Published in final edited form as:
J Am Chem Soc. 2008 June 11; 130(23): 7235–7237. doi:10.1021/ja801269x.

The Catalytic Mechanism of an Aspartic Proteinase Explored with Neutron and X-ray Diffraction

Leighton Coates[†,‡], **Han-Fang Tuan**[†], **Stephen Tomanicek**[†], **Andrey Kovalevsky**[§], **Marat Mustyakimov**[§], **Peter Erskine**[¶], and **Jon Cooper**[¶]

Spallation Neutron Source, Oak Ridge National Laboratory, 1 Bethel Valley Road, Oak Ridge, Tennessee 37831, Department of Chemistry, University of Toledo, Toledo, Ohio 53606, Bioscience Division, Mailstop M888, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, and Laboratory for Protein Crystallography, Centre for Amyloidosis and Acute Phase Proteins, Department of Medicine, University College London (Hampstead Campus), Rowland Hill Street, London NW3 2PF, U.K.

Aspartic proteinases are a class of enzymes widely distributed among fungi, plants, vertebrates, and viruses. They are involved in numerous disease conditions, including hypertension, amyloid disease, malaria, and AIDS.[1] In HIV, the proteinase is essential for the maturation of the virus particle, and inhibitors have a proven therapeutic effect in the treatment of AIDS. There have also been significant recent advances in the development of orally active renin inhibitors which are in clinical trials.[2] Thus inhibitors to this class of enzyme with improved characteristics are therefore much sought after as potential therapeutic agents. The aspartic proteinase class of enzyme comprises two structurally similar domains, each contributing an aspartic acid residue to form a catalytic dyad that cleaves the substrate peptide bond. The two catalytic aspartate carboxyls are held coplanar by a network of hydrogen bonds involving the surrounding main chain and the conserved amino acid side-chain groups. Most aspartic proteinases are inhibited by pepstatin, a microbial peptide that contains the unusual amino acid statine, and a number of crystal structures of complexes containing this naturally occurring analogue have been reported.[3,4] Endothiapepsin, a member of the aspartic proteinase family, is composed of 330 amino acid residues, with roughly 170 residues in each domain. Endothiapepsin is derived from the fungus *Endothia parasitica*, and numerous structures of this enzyme bound to a range of renin inhibitors have been analyzed with the aim of developing improved compounds by structure-based design.[5] Crystallographic studies of all native aspartic proteinases have revealed a water molecule hydrogen bonded to the carboxyl groups of both aspartate residues. This water molecule is within hydrogen-bonding distance of all four carboxyl O atoms and has been implicated in catalysis. It has been suggested[6] that this water is partly displaced upon substrate binding and is polarized by one of the catalytic aspartate residues. It then initiates enzymatic catalysis by attacking the scissile bond carbonyl group of the substrate. The transition state of aspartic proteinase catalysis has been studied most extensively using enzyme—inhibitor complexes which possess a range of different transition state mimics. Most inhibitors mimic one or both hydroxyls in the putative transition state. The statine-based inhibitors contain one hydroxyl that occupies the same position as the water

E-mail: coatesl@ornl.gov.
[†]Oak Ridge National Laboratory.
[‡]University of Toledo.
[§]Los Alamos National Laboratory.
[¶]University College London.

Supporting Information Available: Data collection and refinement statistics for all three structures discussed within this paper, the atomic resolution, the room temperature X-ray, and room temperature neutron. This material is available free of charge via the Internet at http://pubs.acs.org.

molecule in the native enzyme and binds via hydrogen bonds to both catalytic aspartates.[4] Three X-ray structures of endothiapepsin complexed with the statine-based inhibitor have been refined to atomic resolution,[7] and a number of low-barrier hydrogen bonds (LBHBs) were identified within the active site. Another class of transition state mimic, the difluoroketone analogue [—CO—CF_2—], such as that present in the renin inhibitor PD-135,040, [8] is readily hydrated to the *gem*-diol form [—$C(OH)_2$—CF_2—] in water. Upon binding to the enzyme, this mimics the putative tetrahedral intermediate very closely.[9] NMR studies have also shown that the enzyme binds ketone analogues in the hydrated *gem*-diol form.[10] Mechanistic proposals have been made based on X-ray diffraction studies of the difluoroketone *gem*-diol transition state analogues.[9,11,12] The catalytic mechanism proposed in ref [9] is shown in Figure 1.

Results from this study suggest that the transition state is stabilized by a negative charge on Asp32, one of the two catalytic aspartate residues. Previous investigation[13] using neutron diffraction on a complex of endothiapepsin with the hydroxyethylene inhibitor H261 provided evidence that the other catalytic residue, Asp215, is protonated on the outer oxygen. Due to the low optimum pH of the enzyme[14] and the close proximity of the two aspartates, it is likely that one catalytic aspartate is protonated and the other charged at the enzyme's operational pH. As aspartic acid carboxyl bond lengths vary in different protonation states, bond length analysis with atomic resolution X-ray data enables the protonation states of the catalytic aspartates to be revealed indirectly. We successfully obtained a single large protein crystal of the complex between endothiapepsin and the *gem*-diol inhibitor PD-135,040 for neutron diffraction studies and several smaller crystals suitable for X-ray studies. Three data sets were collected on this complex: an atomic resolution X-ray data set, a room temperature X-ray data set, and a room temperature neutron data set.[8] A key advantage of joint X-ray and neutron analysis is the improvement it offers in the data-to-parameter ratio. In neutron diffraction, a large number of atoms are being refined, typically double the number of atoms in an X-ray refinement, due to the sensitivity of neutrons to hydrogen and deuterium atoms. Therefore, great care must be taken to limit the parameters being refined so a respectable data-to-parameter ratio can be maintained. Joint refinement with the additional X-ray data gives a higher data-to-parameter ratio, providing a more robust refinement procedure. The atomic resolution data for this complex are of very high quality with an $I/\sigma(I)$ of 10.3 in the outer resolution shell (1.05–1.0 Å).[8] Stereochemical restraints on the active site aspartates were removed after the refinement had converged to allow the carboxyl bond lengths to refine for bond length analysis (Figure 2). The carboxyl bond lengths allow us to deduce the aspartate protonation states since neutral carboxyls generally have bond lengths of around 1.21 and 1.32 Å for the C=O and C—OH bonds, respectively. In contrast, negatively charged carboxyls are expected to have identical C—O bond lengths of around 1.26 Å due to resonance. The protonation states of the carboxyls were deduced by comparing the difference in carboxyl bond lengths to their associated estimated standard deviation (ESD).[15–17] For aspartates with differences in carboxyl bond lengths over three times greater than the ESD, we designated them as protonated. Where differences in carboxyl bond lengths were less than three times the ESD, we designated them as charged. The carboxyl bond lengths for Asp32 both refined to 1.26 Å (ESD 0.01Å) and those for Asp215 were refined to 1.22 Å on the inner oxygen and 1.28 Å on the outer oxygen (ESD 0.02Å). These distances, which are shown in Figure 2, lead to the conclusion that, when the *gem*-diol inhibitor is bound, Asp32 is charged while Asp215 is protonated on the outer oxygen. Some of the hydrogen bond lengths within the active site are short (~2.50 Å) and can be considered as low-barrier hydrogen bonds.[18] The shortest hydrogen bond (2.54 Å) within the active site occurs between the outer oxygen of Asp32 and the outer inhibitory hydroxyl. In low-barrier hydrogen bonds, the proton is shared by the donor and acceptor atoms to a greater extent than is normal in weaker hydrogen-bonding interactions. As a result, the proton can be in either of two positions that are close to the midpoint between the donor and acceptor atoms.

This delocalization might be expected to make the proton (or deuteron) in the hydrogen bond between the outer oxygen of Asp32 and the hydroxyl of the *gem*-diol less visible in both X-ray and neutron analyses. Occupancy refinement of the H/D backbone exchange pattern (Figure 3) showed that, in general, residues closer to the center of mass of the protein underwent lower levels of H/D exchange. This is likely to be caused by lower levels of solvent accessibility and involvement in β-sheet secondary structure. As might be expected, atoms at the surface of the protein show a higher level of exchange, which is probably due to increased mobility and solvent accessibility.

The initial neutron model contained no hydrogen or deuterium atoms within the active site. This was done to remove any model bias from the structure. Distinct positive peaks in the difference map for deuterium atoms at the active site were observed. Deuterium atoms were added into these peaks, and the resulting nuclear density maps were analyzed (Figure 3b). The occupancy of the deuterium atoms added at the active site was refined. The deuterium atom on the outer oxygen of Asp215 refined to an occupancy of 66% and the deuterium atom on the inner *gem*-diol hydroxyl orientated toward the inner oxygen of Asp32 refined to an occupancy of 100%. The partial occupancy (66%) of the deuterium atom on Asp215 ties in with C—O bond length (1.28 Å) from the atomic resolution X-ray structure. This is shorter than the ideal 1.32 Å value for a fully protonated aspartate, suggesting partial protonation. This has recently been seen in investigations into the catalytic mechanism of h-aldose reductase.[19]

Interestingly, there is no electron density to suggest hydrogen positions at the active site with the atomic resolution X-ray data set. This is consistent with the findings of others where it has been reported that at a resolution of 0.66 Å only 54% of protein hydrogen atoms were visible.[20] With reference to Figures 2 and 3b, it is apparent that the shortest donor-acceptor distance (namely that between the outer carboxyl oxygen of Asp32 and the outer *gem*-diol hydroxyl) corresponds exactly with the deuteron of poorest visibility in the neutron map. As discussed above, this might be expected for a low-barrier hydrogen bond due to delocalization of the proton (or deuteron) which is a feature of this type of interaction. The procedures and methods used for protein crystallization, data collection, and data set statistics were described in detail in ref [8]. SHELX[21] was used to refine the atomic resolution and room temperature X-ray structures. After modeling the initial solvent structure, anisotropic B-factors were refined and riding hydrogens added to the atomic resolution data set. The final R_{factor} was 15.7% with R_{free} of 19.6% for the room temperature X-ray data set, and the final R_{factor} was 11.1% with R_{free} of 13.6% for the atomic resolution data set. A neutron-adapted version of *CNS*, *nCNS*,[22] was used in the joint refinement of the room temperature neutron and X-ray data sets. The final R_{factor} for the neutron data set was 21.9% with an R_{free} of 28.1%. The structure factors and final structure from each refinement has been deposited in the protein data bank with the accession codes 2jji high-resolution X-ray, 2jjj room temperature X-ray, and 2vs2 X-ray/neutron refinement. Although the atomic resolution X-ray structure gave no direct evidence of the protonation states at the active site, it did give significant information via bond length analysis. The neutron diffraction data set, even though it was at much lower resolution (2.0 Å) than the X-ray data, did give direct evidence showing the protonation states at the active site. The nuclear density at the active site clearly (Figure 3b) shows that Asp215 is deuterated on its outer oxygen. The occupancy of the deuterium modeled here refined to 66%. Another deuterium on the inner inhibitor hydroxyl is orientated toward Asp32, and the occupancy of this deuterium atom refined to 100%. The lack of nuclear density for the deuteron involved in the hydrogen bond between the outer oxygen of Asp32 and the outer *gem*-diol hydroxyl is consistent with this being the shortest hydrogen bond in the atomic resolution X-ray analysis. Its length is consistent with that of a low-barrier hydrogen bond in which the proton is diffusely distributed about the midpoint. Low-barrier hydrogen bonds have significantly greater stability than normal hydrogen bonds,[18] and this may be an important feature of stabilizing the transition state. The *gem*-diol hydroxyl which forms this interaction is that which, in the

enzyme-catalyzed reaction, would originate from the carbonyl oxygen of the substrate peptide bond. The other hydroxyl (located almost symmetrically between the two aspartate carboxyls) would most probably originate from the water molecule which is bound in essentially the same position to the native enzyme in all known aspartic proteinase structures. The requirement to stabilize the *gem*-diol by forming a low-barrier hydrogen bond with the hydroxyl that originates from the substrate carbonyl may reflect on the strain which the scissile peptide bond must experience upon binding to the active site. The results from the neutron and X-ray studies are consistent with an earlier neutron diffraction study of endothiapepsin co-complexed with a hydroxyethylene inhibitor, H261,[13] and a number of atomic resolution X-ray studies.[7] They are also in agreement with the mechanism and proton arrangement of the transition state proposed in ref [9].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

Research sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory (ORNL), managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. We also are grateful to the ESRF (Grenoble, France) for beamtime and travel support. A NIH-NIGMS funded consortium (1R01GM071939-01) between LANL and LNBL to develop computational tools for neutron protein crystallography is acknowledged.

References

(1). Cooper JB. Curr. Drug Targets 2002;3:155–174. [PubMed: 11958298]

(2). Wood JM. Biochem. Biophys. Res. Commun 2003;308:698–705. [PubMed: 12927775]

(3). Umezawa H, Aoyagi T, Morishima H, Matsuzaki M, Hamada M. J. Antibiot 1970;23:259–262. [PubMed: 4912600]

(4). Bailey D, Cooper JB, Veerapandian B, Blundell TL, Atrash B, Jones DM, Szelke M. Biochem. J 1993;289:363–371. [PubMed: 8424781]

(5). Bailey D, Cooper JB. Protein Sci 1994;3:2129–2143. [PubMed: 7703859]

(6). Suguna K, Padlan EA, Smith CW, Carlson WD, Davies DP. Proc. Natl. Acad. Sci. U.S.A 1987;84:7009–7013. [PubMed: 3313384]

(7). Coates L, Erskine PT, Crump MP, Wood SP, Cooper JB. J. Mol. Biol 2002;318:1405–1415. [PubMed: 12083527]

(8). Tuan H-F, Erskine P, Langan P, Cooper J, Coates L. Acta Crystallogr 2007;F63:1080–1083.

(9). Veerapandian B, Cooper JB, Sali A, Blundell TL, Rosatti RL, Dominy BW, Damon DB, Hoover DJ. Protein Sci 1992;1:322–328. [PubMed: 1304340]

(10). Rich DH, Bernatowicz MS, Schmidt PG. J. Am. Chem. Soc 1982;104:3535–3536.

(11). James MNG, Sielecki AR, Hayakawa K, Gelb MH. Biochemistry 1992;31:3872–3886. [PubMed: 1567842]

(12). Coates L, Erskine PT, Mall S, Williams PA, Gill RS, Wood SP, Cooper JB. Acta Crystallogr 2003;D59:978–981.

(13). Coates L, Erskine PT, Wood SP, Myles D, Cooper JB. Biochemistry 2001;40:13149–13157. [PubMed: 11683623]

(14). Fruton JS. Adv. Enzymol. Relat. Areas Mol. Biol 1976;44:1–36. [PubMed: 775937]

(15). Deacon A, Gleichmann T, (Gilboa) A. J. Kalb, Price H, Raftery J, Bradbrook G, Yariv J, Helliwell JR. J. Chem. Soc., Faraday Trans 1997;93:4305–4312.

(16). Habash J, Raftery J, Weisgerber S, Cassetta A, Lehmann MS, Hoghoj P, Wilkinson C, Campbell JW, Helliwell JR. J. Chem. Soc., Faraday Trans 1997;93:4313–4317.

(17). Ahmed HU, Blakeley MP, Cianci M, Cruickshank DWJ, Hubbard JA, Helliwell JR. Acta Crystallogr 2007;D63:906–922.

(18). Cleland WW, Frey PA, Gerlt JA. J. Biol. Chem 1998;273:25529–25532. [PubMed: 9748211]

(19). Blakeley MP. Proc. Natl Acad. Sci. U.S.A 2008;105:1844–1848. [PubMed: 18250329]

(20). Hazemann I, Dauvergne MT, Blakeley MP, Meilleur F, Haertlein M, Van Dorsselaer A, Mitschler A, Myles DAA, Podjarny A. Acta Crystallogr 2005;D61:1413–1417.

(21). Sheldrick GM. Acta Crystallogr 2007;A64:112–122.

(22). MustyakimovMAfononePAdamsPDLanganP nCNS, a patch for CNS, Copyright (C-06, 104), 2007, Los Alamos National Security, LL.

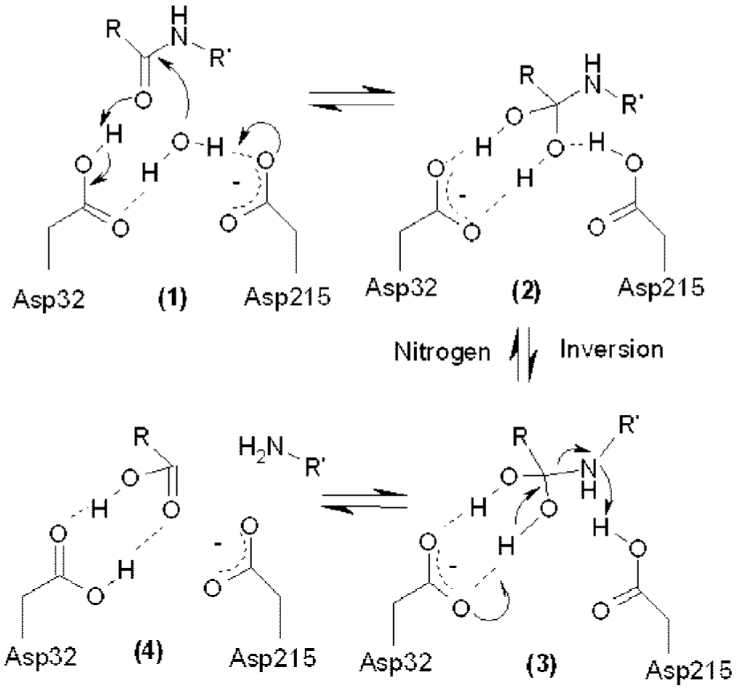

Figure 1.

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 Asp32. Fission of the scissile C—N bond is accompanied by transfer of a proton to the leaving amino group either from Asp215 (with nitrogen inversion) or from bulk solvent. Dashed lines indicate hydrogen bonds or charge sharing (as appropriate).

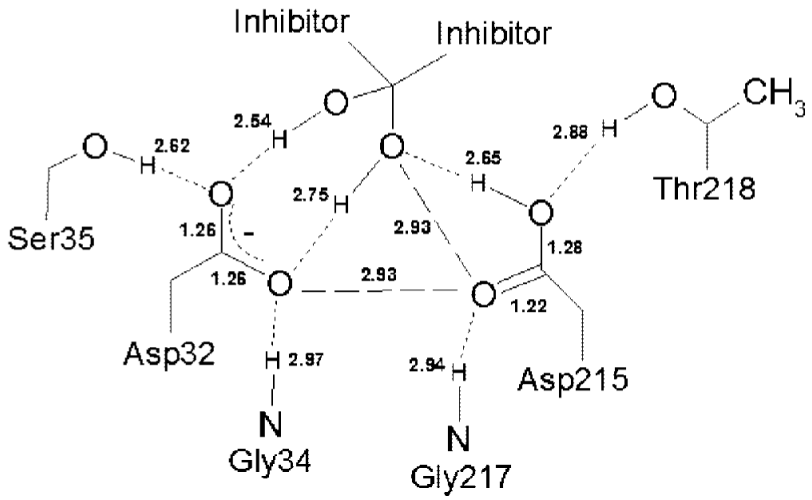

Figure 2.
The bond lengths at the active site for the endothiapepsin—*gem*-diol complex; hydrogen bonds are indicated by the dotted lines. All bond lengths are in Å.

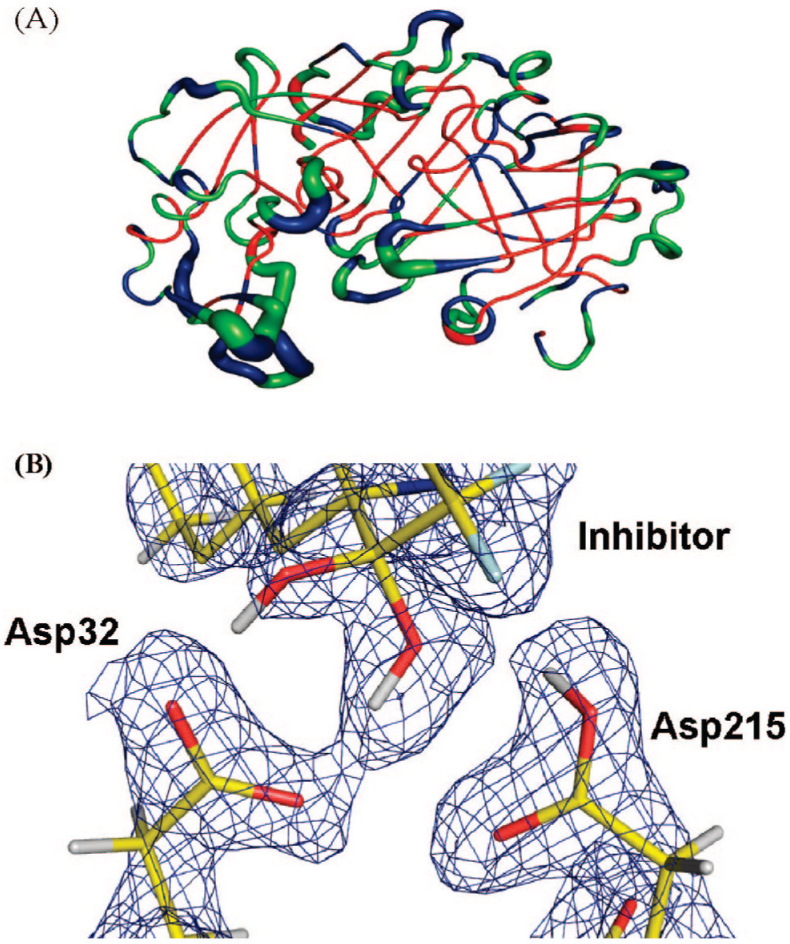

Figure 3.

(a) Backbone H/D exchange and B-factor patterns for the endothiapepsin cocrystallized with PD-135,040. The extent of H/D exchange was divided to three classes based on the refined deuterium occupancy values. Red: no exchange (occupancy = 0–0.15); blue: partially exchanged (occupancy = 0.15–0.70); green: fully exchanged (occupancy = 0.70–1.00). B-factor values are represented by the width of the tube. (b) SigmaA weighted $2F_o — F_c$ nuclear density map at the active site (1.5 rms contour level).