

**J. Guo, J. B. Cooper* and
S. P. Wood**Laboratory of Protein Crystallography, Centre for
Amyloidosis and Acute Phase Proteins, UCL
Division of Medicine (Royal Free Campus),
Rowland Hill Street, London NW3 2PF, England

Correspondence e-mail: jon.cooper@ucl.ac.uk

Received 11 November 2013

Accepted 5 December 2013

PDB reference: endothiapepsin–FRY complex,
4lp9

The structure of endothiapepsin complexed with a Phe-Tyr reduced-bond inhibitor at 1.35 Å resolution

Endothiapepsin is a typical member of the aspartic proteinase family. The catalytic mechanism of this family is attributed to two conserved catalytic aspartate residues, which coordinate the hydrolysis of a peptide bond. An oligopeptide inhibitor ($IC_{50} = 0.62 \mu M$) based on a reduced-bond transition-state inhibitor of mucorpepsin was co-crystallized with endothiapepsin and the crystal structure of the enzyme–inhibitor complex was determined at 1.35 Å resolution. A total of 12 hydrogen bonds between the inhibitor and the active-site residues were identified. The resulting structure demonstrates a number of novel subsite interactions in the active-site cleft.

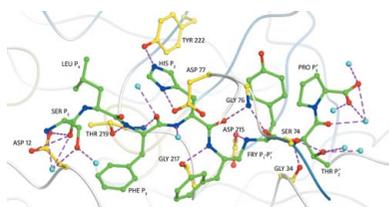
1. Introduction

The aspartic proteinases are a class of enzymes that are distributed broadly in nature. They play important roles in malaria, amyloid disease, hypertension and tumourigenesis, and are also crucial to the maturation of virions (Coates *et al.*, 2002). The successful use of aspartic proteinase inhibitors in HIV/AIDS and hypertension treatments has increased interest in drug discovery against these enzymes.

Endothiapepsin is derived from the fungus *Endothia parasitica* and, like other aspartic proteinases, it catalyzes the hydrolysis of peptide bonds through the action of its two aspartate residues, numbered Asp32 and Asp215 in porcine pepsin, an enzyme that was regarded as the family archetype for many years (Blundell *et al.*, 1990). Endothiapepsin has a similar substrate preference to porcine pepsin A. Both enzymes prefer hydrophobic residues at P_1 – P'_1 in the cleavage site. Williams *et al.* (1972) analyzed the cleavage sites in insulin and found the cleavage rates of endothiapepsin to be Phe-Phe > Tyr-Leu > Gln-His >>> Leu-Val > Asn-Gln. In all aspartic proteinases, the two catalytic aspartate carboxyls are held coplanar by a hydrogen-bond network formed with the surrounding conserved side-chain and main-chain groups. A catalytic solvent water molecule is found hydrogen-bonded tightly between the carboxyls of the two aspartates (Fig. 1). This water molecule, which is within hydrogen-bonding distance of all four carboxyl O atoms, can nucleophilically attack the substrate scissile-bond carbon group and generate a tetrahedral intermediate (Suguna *et al.*, 1987; Veerapandian *et al.*, 1992). The tetrahedral intermediate can collapse and produce the hydrolysis products.

The most successful designed inhibitors of aspartic proteinases are those that mimic the tetrahedral intermediate in the transition state. One hydroxyl group of the intermediate forms hydrogen bonds to both of the aspartate residues in the active site. X-ray crystal structure analysis shows that the water molecule which participates in the catalytic process is displaced by the transition-state mimic moiety (Suguna *et al.*, 1987). Compounds synthesized based on this theory bind much more tightly compared with the substrate. These compounds are named transition-state analogues (Coates *et al.*, 2002).

The inhibitor used in this study is an eight-residue oligopeptide with a reduced peptide bond between a Phe and a Tyr which was designed to act on *Mucor pusillus* pepsin based on a study of its cleavage sites in insulin (Scott, 1979). Its sequence and chemical structure are shown in Fig. 2. The crystal structure of the endothiapepsin complex was determined at 1.35 Å resolution and a comparison with other inhibitor structures was undertaken.



2. Crystallization

Crude endothiapepsin is prepared commercially as a fungal rennet from culture lysates of *Endothia* (or *Cryphonectria*) *parasitica*, the causative agent of American chestnut blight. The enzyme was purified by ammonium sulfate precipitation followed by ion-exchange and gel-filtration chromatography as described by Whitaker (1970) and was concentrated to 3.5 mg ml^{-1} using a Vivaspin centrifugal concentrator. The inhibitor, which was synthesized using solid-phase

methods as described in the Supplementary Material of Bailey *et al.* (2012), was a kind gift from Drs P. Štrop and M. Souček, Institute of Organic Chemistry and Biochemistry, Prague, Czech Republic. A turbidimetric κ -casein cleavage assay confirmed that the inhibitor had extremely high inhibition activity against purified endothiapepsin, with an IC_{50} of $0.62 \mu\text{M}$. The hanging-drop method was used for crystallization of the complex, with a stock solution of 0.1 M sodium acetate pH 4.0–5.2 as buffer and 43–67% saturated ammonium sulfate as precipitant. Milli-Q water was then used to make the final

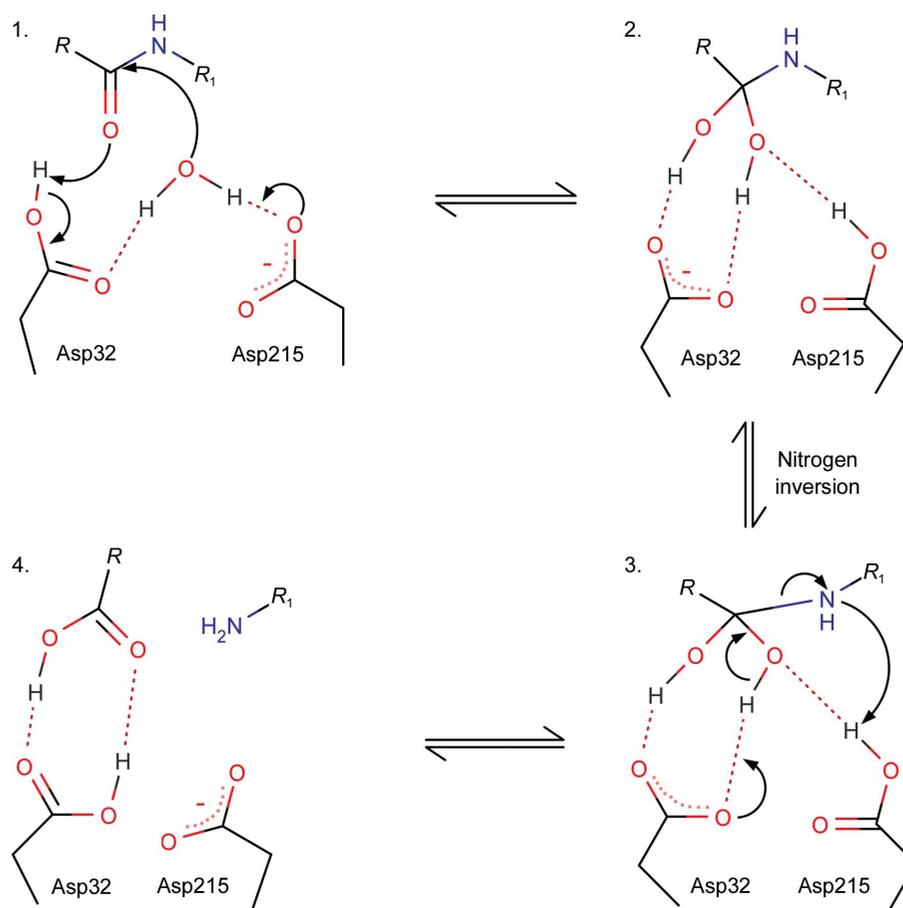


Figure 1

The suggested catalytic mechanism of endothiapepsin. A water molecule bound to Asp32 and Asp215 nucleophilically attacks the scissile-bond carbonyl. A tetrahedral intermediate is then generated and stabilized by hydrogen bonds. The C–N bond is cleaved by transferring a proton to the amino group. Hydrogen bonds are indicated by dotted lines.

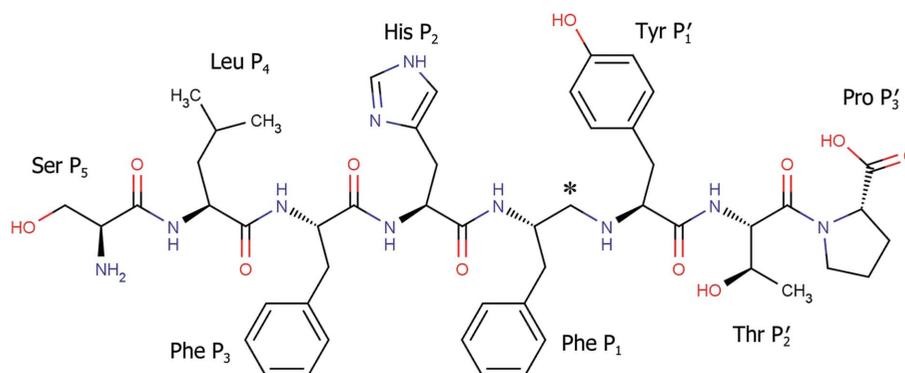


Figure 2

The chemical structure of the inhibitor (FRY) co-crystallized with endothiapepsin. The sequence of the oligopeptide is Ser-Leu-Phe-His-Phe*^{Tyr}-Thr-Pro. There is a reduced peptide bond (indicated by an asterisk) between Phe and Tyr, *i.e.* the carbonyl group is replaced by a methylene $-\text{CH}_2-$ moiety.

Table 1

X-ray statistics for the endothiapepsin-FRY complex structure.

Values in parentheses are for the outer resolution shell.

Data collection	
Beamline	I02, DLS
Wavelength (Å)	1.043
Space group	$P2_1$
Unit-cell parameters	
a (Å)	52.6
b (Å)	72.8
c (Å)	45.1
β (°)	108.5
Mosaic spread (°)	0.22
Resolution range (Å)	28.32–1.35 (1.38–1.35)
R_{merge} (%)	4.7 (57.1)
Data completeness (%)	98.2 (84.1)
Mean $I/\sigma(I)$	14.8 (2.1)
Multiplicity	3.3 (2.3)
No. of observed reflections	230441 (10170)
No. of unique reflections	69622 (4381)
Refinement	
R factor (%)	12.7
R_{free} (%)	16.6
R.m.s.d., bond lengths (Å)	0.022
R.m.s.d., bond angles (°)	2.138

volume in each well 1 ml. The drops consisted of 6 μl protein solution containing 1 mM inhibitor and 4 μl well solution. Crystals started to appear after 5 d and most of the crystals were obtained after 10 d (Fig. 3). The optimal condition was found to be pH 4.9 and 63% saturated ammonium sulfate precipitant. Since the crystals grew as large clusters of plates (Fig. 3), it was necessary to break them into smaller pieces with a needle tip prior to mounting. The selected crystals were cryoprotected by gradual addition of glycerol to a final concentration of 30% (v/v) and mounted in loops before flash-cooling and storage in liquid nitrogen.

3. Data collection and processing

X-ray data were collected on beamline I02 at Diamond Light Source (DLS), Oxfordshire, England with a wavelength of 1.043 Å and an oscillation angle of 1°. A total of 190° of data were collected from a single piece of crystal using a Pilatus 6M-F detector. The exposure time per image was 2 s and the crystal-to-detector distance was 221.9 mm. Data integration was performed with *MOSFLM* (Leslie, 2006; Powell *et al.*, 2013), while data scaling and merging were

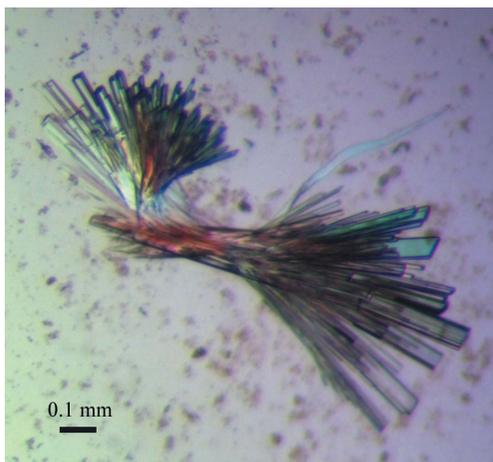


Figure 3

A cluster of crystals of endothiapepsin complexed with the inhibitor obtained by the hanging-drop method. The single-crystal fragment used for data collection had dimensions of approximately $200 \times 100 \times 50 \mu\text{m}$.

performed with *SCALA* (Evans, 2006; Evans & Murshudov, 2013) and other programs in the *CCP4* suite (Winn *et al.*, 2011). Data processing indicated that the space group of the enzyme crystal was $P2_1$, and the unit-cell parameters were $a = 52.6$, $b = 72.8$, $c = 45.1$ Å, $\beta = 108.5^\circ$ with one molecule per asymmetric unit (see Table 1 for details).

4. Structure determination and refinement

Since the crystals of the inhibitor complex are isomorphous to those of native endothiapepsin (Blundell *et al.*, 1990), the inhibitor was built into a difference Fourier map weighted according to Read (1986) that was obtained after refinement of the protein moiety with *REFMAC* (Murshudov *et al.*, 1997, 2011) had been performed using data for the complex. The structure was rebuilt manually following each subsequent refinement round using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). The geometric restraints for refinement of the inhibitor compound were generated using *PRODRG* (Schüttelkopf & van Aalten, 2004). The structure was refined with anisotropic temperature factors at 1.35 Å resolution to an R factor of 12.7% and an R_{free} of 16.6%. All of the amino acids in the enzyme model are within the allowed regions of the Ramachandran plot and the full length of the inhibitor was found to fit the active site of the enzyme and the electron-density map well when contoured at 1 r.m.s. (Fig. 4, generated using *CueMol*; <http://cuemol.sourceforge.jp/en/>).

5. Structure of the complex

The overall structure is similar to those of other endothiapepsin-inhibitor complexes, in particular that with the inhibitor H256

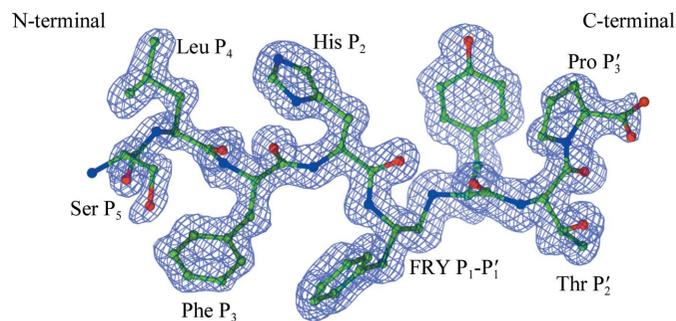


Figure 4

$2F_o - F_c$ map of the refined inhibitor bound in the active site contoured at 1.0 r.m.s. Most parts of the inhibitor fitted the map quite well except for two small parts of the end residues.

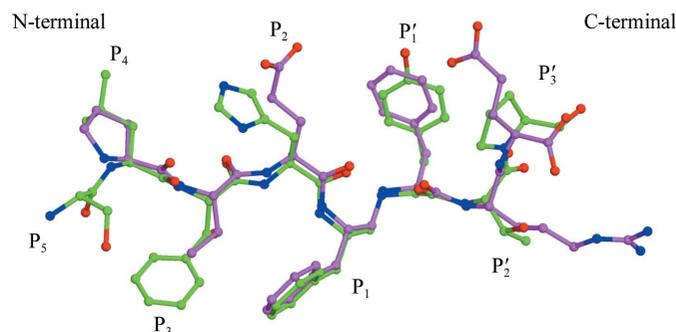


Figure 5

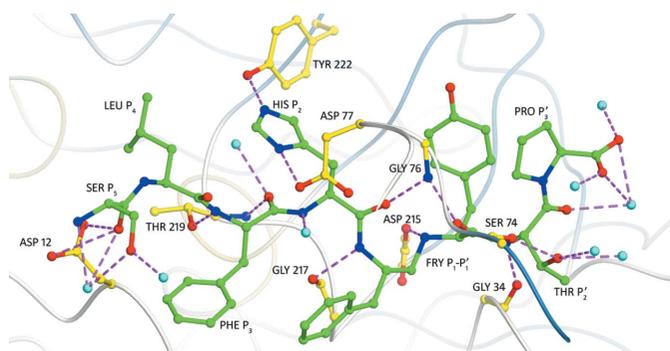
A structural comparison of the inhibitors FRY and H256. The C atoms of the FRY inhibitor are coloured green, while those of H256 are coloured pink.

Table 2

Hydrogen bonds formed by the inhibitors FRY and H256 to the active-site residues.

The amino acids are numbered according to the scheme of Blundell *et al.* (1990).

FRY		H256	
Hydrogen bond	Length (Å)	Hydrogen bond	Length (Å)
Ser P ₅ O–Asp12 OD1	3.42	Pro P ₄ N–Asp12 OD1	3.40
Ser P ₅ O–Asp12 OD2	3.48	Thr P ₃ OG1–Gly219 O	2.86
Phe P ₃ N–Thr219 OG1	2.95	Thr P ₃ N–Thr219 OG1	2.91
Phe P ₃ O–Thr219 N	3.00	Thr P ₃ O–Thr219 N	3.08
His P ₂ O–Gly76 N	3.07	Thr P ₃ O–Thr219 OG1	3.48
His P ₂ ND1–Asp77 OD2	3.13	Glu P ₂ O–Gly76 N	3.03
His P ₂ NE2–Tyr222 OH	3.29	Glu P ₂ N–Asp77 OD2	2.97
FRY P ₁ –P ₁ ' O–Gly76 N	3.04	FRF P ₁ –P ₁ ' O–Gly76 N	3.08
FRY P ₁ –P ₁ ' N–Asp215 OD2	2.71	FRF P ₁ –P ₁ ' N2–Asp215 OD2	2.78
FRY P ₁ –P ₁ ' NAA–Gly217 O	2.88	FRF P ₁ –P ₁ ' N Gly217 O	2.94
Thr P ₂ ' N–Gly34 O	2.98	Arg P ₂ ' N–Gly34 O	2.95
Thr P ₂ ' OG1–Ser74 O	2.77	Arg P ₂ ' NH1–Leu128 O	2.95
		Glu P ₂ ' N–Ser74 O	2.95

**Figure 6**

The enzyme and the residues which form hydrogen bonds to the inhibitor. The inhibitor conformation is stabilized by these 12 hydrogen bonds, which are indicated by purple dotted lines. The C atoms of the inhibitor are coloured green, while the C atoms of the enzyme residues are coloured yellow. The thin tube shows the tertiary structure of the enzyme.

(Foundling *et al.*, 1987; Coates *et al.*, 2002), which is shown superposed in Fig. 5. H256 possesses a Phe-reduced-bond-Phe analogue (FRF) instead of the Phe-reduced-bond-Tyr of the compound reported here (FRY). A total of 12 hydrogen bonds were identified between the inhibitor and the active-site residues of the enzyme (Table 2); nine water molecules were also found to form hydrogen bonds to the inhibitor (Fig. 6). The bound conformation of the inhibitor is well stabilized by these hydrogen bonds and no disordered residues were found within the hydrophobic pockets of the enzyme. Three sulfate ions which made interactions with the enzyme and/or with water molecules were also found and are present in the deposited structure.

Since changing the Phe residue of H256 to Tyr in FRY gives the inhibitor a slightly more polar character at P₁', it was assumed that more interactions of this nature would be found. However, no such interactions were found with the new inhibitor compared with H256 (Table 2), which is also a reduced-bond inhibitor (Foundling *et al.*, 1987). The presence of Ser at the P₅ site gives the inhibitor two more hydrogen bonds to Asp12. It is interesting that the Leu at P₄ of the new inhibitor is accommodated as well in the pocket as the Pro in

H256. Replacement of the Thr at the P₃ site of H256 with a Phe in FRY gives the inhibitor less hydrogen-bonding potential, but two hydrogen bonds to the enzyme are still formed and the hydrophobic S₃ pocket is better filled by the aromatic P₃ side chain of FRY. At the P₂ site, replacement of the Glu of H256 by a His in FRY gives the inhibitor one more hydrogen bond. Both the FRY group in the inhibitor and the FRF group in H256 can form three hydrogen bonds to the active-site residues. As can be seen in the superposition shown in Fig. 6, the Tyr side chain at P₁' seems to be prevented from going as deeply into the S₁' pocket as it might. We attribute this to the bulky side chains of the Ile residues 297, 299 and 301 which reside in the loop located to the upper right of the P₁' Tyr side chain from the view shown in Fig. 5. The same isoleucine-rich loop of the enzyme is thought to be important for its specificity (Dhanaraj *et al.*, 1992). The P₂' Thr is also interesting because its size is more compatible with the compactness of the pocket than the P₂' Arg in H256. The hydrophobic and rigid side chain of the P₃' Pro seems to shield the P₁' Tyr side chain from solvent quite well and presumably contributes to the high affinity of this inhibitor for the enzyme. The refined structure has been deposited in the Protein Data Bank as PDB entry 4lp9.

We gratefully acknowledge the Diamond Light Source for X-ray beam time and travel support during data collection (award MX-7131). We are grateful to Dr Peter Štrop and the late Dr Milan Souček (Institute of Biochemistry and Molecular Biology, Prague, Czech Republic) for synthesis of the inhibitor.

References

- Bailey, D. *et al.* (2012). *Acta Cryst.* **D68**, 541–552.
- Blundell, T. L., Jenkins, J. A., Sewell, B. T., Pearl, L. H., Cooper, J. B., Tickle, I. J., Veerapandian, B. & Wood, S. P. (1990). *J. Mol. Biol.* **211**, 919–941.
- Coates, L., Erskine, P. T., Crump, M. P., Wood, S. P. & Cooper, J. B. (2002). *J. Mol. Biol.* **318**, 1405–1415.
- Dhanaraj, V. *et al.* (1992). *Nature (London)*, **357**, 466–472.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- Foundling, S. I. *et al.* (1987). *Nature (London)*, **327**, 349–352.
- Leslie, A. G. W. (2006). *Acta Cryst.* **D62**, 48–57.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Powell, H. R., Johnson, O. & Leslie, A. G. W. (2013). *Acta Cryst.* **D69**, 1195–1203.
- Read, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
- Schüttelkopf, A. W. & van Aalten, D. M. F. (2004). *Acta Cryst.* **D60**, 1355–1363.
- Scott, R. (1979). *Topics in Enzyme and Fermentation Biotechnology*, edited by A. Wiseman, Vol. 3, pp. 103–169. Harlow: Ellis Horwood.
- Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D. & Davies, D. R. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 7009–7013.
- Veerapandian, B., Cooper, J. B., Šali, A., Blundell, T. L., Rosati, R. L., Dominy, B. W., Damon, D. B. & Hoover, D. J. (1992). *Protein Sci.* **1**, 322–328.
- Whitaker, J. R. (1970). *Methods Enzymol.* **19**, 436–445.
- Williams, D. C., Whitaker, J. R. & Caldwell, P. V. (1972). *Arch. Biochem. Biophys.* **149**, 52–61.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.