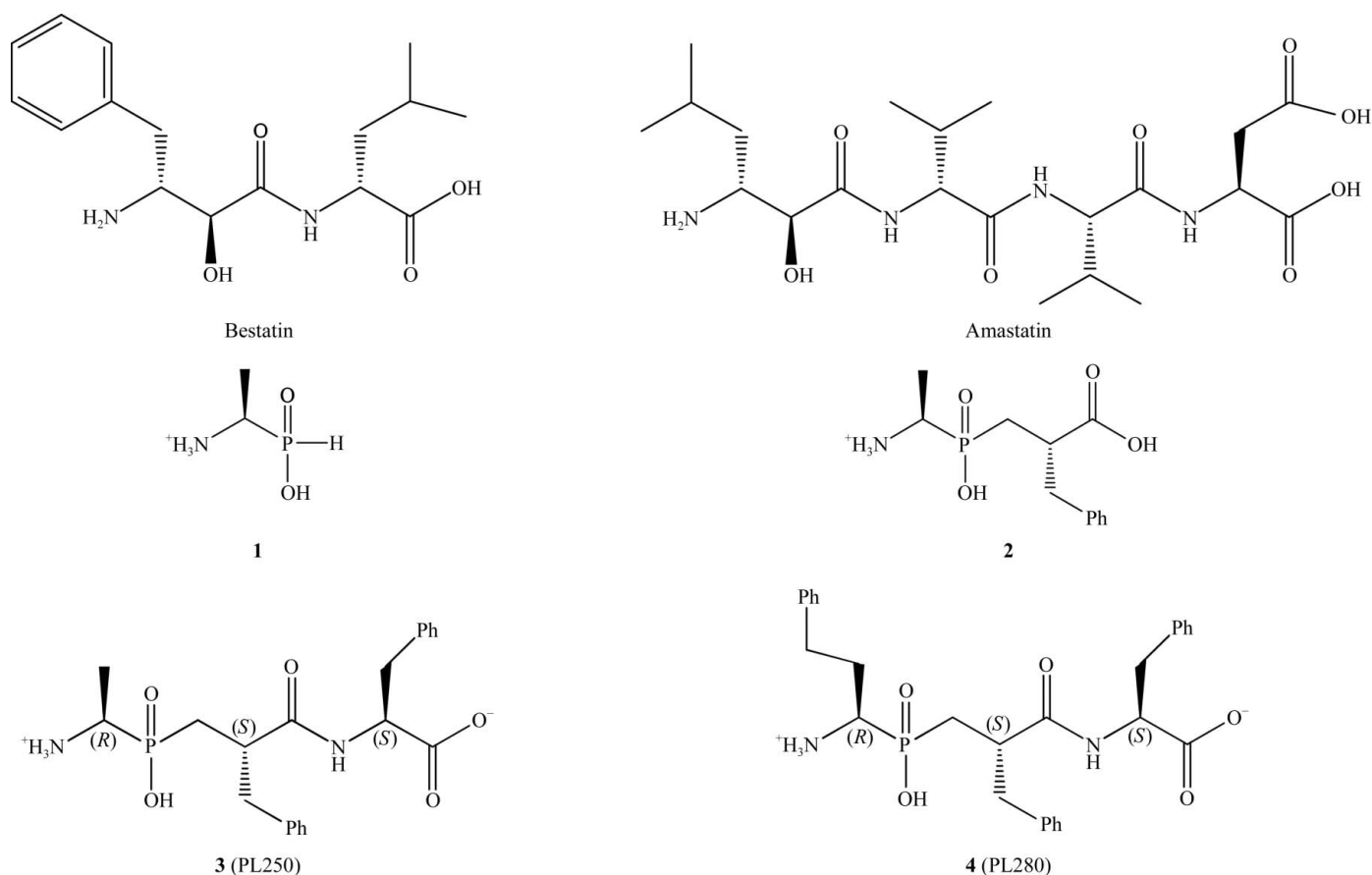

Figure 1
Chemical structures of the generic inhibitory compounds bestatin and amastatin and the α-aminophosphinic inhibitors (see Table 2) compound **1**, compound **2**, PL250 (compound **3**) and PL280 (compound **4**).

3-phenylpropanoate [H$_3$N$^+$-CH(CH$_3$)-P(O)(OH)-CH$_2$-CH(CH$_2$Ph)-CONH-CH(CH$_2$Ph)CO$_2$H] (Fig. 1), which mimics a tripeptide and could be considered to be a 'transition-state analogue', inhibits hog kidney APN with a potency in the nanomolar range and has been used to visualize this enzyme in rat and chick by autoradiography using its radiolabelled form (Jardinaud *et al.*, 2004; Sihn *et al.*, 2006).

Comparison of the three-dimensional structures of the free APN of *E. coli* and of its complexes with bestatin (Ito *et al.*, 2006; Addlagatta *et al.*, 2006) and PL250 (this study) allows the confirmation of the proposed role of Tyr381 in the stabilization of the transition state during the hydrolytic process (Ito *et al.*, 2006) and demonstrates that the highly potent inhibition of APN by PL250 arises from the additional C-terminal amino-acid interactions with the S$_2'$ subsite.

The identical affinities of this phosphinic inhibitor for aminopeptidase N from *E. coli* and from hog kidney suggests that the same organization of the active site occurs in the bacterial and mammalian enzymes and supports the use of these structural data to rationally design various classes of potent human APN inhibitors.

2. Experimental procedures

2.1. Materials

Aminopeptidase from *E. coli* was produced and purified as described previously (Ito *et al.*, 2006). Aminopeptidase N from hog kidney (4.07 mg ml^{-1}; specific activity 5.41 U mg^{-1}) was purchased from Pierce Biotechnology and Ala-β-NA was purchased from Sigma. The four phosphinic inhibitors were synthesized as described by Chen *et al.* (1999).

2.2. Crystallization and data collection

Crystallization was carried out as described previously (Ito *et al.*, 2006; Onohara *et al.*, 2006). Crystals of the complex with PL250 inhibitor were obtained by soaking ligand-free crystals for 3 d in a prepared solution containing 300 µ*M* PL250, 2.2 *M* ammonium sulfate and 100 m*M* MES buffer pH 6.4. These crystals were isomorphous with ligand-free crystals, belonging to space group *P*3$_1$21 with unit-cell parameters *a* = *b* = 120.7, *c* = 171.0 Å, γ = 120°. Diffraction data were collected to 1.55 Å resolution at 100 K using a wavelength of 1.000 Å at the Photon Factory BL-17A station synchrotron-radiation source (Tsukuba, Japan). For data collection under cryogenic conditions, the crystals were soaked for 1 min in a prepared solution containing 30%(*v*/*v*) glycerol, 1.75 *M* ammonium sulfate and 100 m*M* MES buffer pH 6.4. The crystal was mounted in a nylon loop and flash-cooled in a liquid-nitrogen stream at 100 K. These data were processed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The data-collection statistics are summarized in Table 1.

2.3. Structure determination and refinement

Scaling and reduction of the diffraction data was performed using the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). The structure of ligand-free APN

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the last resolution shell.

Data collection	
Space group	*P*3$_1$21
Unit-cell parameters	
$a = b$ (Å)	120.7
c (Å)	171.0
γ (°)	120
Beamline	Photon Factory BL-17A
Wavelength	1.000
Resolution range (Å)	50–1.55 (1.65–1.55)
No. of unique reflections	207284 (20585)
Completeness (%)	99.6 (100)
Redundancy	6.7 (7.9)
R_{merge}†	0.089 (0.288)
Mean $I/\sigma(I)$	35.7 (8.8)
Wilson B factor (Å^2)	15.9
Refinement	
Resolution range (Å)	20–1.55 (1.59–1.55)
R factor	0.178 (0.201)
R_{free}	0.195 (0.213)
R.m.s. deviations	
Bond lengths (Å)	0.009
Bond angles (°)	1.20
Average B factor (Å^2)	
Main chain	15.1
Side chain	16.8
Waters	28.9
Inhibitor	15.6

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity for multiple measurements.

(PDB code 2dq6) was used as the initial model. The structure of the APN–PL250 complex was refined by rigid-body refinement, simulated annealing and energy minimization with the program *CNS* (Brünger *et al.*, 1998) using the data from 20 to 1.55 Å resolution. The structure was scrutinized by inspecting the composite OMIT map using the program *XtalView* (McRee, 1999). The difference Fourier map clearly displayed residual electron density corresponding to the inhibitor bound at the active site. Refinement and model rebuilding were alternately carried out for several cycles using the programs *REFMAC*5 from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) and *Coot* (Emsley & Cowtan, 2004); water molecules were then picked out from the difference map on the basis of peak height and distance criteria. Water molecules whose thermal factors were above 50 Å^2 after refinement were removed from the list. Further model-building and refinement cycles resulted in an R factor of 17.8% and an R_{free} of 19.5% using 196 426 reflections from 20 to 1.55 Å resolution. The maximal thermal factor of the water molecules was 49 Å^2.

2.4. Enzyme-activity assay

The enzyme activity was measured using Ala-β-NA as a substrate for both *E. coli* and hog kidney aminopeptidases N. Hog kidney aminopeptidase (0.2 mU ml^{-1}, 40 µg ml^{-1}) in 50 m*M* Tris–HCl buffer pH 7.5 was preincubated with or without various concentrations of the tested inhibitors in a final volume of 100 µl in 96-well microplates for 10 min at 310 K. The substrate Ala-β-NA (K_m = 50 µ*M* for the hog

kidney enzyme) was added to a final concentration of 50 µM and the emitted fluorescence was measured for 30 min (λ_{ex} = 340 nm, λ_{em} = 405 nm) on a Berthold Twinkle LB 970 fluorimeter. Blank values were obtained by measuring the fluorescence emitted by the same reagents in the presence of heat-denatured enzyme. Assuming competitive inhibition of APN by the four inhibitors, their K_i values were calculated using the Cheng–Prusoff equation $K_i = IC_{50}/(1 + [S]/K_m)$.

For *E. coli* aminopeptidase N, the assay was performed as described by Ito *et al.* (2006). Briefly, the reaction mixture consisted of 0.8 ml 20 mM Tris–HCl buffer pH 7.0, 0.1 ml enzyme solution, 0.1 ml of a 3 mM solution of Ala-β-NA (K_m = 300 µM for the *E. coli* enzyme) and various concentrations of inhibitor. After 10 min incubation at 310 K, the reaction was stopped by the addition of 0.5 ml Fast Garnet GBC solution (1 mg ml^{-1}) containing 10% Triton X-100 in 1 M sodium acetate buffer pH 4.0. The absorbance at 550 nm was measured after 20 min.

3. Results and discussion

3.1. Aminopeptidase inhibition

Four inhibitors (Fig. 1) were tested with the aim of potentially linking their respective potencies to their enzyme active-site occupancy. The present results show that compound **3** binds the S_1, S_1' and S_2' subsites (following the nomenclature of Schechter & Berger, 1967) and compound **4** is very likely to bind *E. coli* APN similarly, whereas we could assume that compound **2** recognizes the S_1 and S_1' subsites and that compound **1** only recognizes the S_1 subsite.

As shown in Table 2, a very significant increase in APN recognition was observed from compound **1** (K_i = 8.4 µM) to **2** (K_i = 0.2 µM) to **3** (K_i = 2.2 nM) and **4** (K_i = 2.3 nM).

Table 2
Inhibitory potency of the phosphinic inhibitors (Fig. 1) towards hog kidney APN and *E. coli* APN.

Each experiment was established in triplicate.

	K_i (M)	
	Hog kidney APN	*E. coli* APN
1	$8.4 \pm 0.8 \times 10^{-6}$	
2	$2.17 \pm 0.06 \times 10^{-7}$	
3 (PL250)	$2.2 \pm 0.5 \times 10^{-9}$	$1.5 \pm 0.2 \times 10^{-9}$
4 (PL280)	$2.3 \pm 0.5 \times 10^{-9}$	$25 \pm 1 \times 10^{-9}$

Moreover, the inhibitory potency of compound **3** (PL250) is not significantly different when measured using the bacterial *E. coli* APN (K_i = 1.5 nM) or the mammalian hog APN (K_i = 2.2 nM). Conversely, for compound **4** the K_i value is ten times better for hog APN (K_i = 2.3 nM) than for *E. coli* APN (K_i = 25 nM).

3.2. Structure of the APN–PL250 complex

The refined model of the APN–PL250 complex comprised 866 residues, a zinc ion, the PL250 inhibitor, a sulfate anion, seven glycerol molecules and 1039 water molecules with an R factor of 17.8% at 1.55 Å resolution. The present structure shows Ala122 in the *cis*-peptide conformation as in the ligand-free form and in the APN–bestatin complex (PDB codes 2dq6 and 2dqm, respectively). This model lacks the four N-terminal residues owing to a lack of interpretable electron density. From the Fourier map, alternative conformations of side chains were found for 28 residues. Most of these are located on the protein surface and two residues, Glu382 and Ser784, are located on the surface of the cavity that is formed inside the protein. The average thermal factors of the main-chain atoms, side-chain atoms, inhibitor atoms, sulfate anion, glycerol molecules and water molecules were 15.1, 16.8, 15.6, 45.3, 25.6 and 28.9 Å^2, respectively. Analysis of the stereochemistry with *PROCHECK* (Laskowski *et al.*, 1993) showed that all the main-chain atoms fell within the allowed region of the Ramachandran plot, with 740 residues (93.7%) in the most favoured region and 50 residues (6.3%) in the additionally allowed region. This model was superimposed onto the structures of the ligand-free form and the APN–bestatin complex (PDB code 2dqm) using the least-squares method at C^α positions with r.m.s. deviations of 0.12 and 0.11 Å, respectively. The overall structure of the APN–PL250 complex was almost identical to those of the ligand-free form and the the APN–bestatin complex.

In wild-type *E. coli* APN, the consensus Zn-binding sequence is ^{297}HEYFH^{301}(X)$_{18}$E^{320}, in which His297, His301 and Glu320, together with a water molecule, bind the zinc ion with a tetrahedral geometry and Glu298 functions as a catalytic base through activation of the water molecule (Wat 2) that
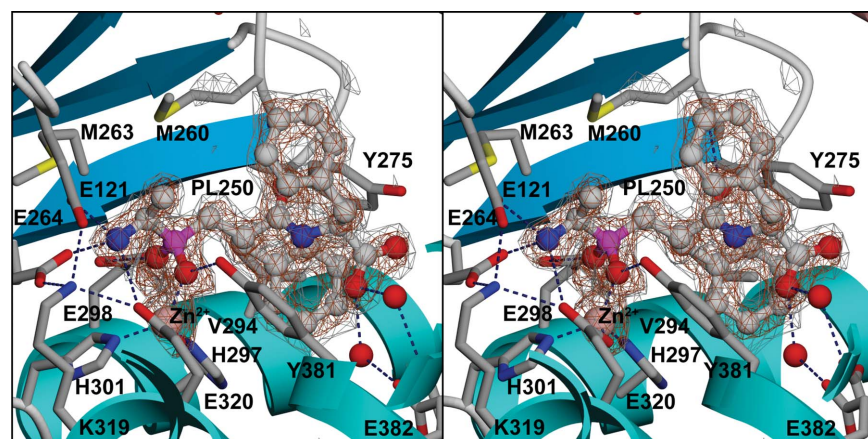

Figure 2
Stereoview of PL250 binding to the active site of *E. coli* APN. The inhibitor, PL250, is shown as a ball-and-stick model with an $F_o - F_c$ OMIT map. The grey and red meshes are contoured at the 3.0σ and 5.0σ levels, respectively. The difference Fourier map was calculated using the program *CNS* (Brünger *et al.*, 1998) without consideration of the structure factors of PL250. The zinc ion is represented as a pink sphere and residues at the active site are shown as stick models. This diagram was produced with the program *POVScript+* (Fenn *et al.*, 2003) and rendered with the program *POV-Ray* (http://www.povray.org/).

attacks the carbon of the amide bond to be cleaved (Ito *et al.*, 2006; Addlagatta *et al.*, 2006).

The PL250 inhibitor was tightly bound to the active site of *E. coli* APN (Figs. 2 and 3). In the complex between *E. coli* APN and PL250, the water molecule acting as a zinc ligand (Wat 1) is replaced by the two O atoms of the phosphinic moiety, as expected for a transition-state analogue, leading to pentacoordination of the zinc ion [$Zn\cdots O1(P) = 2.01$ Å and $Zn\cdots O2(P) = 2.94$ Å]. The geometry of the zinc complex corresponds to a distorted trigonal bipyramid. The O atoms of the phosphinic moiety are also involved in hydrogen bonds to Glu298 ($O2\cdots Glu298 = 2.61$ Å), Glu320 ($O1\cdots Glu320 = 2.96$ Å) and Tyr381 ($O1\cdots HO$ Tyr381 = 2.69 Å).

The active site of uncomplexed *E. coli* APN is also characterized by the presence of a salt-bridge network associating three glutamate residues (Glu121, Glu264 and Glu320) and the ε-NH_3^+ of Lys319 (Ito *et al.*, 2006; Addlagatta *et al.*, 2006). In the complex with PL250, the free amino group of the inhibitor participates in this network by hydrogen bonding to two of these glutamates, *i.e.* Glu264 (2.69 Å) and Glu121 (2.78 Å) (Fig. 3).

The inhibitor backbone is in an extended conformation antiparallel to the β-strand of the ^{261}GAMEN^{265} motif (Luciani *et al.*, 1998; Vazeux *et al.*, 1998) leading to various types of interactions: the carbonyl of the inhibitor amide bond forms hydrogen bonds to the NH group of Gly261 and the CO group of Ala262, which is in close vicinity of the CH_2 group of the phosphinic moiety. This methylene corresponds to the NH of the amide group of bestatin, which forms a hydrogen bond to the carbonyl of Ala262 in its complex with *E. coli* APN (Ito *et al.*, 2006; Addlagatta *et al.*, 2006).

The S_1 subsite of *E. coli* APN was shown to be lined by Glu121, Tyr376, Met263 and Met260, the side chains of which

were found to limit the size of this subsite in the ligand-free form (Ito *et al.*, 2006; Addlagatta *et al.*, 2006). In the complex with PL250, the methyl side chain of the inhibitor binds within this subsite, leading to a small displacement of the lateral chain of Met260.

The $S_1'S_2'$ domain represents an open hydrophobic cleft in which the two benzylic rings of PL250 are located. The S_1' subsite was surrounded by Tyr275, Val294, His297, Val324, Asp327, Tyr381 and the side-chain $C^\gamma H_2$-$C^\beta H_2$ group of Arg293 (Fig. 4). The benzyl side chain of the phosphinic moiety is found in this essentially hydrophobic pocket with the centre of the aromatic ring at about 6 Å from these residues. This structural arrangement is almost identical to that found for the leucine side chain in the APN–bestatin complex. The S_2' subsite, in which the C-terminal benzyl side chain of PL250 is bound, corresponds to a large domain formed by the Met260, Tyr376, Thr377 and Arg825 side chains (Fig. 4). These residues form a hydrophobic pocket with their constituent atoms at about 6 Å from the centre of the C-terminal phenyl ring of the inhibitor. The two benzyl groups in S_1' and S_2' are in an antiparallel orientation (with a distance of 5.8 Å between the rings) with the C-terminal benzyl moiety pointing towards the S_1 methyl group (distance 4.2 Å).

A bridge is formed between the carboxylate of PL250 and the γ-carboxylate of Glu382 *via* two water molecules. Consistent with this relatively weak interaction, the carboxyl group of the inhibitor can be replaced by an uncharged group without loss of affinity (not shown).

The minimal compound structure for the recognition of a zinc aminopeptidase consists of amino-acid derivatives bearing a zinc ligand such as β-aminothiol, α-aminohydroxamate, α-aminophosphonate *etc.* (Fournié-Zaluski & Roques, 2002; Bauvois & Dauzonne, 2006). These compounds, which recognize only the S_1 subsite of the catalytic site, have inhibitory potencies from 10^{-5} to 10^{-8} M, depending on the zinc ligand moiety.

The commonly used but nonspecific inhibitor bestatin (Fig. 1), a dipeptide analogue, inhibits aminopeptidase N with a K_i value of 4×10^{-6} M (Rich *et al.*, 1984), suggesting that all the stabilizing interactions between the inhibitor and the enzyme are not completely satisfied. In this chemical series, the best inhibitor of APN is amastatin (Fig. 1), a tetrapeptide analogue of bestatin with a K_i value of 2×10^{-8} M (Rich *et al.*, 1984). This suggests that, in this latter case, some additional interactions are likely to occur within the S_2' subsite as in PL250 (see below).

Using a homogenous series of α-aminophosphinic inhibitors we have shown (Table 1) an improvement in the inhibitory potency by three orders of magnitude from the amino acid **1** to the pseudotripeptide **3** (PL250), with a K_i value of about 2 nM. In
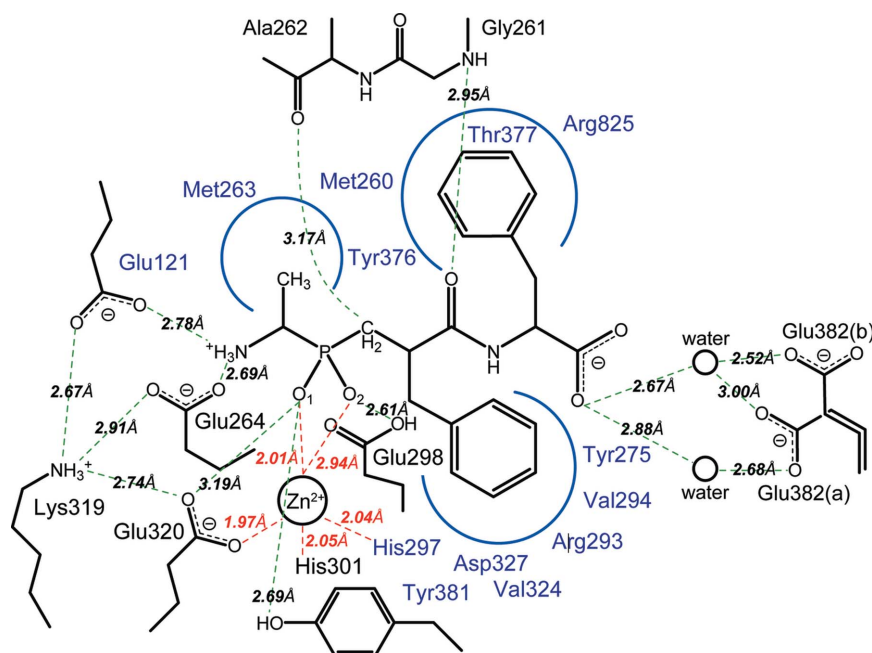

Figure 3
Schematic representation of the interactions occurring in the APN–PL250 complex at the active-site level.
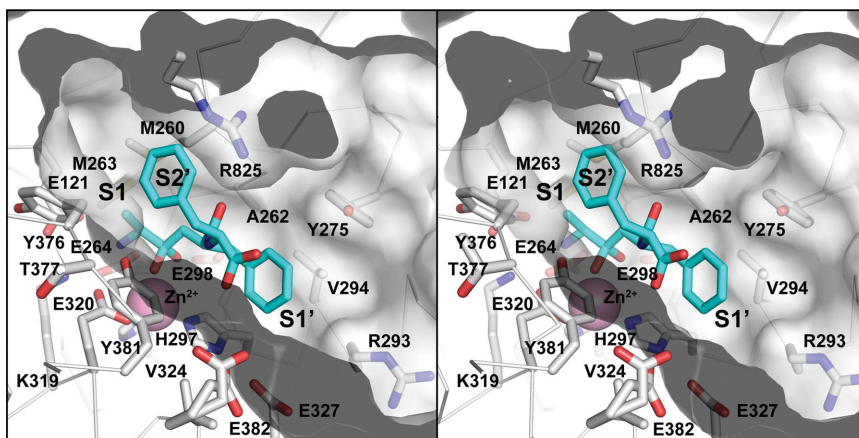
Figure 4
Stereoview of the surface diagram at the active site in the APN–PL250 complex. The S_1, S_1' and S_2' subsites of *E. coli* APN are shown. The zinc ion and main chain are shown as a pink sphere and a C^α trace, respectively. The residues constituting these subsites are represented using stick models. The PL250 stick model binding to the active site is in cyan. This diagram was drawn and rendered with the program *PyMOL* (http://www.pymol.org/).
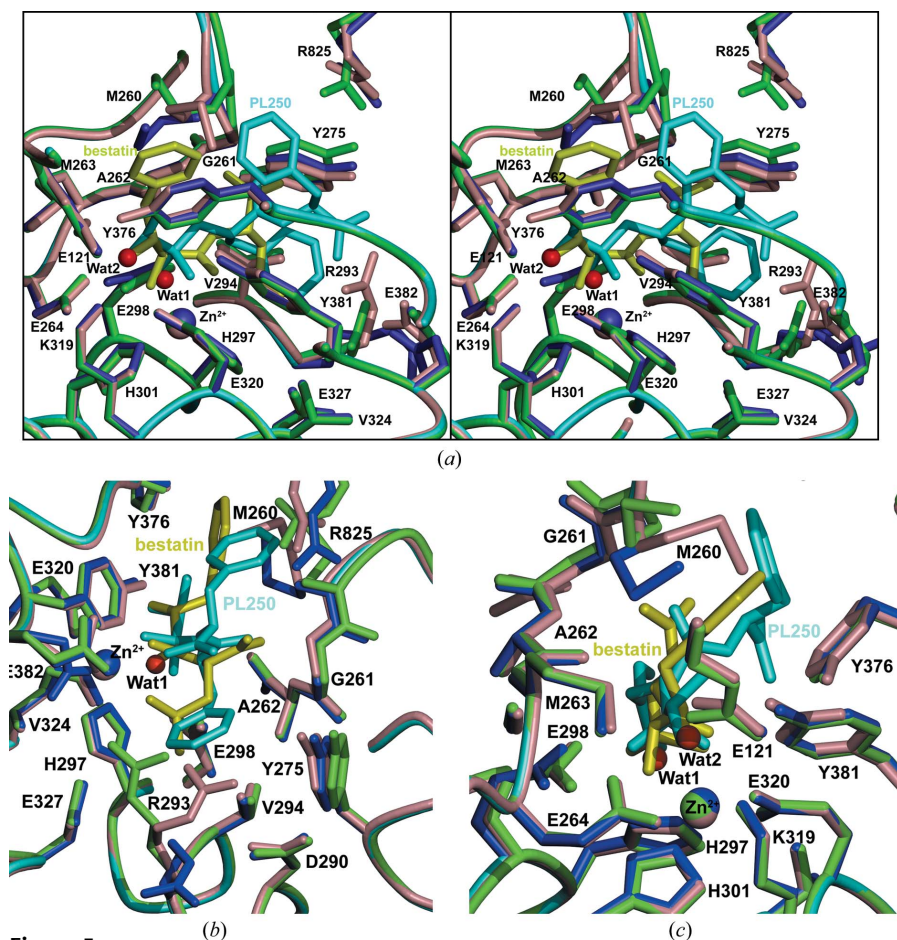

Figure 5
Comparison of the active-site structures of ligand-free APN, the APN–bestatin complex and the APN–PL250 complex. (*a*) Stereoview of a superimposition of the three active-site structures. The structures of APN complexed with bestatin and PL250 were superimposed onto that of ligand-free APN using the least-squares method. Ribbon and stick models of ligand-free APN, the APN–bestatin complex and the APN–PL250 complex are represented in pink, green and blue, respectively. Two water molecules, bestatin and PL250 are represented in red, yellow and cyan, respectively. (*b*) The catalytic centre and the S_1 subsite of APN in its free and complexed states, showing the different conformations of Glu298 and Met260 in the three structures. (*c*) This diagram is focused on the S_1' and S_2' subsites. These diagrams were produced with the program *POVScript+* (Fenn *et al.*, 2003) and rendered with the program *POV-Ray* (http://www.povray.org/).

this series of inhibitors, a further extension of the peptide chain does not improve the inhibitory potency (not shown), indicating that PL250 optimally fits the APN active site.

Comparison of the K_i values of the aminophosphinic compounds **2** ($K_i = 2 \times 10^{-7}$ M) and **3** ($K_i = 2 \times 10^{-9}$ M) (Chen *et al.*, 1999) reflects the important contribution of hydrophobic interactions inside the S_2' subsite to APN inhibition. Likewise, comparison of the K_i values of bestatin ($K_i = 4 \times 10^{-6}$ M) and compound **2** ($K_i = 2 \times 10^{-7}$ M), which are expected to recognize the same APN subsites, shows a significant difference (a factor of 20) between their affinity for the enzyme that could be attributed at least partly to the better zinc-chelating properties of a phosphinic group compared with the hydroxyl amide ligand of bestatin.

3.3. Structural comparison of *E. coli* APN–bestatin and *E. coli* APN–PL250 complexes

Large similarities were observed between the two complexes regarding their common domain of interaction. Thus, in the catalytic site both inhibitors act as bidentate ligands of the zinc ion (using two O atoms from either the phosphinic moiety in PL250 or the hydroxyl amide bond in bestatin), leading to a final stable pentacoordination of the metallic ion. The relative distances between the Zn atom and the O atoms are similar, but the strength of the zinc coordination is likely to be higher for the phosphinic group.

In the APN–PL250 complex, the free amino group of PL250 is hydrogen bonded to the γ-carboxylates of Glu121 and Glu264 located in the negatively charged pocket, which provides the exopeptidase specificity of APN. As shown in Figs. 5(*a*) and 5(*b*), the free amino group of PL250 was located at the same position as the amino group of bestatin in the APN–bestatin complex and the water molecule (Wat2) in the ligand-free APN structure (Ito *et al.*, 2006; Addlagatta *et al.*, 2006). In the APN–PL250 complex, the two O atoms of the phosphinic group take similar places to the carbonyl and hydroxyl

groups of bestatin in the *E. coli* APN complex. All these groups replace the water molecule Wat1 that is present in ligand-free APN (Ito *et al.*, 2006; Addlagatta *et al.*, 2006).

The conformations of the amino-acid side chains around the zinc ion were almost identical in the ligand-free APN, APN–bestatin and APN–PL250 complexes, with the exception of Glu298 (Fig. 5*c*). Indeed, in the APN–PL250 complex, Glu298, which functions as the catalytic base, was shifted by 0.66 Å from its corresponding position in the APN–bestatin complex. This shift of Glu298 is related to the phosphinic group, in which the oxygen (P)O2 was found at 1.04 Å from the position of Wat1 in the free enzyme and at 0.73 Å from the hydroxyl group of bestatin in its complex with APN. In the Michaelis complex, the NH amide group in the P_1' position of the substrate was expected to form hydrogen bonds to the carbonyl group of Ala262 and the carboxyl group of Glu298 (Ito *et al.*, 2006). It was also expected that these hydrogen bonds would be conserved in the tetrahedral transition state mimicked by the phosphinic inhibitor. Consistent with this hypothesis, the CH_2 group of the phosphinic moiety, which corresponds to the NH group in the P_1' position, was in close contact (3.17 Å) with the carbonyl group of Ala262 and was separated from Glu298 by 4.06 Å.

The catalytic process of zinc-dependent metalloproteases requires the involvement of residues in the stabilization of the transition state, although the organization can vary depending on the protease family. For example, in the M4 family the transition state is stabilized by a tyrosine residue and a histidine residue (Tyr157 and His231 for thermolysin TLN; Matthews, 1988; Beaumont *et al.*, 1995), whereas in the neprilysin family (M13 family) only one histidine is involved (His711; Dion *et al.*, 1993; Marie-Claire *et al.*, 1998; Schultz *et al.*, 2009) in the same process. Furthermore, the involvement of a tyrosine in the transition-state stabilization is not restricted to peptidases of the MA clan, since Tyr149 was shown to play this role in astacin (family M12 clan MB; Grams *et al.*, 1996). In the APN–bestatin complex, the formation of a hydrogen bond between the hydroxyl group of Tyr381 and the O atom of the amide group (2.64 Å) of the inhibitor was proposed to mimic the stabilization of the tetrahedral transition state that occurs during the enzymatic process (Ito *et al.*, 2006; Addlagatta *et al.*, 2006; Figs. 5*b* and 5*c*). This was confirmed in this study with PL250, which corresponds to a true transition-state analogue. In this case, one of the P—(O) O atoms forms a hydrogen bond to Tyr381 OH (2.66 Å). Interestingly, this tyrosine is well conserved in the M1 peptidase family, suggesting that this amino acid plays the same stabilizing role in all these enzymes.

The isobutyl side chain of bestatin at the P_1' position was reported to be surrounded by Tyr275, Val294, His297, Val324, Tyr381 and the lateral face of the guanidinium group of Arg293 (Addlagatta *et al.*, 2006). The benzyl group of PL250 at the P_1' position is bound to the same open hydrophobic cleft as the isobutyl group in the APN–bestatin complex (Figs. 3 and 4). This phenyl ring is oriented almost perpendicularly to Tyr275, with a distance of 5.59 Å between the centres of the aromatic rings. One of the carboxyl O atoms of Asp327,

located in the S_1' subsite at a distance of 4.00 Å from the benzyl group of PL250, is involved in a hydrogen bond (2.78 Å) to His297, a ligand of the zinc ion. These various interactions are hypothesized to stabilize the inhibitor inside the active site.

Although APN seems to exhibit a preference for hydrophobic moieties in the P_1' position (Kenny *et al.*, 1987), some peptides with hydrophilic or even charged residues can be found in this position. In this case, stabilization of the complex could be ensured by interactions with APN hydrophilic amino acids located in the S_1' subsite, such as Asp327 in *E. coli* APN. The importance of this residue is confirmed by its isosteric replacement by Glu413 in hog APN and Glu418 in human APN. This suggests that the acidic residue conserved in this subsite may be able to interact with a P_1' side chain containing positively charged groups through hydrogen bonding. In accordance with this assumption, peptides with a lysine or an arginine residue in the P_1' position, such as thymopentin, somatostatin and neurokinin A, were shown to be cleaved by APN (Noren *et al.*, 1997).

Moreover, another important parameter which could explain the binding of non-aromatic side chains is the high flexibility of the Arg293 side chain, which is nicely demonstrated by comparison of its positions in the ligand-free APN, the APN–bestatin complex and the APN–PL250 complex (Fig. 5*b*).

Another interesting point is the flexibility of the Met260 side chain located at the bottom of the S_1 subsite. In the bestatin–APN complex, the displacement of this side chain has been explained by the size of the S_1 benzyl moiety of the inhibitor (Ito *et al.*, 2006; Addlagatta *et al.*, 2006). In the PL250–APN complex, despite the smaller size of the P_1 methyl group of PL250, the Met260 side chain is also displaced from its position in free APN, which is likely to be a consequence of the proximity of the phenylalanine in the P_2' position of PL250 (Fig. 5*c*).

Finally, the carboxylate group of bestatin shows a two-point interaction with the enzyme: one O atom is hydrogen bonded to Gly261 NH and Ala262 NH, while the other O atom interacts with the guanidinium groups of Arg783 and Arg825 through two water molecules (Addlagatta *et al.*, 2006). In the complex with PL250, the hydrogen bond between Gly261 NH and the CO of the amide bond is also observed, but owing to the presence of a P_2' moiety the interaction with the two arginine residues is lacking. In this conformation, Arg825 actually lines the S_2' subsite along with Thr377, Tyr376 and Met260 and all these residues are close to the centre of the aromatic ring of the C-terminal Phe (Fig. 3). One O atom of the carboxylate group of PL250 interacts through two water molecules with the γ-carboxylate of Glu382, which adopts two different positions (*a* and *b*) in the crystal (Fig. 3).

3.4. Structural homologies between *E. coli* APN and mammalian APN

Based on sequence alignment, the main differences between the active sites of human and *E. coli* APN seem to

be associated with the anionic site and the S_1 pocket, as previously outlined (Ito *et al.*, 2006).

In the anionic site of *E. coli* APN, a network involving three glutamates (Glu121, Glu264 and Glu320, which is also a zinc ligand) and the lysine residue Lys319 forms a negatively charged cavity which seems to be relatively well conserved in bacterial and insect aminopeptidases. In mammalian APN, however, this network is not as conserved; in human APN the analogue of Lys319 is Asn410 and Glu121 is replaced by Gln213. Moreover, Gln136, Glu271, Glu318 and Asn317 in human LTA4H, which belongs to the same M1 family, replace the Glu121, Glu264, Glu320 and Lys319 residues of *E. coli* APN, respectively. These residues are identical in human LTA4H and in human APN. In the LTA4H–bestatin complex (PDB code 1hs6; Thunnissen *et al.*, 2001), a water molecule is located at the corresponding position to an ε-amino group of Lys319 in *E. coli* APN and the hydrogen-bonding network is formed through this water molecule. The sequence homology between human LTA4H and APN suggest that a similar network may occur in human APN.

The S_1 subsite of *E. coli* APN is limited by the side chain of Met260, leading to a subsite cavity that is well adapted to small residues such as Ala. However, the structure of the APN–bestatin complex showed that the side chain of Met260 is able to move in order to facilitate the interaction of larger residues (such as the benzyl side chain of bestatin) inside this pocket (Ito *et al.*, 2006; Addlagatta *et al.*, 2006). In mammalian APN, Met260 is replaced by an alanine residue (Ala351 in human enzyme), leading to an extended cavity which could accept large side chains, such as a phenethyl group, without a requirement for side-chain transconformation. This accounts for the identical inhibitory potency of various aminophosphinic derivatives, such as compounds **3** and **4** (Table 2), towards hog APN, whatever the size of the S_1-interacting side chain (Chen *et al.*, 1999). Interestingly, this is different for *E. coli* APN inhibition, where the K_i values of compounds **3** and **4** are 1.5 and 25 n*M*, respectively. This tenfold decrease in affinity could result from an energetically consuming transconformational process involving the side chain of Met260.

Consistent with the high homology in the active site of all these aminopeptidases N, we have found that the K_i values of PL250 for *E. coli* APN ($K_i = 1.5 \times 10^{-9}$ *M*) and hog kidney APN ($K_i = 2.2 \times 10^{-9}$ *M*) were almost identical. As the sequence identity between the mammalian APNs are around 90%, the results obtained with the hog enzyme could very likely be extended to human APN.

One of the peculiarities of APN is the complete occlusion of the active site by the C-terminal domain of the enzyme (domain IV; Ito *et al.*, 2006; Addlagatta *et al.*, 2006). Two mechanisms have been proposed to account for the binding step of substrates into the active site. Large transconformational changes could occur, either constitutively or induced by the substrate, allowing the introduction of the substrate into the catalytic site (Ito *et al.*, 2006; Addlagatta *et al.*, 2006). The peptide could also enter through the small cavity (10 Å) present in the C-terminal α-domain of the enzyme (domain IV).

A superimposition of the three-dimensional structure of domain IV in the complexes of *E. coli* APN with bestatin and PL250 does not show any differences at the level of the hole; the increased size of PL250 compared with bestatin is not sufficient to discriminate between these two proposed mechanisms.

In conclusion, the present study demonstrates the importance of S_2' subsite binding for the optimization of APN inhibitors. This subsite is large, essentially hydrophobic and can accept bulky side chains. However, the optimal size of these chains remains to be determined, since the steric hindrance arising from the domain IV hole could behave as a restrictive factor in the absence of conformational modifications.

References

Addlagatta, A., Gay, L. & Matthews, B. W. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 13339–13344.

Bauvois, B. & Dauzonne, D. (2006). *Med. Res. Rev.* **26**, 88–130.

Beaumont, A., O'Donohue, M. J., Paredes, N., Rousselet, N., Assicot, M., Bohuon, C., Fournié-Zaluski, M. C. & Roques, B. P. (1995). *J. Biol. Chem.* **270**, 16803–16808.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Chen, H., Roques, B. P. & Fournie-Zaluski, M. C. (1999). *Bioorg. Med. Chem. Lett.* **9**, 1511–1516.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Delmas, B., Gelfi, J., L'Haridon, R., Vogel, L. K., Sjöström, H., Noren, O. & Laude, H. (1992). *Nature (London)*, **357**, 417–420.

Dion, N., Le Moual, H., Crine, P. & Boileau, G. (1993). *FEBS Lett.* **318**, 301–314.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Fenn, T. D., Ringe, D. & Petsko, G. A. (2003). *J. Appl. Cryst.* **36**, 944–947.

Fournié-Zaluski, M. C. & Roques, B. P. (2002). *Ectopeptidases*, edited by J. Langner & S. Ansorge, pp. 51–94. New York: Kluwer Academic Publishers/Plenum Press.

Grams, F., Dive, V., Yiotakis, A., Yiallouros, I., Vassiliou, S., Zwilling, R., Bode, W. & Stöcker, W. (1996). *Nature Struct. Biol.* **3**, 671–675.

Hashida, H., Takabayashi, A., Kanai, M., Adashi, M., Kondo, K., Kohno, N., Yamaoka, Y. & Miyake, M. (2002). *Gastroenterology*, **122**, 376–386.

Hersh, L. B., Aboukhair, N. & Watson, S. (1987). *Peptides*, **8**, 523–532.

Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., Ito, T. & Yoshimoto, T. (2006). *J. Biol. Chem.* **281**, 33664–33676.

Jardinaud, F., Banisadr, G., Noble, F., Melik-Parsadaniantz, S., Chen, H., Dugave, C., Laplace, H., Rostene, W., Fournié-Zaluski, M. C., Roques, B. P. & Popovici, T. (2004). *Biochimie*, **86**, 105–113.

Kenny, A. J., Stephenson, S. L. & Turner, A. J. (1987). *Mammalian Ectoenzymes*, edited by A. J. Kenny & A. J. Turner, pp. 169–210. Oxford: Elsevier.

Kyrieleis, O. J. P., Goettig, P., Kiefersauer, R., Huber, R. & Brandstetter, H. (2005). *J. Mol. Biol.* **349**, 787–800.

Larrinaga, G., Callado, L. F., Agirregoitia, N., Varona, A. & Gil, J. (2005). *Neurosci. Lett.* **383**, 136–140.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Look, A. T., Ashmun, R. A., Shapiro, L. H. & Peiper, S. C. (1989). *J. Clin. Invest.* **83**, 1299–1307.

Luciani, N., Marie-Claire, C., Ruffet, E., Beaumont, A., Roques, B. P. & Fournié-Zaluski, M. C. (1998). *Biochemistry*, **37**, 686–692.

Marie-Claire, C., Ruffet, E., Tiraboschi, G. & Fournié-Zaluski, M. C. (1998). *FEBS Lett.* **438**, 215–219.

Maroux, S., Louvard, D. & Baratti, J. (1973). *Biochim. Biophys. Acta*, **321**, 282–295.

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

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.

Montiel, J. L., Cornille, F., Roques, B. P. & Noble, F. (1997). *J. Neurochem.* **68**, 354–361.

Noble, F., Banisadr, G., Jardinaud, F., Popovici, T., Lai-Kuen, R., Chen, H., Bischoff, L., Melik Parsadaniantz, S., Fournié-Zaluski, M. C. & Roques, B. P. (2001). *Neuroscience*, **105**, 479–488.

Nocek, B., Mulligan, R., Bargassa, M., Collart, F. & Joachimiak, A. (2008). *Proteins*, **70**, 273–279.

Noren, O., Sjöström, H. & Olsen, J. (1997). *Cell Surface Peptidases in Health and Disease*, edited by A. J. Kenny & C. M. Boustead, pp. 175–191. Oxford: BIOS Scientific Publishers.

Onohara, Y., Nakajima, Y., Ito, K., Xu, Y., Nakashima, K., Ito, T. & Yoshimoto, T. (2006). *Acta Cryst.* F**62**, 699–701.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rawlings, N. D. & Barrett, A. J. (1995). *Methods Enzymol.* **248**, 183–228.

Rich, D. H., Moon, B. J. & Harbeson, S. (1984). *J. Med. Chem.* **27**, 417–422.

Roques, B. P., Noble, F., Daugé, V., Fournié-Zaluski, M. C. & Beaumont, A. (1993). *Pharmacol. Rev.* **45**, 87–146.

Saiki, I., Fujii, H., Yoneda, J., Abe, F., Nakajima, M., Tsuruo, T. & Azuma, I. (1993). *Int. J. Cancer*, **54**, 137–143.

Sato, Y. (2004). *Biol. Pharm. Bull.* **27**, 772–776.

Schechter, I. & Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157–162.

Schultz, H., Dale, G. E., Karimi-Nejad, Y. & Oefner, C. (2009). *J. Mol. Biol.* **385**, 178–187.

Sihn, G., Savary, K., Michaud, A., Fournié-Zaluski, M. C., Roques, B. P., Corvol, P. & Gasc, J. M. (2006). *Differentiation*, **74**, 119–128.

Thunnissen, M. M. G. M., Nordlund, P. & Haeggström, J. Z. (2001). *Nature Struct. Biol.* **8**, 131–135.

Vazeux, G., Iturrioz, X., Corvol, P. & Llorens-Cortes, C. (1998). *Biochem. J.* **334**, 407–413.

Yeager, C. L., Ashmun, R. A., Williams, R. K., Cardellichio, C. B., Shapiro, L. H., Look, A. T. & Holmes, K. V. (1992). *Nature (London)*, **357**, 420–422.

Zini, S., Fournié-Zaluski, M. C., Chauvel, E., Roques, B. P., Corvol, P. & Llorens-Cortes, C. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 11968–11973.