# Monomeric human cathepsin E

Sylvia D. Fowler[a], John Kay[a,*], Ben M. Dunn[b], Peter J. Tatnell[a]

[a]*School of Molecular and Medical Biosciences, University of Wales College of Cardiff, Museum Avenue, PO Box 911, Cardiff CF1 3US, UK*
[b]*Department of Biochemistry and Molecular Biology, J. Hillis Miller Health Center, University of Florida, Gainesville, FL 32610, USA*

**Abstract** **Cathepsin E is a homodimer, consisting of two monomers linked by an inter-molecular disulphide bond. The cysteine residue involved is located near to the N-terminus of the mature proteinase. By mutating this residue to alanine, a monomeric form of human cathepsin E was engineered and purified. The activity of the resultant enzyme was not altered significantly (in terms of its ability to hydrolyse two chromogenic peptide substrates; and its susceptibility to inhibition by pepstatin). However, the stability of the mutant enzyme to alkaline pH and to temperature was markedly reduced.**

*Key words:* Human cathepsin E; Aspartic proteinase; Homodimer; Mutation; Engineered monomer; Activity; Stability

## 1. Introduction

Aspartic proteinases have been isolated and studied from a wide variety of organisms, varying from vertebrates to plants, fungi, parasites and retroviruses [1]. Of the five known to be present in the human body, three (pepsin, gastricsin and renin) are secretory enzymes while one (cathepsin D) is located within the lysosomes. The fifth proteinase, cathepsin E is neither a secretory enzyme nor is it located in the lysosomes. The crystal structures that have been determined for pepsin [2], renin [3] and cathepsin D [4] indicate that these enzymes are closely similar not just to each other but also to aspartic proteinases of vertebrate/fungal or parasitic origin. In their mature forms, these non-viral aspartic proteinases have $M_r = {\sim}40K$ and are (commonly) single polypeptide chains. In this regard too, cathepsin E is distinct since it has a molecular mass of ~84 kDa. Located close to the N-terminus of the mature enzyme in the sequence Thr-Glu-Ser-Cys-Ser~, cathepsin E has a unique cysteine residue that is not present in any other aspartic proteinase [1]. It has been postulated that by formation of an intermolecular-disulphide bond, this residue brings about dimerisation of two identical 42 kDa subunits to give rise to the ~84 kDa form of the mature enzyme. In this report, it is established for the first time unequivocally that this is indeed the case and the significance of this unique cysteine residue is investigated in terms of its influence on the activity and stability of cathepsin E.

## 2. Materials and methods

Cathepsin E was obtained from human tissues as described previously [5] and is referred to as naturally-occurring enzyme in the text.

The cDNA encoding human procathepsin E was generated as reported in [6]. Synthetic oligonucleotides were purchased from Genosys Biotechnologies Inc., Cambridge, UK.

The Cys[4]-Ala mutation (and a convenient *Nhe*I restriction site) were introduced by overlapping PCR using the 5'-forward
    (5' AG GAA TTC CAT ATG GGA TCC CTT CAC AGG 3' – Primer 1)
and 3'-reverse
    (5' CCC AGA TCT TTA GGG GAC TGC TGG GGC 3' – Primer 2)
oligonucleotides as primers and the complementary pair of primers
    (5' GTC CAT GCT AGC GGA CTC GGT GAA CTG GAT 3' – Primer 3)
    5' GAG TCC GCT AGC ATG GAC CAG AGT GCC AAG 3' – Primer 4.

DNA amplifications were performed using 20 ng of the cDNA encoding procathepsin E [7], 2 units of Vent DNA polymerase (New England Biolabs, Beverly, MA, USA), 2 mM MgSO$_4$, 1 mM of each primer and 10 mM dNTPs in a final volume of 100 ml. Amplification was achieved by using a Biometra Trio (Maidstone, Kent, UK) programmable thermoblock. An initial denaturation step of 94°C for 3 min was followed by 20 cycles of: 94°C for 1 min, 58°C for 1 min and 72°C for 1.5 min, after which a final elongation step of 72°C for 5 min was performed.

Amplifications involving the primer combinations 1 with 3 and 2 with 4 resulted in the generation of two products of approximately 150 and 1000 bp, respectively, which were each cloned into pUC18 using the Sureclone Ligation Kit (Pharmacia Biotech. Ltd., Milton Keynes, UK). Authentication of the nucleotide sequence of both cloned products was achieved by dideoxy sequencing using a Pharmacia ALF automated sequencer. The 150 and 1000 bp fragments were subsequently excised from pUC18 and via a 3-way ligation reaction were sub-cloned into the *Nde*I/*Bam*HI sites of the T7-expression vector pET16b (AMS Biotech. Ltd., Witney, Oxon, UK). Positive recombinant molecules containing the unique *Nhe*I restriction site were sequenced to confirm the continuity of the procathepsin E open reading frame with respect to the initiation methionine residue provided by the expression vector pET16b.

Expression of the Cys[4]-Ala procathepsin E mutant was carried out in *E. coli* BL21(DE3) pLysS as described previously for wild-type protein [6] except for the following modifications which increased the final yield of purified (mutant) enzyme by approximately 4-fold. All inductions were initiated by the addition of 1 mM IPTG. The initial lysed cell debris was subjected to an additional wash of 100 mM Tris-HCl, pH 11.0, supplemented with 50 mM mercaptoethanol prior to the solubilisation step with 6 M urea. Renaturation of the solubilised material was achieved by a 200-fold dilution of the denaturant in 10 mM Tris-HCl buffer, pH 9.5, and the resultant solution was re-concentrated using a Filtron Ultrasette, 5 kDa cut-off, tangential flow concentrator (Flowgen Instruments Ltd., Sittingbourne, Kent, UK).

Analyses of samples by SDS-PAGE and Western blotting and proteolytic activity measurements against haemoglobin as a substrate were carried out as described in [5,6], as was N-terminal sequence analysis by Edman degradation and determination of kinetic parameters for chromogenic peptide substrate hydrolysis and inhibitor binding.

Measurements of proteinase stability were made by incubating samples of the mutant or naturally-occurring enzymes at different temperatures in buffers of a variety of pH values (Tris, pH 7.5 and 8.5; CAPS, pH 9.5 and 10.5). All were used at a final ionic strength of 0.1 M. At appropriate times, aliquots were removed and residual activity was measured at pH 3.1 using the chromogenic substrate Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu as described [5,6]. The influence of glycerol was also assessed by its inclusion in the different pH incubation buffers.

*Corresponding author. Fax: (44) (1222) 87-4116.

## 3. Results and discussion

Following mutation of the cDNA encoding procathepsin E, the presence of the alteration in the clone was verified by dideoxy sequencing in both strands. Subsequently, the $Cys^4$-Ala mutant DNA was manipulated into the expression plasmid pET-16b. The recombinant plasmid was transformed into *E. coli* and induced with IPTG. After 2 h, the cells were harvested, lysed and the mutant zymogen was extracted with urea and re-folded by rapid dilution at pH 9.5, as described in section 2. After concentration, the pH of the resultant protein solution was adjusted to 3.1 and the mixture was incubated for 1 h at 37°C. Following centrifugation, the supernatant was neutralised, centrifuged again and aliquots were analysed by SDS-PAGE (Fig. 1). A single band of protein was observed with mol. wt. close to 40K under both reducing (Fig. 1, Lane 2) and non-reducing (Lane 4) conditions. In contrast, analysis of wild-type cathepsin E under the same conditions revealed that, whereas the protein did migrate as a monomeric species in the presence of reducing agent (Fig. 1, Lane 1), when the SDS-PAGE was performed in the absence of EtSH most of the protein was present in the dimeric form of approximately 84 kDa (Fig. 1, Lane 3). Naturally-occurring (pro)cathepsin E contains one -Asn-X-Thr motif [7,8] and is thus a glycoprotein as isolated from human tissues. The influence of the carbohydrate on mobility is evident on SDS-PAGE under reducing conditions, since the monomeric naturally-occurring enzyme (Lane 1) migrated with a slightly higher mol. wt. than that of the recombinant mutant produced in *E. coli* (Lane 2).

The mutant enzyme thus produced was subjected to 20 cycles of Edman degradation. This confirmed that mature enzyme had been generated by (complete) autoactivation of the re-folded zymogen during the incubation at pH 3.1 (see above). However, a unique sequence was not observed; rather three overlapping sequences were elucidated, and were present in the approximate ratios of 2:1:1. These were:

```
T E S A S M D Q S A K E P L I N Y L D ~
    S A S M D Q S A K E P L I N Y L D ~
          D Q S A K E P L I N Y L D ~
```

This microheterogeneity at the N-terminus of the mutant enzyme is in contrast to the situation observed with recombinant wild-type cathepsin E where a unique sequence of TESC~ was determined [6]. Thus, it would appear that the presence of cysteine-4 and its ability to bring about dimerisation through
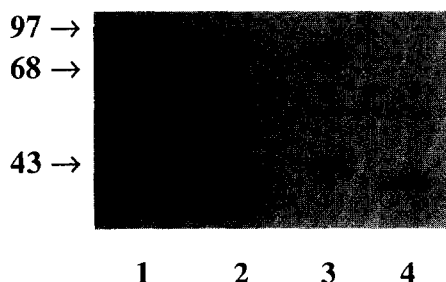


Fig. 1. SDS-PAGE of samples containing cathepsin E. Samples (containing approx. 1 mg of protein) were analysed under reducing conditions (in the presence of 50 mM mercaptoethanol) for Lanes 1 and 2 or in the absence of reducing agent (Lanes 3 and 4). Naturally-occurring human cathepsin E was loaded in Lanes 1 and 3 while the $Cys^4$-Ala mutant enzyme was applied in Lanes 2 and 4. Staining was with Coomassie blue and mol. wt. markers (kDa) migrated as indicated.
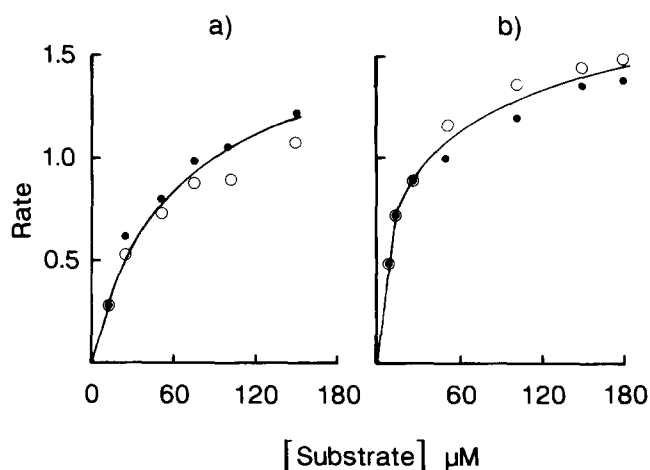


Fig. 2. Influence of reducing agent on the substrate dependence for hydrolysis by naturally-occurring cathepsin E. Measurements of initial velocity (pmol/s × $10^{-2}$) were made at differing concentrations of substrate in the absence (●) and presence (○) of reducing agent. In (a) the substrate was Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu, pH was 3.1 and dithiothreitol (20 mM) was the reducing agent; in (b) the substrate was Pro-Pro-Thr-Ile-Phe*Nph-Arg-Leu, the pH was 5.8, 6.25 mM ATP [15] was included in all assays and mercaptoethanol (50 mM) was the reducing agent.

disulphide bond formation has an influence on the process of autoactivation, restricting the autolytic cleavage to bond(s) upstream from the disulphide bond. In this regard, the naturally-occurring mature form of human cathepsin E is noticeably longer (by ~8 residues) than other aspartic proteinases of vertebrate and fungal origin [6,8,9]. However, when the cysteine is no longer present (replaced by Ala in the mutant), the autoactivation processes by which zymogens of aspartic proteinases are progressively diminished in size from the N-terminal end until the mature enzyme is produced [9], are apparently able to proceed further to produce protein species that are shorter in length at their N-terminal end and are thus more akin in size to other aspartic proteinases, e.g. pepsin and cathepsin D.

To date, (full-length gene) sequences have been elucidated for cathepsin E of human [7], guinea pig [10], rabbit [8] and rat [11] origin and a partial protein sequence has been determined for monkey cathepsin E (B.M.D., J.K., S. Rapundalo and B. Batley, unpublished observations).

| Human: | T E S C S ~ |
|---|---|
| Monkey: | T E S C S ~ |
| Rabbit: | T E T C T ~ |
| Guinea pig: | - - - C S ~ |
| Rat: | S E S C N ~ |

In all cases, the cysteine (at residue 4) in the sequence is conserved and the enzymes exist in the tissues of their respective species predominantly in dimeric form. In order to determine whether the presence of the interchain disulphide link is important for activity, the ability of the $Cys^4$-Ala mutant cathepsin E to hydrolyse two synthetic substrates was examined. The kinetic parameters obtained (Table 1) for each of the two substrates were not substantially different from those reported previously for the hydrolysis of these chromogenic peptides by recombinant wild-type [6] and naturally-occurring [5] human cathepsin E. For example, the corresponding $k_{cat}/K_m$ values measured for the two substrates with wild-type recombinant

Table 1
Kinetic constants for the hydrolysis of chromogenic substrates by the Cys[4]-Ala mutant cathepsin E

|  | $K_m$ ($\mu$M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($mM^{-1} \cdot s^{-1}$) |
|---|---|---|---|
| Pro-Pro-Thr-Ile-Phe*Nph-Arg-Leu | 45 | 70 | 1,550 |
| Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu | 40 | 82 | 2,050 |

Measurements were made in 0.1 M sodium formate buffer, pH 3.1, at 37°C. The estimated precision of the values obtained was in the range 10–15%.

enzyme were 2,300 and 1,850 $mM^{-1} \cdot s^{-1}$ [6] compared to the values of 1,550 and 2,050 $mM^{-1} \cdot s^{-1}$ determined for the mutant (Table 1). Furthermore, the Cys[4]-Ala mutant of cathepsin E was just as susceptible ($K_i < 0.1$ nM) to inhibition by pepstatin as its wild-type recombinant and naturally-occurring counterparts. In further substantiation of this, the influence of reducing agent on the hydrolysis of the two chromogenic substrates by naturally-occurring human cathepsin E was examined (Fig. 2). The shape of each of the $v$ versus [s] curves, measured at pH 3.1 in one case (Fig. 2a) and at pH 5.8 in the other (Fig. 2b), was indistinguishable from the substrate dependence observed in the presence of reducing agent (dithiothreitol in Fig. 2a; mercaptoethanol in Fig. 2b). On these bases, it would appear that the *activity* of cathepsin E is *not* dependent on the presence of the interchain disulphide bond.

Some aspartic proteinases such as pepsin are not stable in alkaline pH and it has been shown that this results from a selective denaturation of the N-terminal domain [12]. Consequently, the stability of the Cys[4]-Ala mutant cathepsin E to temperature and alkaline pH was evaluated. At 25°C and below, the mutant was stable ($t_{1/2} \gg 3$ h) at pH values as high as pH 9.5 so that, for affinity purification purposes, cathepsin E can be recovered from pepstatin-agarose columns by desorption at pH values up to 9.5. When the temperature was elevated to 37°C, the mutant was still stable at pH 8.5 ($t_{1/2} \sim 16$ h) but at pH 9.5, a half-life of 40 min was measured and at pH 10.5, the mutant enzyme was inactivated very rapidly ($t_{1/2} \sim 3$ min). On this basis, the temperature dependence was assessed at pH 8.5. Although the enzyme was stable at 37°C, a half-life of 40 min was measured at 45°C and further elevation to 55°C resulted in a rapid inactivation ($t_{1/2} < 5$ min). This sensitivity to 55°C was used to compare the stabilities of the monomeric Cys[4]-Ala mutant cathepsin E and the dimeric wild-type form of the enzyme. At pH 7.5 and 55°C, the dimeric enzyme had a half-life of 35 min. In contrast, the Cys[4]-Ala mutant was very unstable, with a half-life of only 1.7 min, measured both in the absence and in the presence of glycerol (2.5% final concentration). (The influence of glycerol was examined because for long-term storage of cathepsin E, which is inactivated upon freezing [13], the enzyme is commonly kept at −20°C in 50% glycerol in order to prevent solidification). Thus, it would appear that the presence of the inter-chain disulphide bond in cathepsin E confers stability on the dimeric enzyme that is lost in the Cys[4]-Ala mutant. These data which provide the first direct evidence for this conclusion, are in keeping with earlier findings with naturally-occurring guinea pig cathepsin E [8]. In that report, it was shown that when mercaptoethanol was present in the incubation buffer, monomeric cathepsin E had a half-life of approximately 2 h at pH 7.5 and 28°C whereas dimeric enzyme was essentially stable.

While this manuscript was nearing completion, Finley and Kornfeld [14] described their construction of a variety of chimaeric cDNAs in which regions of human procathepsin E were systematically replaced with the counterpart sequences from pepsinogen (which is a secretory zymogen). These elegant experiments revealed that the region of (pro)cathepsin E that specifies retention of this protein within the endosomal/e.r. network is indeed the sequence of residues in which cysteine-4 is located [14]. It would appear then that for its activity and fulfilment of its physiological roles in that capacity, cathepsin E does not need to exist as a disulphide-linked homodimer. However, since the enzyme is not, unlike pepsin, directed extracellularly into a very acidic environment, nor trafficked into an acidic lysosomal organelle (as is cathepsin D) but rather is retained within a cellular compartment that is at neutral or only slightly acidic pH, it would seem that the device of introducing an interchain disulphide bond at the N-terminus of the molecule not only confers stability upon the protein but also ensures that it is retained within the appropriate cytomorphological compartment of the cell. Structural studies by X-ray crystallographic analysis of the wild-type and Cys[4]-Ala mutant enzymes should provide a valuable insight into the importance of this highly-conserved cysteine residue in a unique region of this aspartic proteinase.

## References

[1] Takahashi, K. (ed.) (1995) Aspartic Proteinases: Structure, Function, Biology and Biomedical Implications, pp. 1–550, Plenum Press, New York.
[2] Sielecki, A.R., Fedorov, A.A., Boodhoo, A., Andreeva, N.S. and James, M.N.G. (1990) J. Mol. Biol. 214, 143–170.
[3] Sielecki, A.R., Hayakawa, K., Fujinaga, M., Murphy, M.E.P., Fraser, M., Muir, A.K., Carilli, C.T., Lewicki, J.A., Baxter, J.D. and James, M.N.G. (1989) Science 243, 1346–1351.
[4] Metcalf, P. and Fusek, M. (1993) EMBO J. 12, 1293–1302.
[5] Jupp, R.A., Richards, A.D., Kay, J., Dunn, B.M., Wyckoff, J.B., Samloff, I.M. and Yamamoto, K. (1988) Biochem. J. 254, 895–898.
[6] Hill, J., Montgomery, D.S. and Kay, J. (1993) FEBS Lett. ]326, 101–104.
[7] Azuma, T., Pals, G., Mohandas, T.K., Couvreur, J.M. and Taggart, R.T. (1989) J. Biol. Chem. 264, 16748–16753.
[8] Kageyama, T. (1993) Eur. J. Biochem. 216, 717–728.
[9] Athauda, S.B.P., Takahashi, T., Kageyama, T. and Takahashi, K. (1991) Biochem. Biophys. Res. Commun. 175, 152–158.
[10] Kageyama, T., Ichinose, M., Tsukada, S., Miki, K., Kurokawa, K., Koiwai, O., Tanji, M., Yakabe, E., Athauda, S.B.P. and Takahashi, K. (1992) J. Biol. Chem. 267, 16450–16459.
[11] Okamoto, K., Hu, H., Misumi, Y., Ikehara, Y. and Yamamoto, K. (1995) in: Intracellular Protein Catabolism (Bond, J.S. and Suzuki, K. eds.) Proc. 10th ICOP Symp., Plenum Press, Tokyo, in press.
[12] Lin, X.-L., Loy, J.A., Sussman, F. and Tang, J. (1993) Protein Sci. 2, 1383–1390.
[13] Valler, M.J. (1986) Ph.D. Thesis, University of Wales.
[14] Finley, E.M. and Kornfeld (1994) J. Biol. Chem. 269, 311259–31266.
[15] Thomas, D.J., Richards, A.D., Jupp, R.A., Ueno, E., Yamamoto, K., Samloff, I.M., Dunn, B.M. and Kay, J. (1989) FEBS Lett. 243, 145–148.