Atomic resolution analysis of the catalytic site of an aspartic proteinase and an unexpected mode of binding by short peptides

PETER T. ERSKINE,¹ LEIGHTON COATES,¹ SANJAY MALL,¹ RAJ S. GILL,¹
STEVE P. WOOD,¹ DEAN A.A. MYLES,² AND JON. B. COOPER¹

¹Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Southampton SO16 7PX, UK

²European Molecular Biology Laboratory (EMBL) Grenoble Outstation, Institut Laue-Langevin, Grenoble F38042, France

(RECEIVED January 31, 2003; FINAL REVISION April 2, 2003; ACCEPTED April 29, 2003)

Abstract

The X-ray structures of native endothiapepsin and a complex with a hydroxyethylene transition state analog inhibitor (H261) have been determined at atomic resolution. Unrestrained refinement of the carboxyl groups of the enzyme by using the atomic resolution data indicates that both catalytic aspartates in the native enzyme share a single negative charge equally; that is, in the crystal, one half of the active sites have Asp 32 ionized and the other half have Asp 215 ionized. The electron density map of the native enzyme refined at 0.9 Å resolution demonstrates that there is a short peptide (probably Ser-Thr) bound noncovalently in the active site cleft. The N-terminal nitrogen of the dipeptide interacts with the aspartate diad of the enzyme by hydrogen bonds involving the carboxyl of Asp 215 and the catalytic water molecule. This is consistent with classical findings that the aspartic proteinases can be inhibited weakly by short peptides and that these enzymes can catalyze transpeptidation reactions. The dipeptide may originate from autolysis of the N-terminal Ser-Thr sequence of the enzyme during crystallization.

Keywords: Aspartic proteinase; atomic resolution; succinimide; inhibition; transpeptidation; protonation state; low-barrier hydrogen bonds; slow binding

A number of aspartic proteinases have received much attention in recent years due to their roles in various disease states. These include renin due to its role in the renin-angiotensin system and the HIV proteinase, which is critical for maturation of the retrovirus (Cooper 2002). The aspartic proteinase β -secretase is involved in Alzheimer's disease (Hong et al. 2000), and the enzymes produced by various parasitic organisms such as *Candida albicans* and *Plasmodium falciparum* play an important role in the pathology of infection by these organisms (Silva et al. 1996b; Pichová et al. 2001). The stomach enzyme pepsin is thought to play a

major role in peptic ulcer disease, and the lysosomal enzyme cathepsin D has been implicated in tumorigenesis (Fusek and Vetvicka 1994).

Most eukaryotic aspartic proteinases are monomeric and consist of a single-chain of ~330 amino acids, which forms two similar domains with the active site located between (Fig. 1; Tang et al. 1978). In contrast, retroviral aspartic proteinases are dimeric, consisting of two identical subunits, each roughly equivalent to one domain of a eukaryotic aspartic proteinase (Davies 1990). The amino acid sequences of the eukaryotic enzymes have signs of an internal repeat relating the two halves of the molecule, their identity being greatest in the vicinity of the active site that involves the two conserved Asp-Thr-Gly sequences. Hence, the eukaryotic aspartic proteinases are thought to have evolved divergently from a primitive dimeric enzyme (resembling the retroviral proteinase) by gene duplication and fusion. In all

Reprint requests to: Jon. B. Cooper, Division of Biochemistry and Molecular Biology, School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton, SO16 7PX, UK; e-mail: J.B.Cooper@soton.ac.uk; fax: 44 (0)-2380-594459.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.0305203>.

Figure 1. A ribbon diagram of the tertiary fold of endothiapepsin. In this view, the N-terminal domain is on the *left*, and the C-terminal domain is on the *right*. The active site cleft can be seen toward the center of the molecule and is occupied by the inhibitor H261 shown in ball-and-stick representation.

aspartic proteinases, the base of the active site cleft is made of β -strands that contain the catalytic aspartate residues (32 and 215 in porcine pepsin). The side-chain carboxyl groups of the aspartates are held coplanar and within hydrogen bonding distance by an intricate arrangement of hydrogen bonds. A solvent molecule is H-bonded between both aspartate carboxyls and is presumed to take part in the catalytic mechanism (Pearl and Blundell 1984). In numerous chemical studies, the failure to trap covalently bound substrate indicated that the reaction involves an intermediate that binds non-covalently to the enzyme (Hofmann et al. 1984). NMR studies using an inhibitor with a ketone analog of the scissile peptide bond (Rich et al. 1982) suggested that it binds to the enzyme in a hydrated *gem*-diol form ($>\text{C}[\text{OH}]_2$). Thus, current proposals for the catalytic mechanism invoke nucleophilic attack of the active site water on the scissile bond carbonyl generating a *gem*-diol intermediate (Suguna et al. 1987; James et al. 1992; Veerapandian et al. 1992; Silva et al. 1996a).

The best synthetic inhibitors are those in which one or both of the hydroxyl groups of the putative tetrahedral intermediate are mimicked by the transition state analog. X-ray analysis has shown that one hydroxyl binds in the same position as the active site water molecule in the native enzyme, and most of the transition state analogs (e.g., hydroxyethylene) possess the corresponding hydroxyl group. Here we report the atomic resolution structure of endothiapepsin bound to one such hydroxyethylene transition state analog (H261; Fig. 2) refined at 1.1 Å resolution and that of the native enzyme refined at 0.9 Å resolution. The latter structure established that both of the active site aspartates share a negative charge in the ground state, and that the enzyme unexpectedly has a dipeptide bound at the active site, which may account for earlier observations of weak inhibition by short peptides and the ability of these enzymes to catalyze transpeptidation.

Results and Discussion

X-ray analysis

The structure of endothiapepsin was originally solved by Blundell et al. (1990). The crystals reported herein were monoclinic $P2_1$ with unit cell dimensions of $a = 42.7$ Å, $b = 74.6$ Å, $c = 42.5$ Å, and $\beta = 97.4^\circ$. This crystal form of endothiapepsin differs from the original form by having a remarkably low solvent content (39%) and diffracts to atomic resolution at 100 K. The X-ray data obtained from the H261 complex extend to 1.1 Å resolution and have an R_{merge} of 12.5%, a completeness of 99.8%, and a mean multiplicity of 3.7. The data set obtained from the native enzyme crystal has been processed to 0.91 Å resolution, giving an R_{merge} of 10.9%, a completeness of 99.4%, and a mean multiplicity of 4.7. Realistically, these data extend to ~ 0.95 Å resolution because the intensities beyond this have a rather high R_{merge} (50%), but the data to 0.91 Å have been included in the current refinement, as the strong terms in the outermost resolution shell have better agreement statistics. For the inhibitor complex data set (1.1 Å), refining anisotropic U values lowered the R_{free} factor by $\sim 2\%$, and use of the “riding hydrogen” model lowered it by a further $\sim 1.5\%$. For the native data set (0.9 Å), anisotropic U value refinement lowered the R_{free} by $>3\%$, and inclusion of “riding hydrogens” lowered R_{free} by $>3\%$. The final R factor and R_{free} value for the 1.1 Å H-261 data set are 14.0% and 17.8%, respectively. For the 0.9 Å native data set, the $R_{\text{refinement}}$ value is 12.1% for the working set and 14.7% for the R_{free} set. Full data collection and refinement statistics are given in Table 1. For the native structure, the percentage of residues within the “most favored” regions of the Ramachandran plot, as assessed by the PROCHECK criteria (Laskowski et al. 1993), is 93.9%. The remaining 6.1% of residues are within the so-called additional allowed boundary. For the H261 complex, 94.6% of residues are within the “most favored” regions, and the remaining 5.4% of residues are within the “additional allowed” boundary. Both of the refined models have mean main-chain B factors close to 10 Å² indicating that the structures are well defined by the atomic resolution data sets. As observed in previous analyses at lower resolution (Bailey and Cooper 1994), the B factors for the active site flap (residues 71–82) are much

Figure 2. The chemical formula of the inhibitor H261. This inhibitor possesses a hydroxyethylene analog of the scissile peptide bond in a non-hydrolyzable group that mimics a Leu-Val dipeptide.

Table 1. X-ray and refinement statistics for both structures

	H261 complex	Native
Unit cell and symmetry		
Space group	$P2_1$	$P2_1$
a (Å)	42.58	42.72
b (Å)	74.51	74.66
c (Å)	42.35	42.55
β (°)	97.56	97.27
Entire data set		
Resolution range of data collected (Å)	15.6–1.1	42.3–0.91
R_{merge} (%)	12.5	10.9
Completeness (%)	99.8	99.4
Overall $\langle I/\sigma(I) \rangle$	9.8	8.6
Multiplicity	3.7	4.7
Outer shell		
Resolution range (Å)	1.14–1.1	0.96–0.91
R_{merge} (%)	32.0	50.3
Completeness (%)	99.5	99.4
$\langle I/\sigma(I) \rangle$	4.8	3.0
Multiplicity	3.5	3.4
Refinement (no $\sigma[F]$ cutoff, 5% test set)		
R factor (%)	14.0	12.1
R free factor (%)	17.8	14.7
Total number of reflections	105,807	186,749
Number of parameters refined	27,154	27,802
Number of restraints	33,073	34,005
RMSD of bond lengths (Å)	0.013	0.015
RMSD of angle distances (Å)	0.028	0.029
RMSD of deviation for bumps (Å)	0.074	0.121
RMSD of planar group chiral volumes (Å ³)	0.079	0.085
RMSD of chiral volumes (Å ³)	0.078	0.100

RMSD indicates root mean square deviation.

lower in the inhibitor complex than in the native enzyme, in which the flap is more disordered.

The quality of the electron density in both maps is very good as exemplified by the region shown in Figure 3. This shows a region of native endothiapepsin at 0.9 Å (contoured at 3.0 root mean square), including Asp 51 and Gly 52. In the original analyses of endothiapepsin and its inhibitor complexes, which were typically at 2.0 Å resolution, this region of the molecule had rather poor agreement with the electron density maps. A rationale for this effect is provided by the current atomic resolution structures, which show that these two residues have condensed to form a succinimide. Biochemical studies have shown that Asp-Gly dipeptides undergo slow succinimide formation when the main-chain nitrogen of the glycine nucleophilically attacks the side-chain carboxyl of the aspartate (see scheme in Fig. 3). Succinimide and β -aspartic acid formation in proteins is a characteristic of aging (Lindner and Helliger 2001). β -Aspartic acid is formed when the succinimide breaks down such that what was the side chain of the Asp becomes the main chain. These modifications are especially significant in proteins that have to function for the entire lifespan of the organism, for example, those in the eye lens, and it has been suggested

that the accumulation of these unusual residues may be a triggering factor in amyloidogenesis.

Active site interactions

The current analysis of the H261 complex confirms an earlier report on the bound structure of this inhibitor determined at 1.6 Å resolution by using X-ray data obtained at room temperature (Veerapandian et al. 1990). Electron density for the bound inhibitor (H261) at 1.1 Å resolution is shown in Figure 4 (contoured at 2.2 root mean square). The inhibitor is tightly bound by a set of hydrogen bonds involving its main-chain groups. These and other interactions made by this inhibitor and others have been reported in detail in earlier analyses (Bailey and Cooper 1994). The exposed histidine side-chains at P₅, P₂, and P₃' are less well defined, and that at P₂ was refined as an Ala due to paucity of the map for this side chain in the current analysis.

In the first electron density map for the native enzyme at 0.9 Å resolution, it was apparent that a small molecule was bound in the active site cleft. It appears to be a dipeptide with its N-terminal nitrogen bound to the outer carboxyl of Asp 215, presumably by a salt bridge. The two amino acids

Figure 3. (A) The electron density of the succinimide group formed by Asp 51 and Gly 52 in the structure of native endothiapepsin refined at 0.9 Å resolution. The map (shown as pale lines) has been σ_A -weighted (Read 1986) and is contoured at 3.0 root mean square. (B) The intramolecular cyclization reaction which leads to succinimide formation in proteins.

of the dipeptide lie in the same positions as the P_1' and P_2' residues of H261. The exact sequence of the dipeptide is not certain, but several efforts at refinement with different amino acids built in to the density lent faith to the conclusion that it is a Ser-Thr dipeptide (Fig. 5). The following dipeptides were synthesized to test whether any possessed

inhibitory activity against endothiapepsin: Ser-Thr, Pro-Thr, and Ser-Val. These dipeptides could fit the electron density for the ligand almost equally well. As can be seen in Figure 6, only the Ser-Thr dipeptide was able to inhibit the enzyme in the κ -casein cleavage assay, which lends further credence to the idea that this is the ligand observed in the active site, although it is conceivable that a mixture of short peptides are bound. Curiously, the other two peptides (Pro-Thr and Ser-Val) appeared to slightly increase the rate of reaction in the clotting assay, which cannot be explained readily, although their greater hydrophobic character may be significant.

If the bound ligand is a Ser-Thr dipeptide, then it probably originated from autolytic cleavage of the NH_2 -Ser-Thr-sequence at the N terminus of the enzyme during the 12-year crystallization period. Overall, the analysis provides a preliminary indication that oligopeptides can inhibit the enzyme by electrostatic interaction at the active site. This is consistent with classical findings that the aspartic proteinases can be inhibited weakly by short peptides, and that these enzymes can also catalyze transpeptidation reactions (Fruton 1976). In this process, the ability of the active site

Figure 4. The electron density for the inhibitor H261 in the complex with endothiapepsin refined at 1.1 Å resolution. The map (shown as pale lines) has been σ_A -weighted (Read 1986) and is contoured at 2.2 root mean square.

Figure 5. The electron density for the dipeptide bound at the catalytic center of the “native” enzyme. The map (shown as pale lines) is at 0.9 Å resolution and has been σ_A -weighted (Read 1986) and contoured at 1.3 root mean square. Putative hydrogen bonds are shown as dashed lines. The N-terminal nitrogen of the dipeptide appears to form a hydrogen bond with the active site water molecule and a salt-bridge interaction with the Asp 215 of the catalytic diad.

cleft to retain one of the products of the hydrolytic reaction is clearly important for the subsequent condensation step involving a second peptide that binds at the active site. Previously, transpeptidation has been cited as evidence for a covalently bound intermediate. However, our results indicate that a covalent intermediate may not be necessary for transpeptidation to occur.

In refinement, the occupancy of the Ser-Thr dipeptide converged to a value of 39%; that is, unlike the tight-binding inhibitor H261, it is not present in all enzyme molecules in the crystal. The putative salt-bridge interaction made by the amino group of the dipeptide with Asp 215 indicates that this aspartate could be negatively charged in the ground state of the enzyme. However, unrestrained refinement of the carboxyl C–O bond lengths (see later) indicates that both aspartates are equally likely to be charged in the ground state; that is, Asp 215 will be negatively charged (and able to make a salt bridge) in only 50% of the active sites in the crystal at any one time. This correlates with the finding that the occupancy of the dipeptide is <50%.

The charge status of catalytic residues

In catalysis, the active site water molecule may become partly displaced upon substrate binding and polarized by one of the aspartate carboxyls (Suguna et al. 1987). The water is then thought to nucleophilically attack the scissile bond carbonyl to form the tetrahedral intermediate. Current proposals for the catalytic mechanism differ in the assignment of protonation states to the catalytic groups during the reaction. Because hydrogen atoms cannot be readily

located by X-ray analysis of proteins, their putative positions usually have to be inferred from the local geometry of surrounding polar atoms. One of the key features of the mechanism proposed by Veerapandian et al. (1992; shown in Fig. 7) is the stabilization of the transition state by a negative charge localized on Asp 32. The assignment of a negative charge to this residue was made on the basis that its hydrogen bonding capacity is satisfied to a greater extent than that of Asp 215 in the complex with a *gem*-diol inhibitor; that is, a negative charge on Asp 32 would be more stable than one on Asp 215 in the transition state. However, the protonation states of the catalytic aspartates in the complexes had not been determined with any certainty at that stage.

Information on the ionization states of the active site carboxyl groups is of crucial importance in understanding the catalytic mechanism of aspartic proteinases. However, recent work has shown that a number of functional groups in proteins, including carboxylates, are particularly vulnerable to radiolytic damage at the large X-ray doses that might be required to collect atomic resolution data (Burmeister 2000; Ravelli and McSweeney 2000; Leiros et al. 2001). Aspartate and glutamate residues, including those involved in buried salt-bridge interactions, appear to suffer from decarboxylation of the side chain. These effects would be expected to complicate the electron density map and make the process of locating weak density due to hydrogen atoms rather unreliable. In addition, active site protons involved in low-barrier hydrogen bonds (frequently found at enzyme active sites) will be distributed between two positions ~0.5 Å apart (Cleland et al. 1998). Thus hydrogens of interest may have high temperature factors that will compound the difficulties of locating them in the electron density map. However, unrestrained refinement by using atomic resolution X-ray data provides an alternative means of defining the protonation states of carboxylate groups in well-ordered parts of the structure. Neutral carboxyl groups have a sig-

Figure 6. The effects of three dipeptides (ST, SV, and PT) on the activity of endothiapepsin. The effect on the rate of reaction was determined by a κ -casein cleavage assay at two concentrations of each peptide. Each rate of reaction was calculated as the reciprocal of the lag-time (sec^{-1}) for the turbidimetric assay. Only the ST peptide has an inhibitory effect.

Figure 7. The catalytic mechanism proposed by Veerapandian et al. (1992). This mechanism is based on the X-ray structure of a difluoroketone (*gem*-diol) inhibitor bound to endothiapepsin. A water molecule tightly bound to the aspartates in the native enzyme is proposed to nucleophilically attack the scissile bond carbonyl. The resulting tetrahedral intermediate (2) is stabilized by hydrogen bonds to the negatively charged carboxyl of aspartate 32. Fission of the scissile C–N bond is accompanied by transfer of a proton to the leaving amino group either from Asp 215 (with nitrogen inversion) or from bulk solvent. Dashed lines indicate hydrogen bonds.

nificant difference between their C–OH and C=O bond lengths (1.21 Å for the C=O bond and 1.32 Å for the C–OH bond), whereas ionized carboxylates have identical C–O bond lengths (typically 1.27 Å) due to resonance. For each carboxyl group, the estimated standard deviation (ESD) of the difference in C–O bond lengths can be calculated from the bond length ESDs obtained by normal matrix inversion and compared with the size of the bond length difference. When the difference in C–O bond length exceeds three times the ESD, it is very likely that the group is protonated. Thus, unrestrained refinement of the carboxyl groups is a very powerful tool for determination of the protonation state of each residue and avoids the difficulties of locating weak electron density due to the hydrogens themselves (Coates et al. 2001; Helliwell et al. 2002). Unrestrained refinement of a number of endothiapepsin inhibitor complexes similar to H261, but at higher resolution, allowed the assignment of charges to a number of the aspartate and glutamate groups in the enzyme (Coates et al. 2002). The consensus charges assigned to the carboxyl groups based on these previous structures and the H261 structure reported herein are given in Table 2.

The protonation state of the catalytic residues has been a long-standing controversy in the aspartic proteinase field. In

the optimal pH range, it is likely that the aspartate diad possesses a single negative charge. Unrestrained refinement of the carboxyl groups of native endothiapepsin at 0.9 Å showed that both aspartates have C–O bond lengths that are almost identical within the errors of measurement (ESDs = 0.01 Å; Fig. 8; Table 2). Although this is characteristic of charged carboxylate groups, it is very unlikely that both aspartates are ionized at the same time due to their proximity and the low pH of the crystallization buffer. Therefore, this effect must be due to the presence of two canonical forms at the catalytic center; that is, in half of the enzyme molecules Asp 32 is ionized, and in the remaining half Asp 215 is ionized (Fig. 8). Thus, our atomic resolution studies of the native enzyme have shown that the distribution of charge at the catalytic center in the absence of substrate is symmetric—an effect that was predicted earlier (Pearl and Blundell 1984) but has not been substantiated

Table 2. The protonation states of the enzyme carboxyl groups deduced from analysis of the C–O bond lengths obtained by unrestrained refinement

	State in native enzyme	Consensus state in complexes
Enzyme aspartic acid residue		
8	Charged	Charged
11	Charged	Protonated
12	Charged	—
30	Protonated	Protonated
32	Charged	Charged
37	Protonated	Protonated
77	Protonated	Charged
83	Charged	Charged
87	Charged	Charged
114	Charged	Split
118	Charged	Charged
140	Split	Charged
147	Charged	Charged
154	Charged	Charged
171	Charged	Charged
211	Charged	Charged
215	Charged	Protonated
271	Charged	Charged
274	Charged	Charged
304	Protonated	—
Enzyme glutamic acid residue		
44	Split	Split
49	Protonated	Protonated
102	Protonated	—
112	Charged	Charged
191	Protonated	Charged

The column showing the charge of each residue in the inhibitor complex is a consensus of the results for several complexes published previously (Coates et al. 2002) in addition to the data for the H261 complex. Dashes indicate residues for which there were no consistent indications of their charge status. Some residues are shown as split because they have dual conformations, and thus, their protonation states are unlikely to be determined accurately by this method.

Figure 8. (A) The covalent- and hydrogen-bond lengths at the catalytic center of the native enzyme obtained from unrestrained refinement by using the 0.9 Å resolution data set. The equivalence of the carboxyl C–O bond lengths indicates the existence of two canonical forms, which are shown on the *right*. Note that the bonding geometry for a proton located directly between the two inner aspartate oxygens is unfavorable. (B) The catalytic center of the H261 inhibitor complex showing the interatomic distances obtained from unrestrained refinement by using the 1.1 Å resolution data set. Note the unusually short hydrogen bonds (putative low-barrier hydrogen bonds) between the inhibitor hydroxyl group and the two catalytic aspartates, shown with asterisks. All distances are shown in Ångstrom units, and hydrogen bond distances refer to the separation of the donor and acceptor oxygen atoms.

until now. However, an asymmetric charge distribution may arise in the presence of substrate.

The inhibitor H261 possesses a hydroxyethylene analog (–CHOH–CH₂–) in place of the scissile peptide bond. X-ray structural studies of the complex of H261 with endothiapepsin (Veerapandian et al. 1990) showed that the hydroxyl replaces the water molecule found at the catalytic center of the native enzyme. The hydroxyl group appears to form short hydrogen bonds with the inner carboxyl oxygen of Asp 32 and the outer carboxyl oxygen of Asp 215, indicating that protons must reside between these atoms allowing them to form hydrogen bonds. The other two possible hydrogen bonds to the hydroxyl group involve the outer carboxyl oxygen of Asp 32 and the inner oxygen of Asp 215, but both have unfavorable geometry (donor-acceptor distances are too long). These findings are confirmed by the atomic resolution analysis of the endothiapepsin-H261 complex. Unrestrained refinement using the 1.1 Å resolution data set shows the pattern of C–O bond lengths (shown in Fig. 8) to agree with those of the inhibitor complexes reported earlier (Coates et al. 2002). These data, together with

the results of a neutron diffraction study (Coates et al. 2001), are consistent with the model of the intermediate shown in Figure 7, in which Asp 32 is negatively charged and Asp 215 is neutral.

Defining the ionization states of carboxyl groups has important implications for understanding many other properties of enzymes, including the variability of their pI values and pH optima. The protonation states of all the enzyme carboxyls, as deduced from the unrestrained refinement of the native enzyme at 0.9 Å resolution, are shown in Table 2, and as expected, the majority of residues have the same deduced protonation state as in the inhibitor complex. One early proposal for the mechanism of aspartic proteinases invoked a charge-relay system involving the strongly conserved residue Asp 304 (Pearl and Blundell 1984). The side-chain of this group is buried beneath the catalytic center, and it interacts with the aspartate diad by a network of hydrogen bonds through which the charge-relay system was proposed to operate. Unrestrained refinement of the native enzyme structure shows that this aspartate has C–O bond lengths of 1.21 and 1.31 Å, and the ESD for the bond length

difference is only 0.01 Å. Thus, the atomic resolution data for the native enzyme indicate that the carboxyl of Asp 304 is protonated in the native enzyme and is unlikely to be involved in a charge-relay system.

Low-barrier hydrogen bonds

The donor-acceptor distances for the hydrogen bonds made between the hydroxyl of H261 and the catalytic aspartates are very short (~2.6 Å). This was suggested by many previous X-ray studies, which were typically at 2.0 Å resolution, and has recently been confirmed by refinement of several complexes using atomic resolution data (Coates et al. 2002). Hydrogen bonds as short as 2.5 to 2.6 Å are referred to as low-barrier hydrogen bonds (LBHBs) because the proximity of the donor and acceptor atoms reduces the energy barrier that normally prevents transfer of the hydrogen atom from the donor to the acceptor group. The LBHB has been proposed as being important in stabilizing the catalytic intermediates of a number of enzymes, including citrate synthase and the serine proteinases (Cleland et al. 1998). LBHB formation upon binding of an inhibitor or a substrate is thought to be due to steric compression in the active site cleft. Recently, it has been proposed that LBHBs also play a role in the aspartic proteinase mechanism by facilitating transfer of protons between the aspartate diad and the catalytic intermediates (Coates et al. 2002). This was based on NMR and atomic resolution X-ray crystallography. In contrast, NMR analysis of the native enzyme demonstrated that these LBHBs were not detectable. However, inspection of the atomic resolution structure of the native enzyme shows that there are very short hydrogen bonds between the outer oxygens of the aspartates and the hydroxyl groups of threonines 35 and 218. The inability of NMR to detect these interactions may be due to their higher H-D exchange rate in the absence of a tight-binding inhibitor in the active site cleft.

One phenomenon that LBHBs may be able to explain is the fact that some classes of aspartic proteinase inhibitor are slow-binding (i.e., require a time-scale of minutes to achieve maximum inhibition). It has been suggested that slow-binding is due to the time-lag required to displace the active-site water molecule. However, crystal structures showed that all classes of inhibitor displace this water molecule, but only some exhibit slow-binding. Our current work raises the possibility that slow-binding of inhibitors may be a consequence of the time required for the transition state analog to form LBHBs by steric compression at the catalytic center. The fact that phosphinic acid analogs require several hours to achieve maximum binding (Bartlett and Kezer 1987) may be due to slow formation of even shorter interactions known as single-well hydrogen bonds (2.4 Å), which are observed in the atomic resolution crystal structure (Coates et al. 2002).

Materials and methods

Crystals of native endothiapepsin and the complex with H261 were obtained by using the batch method with the enzyme at a concentration of 2 mg/mL in 100 mM sodium acetate buffer (pH 4.6; Moews and Bunn 1970). Crystals were left to grow after the slow addition of finely ground ammonium sulphate to a final concentration of 0.35 g/mL (55% saturation). The crystals were ~12 years old at the time of data collection, having been stored in mother liquor at room temperature. Both data sets were obtained by using crystals that were cryoprotected in ~30% glycerol and flash-cooled in liquid ethane. In total, 180° of data at 100 K were collected from each crystal by using a MAR345 image plate at the BW7B beam line at DESY (Hamburg), and one or two additional collections were done, using shorter exposure times and/or an attenuated beam, to collect overloaded reflections. The data sets were processed by using MOSFLM (Leslie 1992) and CCP4 programs (CCP4 1994), and the electron density maps were analyzed by using TURBO-FRODO (Bio-Graphics). Least-squares refinement after each rebuild was performed by using SHELX97-2 (Sheldrick and Schneider 1997), in which refinement of anisotropic *U* values was followed by use of the "riding-hydrogen" model. To determine the protonation states of the aspartate and glutamate side-chains from their bond lengths, unrestrained refinement of all enzyme carboxyl groups was undertaken, followed by blocked-matrix inversion to estimate the bond-length standard deviations of these groups. The structures have been deposited with the Protein Data Bank (PDB) and assigned the following accession codes: 1oew for native endothiapepsin and 1oex for the complex with H261.

To attempt to confirm the identity of the dipeptide found in the active site of the "native" enzyme structure, the following dipeptides were synthesized: Ser-Thr, Pro-Thr, and Ser-Val. A turbidimetric κ -casein cleavage assay was used to assess their inhibitory potency by following the increase in absorbance at 500 nm of a 3 mL solution containing 2 mg/mL skimmed milk, 8 mM CaCl₂ in 0.2 M sodium acetate buffer (pH 5.3) upon addition of 10 μ L of enzyme at a concentration of 2 mg/mL. A 5 mg/mL stock solution (~30 mM) of each peptide was made, and aliquots of 0.1 mL and 0.2 mL were added to the 3 mL assay solution (giving final concentrations of ~1 mM and 2 mM, respectively) to assess their inhibitory effect on the enzyme.

Acknowledgments

We thank the Biotechnology and Biological Sciences Research Council (UK) for project grant support and the EMBL Outstation at Deutsches Elektronen Synchrotron (DESY), Hamburg, for synchrotron beam time and associated travel funding. We thank R. Broadbridge (University of Southampton) and Y. George for synthesis of the dipeptides and the assays. We are grateful to the Engineering and Physical Sciences Research Council (UK) for a Ph.D. studentship to L.C.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Bailey, D. and Cooper, J.B. 1994. A structural comparison of 21 inhibitor complexes of the aspartic proteinase from *Endothia parasitica*. *Protein Sci.* **3**: 2129-2143.
- Bartlett, P.A. and Kezer, W.B. 1987. Phosphinic acid dipeptide analogues: Potent, slow-binding inhibitors of aspartic proteinases. *J. Am. Chem. Soc.* **109**: 4282-4283.

- Blundell, T.L., Jenkins, J.A., Sewell, B.T., Pearl, L.H., Cooper, J.B., Tickle, I.J., Veerapandian, B., and Wood, S.P. 1990. X-ray analysis of aspartic proteinases, I: The three dimensional structure at 2.1 Å of endothiapepsin. *J. Mol. Biol.* **211**: 919–941.
- Burmeister, W.P. 2000. Structural changes in a cryo-cooled protein crystal owing to radiation damage. *Acta Crystallogr. D* **56**: 328–341.
- CCP4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**: 760–763.
- Cleland, W.W., Frey, P.A., and Gerlt, J.A. 1998. The low-barrier hydrogen bond in enzymatic catalysis. *J. Biol. Chem.* **273**: 25529–25532.
- Coates, L., Erskine, P.T., Wood, S.P., Myles, D.A.A., and Cooper, J.B. 2001. A neutron Laue diffraction study of endothiapepsin: Implications for the aspartic proteinase mechanism. *Biochemistry* **40**: 13149–13157.
- Coates, L., Erskine, P.T., Crump, M.P., Wood, S.P., and Cooper, J.B. 2002. Five atomic resolution structures of endothiapepsin inhibitor complexes: Implications for the aspartic proteinase mechanism. *J. Mol. Biol.* **318**: 1405–1415.
- Cooper, J.B. 2002. Aspartic proteinases in disease: A structural perspective. *Curr. Drug Targets* **3**: 155–173.
- Davies, D.R. 1990. The structure and function of the aspartic proteinases. *Ann. Rev. Biophys. Biophys. Chem.* **19**: 189–215.
- Fruton, J.S. 1976. The mechanism of the catalytic action of pepsin and related acid proteases. *Adv. Enzymol.* **44**: 1–36.
- Fusek, M. and Vetvicka, V. 1994. Mitogenic function of human procathepsin-D: The role of the propeptide. *Biochem. J.* **303**: 775–780.
- Helliwell, J.R., Price, H.J., Deacon, A., Raftery, I., and Habash, J. 2002. Complementarity of neutron and ultrahigh resolution synchrotron X-ray protein crystallography studies: Results with concanavalin A at cryo and room temperature. *Acta Physica Polonica A* **101**: 583–588.
- Hofmann, T., Dunn, B.M., and Fink, A.L. 1984. Cryoenzymology of penicillopepsin, with an appendix: Mechanism of action of aspartyl proteinases. *Biochemistry* **23**: 5247–5256.
- Hong, L., Koelsch, G., Lin, X.L., Wu, S.L., Terzyan, S., Ghosh, A.K., Zhang, X.C., and Tang, J. 2000. Structure of the protease domain of memapsin 2 (β -secretase) complexed with inhibitor. *Science* **290**: 150–153.
- James, M.N.G., Sielecki, A.R., Hayakawa, K., and Gelb, M.H. 1992. Crystallographic analysis of transition state mimics bound to penicillopepsin: Difluorostatine- and difluorostatone-containing peptides. *Biochemistry* **32**: 3872–3886.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**: 283–291.
- Leiros, H.K.S., McSweeney, S.M., and Smalas, A.O. 2001. Atomic resolution structures of trypsin provide insight into structural radiation damage. *Acta Crystallogr. D* **57**: 488–497.
- Leslie, A.G.W. 1992. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB newsletter on protein crystallography*, No. 26. Daresbury Laboratories, Warrington, UK.
- Lindner, H. and Helliger, W. 2001. Age-dependent deamination of asparagines residues in proteins. *Exp. Gerontol.* **36**: 1551–1563.
- Moews, P. and Bunn, C.W. 1970. An X-ray crystallographic study of the rennin-like enzyme of *Endothia parasitica*. *J. Mol. Biol.* **54**: 395–397.
- Pearl, L.H. and Blundell, T.L. 1984. The active sites of aspartic proteinases. *FEBS Lett.* **174**: 96–101.
- Pichová, I., Pavlíková, L., Dostál, J., Dolejší, E., Hrušková-Heidingsfeldová, O., Weber, J., Ruml, T., and Souek, M. 2001. Secreted aspartic proteases of *Candida albicans*, *Candida tropicalis*, *Candida parapsilosis* and *Candida lusitanae*: Inhibition with peptidomimetic inhibitors. *Eur. J. Biochem.* **268**: 2669–2677.
- Ravelli, R.B.G. and McSweeney, S.M. 2000. The fingerprint that X-rays leave on structures. *Structure* **8**: 315–328.
- Read, R.J. 1986. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**: 140–149.
- Rich, D.H., Bernatowicz, M.S., and Schmidt, P.G. 1982. Direct ^{13}C NMR evidence for a tetrahedral intermediate in the binding of a pepstatin analogue to porcine pepsin. *J. Am. Chem. Soc.* **104**: 3535–3536.
- Sheldrick, G.M. and Schneider, T.R. 1997. SHELXL: High-resolution refinement. *Methods Enzymol.* **277**: 319–343.
- Silva, A.M., Cachau, R.E., Sham, H.L., and Erickson, J.W. 1996a. Inhibition and catalytic mechanism of HIV-1 aspartic protease. *J. Mol. Biol.* **255**: 321–340.
- Silva, A.M., Lee, A.Y., Gulnik, S.V., Majer, P., Collins, J., Bhat, T.N., Collins, P.J., Cachau, R.E., Luker, K.E., Gluzman, I.Y., et al. 1996b. Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from *Plasmodium falciparum*. *Proc. Natl. Acad. Sci.* **93**: 10034–10039.
- Suguna, K., Padlan, E.A., Smith, C.W., Carlson, W.D., and Davies, D.R. 1987. Binding of a reduced peptide inhibitor to the aspartic proteinase from *Rhizopus chinensis*: Implications for a mechanism of action. *Proc. Natl. Acad. Sci.* **84**: 7009–7013.
- Tang, J., James, M.N.G., Hsu, I.N., Jenkins, J.A., and Blundell, T.L. 1978. Structural evidence that for gene duplication in the evolution of acid proteases. *Nature* **271**: 618–621.
- Veerapandian, B., Cooper, J.B., Šali, A., Blundell, T.L., Rosati, R.L., Dominy, B.W., Damon, D.B., and Hoover, D.J. 1992. Direct observation by X-ray analysis of the tetrahedral “intermediate” of aspartic proteinases. *Protein Sci.* **1**: 322–328.
- Veerapandian, B., Cooper, J., Šali, A., and Blundell, T.L. 1990. Three dimensional structure of endothiapepsin complexed with a transition-state isostere inhibitor of renin at 1.6 Å resolution. *J. Mol. Biol.* **216**: 1017–1029.