The active site of aspartic proteinases

Laurence Pearl and Tom Blundell

Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England

Received 16 May 1984

The active site of the aspartic proteinase, endothiapepsin, has been defined by X-ray analysis and restrained least-squares refinement at 2.1 Å resolution with a crystallographic agreement value of 0.16. The environments of the two catalytically important aspartyl groups are remarkably similar and the contributions of the NH_2- and COOH-terminal domains to the catalytic centre are related by a local 2-fold axis. The carboxylates of the aspartyls share a hydrogen bond and have equivalent contacts to a bound water molecule or hydroxonium ion lying on the local diad. The main chains around 32 and 215 are connected by a novel interaction involving diad-related threonines. It is suggested that the two pK_a values of the active site aspartyls arise from a structure not unlike that in maleic acid with a hydrogen-bonded intermediate species and a dicarboxylate characterised by electrostatic repulsions between the two negatively charged groups.

Active site *Enzyme conformation* *Protein crystallography* *X-ray analysis* *Aspartic proteinase* *Catalytic mechanism*

1. INTRODUCTION

Aspartic proteinases comprise a group of homologous enzymes including pepsin, chymosin, renin, cathepsin D and related microbial enzymes such as endothiapepsin and penicillopepsin [1]. With the exception of renin, their optimal catalytic activities are at acid pH and all have two aspartates involved in the catalytic mechanism. Medium resolution or incompletely refined X-ray analyses [2–8] indicated that the two aspartates lie at the centre of a deep and extended active site cleft with significantly different environments, such that Asp 32 is ionized and Asp 215 protonated (the residues are numbered by homology with porcine pepsin) [9]. These observations were used as the basis of several mechanistic proposals in which the two aspartates played distinct roles [10,11]. However, completion of the refinement at high resolution has indicated a quite different hydrogen bonding at the active site. In endothiapepsin we find the environments of the two catalytically important aspartyl groups to be equivalent and related by a local 2-fold axis. Similar results have been reported for penicillopepsin [12]. Here, we describe the active site symmetry of endothiapepsin and discuss its implications for the mechanism of aspartic proteinases.

2. MATERIALS AND METHODS

The crystallographic data for endothiapepsin have been described for 5.0 and 2.7 Å resolution analyses [2,5,13]. The data have been extended to 2.1 Å and the structure refined using restrained least-squares. Full details will be described elsewhere [14]. The sequence of tryptic fragments determined by Dr V. Pedersen has been aligned in the electron density; only one fragment involving residues 306–327 has some uncertainty resulting from its hydrophobicity and consequent insolubility. The model has a crystallographic agreement value of 0.16 at 2.1 Å resolution.

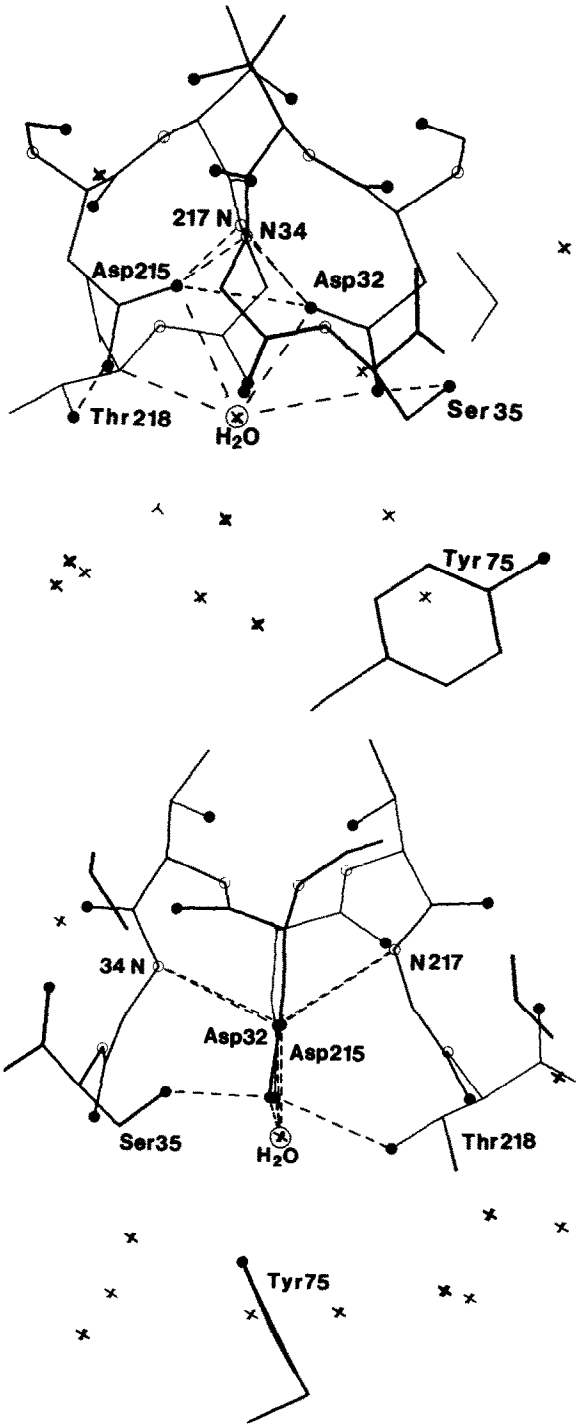
3. RESULTS AND DISCUSSION

The high-resolution X-ray analysis indicates that at pH 4.5 endothiapepsin has an intricate and sym-

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

metrical hydrogen-bond network involving the catalytic aspartyl side chains of 32 and 215 (fig.1a). The aspartyl carboxyl groups lie in a plane with rms deviation <0.1 Å. There is a further peak of electron density lying equidistant to the carboxyls and <0.1 Å out of the plane (fig.1b). This may be a water molecule or, in view of the negative charge shared by the two carboxylates, NH_4^+ or H_3O^+.

The close proximity of the two carboxyls (Oδ2 of Asp 32 is 2.9 Å from Oδ2 of Asp 215) indicates that a proton must lie between them, but the position of the proton in this hydrogen bond cannot be defined by X-ray analysis. The bound water molecule or ion is within hydrogen-bonding distance of all 4 carboxyl oxygens. Both the 'inner' carboxyl oxygens have possible contacts with the peptide nitrogens of two diad-related glycine residues: Gly 34 and Gly 217, although the orientations and distances of these contacts are not favourable for strong hydrogen bonding (see fig.1a). It is possible that the N-H groups of Gly 34 and Gly 217 play an important role in the transition state [15,16]. Although these interacting glycine peptide nitrogens make no other direct contacts, both peptide oxygens, with which they are involved, make hydrogen-bonding contacts: 033 directly to a water molecule, and 0216 to a side chain carboxyl oxygen of Asp 304, the other oxygen of which is hydrogen bonded to a water molecule. The 'outer' carboxyl oxygens make hydrogen-bond contacts with side chain alcohol oxygens: Asp 32 to Ser 35 and Asp 215 to Thr 218.



←

Fig.1. (a) Hydrogen-bonding network in the active site of endothiapepsin involving the similar sequences Asp 32–Thr 33–Gly 34–Ser 35 and Asp 215–Thr 216–Gly 217–Thr 218. Dashed lines indicate possible hydrogen-bond contacts. The view is normal to the plane formed by the carboxyl groups of Asp 32 and Asp 215. The distances (in Å) are:

Oδ1	32	–	water		3.06;
Oδ1	215	–	water		2.98;
Oδ1	32	–	Oγ1	35	2.77;
Oδ1	215	–	Oγ1	218	2.84;
Oδ2	32	–	water		2.77;
Oδ2	215	–	water		3.03;
Oδ2	32	–	N	34	3.27;
Oδ2	215	–	N	34	3.33;
Oδ2	32	–	N	217	3.43;
Oδ2	215	–	N	217	3.01;
Oδ2	32	–	Oδ2	215	2.92

(b) As (a) viewed parallel to carboxyl plane.

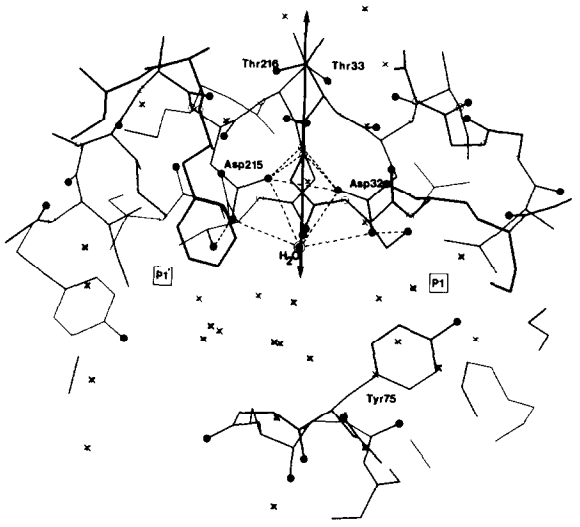


Fig.2. The active site of endothiapepsin, showing the catalytic centre and the groups forming the specificity subsites S_1 and $S_{1'}$. The central line is the best least-squares axis relating the 29 corresponding atoms in residues 31–35 to those in residues 214–218, calculated using the program EZIFIT of Andrew McLachlan. The axis rotation is 178.8° giving an rms fit between the two sides of 0.38 Å. The bound water lies <0.1 Å off the axis.

The symmetrical arrangement of the two aspartates, the surrounding conserved amino acid residues and the bound molecule indicate that the structure may be the average of two structures so that the hydrogen atom occupies positions close to the Oδ2 atoms of Asp 32 and Asp 215 for equal times (see fig.3a,b). There is nothing in the endothiapepsin structure to differentiate between the two aspartates. In a similar way the molecule bound to the carboxylates, while occupying the same positions, may occupy two different hydrogen-bonded orientations (fig.3).

Close by the two aspartyl groups, the side chains of the diad-related threonine groups, Thr 33 and Thr 216, are directed towards the hydrophobic core and are involved in a curious interaction we have called a 'fireman's grip' (where one hand holds the arm whose hand holds its arm). The side chain Oγs (hands) make possible hydrogen-bond contacts with the main chains bearing the corresponding threonines (arms) as shown in fig.4. In the case of Thr 33 the contacts are to the peptide oxygen of 214 and possibly the peptide nitrogen of

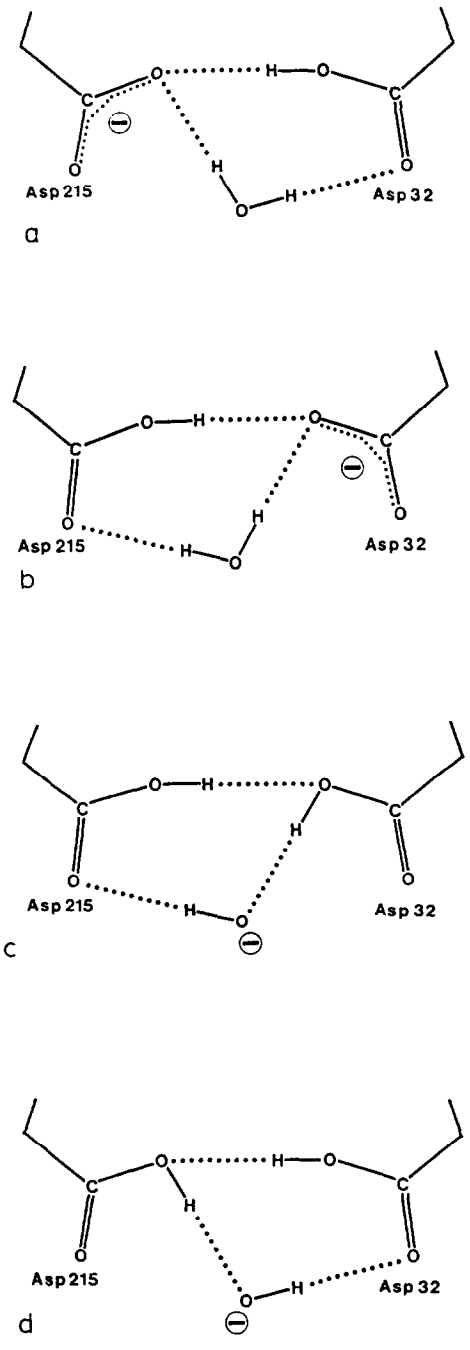


Fig.3. Possible hydrogen-bonded arrangements between the carboxylates of Asp 32 and Asp 215, which are related by a diad, and a water molecule lying on the diad. Although c and d would be expected to make only minor contributions, their effect would be to increase the negative charge on the water oxygen.

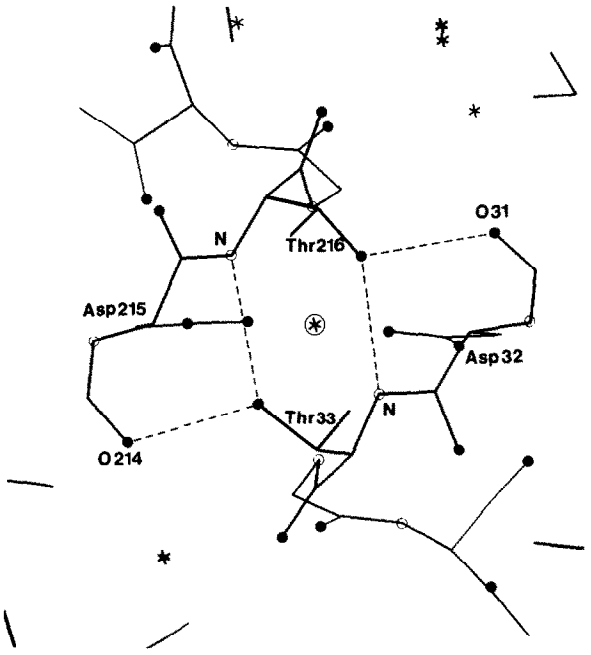


Fig.4. The fireman's grip viewed along the diad. The dotted lines indicate probable hydrogen bonds.

216, while the diad-related Thr 216 contacts 031 and possibly N33. The main chains themselves are out of step and too far apart to make a direct secondary structural interaction of the kind found between strands of β-sheet. In view of the relatively low thermal parameters of the aspartyl groups and their bound solvent, it seems likely that the 'fireman's grip' serves to provide added structural stability for the catalytic centre. The highly symmetrical disposition of the groups in this region of the molecule which is the interface of the two domains supports the suggestion [17] of aspartic proteinase evolution from an ancestral dimer, by a mechanism of gene duplication.

However, aspartic proteinases display bell-shaped pH-activity profiles [18] which have been interpreted as being due to the two catalytic aspartates, one of abnormally high pK_a (~5) and one of abnormally low pK_a (~2). The medium-resolution X-ray analyses [2,3] indicated that this might be due to the fact that Asp 32 was extensively hydrogen bonded and less accessible than Asp 215. Thus Asp 32 was assumed to be deprotonated when the enzyme was active [10]. Even though the hydrogen-bonding arrangement in the highly refined structure of penicillopepsin [12] is quite different from that reported earlier, the authors comment on the 'curious environment' of the peptide bond from Thr 216 to Gly 217 lying between Asp 304 and Asp 215 which must have a 'strong influence on the two catalytic groups'. However, this cannot be general to all aspartic proteinases as neither mouse submaxillary renin nor human renin has a side chain with a carboxyl at residue 304 [19–23].

What then is the origin of the two abnormal pK_a values of aspartic proteinases? The structures of both endothiapepsin and penicillopepsin show conclusively that with a distance of ~2.9 Å between the carboxylate oxygens there must be a hydrogen bond between them. Loss of the proton in this hydrogen bond would result in the close proximity of two negatively charged carboxylates which would be restrained from moving apart by the extensively hydrogen-bonded arrangement. This would increase the free energy of the deprotonated state, make the loss of the proton more difficult and so increase the pK of this proton. With one proton shared in a hydrogen bond between the two carboxylates, the addition of a second proton would be less easy unless the hydrogen bond was broken and the carboxylates displaced in a similar way. This situation parallels that of maleic acid in which pK_1 = 1.91 and pK_2 = 6.33 with a hydrogen-bonded intermediate species and electrostatic repulsion in the dicarboxylate. Therefore we suggest that the pH profile of aspartic proteinases results from abnormal pK_a values of the dicarboxylic acid system as a whole.

If the electron density bound to the two aspartates is an uncharged water molecule, the major contributors to the structure would include those shown in fig.3a,b, with small contributions by those in fig.3c,d. This will have the effect of delocalising the single charge from the two carboxylates onto the water oxygen as long as a proton is bound between the carboxylates, or in other words, when the pH is close to the range of the two pK values of the system. A water bound in this position would then have rather unusual properties. It could be an effective nucleophile or alternatively bind a proton to become an oxonium ion. A distinction between these two possibilities cannot be made on the basis of an X-ray analysis; a

neutron diffraction analysis is required to resolve the ambiguity.

What implications do these observations have for the catalytic mechanism of aspartic proteinases?

Firstly, the replacement of the water or hydroxonium ion capable of hydrogen bonding with the negatively charged aspartates by the carbonyl oxygen of the substrate peptide [24] would be unfavourable. On the other hand, replacement by the hydroxyl of a pepstatin inhibitor which could form a hydrogen bond analogous to that of water would cause no difficulty. There then remain two possible roles for the hydrogen-bonded aspartyl system. If the bound molecule is a water, it can be involved in the nucleophilic attack on the carbonyl carbon of the scissile bond. Alternatively, if it is an oxonium ion it might constitute the electrophilic centre contributing a proton to the amino nitrogen or carbonyl oxygen of the peptide during hydrolysis.

Secondly, the extensive hydrogen bonding of the two threonines of the fireman's grip in a hydrophobic environment implies a high activation energy for a large conformational change such as suggested by authors in [10] in their charge relay mechanism. This mechanism is not supported by the arrangement of Asp 304 and Arg 308 found in the high-resolution studies; in endothiapepsin, like penicillopepsin, they are hydrogen bonded to solvent and the mechanism has now been withdrawn [12]. It is probable that any large conformational changes occurring in the enzyme are confined to the loops on the edge of the active site. The evidence for a conformational change has largely derived from the observation of an increased catalytic efficiency, k_{cat}/K_m, with increased substrate length (the increase being due to k_{cat} rather than to K_m [25]. This may be partly explained by conformational changes in the loops such as the hairpin containing Tyr 75, although this tyrosine is hydrogen bonded to Trp 39 in endothiapepsin and is found in a similar position in the penicillopepsin inhibitor complex [26]. Conformational changes in the substrate could affect the scissile bond; the binding of longer substrates may distort the substrate peptide from planarity. This would have the effect of increasing the single-bond character and decreasing the delocalisation of the lone pair of the peptide nitrogen, features which would favour attack at the carbonyl [27,28].

ACKNOWLEDGEMENTS

We should like to thank Dr M.R. Hollaway of University College, London, for criticism and advice, Dr J.A. Jenkins and Dr B.T. Sewell for their contributions at earlier stages in the project, and Dr I.J. Tickle for the development of programs for use on the Evans and Sutherland interactive computer graphics system. We thank the Science and Engineering Research Council for its financial support.

REFERENCES

[1] Tang, J. (1977) The Acid Proteinases: Structure, Function and Biology, Plenum, New York.
[2] Subramanian, E., Liu, M., Swan, I.D.A., Davies, D.R., Jenkins, J.A., Tickle, I.J. and Blundell, T.L. (1977) Proc. Natl. Acad. Sci. USA 74, 556–559.
[3] Hsu, I.-N., Delbaere, L.T.J. and James, M.N.G. (1977) Nature 266, 140–145.
[4] Andreeva, N.S., Gustchina, A.E., Federov, A.A., Shutzkever, N.E. and Volnova, T.V. (1977) in: Acid Proteinases: Structure, Function and Biology (Tang, J. ed.) pp.23–31, Plenum, New York.
[5] Blundell, T.L., Jenkins, J.A., Khan, G., Roychoudhury, P., Sewell, B.T., Tickle, I.J. and Wood, E.A. (1979) Fed. Eur. Biochem. Soc. 52, 81–94.
[6] Jenkins, J.A., Tickle, I.J., Sewell, B.T., Ungaretti, L., Wollmer, A. and Blundell, T.L. (1977) in: Acid Proteinases: Structure, Function and Biology (Tang, J. ed.) pp.43–60, Plenum, New York.
[7] James, M.N.G. (1980) Can. J. Biochem. 58, 251–271.
[8] James, M.N.G., Hsu, I.-N., Hofmann, T. and Sielecki, A. (1981) in: Structural Studies on Molecules of Biological Interest (Dodson, G., Glusker, J.P. and Sayer, D. eds) pp.350–389, Clarendon, Oxford.
[9] Tang, J., Sepulveda, P., Marciniszyn, J., Chen, K.C.S., Huang, W.-Y., Tao, N., Liu, D. and Lanier, J.P. (1973) Proc. Natl. Acad. Sci. USA 70, 3437–3439.
[10] James, M.N.G., Hsu, I.-N. and Delbaere, L.T.J. (1977) Nature 267, 808–813.
[11] Blundell, T.L., Jones, H.B., Khan, G., Taylor, G., Sewell, B.T., Pearl, L.H. and Wood, S.P. (1980) in: Enzyme Regulation and Mechanism of Action (Mildner, P. and Ries, B. eds) pp.281–288, Pergamon, Oxford.
[12] James, M.N.G. and Sielecki, A. (1983) J. Mol. Biol. 163, 299–361.

[13] Jenkins, J.A., Blundell, T.L., Tickle, I.J. and Ungaretti, L. (1977) J. Mol. Biol. 99, 583–590.
[14] Pearl, L.H., Jenkins, J.A., Sewell, B.T., Pedersen, V. and Blundell, T.L. (1984) in preparation.
[15] Kirkwood, J.G. and Westheimer, F.H. (1938) J. Chem. Phys. 6, 506–513.
[16] Hollaway, M. (1984) personal communication.
[17] Tang, J., James, M.N.G., Hsu, I.-N., Jenkins, J.A. and Blundell, T.L. (1978) Nature 271, 618–521.
[18] Fruton, J.S. (1981) in: The Enzymes (Boyer, P.D. ed.) vol.3, pp.119–164, Academic Press, London.
[19] Misono, K.S., Chang, J.-J. and Inagami, T. (1982) Proc. Natl. Acad. Sci. USA 79, 4858–4862.
[20] Panthier, J.-J., Foote, S., Chambraud, B., Strosberg, A.D., Corvol, P. and Rougeon, F. (1982) Nature 298, 90–92.
[21] Imai, T., Miyazaki, H., Hirose, S., Hori, H., Hayashi, T., Kageyama, R., Ohkubo, H., Nakanishi, S. and Marukami, K. (1983) Proc. Natl. Acad. Sci. USA 80, 7405–7409.
[22] Soubrier, F., Panthier, J.-J., Corvol, P. and Rougeon, F. (1983) Nucl. Acid Res. 11, 7181–7190.
[23] Hobart, P.M., Fogliano, M., O'Connor, B.A., Schaefer, I.M. and Chirgwin, J.M. (1980) Proc. Natl. Acad. Sci. USA, in press.
[24] Bott, R., Subramanian, E. and Davies, D.R. (1982) Biochemistry 21, 6956–6962.
[25] Fruton, J.S. (1976) Adv. Enzymol. Relat. Areas Mol. Biol. 44, 1–36.
[26] James, M.N.G., Sielecki, A., Salituro, F., Rich, D.H. and Hofmann, T. (1982) Proc. Natl. Acad. Sci. USA 79, 6137–6141.
[27] Bernhard, S.A. and Lau, S.-J. (1971) Cold Spring Harb. Symp. Quant. Biol. 36, 75–83.
[28] Kuo, L.C., Fukuyama, J.M. and Makinen, M.W. (1983) J. Mol. Biol. 163, 63–105.