Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural study of the complex between human pepsin and a phosphorus-containing peptidic transition-state analog

Masao Fujinaga,[a] Maia M. Cherney,[a] Nadya I. Tarasova,[b] Paul A. Bartlett,[c] John E. Hanson[c] and Michael N. G. James[a]*

[a]Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada, [b]Molecular Aspects of Drug Design Section, ABL-Basic Research Program, National Cancer Institute, FCDRC, PO Box B, Frederick, Maryland 21702, USA, and [c]Department of Chemistry, University of California, Berkeley, California 94720, USA

Correspondence e-mail: michael.james@ualberta.ca

Received 2 September 1999
Accepted 14 December 1999

PDB Reference: pepsin–inhibitor complex, 1qrp.

The refined crystal structure of the complex between human pepsin and a synthetic phosphonate inhibitor, Iva-Val-Val-Leu^{P}-(O)Phe-Ala-Ala-OMe [Iva = isovaleryl, Leu^{P} = the phosphinic acid analog of L-leucine, (O)Phe = L-3-phenyllactic acid, OMe = methyl ester], is presented. The structure was refined using diffraction data between 30 and 1.96 Å resolution to a final R factor ($\sum \big||F_o| - |F_c|\big| / \sum |F_o|$, where $|F_o|$ and $|F_c|$ are the observed and calculated structure-factor amplitudes, respectively) of 20.0%. The interactions of the inhibitor with the enzyme show the locations of the binding sites on the enzyme from S4 to S3′. Modeling of the inhibitor binding to porcine pepsin shows very similar binding sites, except at S4. Comparison of the complex structure with the structures of related inhibitors bound to penicillopepsin helps to rationalize the observed differences in the binding constants. The convergence of reaction mechanisms and geometries in different families of proteinases is also discussed.

1. Introduction

Human pepsin is an aspartic proteinase produced by the gastric mucosa and is involved in the initial digestion of food. The related enzyme from the pig has been extensively studied. Human pepsin is produced as seven different zymogen isoforms. Structural studies of pepsin A (pepsin I; Fujinaga *et al.*, 1995), pepsinogen I (Bateman *et al.*, 1998), progastricsin (pepsinogen C, pepsinogen II; Moore *et al.*, 1995) and gastricsin intermediate (Khan *et al.*, 1997) have been reported.

Phosphorus-containing compounds have been designed as transition-state mimics to inhibit a variety of different proteinases. The target enzymes include metalloproteinases, serine proteinases and aspartyl proteinases. In particular, interactions of three different phosphorus-containing inhibitors bound to penicillopepsin, a fungal aspartyl proteinase, have been described (Fraser *et al.*, 1992). Here, we study the binding of a phosphonate inhibitor, Iva-Val-Val-Leu^{P}-(O)Phe-Ala-Ala-OMe [Iva = isovaline, Leu^{P} = the phosphinic acid analogue of leucine, (O)Phe = L-3-phenyllactic acid, OMe = methyl ester], bound to human pepsin A. Kinetic studies of this inhibitor with porcine pepsin have been performed (Bartlett *et al.*, 1990). However, attempts to crystallize the inhibitor with porcine pepsin were unsuccessful. The close similarity between the human and porcine pepsins allows a modeling of the inhibitor binding to the latter enzyme. The long length of the inhibitor allows an extensive characterization of the binding sites on the enzyme.

2. Materials and methods

2.1. Source of materials

Human pepsin was obtained from stomach mucosa as previously described (Fujinaga *et al.*, 1995; Tarasova *et al.*, 1995). The synthesis of the phosphonate inhibitor has also been reported previously (Bartlett *et al.*, 1990).

2.2. Crystallization

A solution of the pepsin–inhibitor complex was produced using 10 µl pepsin solution in water (20 mg ml^{-1}) and 2 µl phosphonate-inhibitor solution (5 m*M*) in ethanol. Drops of the well solution [8 µl, 40% saturated $(NH_4)_2SO_4$ in 100 m*M* sodium acetate pH 5.0, 5% 2-methyl-2,4-pentanediol] were mixed with 2 µl pepsin–inhibitor complex solution and seeded with native pepsin crystals produced as described previously (Fujinaga *et al.*, 1995). Seeds grew to dimensions of 1 × 0.4 × 0.2 mm and were isomorphous with native crystals. The space group is $P2_12_12_1$, with unit-cell parameters $a = 71.90$, $b = 151.27$, $c = 40.79$ Å.

2.3. Data collection

Diffraction data were collected on a San Diego multiwire detector system (Xuong *et al.*, 1985) with a Rigaku RU-200 rotating-anode X-ray generator operating at 40 kV, 150 mA. The incident Cu $K\alpha$ X-ray beam was passed through a graphite monochromator and an 0.7 mm collimator. A total of 188 324 reflections to 1.96 Å were collected. Of these, 31 990 were unique reflections, with an $R_{\rm merge}$ $[\sum(I_i - \langle I\rangle)/\sum I_i]$ of 10.8%. The data were 98% complete for all reflections and were 89% complete for reflections with $F/\sigma(F) > 3$.

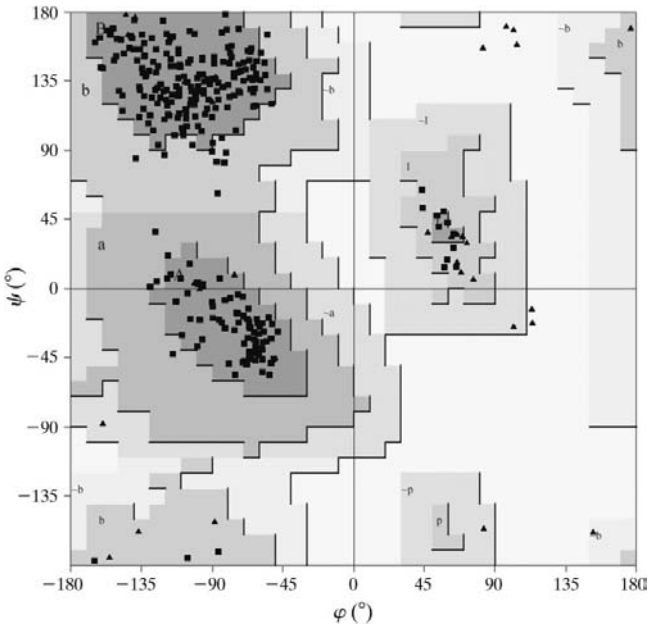


Figure 1
A φ–ψ plot as produced by the program *PROCHECK* (Laskowski *et al.*, 1993).

2.4. Structure solution

The structure was resolved from a difference Fourier synthesis using amplitudes $|F_{\rm complex}| - |F_{\rm native}|$ and the phases from the isomorphous structure of native human pepsin (Fujinaga *et al.*, 1995). The refinement was performed using the crystallographic extension (Fujinaga *et al.*, 1989) to the molecular-simulation package *GROMOS* (van Gunsteren & Berendsen, 1987) interspersed with manual model fitting using *FRODO-TOM* (Chirino & Israel, 1992; Jones, 1978). The final coordinates were deposited with the Protein Data Bank (Bernstein *et al.*, 1977).

3. Results and discussion

3.1. Quality of the model

The final model consists of 2438 protein atoms, 52 inhibitor atoms and 135 solvent atoms. The *R* factor $(\sum|F_o - F_c| / \sum|F_o|)$ is 20.0% for all data in the resolution range 30–1.96 Å. The φ–ψ plot (Fig. 1) shows that 89% of the residues fall in the most favored region of the plot as defined by *PROCHECK* (Laskowski *et al.*, 1993). The root-mean-square (r.m.s.) coordinate error was estimated to be 0.2 Å using a σ_A plot (Read, 1986; Fig. 2). The r.m.s. deviations from ideality for bond lengths, bond angles and improper dihedral angles are 0.01 Å, 3.8 and 2.7°, respectively. The quality of the structure can also be appreciated from the electron-density map of the active-site region shown in Fig. 3. The region around the catalytic aspartate residues Asp32 and Asp215 is shown. There is continuous electron density between Asp32 and the phosphonate group of the inhibitor owing to the short hydrogen bond formed between the two groups (see below). Fig. 4 shows the average *B* factor along the polypeptide chain of pepsin complexed to the phosphonate inhibitor as well as the average *B* factors for the native human pepsin. The disordered regions of the molecule can be seen as spikes in the plot that have *B* factors exceeding about 50 Å^2. These regions are associated with little or no electron density. In addition, the flexible loop around residue 75, known as the 'flap', becomes ordered in the

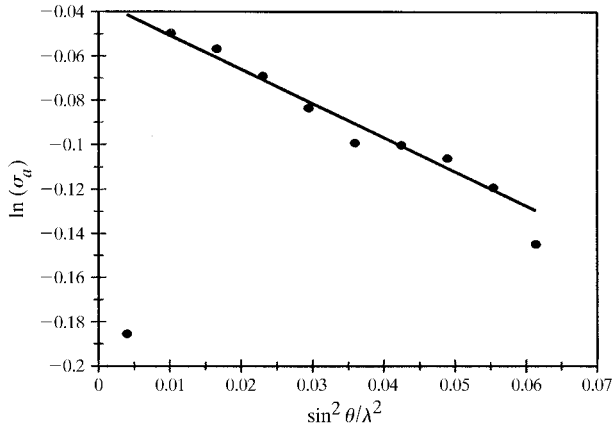


Figure 2
Estimation of coordinate error using a σ_A plot (Read, 1986). The mean-square coordinate error is given by −slope/26.3189.

complex structure. This is a common observation for the aspartyl proteinases, as the flap is intimately involved in binding of a substrate or an inhibitor. The solvent molecules range in B factor from 8.5 to 60.0 Å^2 and have a mean value of 37.2 Å^2.

3.2. Active-site geometry

The hydrogen bonds between the enzyme and the inhibitor are essentially the same as those found between human pepsin and pepstatin (Fig. 5). This is surprising for the hydrogen bonds formed with the prime site residues (S1′, S2′, S3′; nomenclature according to Schechter & Berger, 1967) because pepstatin has two extra atoms inserted in the main chain [$-$CH(OH)CH$_2-$]. As in the complexes of phosphonate inhibitors with penicillopepsin, there is a short hydrogen bond (2.4 Å) from the carboxyl O atom of Asp32 to the phosphonate *pro-S* O atom. At the present resolution of the structure determination, it is impossible to determine the H-atom positions or the protonation state of chemical groups. However, inferences can be made based on the possible hydrogen-bonding distances around the phosphonate group. It would seem that Asp32 is protonated, with the hydrogen shared in a short hydrogen bond to the phosphonate *pro-S* O atom, and the *pro-R* O atom of the phosphonate is protonated and donates a bifurcated hydrogen bond to Asp215.

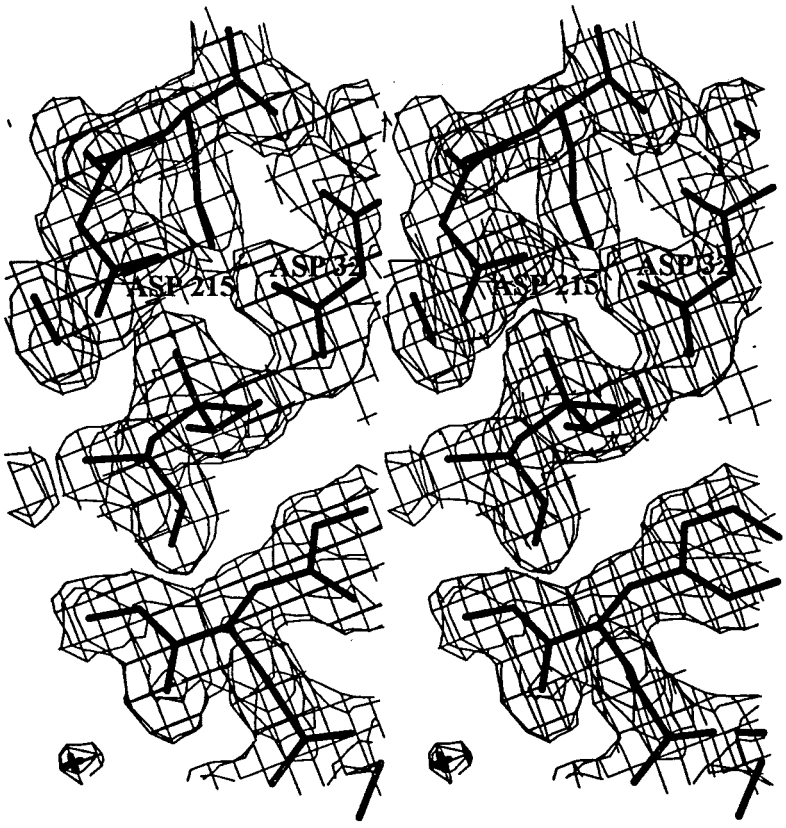


Figure 3
Stereogram showing the electron density in the active site of the refined structure. The map was calculated with coefficients $(2m|F_o| - D|F_c|, \alpha_c)$, calculated with the program *SIGMAA* (Read, 1986) and contoured at 1σ.

3.3. Binding sites

The binding sites from S4 to S3′ are defined by the interactions of the residues P4 to P3′ of the inhibitor with the enzyme (Figs. 6, 7*a*). Indeed, it is unlikely that there are additional binding sites beyond these sites.

The S4 site is flat, very accessible to solvent and not well defined. The main-chain carbonyl group of P4 Iva points towards the solvent. The Iva group packs against a relatively flat hydrophobic surface on the enzyme (residues Met12 and Leu220 contribute to S4).

The sites S1 and S3 are contiguous, with the carbonyl O atom of Gly217 providing some separation of the two sites. The S1 site tends to be hydrophobic in nature, whereas the S3 site is mainly polar, except for the hydrophobic surface provided by the aromatic ring of Phe111. Larger side chains than those in the inhibitor (P1, Leu; P3, Val) would not be easily accommodated at these sites without some rearrangement of the groups involved. Main-chain hydrogen bonds between the inhibitor and the enzyme are made to Ser219, Gly217 and Thr218 and to the catalytic site residues Asp32 and Asp215.

The S2 site is mainly hydrophobic. However, beyond the β position, the side chain can either extend towards S4 or towards S1′. This duality of interaction has been observed with inhibitors bound to other aspartyl proteinases (Foundling *et al.*, 1987; Suguna *et al.*, 1992). The area between S4 and S2 is polar, whereas the junction between S1′ and S2 is non-polar. There is sufficient space in the S2–S1′ zone to accommodate a tyrosine side chain. The main-chain hydrogen bonds are made to Thr77 on the flap.

The S1′ site is mainly hydrophobic in nature. The Phe at the P1 position packs mainly against Tyr189. Leu213, Leu291 and Leu298 form other parts of the S1′ pocket and tend to separate the S1′ site from the S2 site. A hydrogen bond exists between the main-chain carbonyl O atom of P1′ Phe and Gly76. However, there are no potential hydrogen-bond acceptors for the P1′ main-chain amino group that would exist in a normal substrate in place of the phosphoester O atom present in the inhibitor.

The S2′ site is clearly defined by the P2′ alanine residue. The main chain of the inhibitor makes hydrogen bonds to the Gly34 O atom and accepts a hydrogen bond from Tyr189 OH. The side chain points to a relatively large polar site.

The main-chain N of the P3′ residue donates a hydrogen bond to Thr74 O. The side chain points mainly toward the solvent; however, larger residues may be able to make interactions to the mixed hydrophilic/hydrophobic surface of the flap residues.

3.4. Comparison with porcine pepsin

The structures of human and porcine pepsins are very similar. There is 85% sequence identity

between the two enzymes. If the structure of porcine pepsin (Sielecki *et al.*, 1990) is superposed (Guex & Peitsch, 1997) onto the structure of human pepsin from the current study, the r.m.s. deviation for the 322 equivalent C atoms is 0.6 Å. The sequence changes going from porcine to human pepsin in the binding sites are Thr12Met, Ile30Val, Glu287Gln and Val291Leu (Fig. 6). The most significant change is the Thr12Met substitution; it leads to a more extensive hydrophobic binding surface for P4 in the human enzyme (Fig. 7).

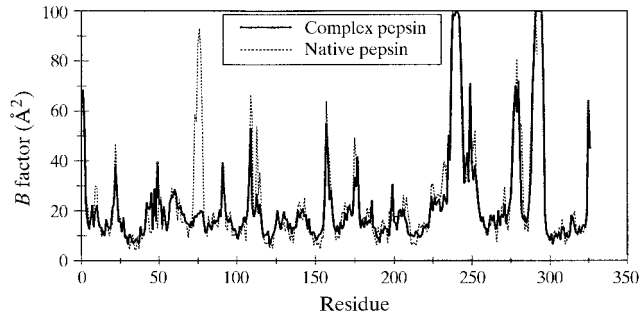


Figure 4
The average main-chain *B* factor along the polypeptide chain. The *B* factors for pepsin bound to the phosphonate inhibitor (solid line) is compared with those for native pepsin (dashed line; Fujinaga *et al.*, 1995).

3.5. Comparison with other phosphonate inhibitors bound to penicillopepsin

The crystal structure of a similar inhibitor, Iva-Val-Val-Leu^{P}-(O)Phe-OMe [Leu^{P} = the phosphinic acid analog of Leu, (O)Phe = L-3-phenyllactic acid, OMe = methyl ester], bound to the fungal aspartyl proteinase penicillopepsin has been reported (Fraser *et al.*, 1992). The conformation of the inhibitor is very similar to the one found in the present study except at two places (Fig. 8*a*). The P1′ residue in the penicillopepsin complex points toward the P2 residue. This is possible because of the much larger S2–S1′ combined site in penicillopepsin. In fact, the S3–S1 site is also much larger in penicillopepsin, so that larger residues could be accommodated at these positions. Another difference in the penicillopepsin binding surface compared with that in pepsin is found at the S4 subsite. This site is much more restricted in penicillopepin than in pepsin. In the case of porcine pepsin, this site is almost non-existent. The restrictive nature of the S4 site in penicillopepsin could explain the 300-fold to 600-fold decrease in binding (Bartlett *et al.*, 1990) of inhibitors containing a Cbz (benzyloxycarbonyl) group at the P4 position relative to inhibitors with Iva at this position (Fig. 8*b*). In addition, the hydrophobic benzyl group buries polar atoms on forming the complex. The open S4 site in porcine pepsin (Fig. 7*b*) would be consistent with only a twofold to fourfold decrease in binding (Bartlett *et al.*, 1990) for the same change of inhibitors.

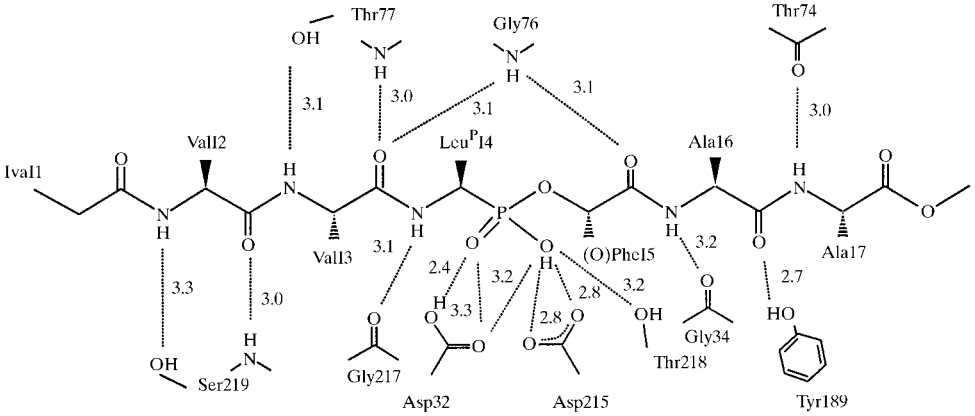


Figure 5
Possible hydrogen-bonding interactions between the enzyme and the inhibitor. Interactions with donor–acceptor distances less than or equal to 3.5 Å are considered to be possible hydrogen bonds. The protonation states of the phosphonate, Asp32 and Asp215 are based on the possible hydrogen-bonding distances.

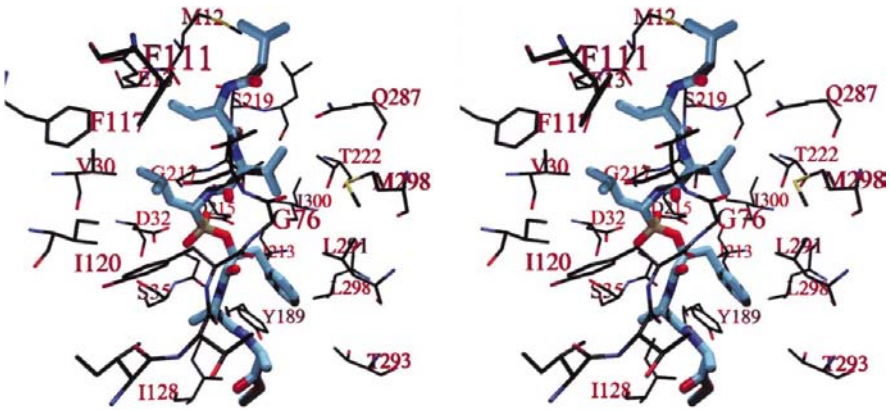


Figure 6
Stereographic view of the binding sites. Residues of the enzyme within van der Waals contact of the inhibitor are shown. The drawing was generated using *VMD* (Humphrey *et al.*, 1996) and rendered using *POV-RAY* (Persistence of Vision Development Team, 1997).

3.6. Short hydrogen bonds

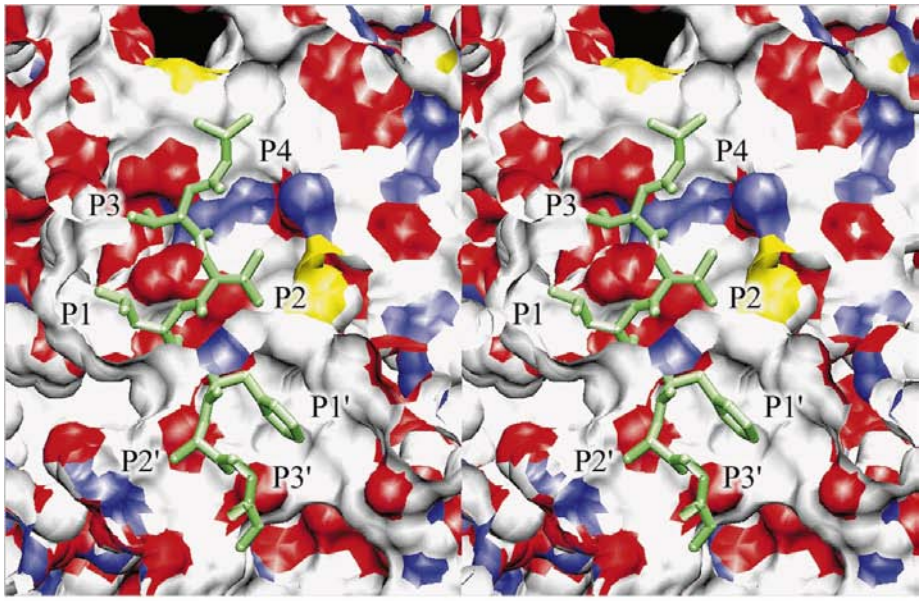
As mentioned above, there is a short hydrogen-bonding distance (2.4 Å) from the *pro-S* phosphonate O atom to the carboxyl group of Asp32. These short hydrogen bonds are sometimes called low-barrier hydrogen bonds, as the barrier for proton transfer from the donor atom to the acceptor atom is thought to be small. There is a controversy (Cleland & Kreevoy, 1994; Frey *et*

al., 1994; Warshel *et al.*, 1995) over the energetics of these short hydrogen bonds and their role in catalysis. However, there seems to be no clear-cut resolution, as assumptions have to be made from both theoretical and experimental points of view. In the present structure, the short hydrogen bond probably does not contribute to a large enhancement in binding energy. A similar inhibitor bound to penicillopepsin also has a short hydrogen bond (Fraser *et al.*, 1992), yet the binding constant is not significantly different from that of a related difluorostatone inhibitor bound to penicillopepsin (James *et al.*, 1992) lacking such a short hydrogen bond.

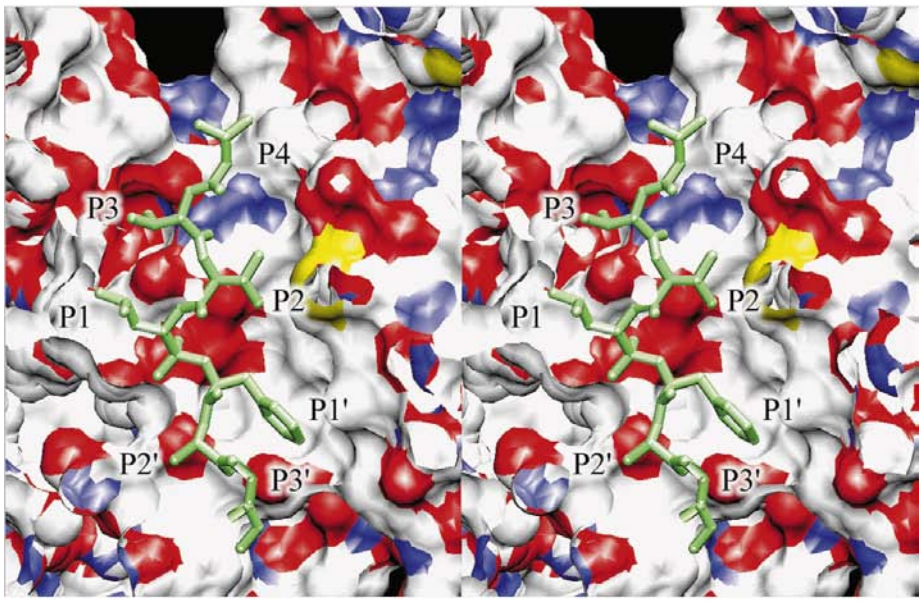
3.7. Convergence of reaction mechanisms

Aspartic proteinases hydrolyze peptide bonds by a general base assisted nucleophilic mechanism (Davies, 1990; James *et al.*, 1992; Suguna *et al.*, 1987). Asp215 serves as a general base, abstracting a proton from a water molecule bound at the active site between Asp32 and Asp215. The resulting nucleophilic hydroxide ion attacks the carbonyl C atom of the scissile peptide bond, leading to a tetrahedral transition state with a negatively charged oxyanion. This negative charge is stabilized by a hydrogen bond from the protonated Asp32 and from the edge of the phenolic ring of Tyr75 (Blundell *et al.*, 1987). Asp215 may serve as the proton donor, transferring its hydrogen to the leaving-group N atom. The collapse of the tetrahedral state leads to the cleavage of the peptide bond.

It turns out that other proteinases have converged to similar reaction mechanisms (Garavito *et al.*, 1977; James, 1993). The metalloproteinases use a reaction mechanism that is closely related to that of the aspartic proteinases. As a representative case, the mechanism of thermolysin has been described by Holden *et al.* (1987). As in the case of the aspartic proteinases, the γ-carboxylate of Glu143 acts as a general base, abstracting a proton from a bound water molecule and converting it into a strong nucleophile. The Zn^{2+} ion may stabilize the OH^{-} ion, facilitating its formation. There appears to be no suitable counterpart in the aspartyl proteinases to stabilize the nucleophile. The electrophilic stabilization of the tetrahedral intermediate is provided by Zn^{2+} and a hydrogen bond from His231. As in the case of the aspartic proteinases, the general base Glu143 can act as the proton donor for the leaving-group N atom. Fig. 9 shows a superposition of the active sites of thermolysin and pepsin, both with phosphorus-containing inhibitors. It shows a similar disposition of the catalytically important groups; the nucleophile, general base and the electrophilic components. It is also clear that not everything can work in



(*a*)



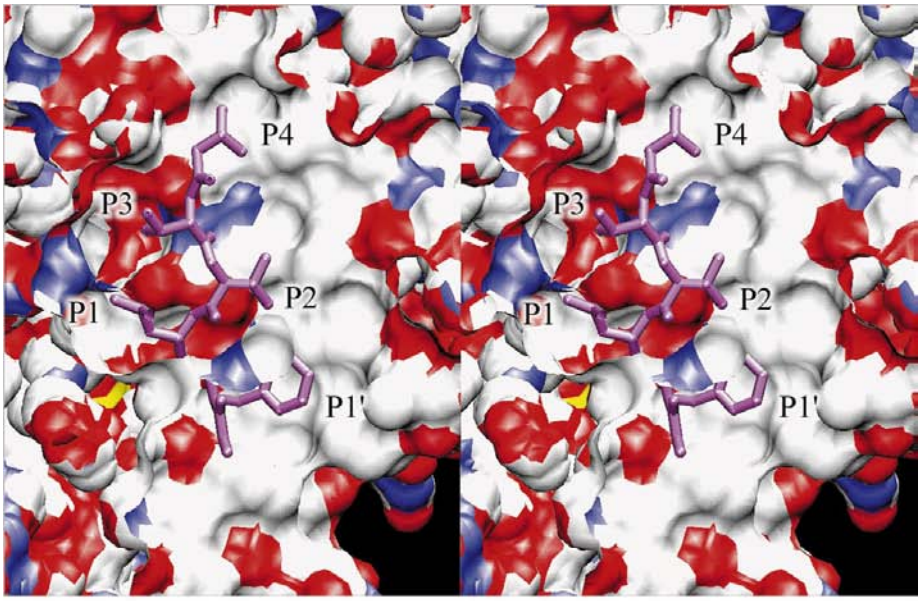
(*b*)

Figure 7
Stereogram of the accessible surfaces of (*a*) human pepsin and (*b*) porcine pepsin showing the binding sites. The surface is colored according to atom type, where carbon is white, nitrogen is blue, oxygen is red and sulfur is yellow. The phosphonate inhibitor is shown as a stick model. The binding of the inhibitor to porcine pepsin was modeled by superposing the human pepsin–inhibitor complex structure onto the native porcine pepsin structure (Sielecki *et al.*, 1990) using *SwissPdbViewer* (Guex & Peitsch, 1997). The image was produced using *VMD* (Humphrey *et al.*, 1996). The atoms in the foreground have been clipped away to show the interactions in the binding sites.

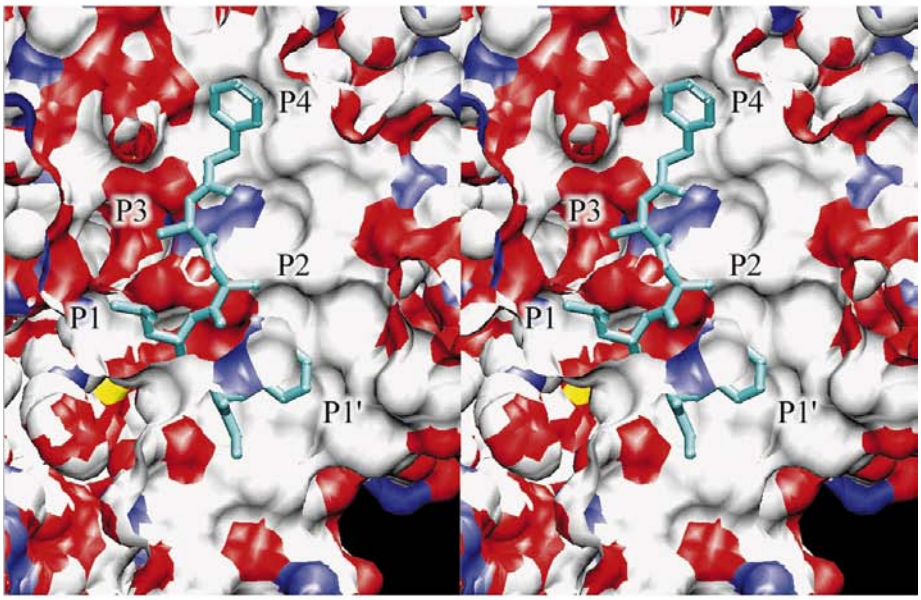
the same way. A Zn^{2+} ion cannot always act like an aspartic acid. Even though they may both be stabilizing the negative charge on the tetrahedral transition state, their interactions with the nucleophilic water will be completely different. The positively charged Zn^{2+} ion will interact with one of the negatively charged lone-pair orbitals on the water molecule. In the case of pepsin, the carboxyl O atom of Asp32 that stabilizes the transition state is protonated. Consequently, the other O atom will have a partial negative charge and will accept a hydrogen bond from the nucleophilic water molecule. In both enzymes, these groups probably help position and orient the nucleophile.

The case of the serine proteinases is a little more complicated, as these enzymes use a two-step mechanism going through a covalent acyl-enzyme intermediate. However, each part of the reaction has the same components as for the other classes of proteinases. The basic mechanism has been described and analyzed theoretically by Warshel *et al.* (1989). Taking chymotrypsin as an example, the general base is provided by His57, which abstracts a proton from the hydroxyl group of Ser195, which acts as the nucleophile. The tetrahedral transition state is stabilized by hydrogen bonds from the main-chain NH groups of Gly193 and of Ser195. The cleavage of the peptide bond results from the collapse of the transition state, assisted by a proton transfer from the general base His57 to the leaving-group N atom. At this point, a covalent acyl-enzyme intermediate is formed. The second half of the reaction proceeds in an analogous fashion, except that now it is a water molecule instead of Ser195 that acts as the nucleophile (Blanchard & James, 1994). Fig. 10 shows the superposition of the active sites of human pepsin in complex with a phosphonate inhibitor (present structure) and of chymotrypsin in complex with the ovomucoid inhibitor domain 3 from turkey (OMTKY3; Fujinaga *et al.*, 1987). The chymotrypsin–OMTKY3 complex is a good mimic of an enzyme–substrate complex. As in the case of the metalloproteinases, the catalytically important groups in the two enzymes have similar dispositions around the scissile peptide bond. In the case of the serine proteinases, the position and the orientation of the attacking nucleophile is assured by the hydrogen bond to the general base His57 $N^{\varepsilon 2}$ and the covalent bond to the C^{β} atom of Ser195.

The cysteine proteinases also use a general base assisted nucleophilic reaction mechanism going through a tetrahedral transition state (Storer & Menard, 1994). As in the case of the serine proteinases, the enzyme uses a two-step mechanism with a covalent acyl-enzyme intermediate. In the case of papain, the nucleophile is the S^{γ} atom of Cys25, the general base is the $N^{\delta 1}$ atom of His159 and Gln19 $N^{\varepsilon 2}$, and Cys25 NH are the electrophiles stabilizing the negative charge on the transition state. However, the active sites of cysteine proteinases cannot be superposed on those of the other groups of



(*a*)



(*b*)

Figure 8
Stereogram of the accessible surfaces of penicillopepsin bound to (*a*) Iva-Val-Val-Leu^{P}-(O)Phe-OMe and (*b*) to Cbz-Ala-Ala-LeuP-(O)-Phe-OMe (Fraser *et al.*, 1992). The surfaces are produced and colored as in Fig. 7.

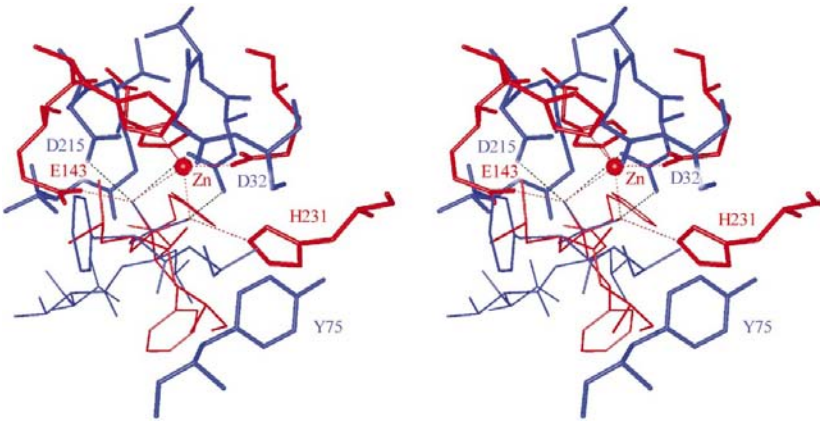


Figure 9
Stereogram of the superposition of the active-site residues of human pepsin–inhibitor complex (blue) with those of a metalloproteinase, thermolysin, in complex with a phosphonamidate inhibitor (red) (Holden *et al.*, 1987; PDB code 4tmn). The superposition was performed by aligning the C^{α} atoms of the P1 and P1′ residues and the atom that corresponds to the carbonyl O atom of a substrate.

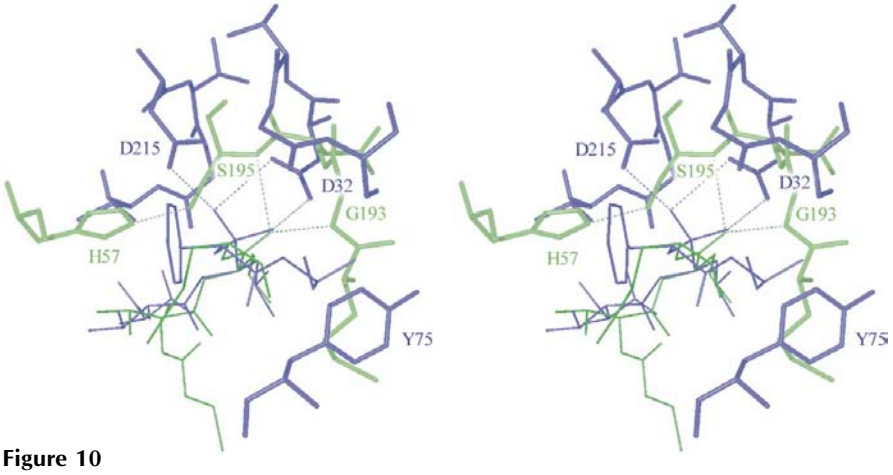


Figure 10
Stereogram of the superposition of the active-site residues of human pepsin–inhibitor complex (blue) with those of a serine proteinase, chymotrypsin, in complex with a protein inhibitor, OMTKY3 (green) (Fujinaga *et al.*, 1987; PDB code 1cho). The superposition was performed as in Fig. 9.

proteinases discussed above. This is because the nucleophilic attack in the case of the cysteine proteinases occurs from the *Si* face of the scissile peptide bond, whereas in the other cases the attack is from the *Re* face. One may think that cysteine proteinases are different from the other enzymes discussed above, since the ion-pair His^{+} S^{-} state is the ground state. However, there are viral cysteine proteinases with a chymotrypsin-like fold (Allaire *et al.*, 1994) which presumably attack from the *Re* face. Cysteine proteinases are not unique in this mode of attack of the peptide bond. Wheat serine carboxypeptidase II (Bullock *et al.*, 1996) has also been shown to hydrolyze peptide linkages *via* nucleophilic attack from the *Si* face.

This research was sponsored by the Medical Research Council of Canada in a grant to the MRC Group in Protein Structure and Function, by the National Institutes of Health (CA-22747) and by the National Cancer Institute, DHHS, under contract NO1-CO-46000 with ABL.

References

Allaire, M., Chernaia, M. M., Malcolm, B. A. & James, M. N. (1994). *Nature (London)*, **369**, 72–76.
Bartlett, P. A., Hanson, J. E. & Giannousis, P. P. (1990). *J. Org. Chem.* **55**, 6268–6274.
Bateman, K. S., Chernaia, M. M., Tarasova, N. I. & James, M. N. (1998). *Adv. Exp. Med. Biol.* **436**, 259–263.
Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Blanchard, H. & James, M. N. (1994). *J. Mol. Biol.* **241**, 574–587.
Blundell, T. L., Cooper, J., Foundling, S. I., Jones, D. M., Atrash, B. & Szelke, M. (1987). *Biochemistry*, **26**, 5585–5590.
Bullock, T. L., Breddam, K. & Remington, S. J. (1996). *J. Mol. Biol.* **255**, 714–725.
Chirino, A. & Israel, M. (1992). *FRODO-TOM*. Pasadena/Edmonton.
Cleland, W. W. & Kreevoy, M. M. (1994). *Science*, **264**, 1887–1890.
Davies, D. R. (1990). *Annu. Rev. Biophys. Biophys. Chem.* **19**, 189–215.
Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., Sibanda, B. L., Hemmings, A., Wood, S. P., Blundell, T. L., Valler, M. J., Norey, C. G., Kay, J., Boger, J., Dunn, B. M., Leckie, B. J., Jones, D. M., Atrash, B., Hallet, A. & Szelke, M. (1987). *Nature (London)*, **327**, 349–352.
Fraser, M. E., Strynadka, N. C., Bartlett, P. A., Hanson, J. E. & James, M. N. (1992). *Biochemistry*, **31**, 5201–5214.
Frey, P. A., Whitt, S. A. & Tobin, J. B. (1994). *Science*, **264**, 1927–1930.
Fujinaga, M., Chernaia, M. M., Tarasova, N. I., Mosimann, S. C. & James, M. N. (1995). *Protein Sci.* **4**, 960–972.
Fujinaga, M., Gros, P. & van Gunsteren, W. F. (1989). *J. Appl. Cryst.* **22**, 1–8.
Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M. Jr, & James, M. N. (1987). *J. Mol. Biol.* **195**, 397–418.
Garavito, R. M., Rossmann, M. G., Argos, P. & Eventoff, W. (1977). *Biochemistry*, **16**, 5065–71.
Guex, N. & Peitsch, M. C. (1997). *Electrophoresis*, **18**, 2714–2723.
Gunsteren, W. F. van & Berendsen, H. J. C. (1987). *GROMOS*. BIOMOS b.v., Groningen, The Netherlands.
Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H. & Matthews, B. W. (1987). *Biochemistry*, **26**, 8542–8553.
Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.
James, M. N. G. (1993). *Proteolysis and Protein Turnover*, edited by J. S. Bond & A. J. Barrett, pp. 1–8. London: Portland Press.
James, M. N. G., Sielecki, A. R., Hayakawa, K. & Gelb, M. H. (1992). *Biochemistry*, **31**, 3872–3886.
Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
Khan, A. R., Cherney, M. M., Tarasova, N. I. & James, M. N. (1997). *Nature Struct. Biol.* **4**, 1010–1015.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Moore, S. A., Sielecki, A. R., Chernaia, M. M., Tarasova, N. I. & James, M. N. (1995). *J. Mol. Biol.* **247**, 466–485.
Persistence of Vision Development Team (1997). *POV-RAY*. http://www.povray.com.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Schechter, I. & Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157–162.
Sielecki, A. R., Fedorov, A. A., Boodhoo, A., Andreeva, N. S. & James, M. N. G. (1990). *J. Mol. Biol.* **214**, 143–170.
Storer, A. C. & Menard, R. (1994). *Methods Enzymol.* **244**, 486–500.
Suguna, K., Padlan, E. A., Bott, R., Boger, J., Parris, K. D. & Davies, D. R. (1992). *Proteins*, **13**, 195–205.
Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D. & Davies, D. R. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 7009–7013.
Tarasova, N., Denslow, N. D., Parten, B. F., Tran, N., Nhuyen, H. P., Jones, A., Roberts, N. B. & Dunn, B. M. (1995). *Adv. Exp. Med. Biol.* **362**, 77–81.
Warshel, A., Naray-Szabo, G., Sussman, F. & Hwang, J. K. (1989). *Biochemistry*, **28**, 3629–3637.
Warshel, A., Papazyan, A. & Kollman, P. A. (1995). *Science*, **269**, 102–106.
Xuong, N. H., Sullivan, D., Neilsen, C. & Hamlin, R. (1985). *Acta Cryst.* B**41**, 267–269.