THE STRUCTURE AND FUNCTION OF THE ASPARTIC PROTEINASES[1]

David R. Davies

Section on Molecular Structure, Laboratory of Molecular Biology, National Institutes of Health, 9000 Rockville Pike, Building 2, Room 316, Bethesda, Maryland 20892

KEY WORDS: pepsins, X-ray diffraction, acid proteases, mechanisms, retroviral proteases.

CONTENTS

PERSPECTIVES AND OVERVIEW

In the past dozen years, our understanding of the structure and function of the aspartic proteinases has increased greatly. We have made spectacular progress in determining precise crystal structures of many of these enzymes so that the mechanistic proposals of the enzymologists could be interpreted in three-dimensional terms. In this review, I describe the structures of these

[1] The US Government has the right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

enzymes as determined by X-ray diffraction and the results of inhibitor binding studies, and discuss how these data have helped to clarify the catalytic mechanism.

The aspartic proteinases make up a widely distributed class of enzymes. They are found in vertebrates, fungi, plants, and, more recently, retroviruses. They are characterized by pH optima in the acid range; inhibition by pepstatin, a hexapeptide from *Streptomyces* (101); the use of two aspartic acid side chains in the catalytic mechanism (29, 58); and specificity for extended peptide substrates (28, 29, 43). Hofmann (39) has noted that aspartic proteinases have a long history: chymosin, in the form of rennet, has been used for millenia in cheese making and aspartic proteinases are also used in making soy sauce, which reportedly originated during the Zhou dynasty (60).

Recently, a novel form of aspartic proteinase has been discovered in retroviruses. Miller et al (67) reported crystal structures for the proteinase from Rouse Sarcoma virus, while others (69, 105) have determined those for the proteinase from the human immunodeficiency virus, HIV-1. These are the first retroviral proteins to be analyzed by X-ray diffraction. Coincidentally, in 1934 Bernal & Crowfoot (Hodgkin) (7) analyzed an aspartic proteinase, pepsin, the first crystalline protein ever analyzed by X-ray diffraction.

As in the other proteinase families, the individual enzymes perform many functions. The best known gastric enzymes, such as pepsin, gastricsin, and chymosin (formerly rennin), are involved in digestion. Cathepsin D, a major lysozomal enzyme, also degrades proteins, but does so intracellularly. Other enzymes, such as renin, produce very specific cleavage in a single protein to produce a decapeptide that is the precursor of angiotensin II, an octapeptide that is a major factor in the control of blood pressure. In the fungi, the aspartic proteinases may play a role in sporulation (39). In the retroviruses, many of the viral proteins are synthesized as polyproteins, which during activation of the virus are cleaved by the retroviral proteinase (21).

All of the vertebrate and probably most of the fungal aspartic proteinases are synthesized as inactive zymogens, and contain an additional N-terminal segment, approximately 45 amino acid residues long, that gets cleaved and separated upon activation (104). Recently, a crystal structure for pepsinogen has been reported (50) along with a proposal for the mechanism of activation. The Protein Data Bank (8) also contains coordinates for an independently determined high resolution refined structure of pepsinogen (36).

Three major meetings have been devoted to the aspartic proteinases. The first, organized by Jordan Tang, was held in 1976, and resulted in a

book (96). The second meeting was in Prague in 1984 and also yielded a book (61). A book of abstracts (25) came out of the third meeting, in Elsinore 1988, organized by Bent Foltmann and colleagues. The proceedings of these three meetings, Prague in particular (61), provide an excellent introduction to this field. Other useful reviews include those by Tang & Wong (99) on sequence, James & Sielecki (51) on structure and mechanism, and Hofmann (39) on general aspects.

PRIMARY STRUCTURE

Most of the mammalian, plant, and fungal aspartic proteinases are single-chain enzymes with a molecular weight of $\sim$ 35 kd. Porcine pepsin was the first to be sequenced (68, 98), and by now there are sequences for 22 proteinases [for review see (25, 99)]. The enzymes are approximately 327 amino acids long, with $\sim$5% sequence identity between all members of the family. The aspartic proteinases have characteristic sequences in the region of the two catalytic aspartyl residues: (hydrophobic, generally Phe)-Asp32-Thr-Gly-Ser in the N-terminal domain, and a corresponding (hydrophobic)-Asp215-Thr-Gly-Ser/Thr in the C-terminal domain (pepsin numbering).

The zymogens have an N-terminal propeptide of up to 50 amino acids long (44 in pepsinogen), which is cleaved upon activation (see below). With one exception they contain an invariant lysine at position 36 in the propeptide (25, 99).

The retroviral proteinases are considerably smaller than the pepsinlike enzymes; they are less than 130 amino acids long. They also contain the sequence Asp-Ser/Thr-Gly, on which grounds Toh et al (100) proposed these proteinases belonged to the general aspartic proteinase superfamily. Pearl & Taylor (74) aligned the sequences of 16 retroviral proteinases with each of the two domains of the pepsinlike proteinases, based on matching patterns of residue conservation and secondary structure predictions. For both criteria, they found a pattern of equivalence between the three sequences. They concluded that the retroviral proteinases could be modeled to be "parsimonious versions" of the domains of the pepsinlike structures and associated in pairs to give an active enzyme.

THE PEPSINLIKE SINGLE-CHAIN PROTEINASES

Three-Dimensional Structure

High-resolution crystal structures have been reported for penicillopepsin (48), rhizopuspepsin (90), endothiapepsin (72), porcine pepsin (3), pep-

sinogen (50), human renin (85), natural and recombinant chymosin (31, 81, 102), and mucorpepsin (Table 1) (102). Comparison of all these proteins shows that they have a considerable degree of structural similarity. In this review, space limitations preclude a separate discussion of each structure. Although many of the figures shown here apply to rhizopuspepsin (90), the general features are applicable to all members of the family.

Figure 1 is a stereo drawing of the Cα atoms of rhizopuspepsin, showing its tertiary structure. The overall secondary structure consists almost entirely of pleated sheet with very little α-helix. The molecule is bilobal with two domains of similar structure related by an approximate dyad rotation axis originally observed in endothiapepsin and penicillopepsin by Tang et al (97). Within the N- and C-terminal domains, less precise two-fold symmetry has been noted (1, 11). A large cleft, about 40 Å long, runs entirely across the molecule, and separates the two domains. Oligopeptide inhibitors have been observed to bind within this cleft. Beneath the two domains, an extensive six-strand antiparallel pleated sheet forms the base of the molecule. Figure 2 shows a schematic representation of the folding. It does not include the residues in the loops, i.e. those not contained within alpha helices or in beta sheets. The central six segments make up the basal pleated sheet; each domain contributes three strands.

The Catalytic Site

The catalytic aspartic acid residues are located on the ends of the loops of two Greek psi–like loops, which extend from each domain and are related to each other by the pseudodyad axis. Here, in the middle of the extensive

Table 1   Crystal structures for the pepsin class of aspartic proteinases

Enzyme	Resolution (Å)	R-value	References
Porcine pepsin	2.0		2–4
Endothiapepsin	2.1	0.16	10, 56, 72, 89
Chymosin, recombinant[a]	2.3		31
Porcine pepsinogen (b)[a]	1.7	0.173	36
Penicillopepsin	1.8	0.136	46, 47, 48
Porcine pepsinogen (a)	1.8	0.167	50
Chymosin[a]	2.2		81, 102
Human recombinant renin	2.5	0.236	85
Rhizopuspepsin	1.8	0.143	88, 89, 90
Mucorpepsin[a]	2.6	—	102

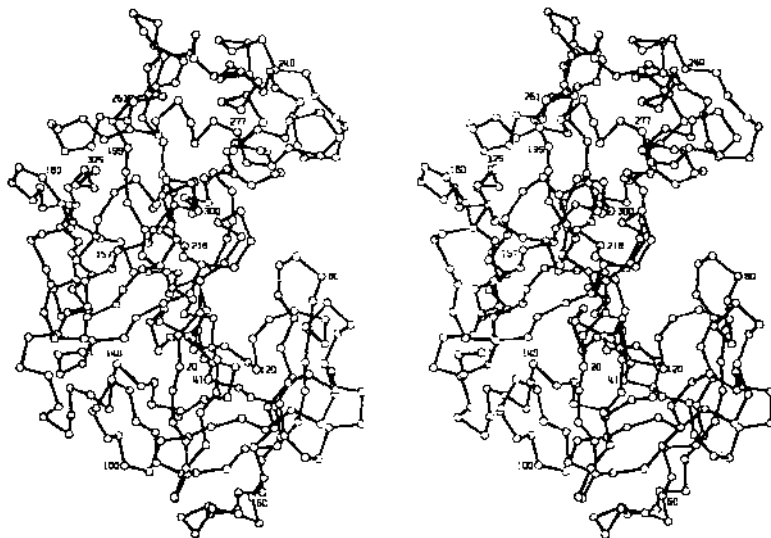[a] Indicates very preliminary announcement in Ref. 25.

Figure 1 Stereo view of the Cα backbone of rhizopuspepsin. Note the bilobal nature of the structure and the extensive substrate binding cleft. The lower and upper lobes correspond to the N- and C-terminal domains, respectively. The flap is visible at the bottom right. Rhizopuspepsin numbering.

cleft, the structures of all these enzymes are very similar. Sielecki et al (85) have superimposed this core region of the active site for the three fungal proteinases, residues 30–35, 120–124, 213–218, and 301–305, that make up the two psi's. They report root mean square (rms) differences for 105 main-chain atoms of approximately 0.24 Å and that human renin differs from these three enzymes by only 0.45 Å.

Figure 3 represents a superposition of penicillopepsin and rhizopuspepsin. The rms displacement of 294 carbon alpha atoms in the two structures is only 1.21 Å, and the figure shows that the active site atoms are almost completely superimposable. The longest stretch of identical residues in penicillopepsin and rhizopuspepsin is from 72 to 85 and includes a hairpin loop known as the "flap," which projects out over the cleft at the active site. In these two structures, however, the positions of the tips of the flaps differ significantly, at least partially because of different crystal contacts in this region.

The two carboxyl groups at the active site are connected to one another through a complex network of hydrogen bonds that makes use of the Asp-
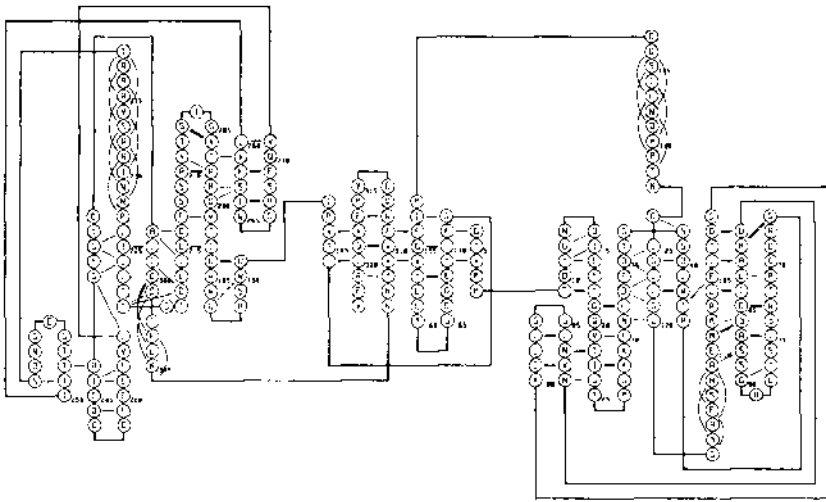
Figure 2 The rhizopuspepsin structure shown schematically. Both domains consist almost entirely of beta-sheet structure with just a few short helical segments. The middle six segments form the basal pleated sheet; the rest of the two domains are on either side. The approximate dyad axis would be located in the center of the figure, perpendicular to the paper. The two χ structures containing the Asp-Thr-Gly-residues occur in the middle of each domain. The flap is the hairpin loop shown at the extreme right.

Thr-Gly-Ser sequence (Figure 4). The two threonines play a crucial role; the hydroxyl groups each form hydrogen bonds to the other threonine's amide nitrogen. They are also within hydrogen-bond distance of the carbonyl oxygen atoms of the two hydrophobic residues preceding the catalytic aspartics (47, 69, 86). See Figure 4.

One might expect this network of hydrogen bonds to make the active-site region of the molecule quite rigid, and the low crystallographic B factors observed in this region confirm this idea. One exception is the flap, which in the native enzyme has a high B factor, indicating high mobility. This flap closes down over inhibitors bound in the cleft, further excluding solvent, while becoming considerably less mobile (18, 54, 92).

The binding pockets for each residue of substrate, up to four on the N-terminal side of the scissile bond and two on the C-terminal side, have been described for penicillopepsin (51, 52), rhizopuspepsin (92), and endothiapepsin (82). The number of enzyme residues involved at each subsite varies with the size of the substrate-inhibitor side chain and with the criteria used for defining a contacting residue, but extensive contact is made in all six subsites.
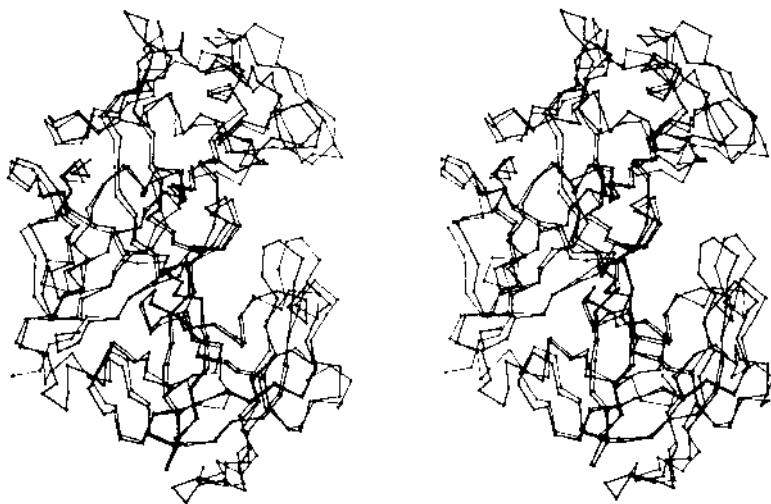
Figure 3 A superposition of penicillopepsin and rhizopuspepsin. Same view as in Figure 1. Rhizopuspepsin is drawn with a thicker line. Note the high degree of similarity in the vicinity of the active site.

The Zymogens and the Mechanism of Activation

The aspartic protcinases arc mostly produced in zymogen form. The X-ray analyses of one zymogen, pepsinogen, have led to a structural interpretation for the lack of activity of the zymogens and for their mechanism of activation. Two studies (36, 50) have determined the crystal structure of pepsinogen.

Pepsinogen, the inactive precursor of pepsin, has an additional 44 amino acid residues at the N-terminus (50). This propeptide is quite basic and contains nine lysine, two arginine, and two histidine residues (Figure 5). The propeptide's release, which activates the zymogen, is carried out at low pH in an autocatalytic manner (38). Spectroscopic analysis showed that electrostatic forces were important in stabilizing the zymogen and that some loss of helical structure accompanied activation (75). McPhie demonstrated that if the pH is reduced to 2.85, a rapid change of conformation occurs that is reversible if the pH is quickly returned to neutrality (66).

James & Sielecki (50) report that the N-terminal 44-residue prosegment folds into a compact domain that occupies the active site/cleft region of pepsin (Figure 5). The conformation of the polypeptide chain in the vicinity
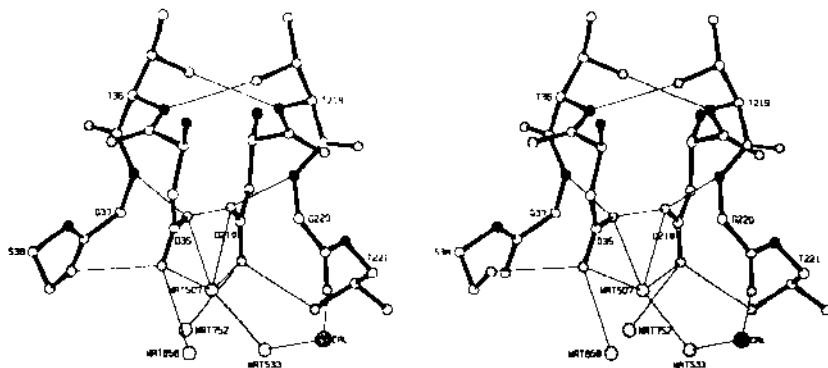
Figure 4   The conformation of the residues Asp-Thr-Gly-Thr-Ser from the two domains at the active site of rhizopuspepsin (90). The thin lines indicate the distances short enough to correspond to potential hydrogen bonds. (Since the hydrogen atoms cannot be seen directly in the normal protein electron density map, the assignment of hydrogen bonds is subject to debate.) The approximate dyad axis passes vertically between the two domains and through the water molecule 507. The two aspartic residues D35 and D218 correspond to D32 and D215 of pepsin. A calcium ion is observed in the vicinity of the catalytic site of rhizopuspepsin (*lower right*), resulting from the crystallization conditions. Nitrogen atoms are black, oxygen atoms are stippled.
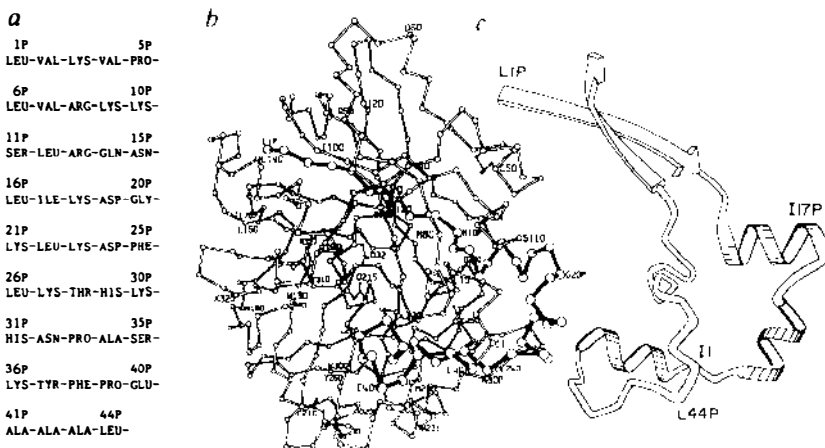

a

1P				5P
LEU-VAL-LYS-VAL-PRO-				
6P				10P
LEU-VAL-ARG-LYS-LYS-				
11P				15P
SER-LEU-ARG-GLN-ASN-				
16P				20P
LEU-ILE-LYS-ASP-GLY-				
21P				25P
LYS-LEU-LYS-ASP-PHE-				
26P				30P
LEU-LYS-THR-HIS-LYS-				
31P				35P
HIS-ASN-PRO-ALA-SER-				
36P				40P
LYS-TYR-PHE-PRO-GLU-				
41P				44P
ALA-ALA-ALA-LEU-				

Figure 5   The propeptide fragment of bovine pepsinogen. (*a*) The sequence of the propeptide. (*b*) The structure of pepsinogen with the propeptide backbone bonds in black. (*c*) A ribbon drawing of the secondary structure of the propeptide. Reproduced from James & Sielecki (50), with permission.

of the two catalytic aspartates is similar to that observed in penicillopepsin (0.36 Å for 48 common atom pairs). Also, the overall folding pattern of the zymogen's active part resembles the common fold of the aspartic proteinases. However, the position of the first 12 residues of pepsin differs in the two structures, and the side chain, Ile 1, of the zymogen is located ~40 Å away from its position in the active enzyme. The first six residues of the propeptide of pepsinogen, Leu-1P to Leu-6P, occupy the same position in the zymogen as the first six residues of pepsin in the active structure.

James & Sielecki (50) suggest that inhibition presumably occurs because residues Ser-11P to Leu-44P in the propeptide block access to the catalytic aspartates in the active site of the enzyme. In particular, the amino group of the almost invariant Lys-36P is located close to the two carboxyl groups of these aspartates, ~0.5 Å from the position of a bound water molecule observed in most of the active enzyme crystal structures.

The propeptide makes a number of electrostatic interactions with the enzyme structure; six of the basic amino acid side chains form ion pairs with carboxylate side chains of the pepsin structure. Lowering the pH possibly protonates these carboxylates, causing a breakup of the complex, followed by cleavage of the bond between Leu-16P and Ile-17P, with further proteolysis leading to formation of the active enzyme (50).

Inhibitor Binding Studies

The catalytic aspartic acids were originally identified by the active site reagents diazoacetylnorleucine methyl ester (78) and 1,2-epoxy 3-(*p*-nitrophenoxy) propane (23). Pepstatin (101), which contains the unusual amino acid statine, (3S,4S)-4-amino-3-hydroxy-6-methylheptanoic acid, and is proposed to be a transition state analog (63), universally inhibits the members of the aspartic proteinase family, evidence that supports the view that these enzymes all have a similar mechanism of action.

A detailed description of the inhibitors of the aspartic proteinases is beyond the scope of this review but is reviewed elsewhere (13, 79, 93). Here, I will describe two classes of inhibitor that bind to one of the three fungal proteinases. They are (*a* below) those containing a statine residue in the P_1 position (13, 14, 79) and (*b, c, d, e* below), those that are transition state isosteres and mimic the tetrahedral intermediate of the transition state (79, 93–95). [The notation of Schechter & Berger (83) is used here, in which the residues of the substrate are numbered away from the scissile bond, $P_1 - - - P_n$ on the N-terminal side and $P_1' - - - P_n'$ on the C-terminal side. Corresponding enzyme subsites are labeled S_n.]

$$-CHR-C(OH)_2- \text{ -- } N^+H_2-CHR'- \qquad \text{Tetrahedral intermediate.}$$

(a) $-CHR-CHOH- \text{ - - - } -CH_2-CO-NH-CHR'-$ Statine analog

(b) $-CHR-CH_2 \text{ - - } -N^+H_2-CHR'- \qquad$ Reduced peptide

(c) $-CHR-CHOH \text{ - - - } -CH_2-CHR'- \qquad$ Hydroxy isostere

(d) $-CHR-CO \text{ - - - - } -CH_2-CHR'- \qquad$ Keto isostere

(e) $-CHR-C(OH)_2- \qquad CH_2-CHR'- \qquad$ Ketone hydrate

The keto isostere (d) has been shown by ^{13}C NMR analysis to form the gem diol tetrahedral adduct with pepsin (45), acquiring a water molecule from the solvent, thus providing support for the general base–general acid type of mechanism.

These inhibitors have been extensively used to explore the subsite specificity of particular aspartic proteinases, but there has also been a focus on inhibitors directed specifically against human renin, with the expectation that some will prove useful in the treatment of hypertension (32, 93). Since these inhibitors also effectively inhibit most of the other aspartic proteinases that have much broader specificity than renin, they have been examined by X-ray diffraction complexed to the fungal enzymes. The structures of these complexes have provided information critical to evaluating the plausibility of the mechanistic proposals as well as providing models for the inhibitor's binding to renin. Now that the structure of renin itself has been determined (85), we will be able to see how well these results, and, in particular, how the modeling of these inhibitors to models of renin compare with the actual binding of these inhibitors directly to renin (12). However, comparison with models will be hampered by the observation that the regions of the renin molecule that provide its specificity have been difficult to model accurately (85).

Renin cleaves the 10–11 bond of angiotensinogen to release the N-terminal decapeptide angiotensin I (Table 2). Cleavage of the two angiotensin I C-terminal residues by angiotensin converting enzyme (ACE) produces the octapeptide angiotensin II, which plays an important role in the regulation of blood pressure, water, and electrolyte balance. Peptides taken from the first 13 amino acids of angiotensinogen were demonstrated to act as weak competitive inhibitors of renin even though they were hydrolyzed slowly (32, 33, 59, 87). These peptides, in particular the octapeptide His-Pro-Phe-His-Leu-Leu-Val-Tyr-OH (86, 87), have been used to design more effective inhibitors for renin, both for the statine-containing class of inhibitors and for the transition-state isosteres.

Table 3 shows those inhibitors complexed with one of the fungal enzymes that have been examined by high-resolution X-ray diffraction. Some bind

Table 2 Angiotensinogen and products

	Amino acid number													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14

RENIN
|
Angiotensinogen, human   Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Val-Ile-His

Angiotensinogen, horse   Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Leu-Val-Tyr-Ser

ACE
|
Angiotensin I   Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu

Angiotensin II   Asp-Arg-Val-Tyr-Ile-His-Pro-Phe

very tightly to the respective proteinase: e.g. the pepstatin fragment binds to penicillopepsin with a K_i value of approximately 1.6 nM (54), and the L-363,564 inhibitor has a K_i to endothiapepsin of 16 nM (26, 27).

GENERAL MODE OF BINDING   All of these inhibitors bind in the cleft so that similar hydrogen bonds form from the enzyme to the backbone of the inhibitor on the N-terminal side of the scissile bond. Figure 6 shows the conformations of three inhibitors of endothiapepsin in two superpositions (26). The reduced inhibitors H-142 and H-256 superpose very well, while H-142 and L-363,564 superpose surprisingly well, especially on the C-terminal side of the scissile bond region, considering that L-363,564 contains a statine residue. The two additional atoms brought in by the statine backbone cause this residue to project beyond P_1, introducing a "frame shift" (13) in the subsequent amino acids. This alignment of pepstatin, originally conceived by Powers et al (77) on the basis of an analysis of pepsin inhibitors and subsequently based directly on the structure of pepstatin bound to rhizopuspepsin (15, 16), facilitated the design of better renin inhibitors (13, 14, 93).

Some of the inhibitor side chain locations beyond the $C\beta$ atom, even on the N-terminal side of the scissile bond, can differ widely, especially in P_2 where the two inhibitors of endothiapepsin in Figure 6b have their histidine side chains oriented in two different ways. The mode of binding for the other two fungal proteinases follows a similar pattern. Rhizopuspepsin has similar binding differences of the side chain of P_2 (91). This presumably reflects the lack of high specificity at these positions in the fungal proteinases.

Sali et al have recently analyzed (82) the binding to endothiapepsin of an oligopeptide analog inhibitor with a (S)-hydroxyethylene moiety

Table 3   Enzyme inhibitor complexes studied crystallographically

Enzyme[a]	Inhibitor	P_8 P_7 P_6 P_5 P_4 P_3 P_2 P_1 ↓[c]	P'_1 P'_2 P'_3 P'_4	References
	Renin substrate	Val-Tyr-Ile-His-Pro-Phe-His-Leu-CO-NH-	Val-Ile-His	
EP	H-142	Pro-His-Pro-Phe-His-Leu-CH_2-NH-	Val-Ile-His-Lys	9, 26, 27, 34
EP	H-256	Pro-Thr-Glu-Phe-CH_2-NH-	Phe-Arg-Glu	9, 18, 26, 27
RP	Reduced peptide	D His-Pro-Phe-His-Phe-CH_2-NH-	Phe-Val-Tyr	92
EP	CP-69,799[b]	Boc-Phe-His-Phe-CH(OH)-CH_2-N-CO-	Lys-Phe	82
RP	Pepstatin	Iva-Val-Val-Leu-CH(OH)-CH_2-CO-NH-	Ala-Sta	15, 16, 88, 91
PP	Pepstatin fragment	Iva-Val-Val-Leu-CH(OH)-CH_2-CO-NH-	OET	52, 53
PP	Lysyl pepstatin fragment	Iva-Val-Val-Lys-CH(OH)-CH_2-CO-NH-	OET	44, 51, 52
RP	Statine peptide 1	Iva-His-Pro-Phe-His-Leu-CH(OH)-CH_2-CO-NH-	Leu-Phe-NH_2	91
RP	Statine peptide 2 (Cyclic disulfide)	Ibu-His-Pro-Phe-Cys-Leu-CH(OH)-CH_2-CO-NH-	Leu-Phe-Thioamide	91
EP	H-261	Boc-His-Pro-Phe-His-Leu-CH(OH)-CH_2-CO-NH-	Val-Ile-His	9
EP	L-363,564	Boc-His-Pro-Phe-His-Leu-CH-CH_2-CO-NH-	Leu-Phe-NH_2	9, 26, 27

[a] EP = endothiapepsin; PP = penicillopepsin; RP = rhizopuspepsin; Iva = isovaleryl; OET = the ethylester of statine heptanoic acid; Ibu = isobuteryl; Boc = N-tertiary butoxycarbonyl.

[b] The middle statine residue is written as Leu-CH-CH_2-CO-NH to show the insertion of two additional atoms in the peptide backbone, producing a "frame shift."
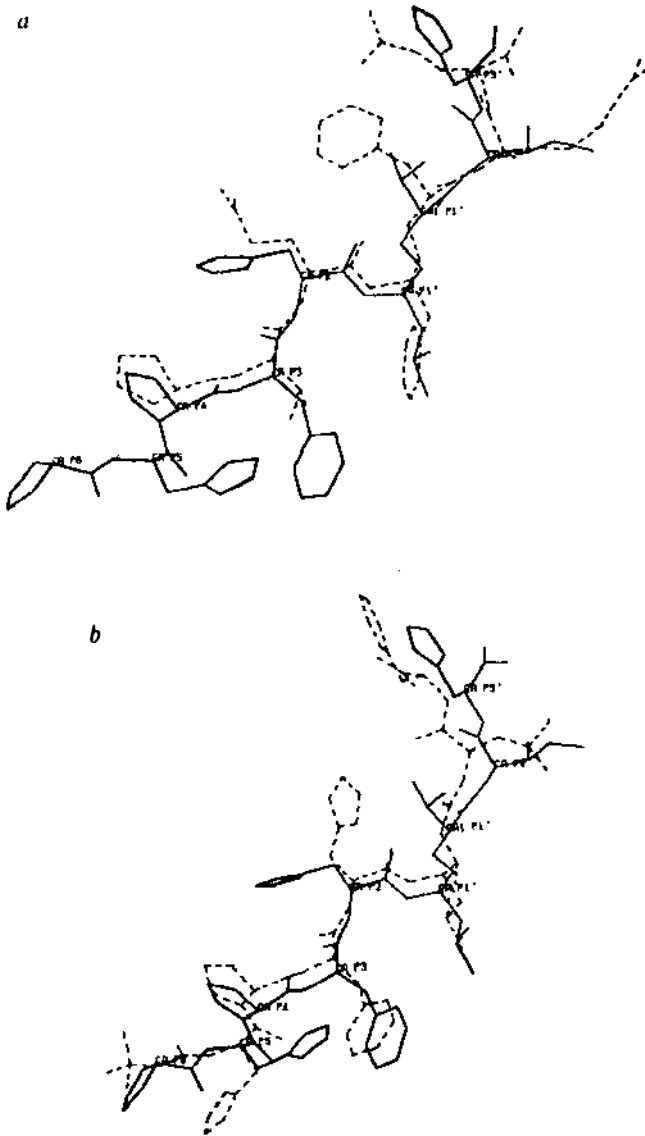
[c] Indicates position of scissile bond.

Figure 6   The conformations of three inhibitors of endothiapepsin shown superposed. (*a*) H-142 (*solid lines*) and H-256 (*broken lines*); (*b*) H-142 (*solid lines*) and L-363,564 (*broken lines*). Reproduced from Foundling et al (26) with permission.

replacing the scissile peptide unit and a nitrogen atom in place of the P_1' Cα atom. In this complex, which crystallizes in a form different from that of the uncomplexed endothiapepsin, both side chain positions are occupied equally by the His in P_2. The authors attribute this phenomenon to the large size and shallowness of the S_2 pocket in this enzyme, which extends to the S_1' pocket and permits binding promiscuity. Also the P_2-P_1 peptide bond twists from its normal planar conformation by an omega rotation of 12 degrees.

THE BACKBONE BINDING   Unlike the variability observed in the positions of the side chains in different inhibitors, the hydrogen bonds to the backbone remain the same probably for all the inhibitors on the N-terminal side of the scissile bond, and on both sides of it for inhibitors of a given class. Figure 7 illustrates the hydrogen bonding scheme for the two classes of inhibitor (pepstatin and isostere) bound to endothiapepsin. The same hydrogen bonds are also observed for these two classes of inhibitor in rhizopuspepsin (91).

In all of these inhibitor complexes, the water molecule observed between the two active site carboxylates in the native enzyme is displaced. In the inhibitors containing statine in P_1, the hydroxyl group of the statine, which is located within 0.5 Å of the water position, effectively replaces the water molecule. Similarly, in the (S)-hydroxyethylene inhibitor CP-69,799, the hydroxyl group lies symmetrically between the two catalytic aspartates (82). The methylene group also is close enough to displace the water molecule in the reduced inhibitors. Inhibitors displace many other water molecules within the cleft of the native enzyme. These molecules include the water molecule (92) assigned a nucleophilic role in the mechanism of action (#507 in rhizopuspepsin) proposed by James & Sielecki (49).

CONFORMATIONAL CHANGES   Some structural changes accompany inhibitor binding. Both Bott et al (16), using pepstatin binding to rhizopuspepsin, and James et al (54), using a pepstatin fragment bound to penicillopepsin, observed that one region of the enzymes changed when the inhibitor was bound. This region is the flap, residues 72–81, a beta hairpin bend that folds over the inhibitor in the active site and contains an invariant Tyr-75 and almost invariant Gly-76. Foundling et al (26) also observed a "slight repositioning of the residues in the flap" in endothiapepsin. The position of the flap is probably the same in all three enzymes when complexed with inhibitor. The observed differences in movement probably reflect differences in the original positions of the flap caused by interaction with neighboring molecules in the crystal. Suguna et al (92) examined the flap changes in a highly refined analysis of a reduced inhibitor bound to rhizopuspepsin. They observed a striking decrease in the B values of the
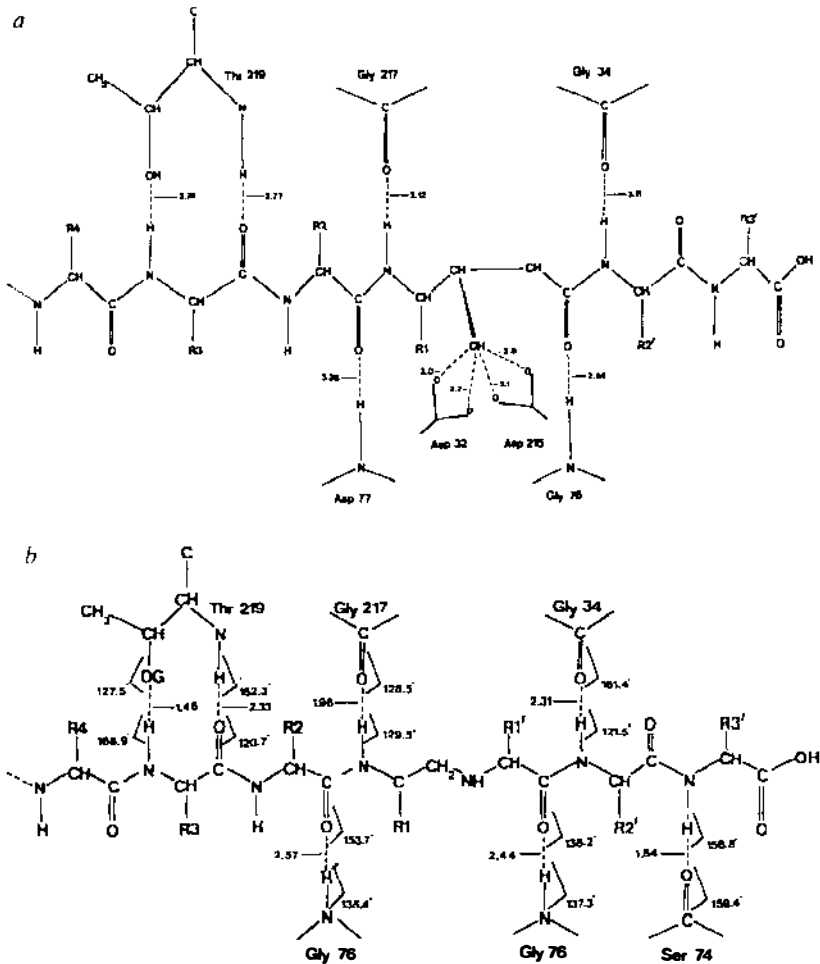
Figure 7 A schematic representation of the hydrogen bonding between endothiapepsin and (*a*) L-363,564 and (*b*) H-142. In (*a*) the distances of the hydrogen bonds are listed. In (*b*) the distance between the proton and the oxygen is given in possible hydrogen bonds. Reproduced from Foundling et al (26) with permission.

flap, which is quite mobile in the native structure, indicating that the flap stabilizes in the complex. This stabilization must partly result from hydrogen bond formation between the glycine and aspartate residues at the tip of the flap and the backbone atoms of the inhibitor, aided by van der Waals contacts (Figure 7). Experiments on endothiapepsin (18) and penicillopepsin (52) yielded similar observations.

The flap's closing over the inhibitor-substrate serves to remove the P_1-P_1' residues from effective contact with solvent. Sali et al (82) computed the loss of solvent accessibility for the inhibitor residues in the CP-69,799 complex and have observed almost complete exclusion of solvent for the middle pair.

Other changes accompany inhibitor binding. In rhizopuspepsin, the side chain of Asn-119 (Phe-117 in pepsin) rotates about the carbon–alpha-carbon–beta-carbon bond. This change could result from the displacement of a water molecule from the cleft by the inhibitor, causing a rearrangement of the side chain that enables the side chain to adopt different hydrogen bonding interactions (92).

The rest of the enzyme has the same three-dimensional structure after inhibitor binding, within the limits of the X-ray analyses. The active site carboxyl groups of these enzymes show no indication of movement despite the displacement of the centrally located water molecule.

In the (S)-hydroxyethyl inhibitor, CP-69,799, binding to endothiapepsin, however, a rigid body rotation of the C-terminal domain, residues 190–303, of 4.1 degrees with a translation of 0.3 Å (82) has been observed. Whether this change in structure results from the inhibitor binding or from the crystal packing forces is not clear in the altered crystal form. No such change has been observed in other bound inhibitors [e.g. there was no indication of such a change in rhizopuspepsin when bound to the reduced inhibitor (92)]. Nevertheless, the fact that twisting of one domain relative to the other does occur opens up many new possibilities in understanding the effect of substrate length on the catalytic rate of these enzymes and the significance of strain on catalysis (29, 70).

The Mechanism of Action

THE CATALYTIC ASPARTIC ACIDS   The mechanism of action of the aspartic proteinases has gradually become the subject of many reviews and discussions (17, 28, 29, 39, 49, 51, 70, 71, 76, 92). Several studies unsuccessfully attempted to demonstrate the existence of covalent intermediates in the reaction (22, 29, 41). Isotope exchange experiments with ^{18}O water (5, 6) also provided strong support for the lack of covalent intermediate formation in transpeptidation reactions. Most now believe the reaction probably proceeds by general base catalysis in which the carboxylate complex activates a water molecule. The effect of strain energy in twisting and peptide bond from its planar conformation has also been discussed (29, 70).

Figure 8 illustrates the conformation of the two carboxyl groups of the catalytic aspartates in three fungal enzymes. The agreement between these three structures is good, suggesting that they represent similar arrange-
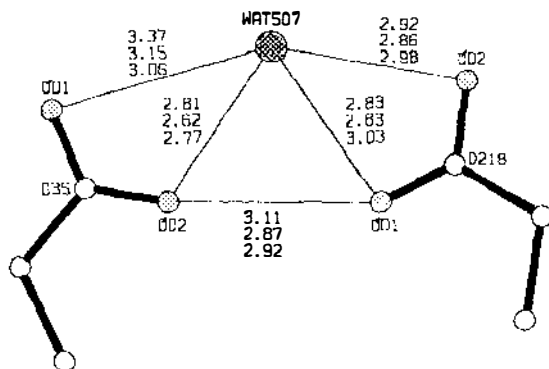
Figure 8 The two carboxyl groups of Asp-35(32) and Asp-218(215) with their associated water molecule. The three sets of distances come from rhizopuspepsin (*top*) (90), penicillopepsin (48), and endothiapepsin (72), respectively. The two oxygens, OD2 of Asp-35(32) and OD1 of Asp-218(215), are nearer to the interior of the molecule and make additional hydrogen bonds with the NH's of Gly-37(34) and Gly-220(217), respectively. The other two oxygens point into the cleft and, in addition to making hydrogen bonds to Ser-38(35) and Thr-221(218), each make hydrogen bond contacts with a water molecule, as in Figure 4. The amino acid numbering is for rhizopuspepsin; the number in parentheses is the corresponding number for pepsin.

ments of hydrogen bonds. Since both penicillopepsin (49) and endothiapepsin (72) crystallize from ammonium sulfate solution at about pH 4, it was suggested that the central peak between the two carboxyl groups might be an ammonium ion or a hydronium ion rather than a water molecule. However, the existence of a similar peak in the same position in rhizopuspepsin, which is crystallized at pH 6 in the absence of ammonium ion, suggests that the peak probably represents a water molecule (92).

In addition to the hydrogen bonds made to the central water molecule, each of these four oxygens of the two catalytic aspartates also takes part in hydrogen bonds to the surrounding protein and, in the case of the two external oxygens, to two water molecules (Figure 3). The state of protonation of the two carboxyl groups is probably such that they share one proton and one negative charge between them. Proton location cannot be determined directly by X-ray diffraction at the resolution of these structures, and where the protons associated with the two carboxyls and the water should be located remains controversial. Most authors have assigned the single carboxyl proton to one of the two interior proximal oxygens (49, 72), although an alternative arrangement has been based on the observed locations of the carboxyl groups that have no proton linking these two oxygen atoms (92).

THE EXTENDED SUBSITE SPECIFICITY   A characteristic feature of peptide hydrolysis by the aspartic proteinases is the extended subsite specificity on either side of the scissile bond (43). The observed active site cleft in the three-dimensional structures is long enough to accommodate at least a heptapeptide substrate. Suguna et al (92) have reported interpretable electron density for six amino acids of an octapeptide inhibitor in rhizopuspepsin, and Foundling et al (9, 26, 28) fit nine residues of the decapeptide H-142 within the cleft.

Fruton and his colleagues demonstrated that the elongation of peptide substrates on either side of the peptide bond would increase k_{cat} by three orders of magnitude, without a significant change in K_m, upon addition of four to six amino acids (28, 29). Hofmann et al (40) have taken advantage of the fungal proteinases' specific binding site S_1 for lysine (44) and designed more soluble substrates that could be used to investigate the specificity of penicillopepsin for long peptides. They synthesized 12 peptides of the form Ac-Ala$_m$-Lys-(NO$_2$)Phe-Ala$_n$-amide, where (NO$_2$)Phe is nitrophenylalanine and m and n each vary between 0 and 3. The K_m values were the same at pH 6.0 for all 12 peptides examined, while k_{cat} increased with increasing chain length. They observed small increases for positions P_2, P'_3, P_4, and P'_4 but obtained large increases averaging 37-fold when they added alanine residues in P_3 and P'_2. The X-ray data seem to rule out the possibility of any major conformational change in the inhibited enzyme upon addition of these residues [although Sali et al obtained different results (82)], yet changes in the electronic state of the catalytically active residues as a result of substrate binding could occur, since these changes would not be detectable by X-ray diffraction. Binding in the subsites could facilitate the distortion of the scissile peptide bond that Fruton (29) and others (70, 71, 92) suggested. Hofmann et al (40) propose that specific interactions, such as hydrogen bonds in subsites S'_2 and S_3, assist in the formation of the productive enzyme-substrate complex. They note that the backbone NH of P_3 forms a hydrogen bond with the hydroxyl group of Thr-219(215), which is a highly conserved residue, and that in P'_2 the backbone oxygen carbonyl forms a hydrogen bond with the terminal amino group of Trp-194(190), an invariant residue (92). These hydrogen bonds could stabilize the distortion of the peptide bond and the orientation of the substrate in its productive mode prior to the formation of the tetrahedral intermediate. However, no such corresponding bond to P'_2 has been reported for inhibitors binding to endothiapepsin (26, 82).

SUBSTRATE MODELING AND THE MECHANISM OF ACTION   The analysis of these inhibitors bound to the fungal proteinases provides a basis for the modeling of substrate, which in turn can lead to an evaluation of possible

mechanisms of action in three-dimensional terms. Recent detailed discussions of the mechanism have come from crystallographers engaged in high resolution structure analyses of these complexes. These evolved from the many previous mechanistic studies (reviewed above). The proposals discussed here come from three research groups engaged in analyzing the structures of penicillopepsin (49), endothiapepsin (70, 72), and rhizopuspepsin (92). Although earlier proposals have also been based on the crystal structures, these have been superceded as the structures became better refined and as more inhibitors were studied.

Most researchers agree that the mechanism of peptide bond hydrolysis by these proteinases proceeds by general base catalysis (39, 57). James & Sielecki (49) based their detailed stereochemical analysis of a possible mechanism on the refined structures of penicillopepsin and of the enzyme complexed to a pepstatin fragment, Iva-Val-Val-Sta-OEt. They modeled a tetrapeptide substrate bound to the enzyme, making use of the observed positions of the backbone and side chains for P_2 and P_1 of the inhibitor. They placed the carbonyl oxygen of the scissile bond in the same location as observed in the pepstatin fragment, i.e. approximately in the location of the central water. Finally, they assumed positions for P'_1 and P'_2 of the substrate, because they could not observe the positions directly from the statine-containing inhibitor upon which the modeling was based. The carbonyl group was then well disposed for attack by the water molecule 284, which in the native structure is hydrogen bonded to the external oxygen of Asp-33(35). The first step in the proposed mechanism is the protonation of the carbonyl oxygen of the substrate by the proton shared by the two catalytic aspartic acids. The nucleophilic attack on the carbonyl carbon by water-284 with transfer of a proton to Asp-32(35) follows, resulting in the formation of the tetrahedral intermediate. Products form through the protonation of the nitrogen atom, either from a solvent molecule or from the catalytic carboxyls.

Pearl & Blundell (72) discussed more active roles for the central "water" molecule, possibly a nucleophilic attack on the carbonyl carbon of the scissile bond if it actually was a water molecule. Alternatively, if the central peak represented an oxonium ion, they speculated that it might contribute a proton to the amide nitrogen or to the peptide's carbonyl group during hydrolysis.

Fruton (29) has discussed the role of the enzyme substrate interaction in contributing to the energy required to attain the transition state. He states that the estimated K_s or K_m values possibly correspond to a fraction of the total binding energy in the productive enzyme substrate interaction and that the remainder fuels the attainment of the transition state in the bond-breaking step (29, 55). Changes in the conformation of both the

enzyme and substrate and the resulting strain at the scissile peptide bond at the catalytic site can bring on the transition state. Pearl (70), who also discussed the cleavage of the peptide bond in terms of strain, argues that the extended subsite specificity of these enzymes produces strain energy that helps to twist the peptide bond out of planarity.

Suguna et al (92) questioned the previous proton arrangements about the two carboxyl groups. They note that many hydrogen bonds connect external groups to the carboxyl oxygens and these would hold the two carboxyls in their observed positions, even in the absence of any direct hydrogen bond between them. The stereochemistry is not the same as that of maleic acid and so the shared proton positions could be somewhere other than between the two proximal oxygens of the two carboxyl groups. They place the central water molecule so that it makes one hydrogen bond with one external oxygen atom and another bond with an internal oxygen atom of the other carboxyl group (Figure 8). Steric hindrance will then prevent the two proximal oxygens from sharing a proton; instead a proton is assigned to the external oxygen that is not hydrogen bonded to the central water molecule. Figure 9 shows the proposed mechanism of peptide bond hydrolysis, with the substrate modeling based on the observed binding of a reduced methylene isostere inhibitor to rhizopuspepsin (92). If the substrate binds with side chain interactions and backbone hydrogen bonds, like those of the inhibitor, then, as observed in the reduced peptide bond of the inhibitor, the scissile peptide bond will twist out of planarity. This twisting reduces the double-bond character of the peptide bond, causing the nitrogen atom to adopt the more pyramidal arrangement of a secondary amine and making it more likely to accept a proton. Then the mechanism proceeds: the central water molecule, which is made more nucleophilic by the hydrogen bonds from it to the carboxyl oxygens, attacks the peptide's carbonyl group. The nitrogen atom can be either simultaneously or subsequently protonated, possibly by the proton from Asp-215, followed by charge rearrangement, peptide bond cleavage, and departure of the products, not necessarily simultaneously. This mechanism differs from that of James & Sielecki (49) mainly in that the carbonyl oxygen of the scissile peptide bond is not placed in the middle of the four carboxyl oxygens and that a different water molecule carries out the nucleophilic attack. Other differences involve the locations of the protons, which cannot be determined by X-ray diffraction.

These mechanisms resemble that proposed for thermolysin (35, 64), a zinc endopeptidase. Here also, the mechanism also involves general base catalysis, the initial step consisting of nucleophilic attack on the carbonyl carbon by a water molecule or hydroxide ion with transfer of a proton to the carboxyl group of Glu-143. In thermolysin, however, the zinc ion
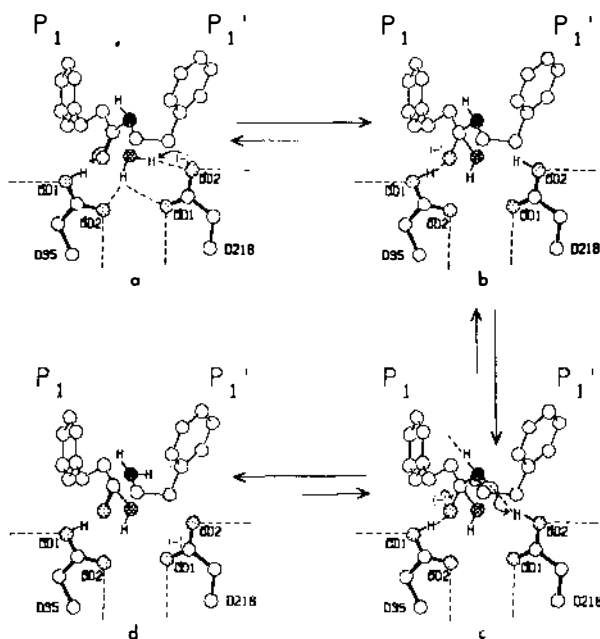
Figure 9 A proposed catalytic mechanism for rhizopuspepsin (90). The two aspartic acid side chains D35 and D218 correspond to D32 and D215 of pepsin. (*a*) The substrate, represented by a Phe-Phe dipeptide, with a peptide group twisted out of planarity, is attacked by the central water molecule with transfer of a proton to Asp-218. (*b*), (*c*) This proton is transferred to the amide nitrogen. (*d*) Rearrangement results in cleavage of the peptide bond followed by departure of the products.

helps in the polarization of the carbonyl group of the peptide bond. The oxyanion hole effects similar polarization in the serine proteinases with two strong amide NH hydrogen bonds to the carbonyl oxygen (80). No equivalent positively charged interaction with the carbonyl oxygen has been observed in the aspartyl proteinases, although a weak hydrogen bond from the carboxyl oxygen of Asp-35 (32 in pepsin) to the carbonyl oxygen has been proposed (92). Blundell et al (9) proposed an alternative polarizing effect through interaction with the ring system of a tyrosine residue on the flap. This Tyr-75 is invariant in the pepsins but absent in the retroviral proteinases (25, 69, 99, 103).

THE RETROVIRAL PROTEINASES

In illuminating illustrations of evolutionary development, the newly published structures of two retroviral proteinases (65, 67, 69, 105) have been

shown to be closely related to the structures of the aspartic proteinases of the pepsin family. Whereas the pepsins are single chain, two-domain structures with the two domains related by approximate two-fold symmetry, the retroviral proteinases are homodimers with two identical subunits related by an exact two-fold axis, and could be regarded as precursors (fossils?) of the pepsin proteinases. Tang et al (97) hypothesized that such homodimers might exist, based on the observed structures of the pepsins.

Many retroviral proteins are initially synthesized as polyproteins that have to be cleaved during the maturation process of the virus (21). A retroviral encoded proteinase, which contains the sequence Asp-Thr-Gly, suggesting that it might be a member of the aspartic proteinase family (100), carries out this cleavage. A more detailed alignment of the retroviral proteinase sequences with the sequences of the two domains of the pepsin family of proteinases led to a model of the dimeric active enzyme (73).

The crystal structures of the proteinases from Rouse sarcoma virus (RSV) (67) and from the human immunodeficiency virus HIV-1 (69, 105) have been reported. The molecules are dimeric and contain a fold similar to that of the other aspartic proteinases (Figures 10 and 11). The active
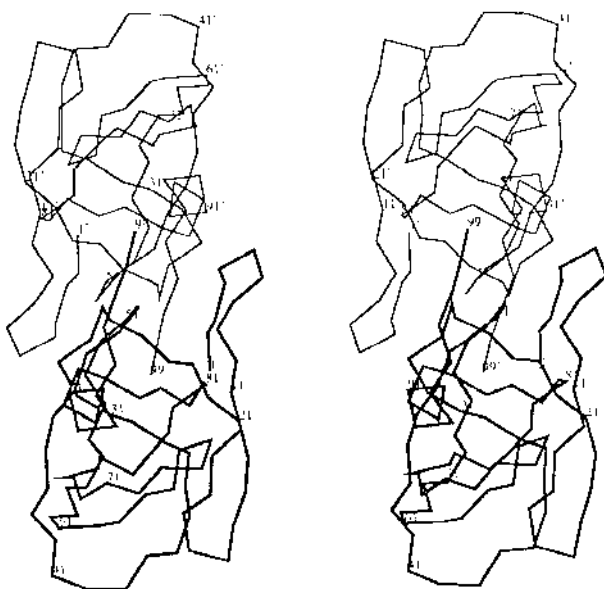

Figure 10   The carbonalpha backbone of the HIV-1 proteinase dimer as viewed looking down the dyad axis. Reproduced from Wlodawer et al (105) with permission. The catalytic Asp-25 residues are part of the broad loops that form part of the interface near the central dyas axis.
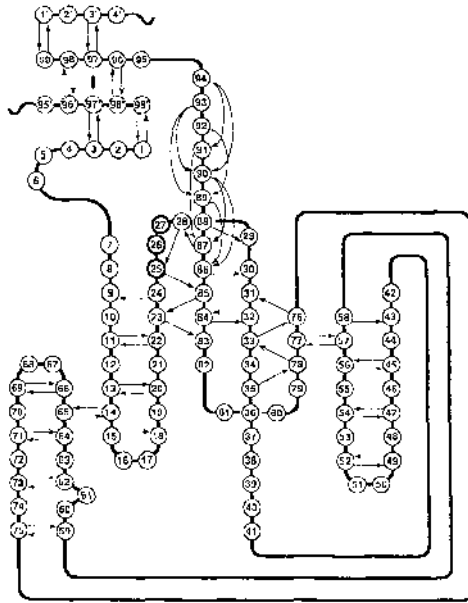
Figure 11   A schematic version of one subunit of the HIV-1 proteinase structure. Compare this with the *N*-terminal domain of rhizopuspepsin in Figure 2. The catalytic residues Asp-25-Gly-27 are in the loop that forms part of the ψ structure. The loop containing residues 44–57 forms the flap. Reproduced from Wlodawer et al (105) with permission.

site regions also closely resemble those of the pepsins; the two aspartic acids are situated at the bottom of an extended cleft that runs the width of the molecule between the two subunits. Each subunit contributes an extended flap to the active site, although for the RSV structure, eight amino acids in this region cannot be traced and may be disordered in the crystal. Presumably, interactions with neighboring molecules, as in penicillopepsin, stabilize the flap in the HIV-1 crystals (47, 48).

The structure initially reported for the HIV-1 proteinase differed from the RSV proteinase in the *N*- and *C*-terminal regions, mainly on the side of the molecule opposite from the active site. Subsequent refinement of the original HIV-1 structure (M. A. Navia, personal communication), together with the X-ray analysis of a chemically synthesized HIV-1 proteinase (84, 105), have eliminated these differences and established the similarity of these two retroviral enzymes.

The HIV-1 proteinase has 99 amino acid residues vs 124 in the RSV enzyme. The additional residues in the RSV structure can be located in three surface turns between β-strands. Similar deletions occur in these regions when the RSV structure is compared with rhizopuspepsin (103).

Both domains of the rhizopuspepsin have been compared with the HIV proteinase (105); for the N-terminal domain, 57 $C\alpha$ could be superimposed with an rms difference of 1.4 Å, while 56 alpha carbon pairs in the C-terminal domain superposed within the same limits. For 88 atom pairs in the active site region, the superposition yielded an rms difference of 0.59 Å.

Both the RSV and HIV-1 proteinase have dimeric active forms, but they are initially expressed as isolated single domains within a polyprotein from which they have to be cleaved. The mechanism of this cleavage and activation is believed to be autocatalytic (20, 24, 30), but the details of this process are not known.

One unusual feature of these enzyme structures is the two-fold symmetry that relates the two subunits. For the HIV-1 protease the packing in the crystal dictates this symmetry, but even in the RSV proteinase, where the two subunits make up the asymmetric unit of the crystal, almost complete two-fold symmetry exists. Since the substrate is asymmetric, complexing it to the enzyme might introduce asymmetry through conformational change in the enzyme. Crystal structure investigations of inhibitor enzyme complexes will help determine what happens to the two subunits when inhibitors bind. For example, do both flaps close down over the inhibitor (substrate), in which case the properties of the complex might differ a little from the pepsins, or does only one flap close, with the direction of the polypeptide enzyme inhibitor (substrate) determining the choice of the flap?

For effective catalysis, the retroviral proteinases seem to require a peptide of at least seven residues (19). These enzymes show unusual specificity in their ability to cleave a peptide bond of polyprotein substrates containing proline in P'_1 (73). They are active against oligopeptides containing the sequence Tyr-Pro (19, 62), although they also show some activity against peptides with other amino acid residues in P_1 and in P'_1; the sequence Met-Mct appears particularly effective (19).

The expertise acquired in many laboratories designing inhibitors for renin should greatly facilitate the rational design of inhibitors for the HIV-1 proteinase. In addition to inhibitors that act directly on the active site, the dimeric nature of the HIV-1 proteinase makes the design of a different class of inhibitor that blocks the formation of the dimer possible.

Literature Cited

1. Andreeva, N. S., Fedorov, A. A., Gustchina, A. E., Riskulov, R. R., Schutzkever, N. E., Safro, M. G. 1978. *Mol. Biol. (USSR)* 12: 704–16
2. Andreeva, N. S., Gustchina, A. E., Fedorov, A. A., Schutzkever, N. E., Volnova, T. V. 1977. See Ref. 96, pp. 23–32
3. Andreeva, N. S., Zdanov, A. S., Gustchina, A. E., Fedorov, A. A. 1984. *J. Biol. Chem.* 259: 11353–65
4. Andreeva, N. S., Zdanov, A. S., Gustchina, A. E., Fedorov, A. A. 1985. See Ref. 61, pp. 137–50
5. Antonov, V. K. 1985. See Ref. 61, pp. 203–20
6. Antonov, V. K., Ginodman, L. M., Rumsh, L. D., Kapitannikov, Yu. V., Barshevskaya, T. N., et al. 1981. *Eur. J. Biochem.* 117: 195–200
7. Bernal, J. D., Crowfoot, D. 1934. *Nature* 133: 794–95
8. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr., Brice, M. D., et al. 1977. *J. Mol. Biol.* 112: 535
9. Blundell, T. L., Cooper, J., Foundling, S. I., Jones, D. M., Atrash, B., Szelke, M. 1987. *Biochemistry* 26: 5585–90
10. Blundell, T. L., Jenkins, J., Pearl, L., Sewell, B. T. 1985. See Ref. 61, pp. 151–61
11. Blundell, T. L., Sewell, B. T., McLachlan, A. D. 1979. *Biochim. Biophys. Acta* 580: 24–31
12. Blundell, T. L., Sibanda, B. L., Hemmings, A., Foundling, S. F., Tickle, I. J., et al. 1986. In *Molecular Graphics and Drug Design*, ed. A. S. V. Burgen, G. C. K. Roberts, M. S. Tute, pp. 323–34. Elsevier
13. Boger, J. 1985. See Ref. 61, p. 401
14. Boger, J., Lohr, N. S., Ulm, E. H., Po, E. M., Blaine, E. H., et al. 1983. *Nature* 303: 81–84
15. Bott, R., Davies, D. R. 1983. *Proceedings of the 8th American Peptide Symposium*, ed. V. J. Hruby, D. H. Rich, 8: 531–40. Rockford, IL: Pierce Chemical Co
16. Bott, R., Subramanian, E., Davies, D. R. 1982. *Biochemistry* 21: 6956–62
17. Clement, G. E. 1973. *Prog. Bioorg. Chem.* 2: 177–238
18. Cooper, J., Foundling, S., Hemmings, A., Blundell, T. 1987. *Eur. J. Biochem.* 169: 215–21
19. Darke, P. L., Nutt, R. F., Brady, S. F., Garsky, V. M., Ciccarone, T. M., et al. 1988. *Biochem. Biophys. Res. Commun.* 156: 297–303
20. Debouck, C., Gorniak, J. G., Strickler, J. E., Meek, T. D., Metcalf, B. W., Rosenberg, M., 1987. *Proc. Natl. Acad. Sci. USA* 84: 8903
21. Dickson, C., Eisenman, R., Fan, H., Hunter, E., Teich, N. 1984. *RNA Tumor Viruses*, ed. R. Weiss, N. Teich, H. Varmus, J. Coffin, pp. 513–648. New York: Cold Spring Harbor Lab. 2nd ed.
22. Dunn, B. M., Fink, A. L. 1984. *Biochemistry* 23: 5241–47
23. Dunn, B. M., Jimenez M., Parten, B. F., Valler, M. J., Rolph, C. E., Kay, J. 1986. *Biochem. J.* 237: 899–906
24. Farmerie, W. G., Loeb, D. D., Casavant, N. C., Hutchinson, C. A. III, Edgell, M. H., Swanstrom, R. 1987. *Science* 236: 305
25. Foltmann, B. 1988. *Abstr. 18th Linderstrom-Lang Conf.*
26. Foundling, S. I., Cooper, J., Watson, F. E., Cleasby, A., Pearl, L. H., et al. 1987. *Nature* 327: 349–52
27. Foundling, S. I., Cooper, J., Watson, F. E., Pearl, L. H., Hemmings, A., et al. 1987. *J. Cardiovasc. Pharmacol.* 10: S59–S68
28. Fruton, J. S. 1970. *Adv. Enzymol.* 33: 401–43
29. Fruton, J. S. 1976. *Adv. Enzymol.* 44: 1–36
30. Giam, C.-Z., Boros, I. 1988. *J. Biol. Chem.* 263: 14617
31. Gilliland, G. L., Winborne, E. L., Nachman, J., Wlodawer, A. 1988. See Ref. 25
32. Haber, E. 1984. *Hypertension* 2: 223
33. Haber, E., Burton, J. 1979. *Fed. Proc.* 38: 2768
34. Hallett, A., Jones, D. M., Atrash, B., Szelke, M., Leckie, B. J., et al. 1985. See Ref. 61, pp. 467–78
35. Hangauer, D. G., Monzingo, A. F., Matthews, B. W. 1984. *Biochemistry* 23: 5730–41

36. Hartsuck, J., Remington, J. 1988. See Ref. 25
37. Hemmings, A. M., Foundling, S. I., Sibanda, B. L., Wood, S. P., Pearl, L. H., Blundell, T. L. 1985. *Biochem. Soc. Trans.* 13: 1036–41
38. Herriott, R. M. 1939. *J. Gen. Physiol.* 22: 65–78
39. Hofmann, T. 1989. *Zoological Research.* Beijing (In Chinese)
40. Hofmann, T., Allen, B., Bendiner, M., Blum, M., Cunningham, A. 1988. *Biochemistry* 27: 1140–46
41. Hofmann, T., Dunn, B. M., Fink, A. L. 1984. *Biochemistry* 23: 5247–56
42. Deleted in proof
43. Hofmann, T., Hodges, R. S. 1982. *Biochem. J.* 203: 603 10
44. Hofmann, T., Hodges, R. S., James, M. N. G. 1984. *Biochemistry* 23: 635–43
45. Holladay, M. W., Salituro, F. G., Schmidt, P. G., Rich, D. H. 1985. *Biochem. Soc. Trans.* 13: 1046–48
46. Hsu, I.-N., Delbaere, L. T. J., James, M. N. G. 1977. *Nature* 266: 140–45
47. Hsu, I.-N., Delbaere, L. T. J., James, M. N. G., Hofmann, T. 1977. See Ref. 96, pp. 61–81
48. James, M. N. G., Sielecki, A. R. 1983. *J. Mol. Biol.* 163: 299–361
49. James, M. N. G., Sielecki, A. R. 1985. *Biochemistry* 24: 3701–13
50. James, M. N. G., Sielecki, A. R. 1986. *Nature* 319: 33–38
51. James, M. N. G., Sielecki, A. R. 1987. *Biological Macromolecule and Assemblies,* ed. J. Jurnak, A. McPherson, 3: 413–82. New York: Wiley
52. James, M. N. G., Sielecki, A. R., Hofmann, T. 1985. See Ref. 61, pp. 163–77
53. James, M. N. G., Sielecki, A. R., Moult, J. 1983. In *Peptides: Structure and Function,* ed. V. J. Hruby, D. H. Rich, 1: 521–30. Rockford, IL: Pierce Chemical Co
54. James, M. N. G., Sielecki, A. R., Salituro, F., Rich, D. H., Hofmann, T. 1982. *Proc. Natl. Acad. Sci. USA* 79: 6137–41
55. Jencks, W. P. 1969. *Catalysis in Chemistry and Enzymology.* New York: McGraw-Hill
56. Jenkins, J., Tickle, I., Sewell, B. T., Ungaretti, L., Wollmer, A., Blundell, T. 1977. See Ref. 96, pp. 43–60
57. Kay, J. 1985. See Ref. 61, p. 1
58. Knowles, J. R. 1970. *Philos. Trans. R. Soc. London Ser. B* 257: 135–46
59. Kokubu, T., Ueda, E., Fujimoto, S., Hiwada, K., Kato, A., et al. 1968. *Nature* 217: 456
60. Komiya, A. 1964. *Soybean Dig.* 24: 43

61. Kostka, V., ed. 1985. *Aspartic Proteinases and Their Inhibitors.* Berlin/New York: de Gruyter
62. Kotler, M., Katz, R. A., Waleed, D., Leis, J., Skalka, A. M. 1988. *Proc. Natl. Acad. Sci. USA* 85: 4185–89
63. Marciniszyn, J., Hartsuck, J. A., Tang, J. 1976. *J. Biol. Chem.* 251: 7088–94
64. Matthews, B. W. 1988. *Acc. Chem. Res.* 21: 333–40
65. McKeever, B. M., Navia, M. A., Fitzgerald, P. M. D., Springer, J. P., Leu, C.-T., et al. 1989. *J. Biol. Chem.* 264: 1919–21
66. McPhie, P. 1972. *J. Biol. Chem.* 247: 4277–81
67. Miller, M., Jaskolski, M., Mohana, R., Leis, J., Wlodawer, A. 1989. *Nature* 337: 576–79
68. Moravek, L., Kostka, V. 1974. *FEBS Lett.* 43: 207–12
69. Navia, M. A., Fitzgerald, P. M. D., McKeever, B. M., Leu, C.-T., Heimbach, J. C., et al. 1989. *Nature* 337: 615–20
70. Pearl, L. H. 1985. See Ref. 61, pp. 189–95
71. Pearl, L. H. 1987. *FEBS Lett.* 214: 8–12
72. Pearl, L. H., Blundell, T. 1984. *FEBS Lett.* 174: 96–101
73. Pearl, L. H., Taylor, W. R. 1987. *Nature* 329: 351–54
74. Pearl, L. H., Taylor, W. R. 1987. *Nature* 328: 482
75. Perlmann, G. E. 1963. *J. Mol. Biol.* 6: 452–64
76. Polgar, L. 1987. *FEBS Lett.* 219: 1–4
77. Powers, J. C., Harley, A. D., Myers, D. V. 1977. See Ref. 96, pp. 141–57
78. Rajagopalan, T. G., Stein, W. H., Moore, S. 1966. *J. Biol. Chem.* 241: 4295–97
79. Rich, D. H. 1985. *J. Med. Chem.* 28: 263
80. Robertus, J. D., Kraut, J., Alden, R. A., Birktoft, J. J. 1972. *Biochemistry* 11: 4293–4303
81. Safro, M. G., Andreeva, N. S., Zdanov, A. S. 1985. See Ref. 61, pp. 183–87
82. Sali, A., Veerapandian, B., Cooper, J. B., Foundling, S. I., Hoover, D. J., Blundell, T. L. 1989. *EMBO J.* 8: 2179–88
83. Schechter, I., Berger, A. 1967. *Biochem. Biophys. Res. Commun.* 27: 157–62
84. Schneider, J., Kent, S. 1988. *Cell* 54: 363–68
85. Sielecki, A. R., Hayakawa, K., Fujinaga, M., Murphy, M. E. P., Fraser, M., Muir, A. K., et al. 1989. *Science* 243: 1346–51
86. Skeggs, L. T., Dorer, F. E., Levine, M.,

Lentz, K. E., Kahn, J. R. 1980. *The Renin-Angiotensin System*, ed. J. A. Johnson, R. R. Anderson, 1: 1–27. New York: Plenum

87. Skeggs, L. T., Lentz, K. E., Hochstrasser, H., Kahn, J. R. 1964. *Can. Med. Assoc. J.* 90: 185

88. Subramanian, E., Liu, M., Swan, I. D. A., Davies, D. R. 1977. See Ref. 96, pp. 33–60

89. Subramanian, E., Liu, M., Swan, I. D. A., Davies, D. R., Jenkins, J. A., et al. 1977. *Proc. Natl. Acad. Sci. USA* 74: 556–59

90. Suguna, K., Bott, R. R., Padlan, E. A., Subramanian, E., Sheriff, S., et al. 1987. *J. Mol. Biol.* 196: 877–900

91. Suguna, K., Padlan, E. A., Bott, R., Boger, J., Davies, D. R. 1989. *Biochemistry.* Submitted

92. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D., Davies, D. R. 1987. *Proc. Natl. Acad. Sci. USA* 84: 7009–13

93. Szelke, M. 1985. See Ref. 61, pp. 421–41

94. Szelke, M., Jones, D. M., Atrash, B., Hallett, A., Leckie, B. J. 1983. In *Peptides: Structure and Function*, ed. V. J. Hruby, D. H. Rich, 1: 579–82. Rockford, IL: Pierce Chemical Co

95. Szelke, M., Leckie, B. J., Hallett, A., Jones, D. M., Suerias-Diaz, J., et al. 1982. *Nature* 299: 555

96. Tang, J., ed. 1977. *The Acid Proteases: Structure, Function and Biology.* New York: Plenum

97. Tang, J., James, M. N. G., Hsu, I. N., Jenkins, J. A., Blundell, T. L. 1978. *Nature* 271: 618–21

98. Tang, J., Sepulveda, T., Marciniszyn, J., Chen, K. C. S., Huang, W. Y., et al. 1973. *Proc. Natl. Acad. Sci. USA* 70: 3437–39

99. Tang, J., Wong, R. N. S. 1987. *J. Cell. Biochem.* 33: 53–63

100. Toh, H., Ono, M., Saigo, K., Miyata, T. 1985. *Nature* 315: 691

101. Umezawa, H., Aoyagi, T., Moroshima, H., Matzusaku, M., Hamada, H., Takeuchi, T. 1970. *J. Antibiot.* 23: 259–62

102. Watson, F., Wood, S. P., Tickle, I. J., Shearer, A., Sibanda, B. L., et al. 1988. See Ref. 25

103. Weber, I. T., Miller, M., Jaskolski, M., Leis, J., Skalka, A. M., Wlodawer, A. 1989. *Science* 243: 928–31

104. Welinder, K. G., Mikkelsen, L., Foltmann, B. 1985. See Ref. 61, pp. 197–202

105. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., et al. 1989. *Science* 245: 616–21