Design of sensitive fluorogenic substrates for human cathepsin D

Sergei V. Gulnik[a,*], Leonid I. Suvorov[a], Pavel Majer[a], Jack Collins[a], Bradley P. Kane[b], Donald G. Johnson[b], John W. Erickson[a]

[a]*Structural Biochemistry Program, SAIC Frederick, National Cancer Institute – Frederick Cancer Research and Development Center, Building 322, P.O. Box B, Frederick, MD 21702-1201, USA*
[b]*AIDS Vaccine Program, SAIC Frederick, National Cancer Institute – Frederick Cancer Research and Development Center, Frederick, MD 21702-1201, USA*

Received 2 May 1997; revised version received 30 June 1997

Abstract Cathepsin D is a lysosomal aspartic proteinase that has been implicated in several pathological processes such as breast cancer and Alzheimer's disease. We designed and synthesized a number of quenched fluorogenic substrates with P2 variations in the series AcEE(EDANS)KPIXFFRLGK(DABCYL)E-NH_2, where X = cysteine, methylcysteine, ethylcysteine, *tert*-butylcysteine, carboxymethylcysteine, methionine, valine or isoleucine. Most of the fluorogenic substrates exhibited greater k_{cat}/K_m ratios than the best cathepsin D substrates described so far. Differences in kinetic constants, which were rationalized using structure-based modeling, might make certain substrates useful for particular applications, such as active site titrations or initial velocity determination using a fluorescent plate reader.

Key words: Cathepsin D; Fluorogenic substrate; Structure-based modeling

1. Introduction

Cathepsin D (EC 3.4.23.5) is a highly abundant lysosomal aspartic proteinase widely distributed in various mammalian tissues. Inhibition studies along with amino acid sequence determinations of cathepsin D from different sources [1–4] indicated that the enzyme is a member of the pepsin family of aspartic proteinases. Due to its rather broad substrate specificity, cathepsin D is thought to play a major role in normal degradation of intracellular and extracellular endocytosed proteins [1]. In addition to normal physiological functions, cathepsin D has been associated with several biological processes of possible therapeutic significance. The enzyme seems to be involved in antigen processing and presentation of peptide fragments to class II MHC molecules [5–7]. Cathepsin D mediated proteolysis has also been linked to connective tissue diseases [8]. The enzyme can degrade extracellular matrix components [9] and may facilitate the spread of tumor cells via diverse mechanisms by its unregulated activity in the extracellular environment [10–12]. The observation of high levels of active cathepsin D within senile plaques in brains of Alzheimer's patients suggests that this enzyme may play a role in proteolytic cleavage of the amyloid precursor protein [13]. Recently, the major β-amyloid precursor protein processing activity from Alzheimer's disease brain extracts was identified and found to be indistinguishable from the activity of cathepsin D [14,15].

To facilitate the search for specific inhibitors of cathepsin D and to facilitate studies of its enzymology, it is necessary to have a sensitive and continuous assay for enzyme activity. Since aspartic proteinases do not display esterase or amidase activities, which are most commonly exploited for proteinase assays, different approaches have been developed for continuous monitoring of enzyme activity. In one technique, the substrate contains a chromophore, such as a nitrophenylalanine (F*) residue, in the P1′ position. Below pH 6.5, protonation of the N-terminal F* generated during cleavage shifts its absorbance maximum from 278 to 272 nm, which makes it possible to continuously follow product formation by monitoring the absorbance decrease at 300–310 nm by UV spectroscopy [16–18].

A second method is based on the use of fluorogenic peptide substrates which contain fluorophor and quencher groups in the same molecule. The fluorescent signal in the uncleaved substrate is quenched due to resonance energy transfer between the fluorophor and quencher groups. The increase in fluorescence during substrate cleavage can be continuously monitored using fluorescence spectroscopy [19]. The use of EDANS[1] (5-[(2-aminoethyl)amino]naphthalene-1-sulfonic acid) and DABCYL (4-(4-dimethylaminophenylazo) benzoic acid) as respective fluorophor and quencher was initially shown to provide good fluorogenic substrates for HIV-1 and AMV proteinases [20]. Fluorogenic substrates for aspartic proteinases with *o*-aminobenzoic acid and *p*-nitroanilide as the fluorophor and quencher pair were recently described by Filippova et al. [21].

In this report, we describe the design, synthesis and catalytic properties of a series of fluorogenic substrates for human cathepsin D. Different residues in the P2 position were examined and their effects on substrate binding were rationalized using the crystal structure of human cathepsin D [22].

2. Materials and methods

Cathepsin D was purified from human liver as described [23], lyophilized and stored at −70°C prior to use. Glycine, sodium acetate, hydrochloric acid and pepstatin A were purchased from Sigma Chemical Co., dimethylsulfoxide (DMSO) and Triton X-100 were purchased from Aldrich Chemical Co., Inc.

*Corresponding author. Fax: (1) (301) 846-6066.
E-mail: gulnik@ncifcrf.gov

Abbreviations: EDANS, 5-(2-aminoethyl) aminonaphthalene-1-sulfonic acid; DABCYL, 4′-dimethylaminoazobenzene-4-carboxylic acid; DIC, diisopropyl carbodiimide; DMF, dimethylformamide; DMSO, dimethylsulfoxide; Fmoc, 9-fluorenylmethoxycarbonyl; HBTU, 2-(1H-benzotriazol-1-yl) 1,1,3,3-tetramethyluronium hexafluorophosphate; HOBt, *N*-hydroxybenzotriazole; NMP, *N*-methyl-2-pyrrollidone; PAL, 5-(4-aminomethyl-3,5-dimethoxyphenoxy) pentanoic acid; TFA, trifluoroacetic acid

PII S0014-5793(97)00886-7

2.1. Modeling

Modeling and energy minimizations were performed with Biosym's InsightII/Discover software on a Silicon Graphics workstation. Parameters and charges for the diazo bond in DABCYL were obtained from Density Functional calculations performed with DGauss on a Cray-YMP as implemented in the Unichem software package (Cray Research). The substrates were modeled in the binding site of cathepsin D in several steps. In the first step, the peptide K-P-I-C-F-F-R-L-G was superimposed onto the P1–P3 sites of pepstatin A using the crystal structure of the cathepsin D/pepstatin A complex [22]. After making the appropriate side chain modifications, the peptide was energy minimized in the binding site, and the P1′–P3′ residues were then manipulated manually to optimize the van der Waals packing and polar interactions. A water molecule was then modeled to fit between the two catalytic aspartates and the scissile bond of the peptide. Finally, the rest of the substrate was modeled and the EDANS and DABCYL groups were added. The added residues were then manually optimized to remove steric contacts. Energy minimizations were performed after each step in the procedure.

2.2. Synthesis

The N_{α}-Fmoc-N_{ε}-DABCYL-L-lysine and *N*-Fmoc-L-glutamic acid-γ-EDANS were prepared using the modified procedure of [24]. All substrates were synthesized by the solid phase method using the Fmoc group for α-amino protection and acid labile groups for side chain protection of trifunctional amino acids. Aminomethyl polystyrene resin modified with PAL linker to a substitution of 0.3–0.4 mmol/g was used as the solid support. The FmocLys(DABCYL)OH and FmocGlu(EDANS)OH were coupled overnight in DMF as HOBT esters (DIC activation) using 3-fold molar excess of amino acid. All other steps were performed on a peptide synthesizer (Applied Biosystems, Model 430A) using a 10-fold molar excess of amino acid, HBTU and HOBt as coupling reagents, and NMP as solvent. The final cleavage of substrates from resin was accomplished by treatment with a mixture of TFA, thioanisole, ethanedithiol, and anisole (90:5:3:2). The crude substrates were purified by reverse phase HPLC on a VYDAC C18 column (25×2.5 cm) using 0.05% aq. TFA-acetonitrile mixtures in a gradient elution. Purity of all compounds was checked by analytical HPLC (column VYDAC C18, 25×0.4 cm) in 0.05% aq. TFA-methanol mixtures. All compounds gave correct amino acid analysis and correct molecular peaks using ion spray MS. Substrates were dissolved in 100% DMSO prior to use.

2.3. Kinetics

The increase of fluorescence during substrate cleavage was monitored using an Aminco Bowman-2 luminescence spectrometer (SLM Instruments, Inc.). Typically, 489 μl of 50 mM Gly-HCl buffer, pH 3.5, was mixed with 5 μl of DMSO and 1 μl of cathepsin D (0.14–0.3 pmol). The mixture was pre-incubated for 4 min at 37°C and reaction was initiated by the addition of 5 μl of substrate in DMSO. The final substrate concentration range was 1–20 μM; final DMSO concentration was 2% for all assays. Increase in fluorescence intensity at the emission maximum of 487 nm with excitation at 349 nm was recorded as a function of time. The initial velocity, measured as fluorescence intensity per unit time (V_f), was calculated from the slope during the linear phase of cleavage, usually 50–100 s, by a first degree polynomial fit using the SLM AB2 2.0 operating software. The inner filter effect [25] was negligible in the range of substrate concentration used, as measured by the linearity of fluorescence increase with substrate concentration after complete hydrolysis. Initial velocities were converted from units of fluorescence (V_f) to concentration per unit time (V_c) using the equation:

$$V_c = V_f*[S]/(F_{100}-F_0)$$

where [S] is substrate concentration, and F_0 and F_{100} are the fluorescence intensities of substrate solution before and after complete hydrolysis. Initial velocity data were fitted to the Michaelis-Menten equation using the program Enzfitter to calculate V_{max} and K_m values. Enzyme active site concentrations used for the k_{cat} calculations were determined by titration of cathepsin D using the tight-binding inhibitor pepstatin A.

The cathepsin D stock solution used for the kinetic experiments contained 0.05% Triton X-100 such that the final concentration of detergent in the reaction mixture was 0.0001%. Under these conditions, no change of Cathepsin D activity in the enzyme stock solution was detected during the total length of the experiment (up to 3 h at 0°C). In the absence of Triton X-100, we observed up to 50% enzyme inactivation over the same period of time. This inactivation is most likely a result of non-specific surface adsorption of cathepsin D rather than a result of autolysis since no significant change in enzyme activity was detected after 100 min incubation of 14 μM cathepsin D over a broad pH range (3.0–7.0) at 37°C (S.V.G., unpublished results).

2.4. HPLC separation of cleavage products

To confirm the P1–P1′ cleavage site, 2.5 ml of 10 μM substrate **5** (Table 1) solution in 0.1 M sodium acetate buffer, pH 3.3, containing 1% DMSO and 4 nM cathepsin D was incubated overnight at room temperature. Two product peptides were separated on a μ-Bondapak *C18* 3.9×300 *mm* reverse phase column using a 60 *min* linear 0–60% gradient of acetonitrile in 0.05% TFA. Cleavage products were subjected to both amino acid and sequence analysis. As expected, cathepsin D cleaved the substrate between the two phenylalanine residues. Cathepsin D cleavage of substrates **6**, **8**, **10** similarly yielded only two products that could be detected by HPLC. In each case, the elution time of one of the products was identical to that of the (FRLGK(DABCYL)E-NH_2) product from substrate **5**, confirming the existence of a single cleavage site between the two phenylalanine residues.

3. Results and discussion

In this study, we evaluated a series of novel, intramolecularly quenched fluorogenic substrates for human cathepsin D. The peptide part of the substrates was derived from the sequence K-P-I-L-F-F*-R-L as described for a cathepsin D sub-

Table 1
Kinetic constants of cathepsin D hydrolysis of different fluorogenic substrates

#	Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1}\ s^{-1}$)
1	KPILFF*RL[a]	28.0	45.0	1.6
2	KPIVFF*RL[a]	46.0	25.0	0.54
3	Ac-EE(Edans)KPILFFRLGK(Dabcyl)E-NH_2	5.7	39.6	7.0
4	Ac-EE(Edans)KPIVFFRLGK(Dabcyl)E-NH_2	7.4	27.0	3.7
5	Ac-EE(Edans)KPICFFRLGK(Dabcyl)E-NH_2	4.5	63.3	14.0
6	Ac-EE(Edans)KPIC(Me)FFRLGK(Dabcyl)E-NH_2	3.3	27.8	8.6
7	Ac-EE(Edans)KPIC(Et)FFRLGK(Dabcyl)E-NH_2	27.5[b]	196.0	7.1
8	Ac-EE(Edans)KPIC(tBu)FFRLGK(Dabcyl)E-NH_2	2.5	29.2	11.5
9	Ac-EE(Edans)KPIC(CbMe)FFRLGK(Dabcyl)E-NH_2	3.9	5.8	1.5
10	Ac-EE(Edans)KPIMFFRLGK(Dabcyl)E-NH_2	2.9	15.5	5.4

[a]Scarborough et al. [18].
[b]Maximum substrate concentration for determination of this value was 20 μM.P1 and P1′ residues are shown in bold.

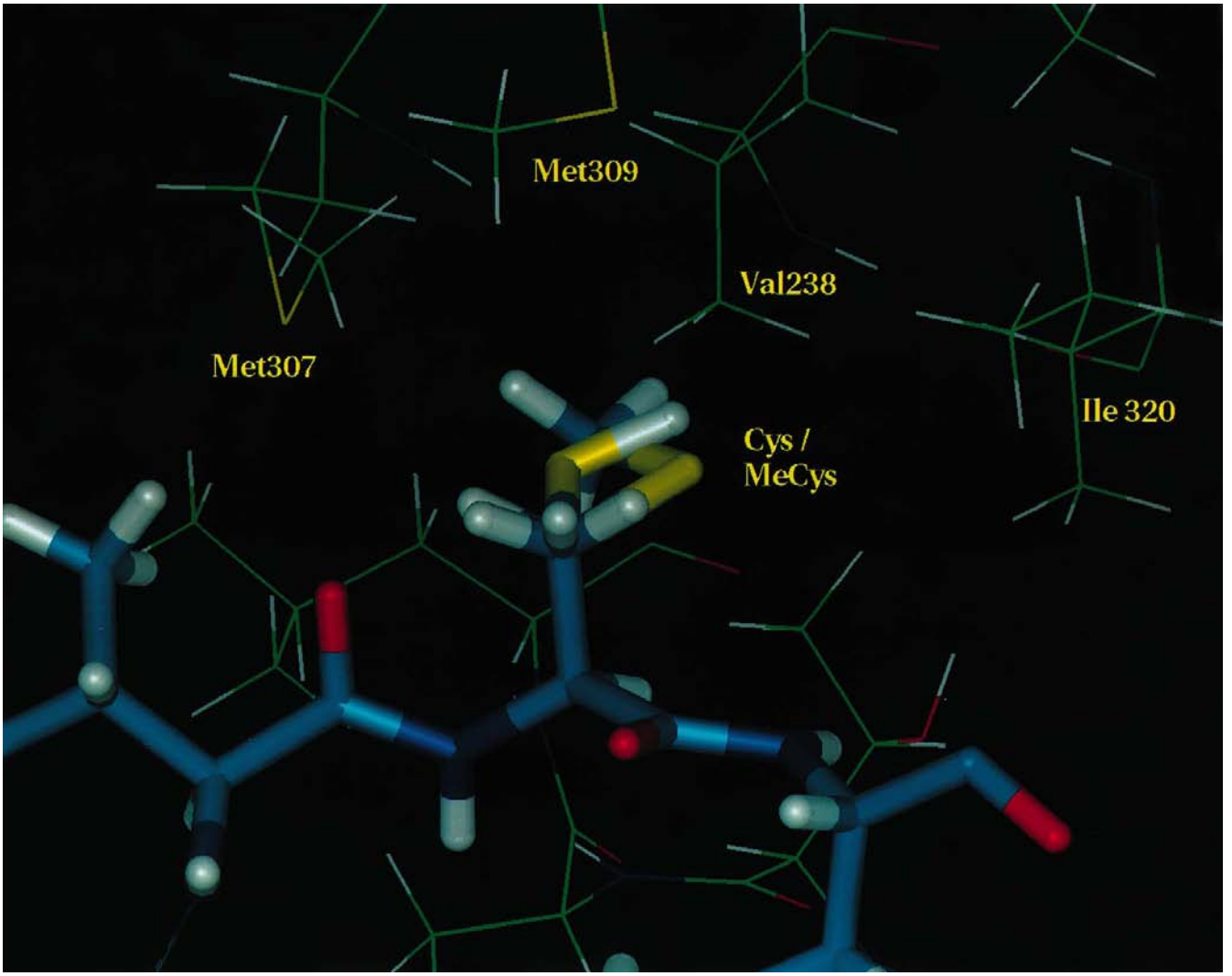

Fig. 1. The model of side chains of cysteine (**5**) and methylcysteine (**6**) in S2 pocket of cathepsin D.

strates by Scarborough, et al. [18]. Several charged amino acids were added to the N- and C-termini of the peptide to attach fluorophor and quencher molecules and to increase solubility of the substrate. The F* residue at P1′ was replaced by the natural phenylalanine residue since the nitro group is not necessary for substrate turnover. EDANS and DABCYL were chosen as the fluorophor and quencher groups, respectively. This donor-acceptor pair is characterized by a number of advantages such as extremely good spectral overlap of EDANS emission with the strong absorption band of DABCYL and good quantum yield, stability against photobleaching, and Stokes shift of over 100 nm for EDANS [20].

Kinetic parameters for cleavage of various fluorogenic substrates are presented in Table 1. Introduction of the fluorophor and quencher groups into substrates with the same P5-P3′ core sequence as the F-F* substrates resulted in a 5–6 fold decrease in K_m value (compare **1** and **3**, and **2** and **4**). The added binding potency may be due to the increased length of the fluorogenic substrates (P7-P6′) or to the presence of the bulky DABCYL and EDANS groups. The k_{cat} values for the chromogenic and fluorogenic substrates were similar, however, suggesting that the donor-acceptor groups at distal positions did not significantly influence peptide catalysis. The ratio of K_m values for the substrates with Val and Leu in P2 (**4** and **3**) is 1.3, which is also in good agreement with the data of Scarborough et al. (1993). Introduction of sulfur-containing side chains in P2, such as Cys (**5**), modified cysteines (**6**, **8**, **9**) and Met (**10**) slightly improved binding, with the exception of ethylcysteine (**7**). Modeling analysis indicated that the side chains of Cys and methylcysteine fit nicely into the hydrophobic S2 pocket of cathepsin D, making favorable hydrophobic interactions with the side chains of Met^{307}, Met^{309}, Val^{238} and Ile^{320} (Fig. 1).

Interestingly, there is nearly a 10 fold difference in K_m between substrates **7** and **10**. Yet, the side chains of Met and ethylcysteine contain the same number of atoms, and differ only by the position of the sulfur atom in the chain. Modeling analysis revealed a substantial geometric difference that would result from location of sulfur atom in the side chain (Fig. 2). The C-S-C bond angle, which is close to 90 deg, is located further from the substrate backbone in the Met side chain, and it allows it to adopt a favorable compact conformation in the S2 pocket. In the case of ethylcysteine, the sulfur atom is

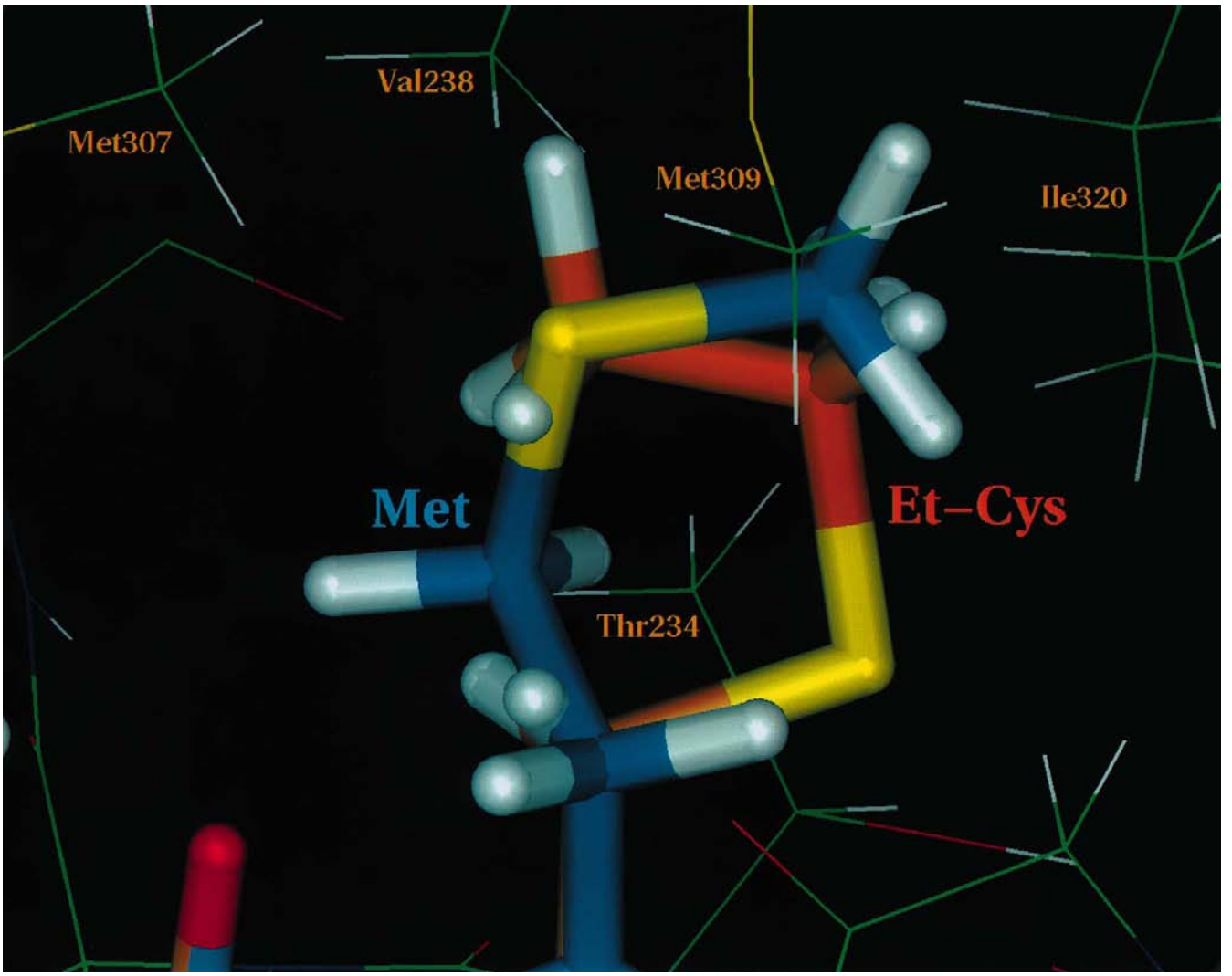

Fig. 2. Different conformation of Met (**10**) and ethylcysteine (**7**) in S2 subsite depends on the position of the sulfur atom of the side chain

too close to the backbone to allow for an optimal compact conformation. We were unable to find a good fit for ethylcysteine without bumping either into itself or to the side chains forming the S2 subsite and conclude that substrate **7** has to alter its backbone conformation in order to fit into the S2 pocket of cathepsin D. The resulting strain introduced into the scissile amide bond might lower the activation energy of cleavage, since the initial bound intermediate would be less stable, and may explain the unusually high k_{cat} value observed for this substrate (196 s^{-1}) which is the highest value ever described for a cathepsin D substrate.

The P2 side chain of carboxymethylcysteine (**9**) makes a unique set of interactions in the S2 pocket with its carboxy group at the position to form hydrogen bonds with the amide NH groups of Ser^{235} and Leu^{236} and the OH group of Ser^{235} (Fig. 3). This carboxy group is also partially exposed to solvent and therefore should contribute to the solvation energy of binding. These reasons may make the interactions in S2 quite favorable and explain the low K_m of 3.9 μM for **9**. On the other hand, the favorable hydrogen bonds made by the carboxymethyl group with the enzyme backbone may interfere with product release and account for the low k_{cat} for this substrate.

Most of the fluorogenic substrates described here have greater k_{cat}/K_m ratios than the best cathepsin D substrates described so far [18,21]. They have low K_m values and a rather wide range of k_{cat} values (5.8–196 s^{-1}) which makes them suitable for different kinetics applications. For example, the substrate with X = ethylcysteine (K_m = 27.5 μM, k_{cat} = 196 s^{-1}) might be convenient for active site titrations since it permits the use of substrate concentrations far below the K_m. The substrate with X = carboxymethylcysteine which has a low K_m with a slow k_{cat} (K_m = 3.9 μM, k_{cat} = 5.8 s^{-1}) could be used in an assay format that requires the reaction velocity to remain constant during a prolonged period, such as the cycling time of a fluorescent plate reader. Moreover, the increase in assay sensitivity afforded by these substrates allowed us to decrease the enzyme concentration used for the assay from 1–6 nM [18] to 0.2–0.3 nM. For routine screening of inhibitors, it is possible to use as low as 0.05–0.1 nM of cathepsin D which permits more precise determination of subnanomolar K_i's.

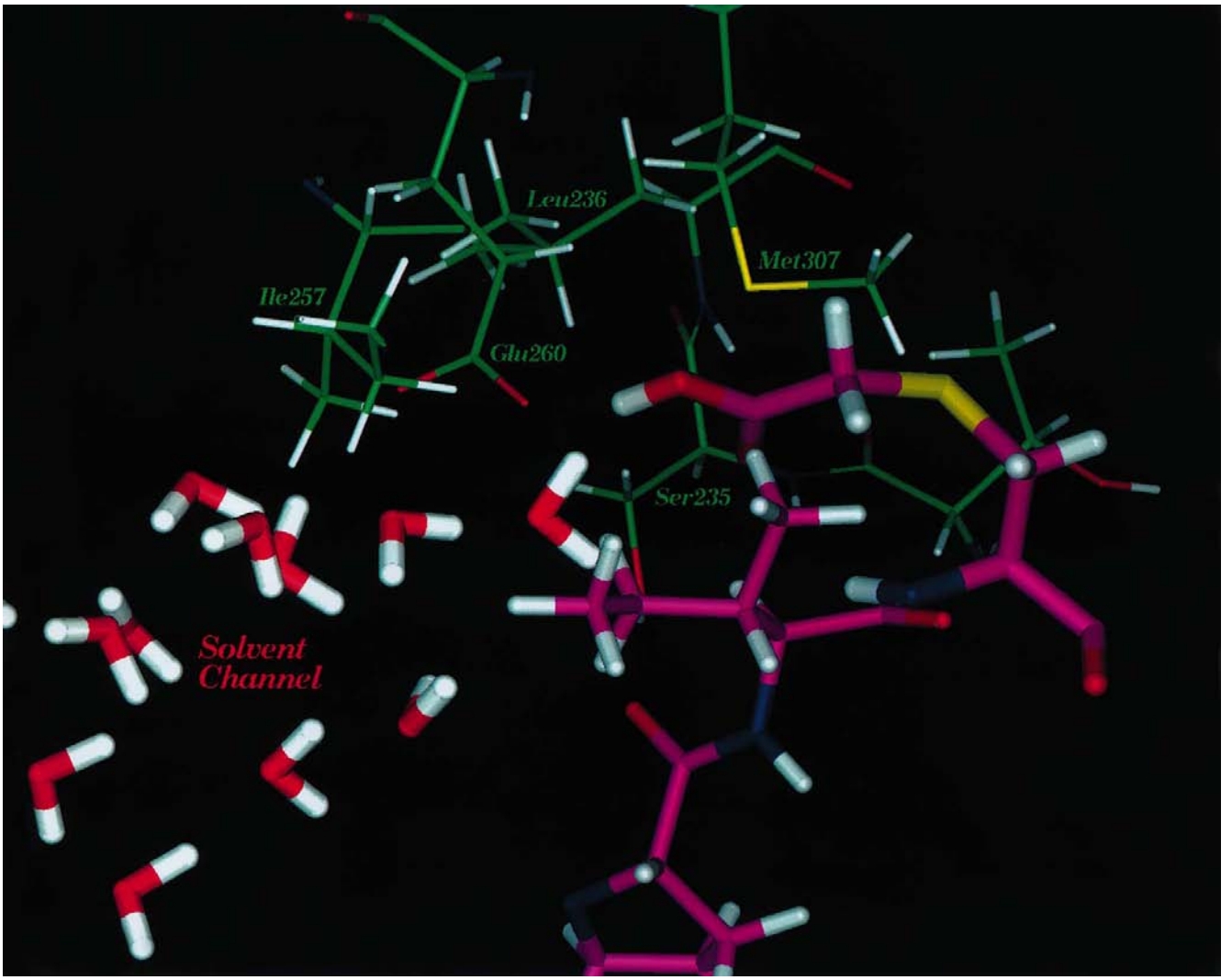

Fig. 3. Hydrogen bonds made by partially solvent exposed carboxyl group of carboxymethylcysteine may stabilize Michaelis complex and interfere with product release.

Acknowledgements: We thank Dr. Grant A. Krafft and Dr. Gary T. Wang for providing initial quantities of substrate **5–9**, Dr. Ben Dunn for helpful discussion, and Ms. Christine Ray for manuscript preparation. The content of this publication do not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

[1] Barrett, A.J. (1977) in: Proteinases in Mammalian Cells and Tissues (Barrett, Ed.), pp. 209–248, Elsevier/North-Holland Biomedical Press, Amsterdam.
[2] Faust, P.L., Kornfeld, S. and Chirgwin, J.M. (1985) Proc. Natl. Acad. Sci. USA 82, 4910–4914.
[3] Fujita, H., Tanaka, Y., Noguchi, Y., Kono, A., Himeno, M. and Kato, K. (1991) Biochem. Biophys. Res. Commun. 179, 190–196.
[4] Shewale, J.G. and Tang, J. (1984) Proc. Natl. Acad. Sci. USA 81, 3703–3707.
[5] van Noort, J.M. and van der Drift, A.C.M. (1989) J. Biol. Chem. 264, 14159–14164.
[6] Guagliardi, L.E., Koppelman, B., Blum, J.S., Marks, M.S., Cresswell, P. and Brodsky, F.M. (1990) Nature 343, 133–139.
[7] Peters, P.J., Neefjes, J.J., Oorschot, V., Ploegh, H.L. and Geuze, H.J. (1991) Nature 349, 669–676.
[8] Woessner, J.F., Jr. (1971) in: Tissue Proteinases, pp. 291–311 (Barrett, A.J. and Dingle, J.T., Eds.), North-Holland, Amsterdam.
[9] Briozzo, P., Morisset, M., Capony, F., Rougeot, C. and Rochefort, H. (1988) Cancer Res. 48, 3688–3692.
[10] Leto, G., Gebbia, N., Rausa, L. and Tumminello, F.M. (1992) Anticancer Res. 12, 235–240.
[11] Garcia, M., Platet, N., Liaudet, E., Laurent, V., Derocq, D., Brouillet, J.P. and Rochefort, H. (1996) Stem Cells 14, 642–650.
[12] Rochefort, H. (1994) Eur. J. Cancer 30A, 1583–1586.
[13] Castaldo, A.M. and Nixon, R.A. (1990) Proc. Natl. Acad. Sci. USA 87, 3861–3865.
[14] Dreyer, R.N. et al. (1994) Eur. J. Biochem. 224, 265–271.
[15] Ladror, W.S., Snyder, S.W., Wang, G.T., Holzman, T.F. and Krafft, G.A. (1994) J. Biol. Chem. 269, (28) 18422–18428.
[16] Dunn, B.M., Jimenez, M., Parten, B.F., Valler, M.J., Rolph, C.E. and Kay, J. (1986) Biochem. J. 237, 899–906.
[17] Scarborough, P.E., Richo, G.R., Kay, J., Conner, G.E. and Dunn, B.M. (1991) in: Structure and Function of Aspartic Proteinases: Genetics, Structure and Mechanisms, pp. 343–347, Plenum Press, New York.
[18] Scarborough, P.E., Guruprasad, K., Topham, C., Richo, G.R., Conner, G.E., Blundell, T.L. and Dunn, B.M. (1993) Protein Sci. 2, 264–276.

[19] Yaron, A., Carmel, A. and Katchalski-Katzir, E. (1979) Anal. Biochem. 95, 228–235.
[20] Matayoshi, E.D., Wang, G.T., Krafft, G.A. and Erickson, J.W. (1990) Science 247, 954–958.
[21] Filippova, I.Y., Lysogorskaya, E.N., Anisimova, V.V., Suvorov, L.I., Oksenoit, E.S. and Stepanov, V.M. (1996) Anal. Biochem. 234, (2) 113–118.
[22] Baldwin, E.T. et al. (1993) Proc. Nat. Acad. Sci. USA 90, 6796–6800.
[23] Gulnik, S., Baldwin, E.T., Tarasova, N. and Erickson, J. (1992) J. Mol. Biol. 227, 265–270.
[24] Wang, G. and Krafft, G. (1992) Bioorg. Med. Chem. Lett. 2, 1665–1668.
[25] Parker, C.A. and Rees, W.T. (1962) Analyst 87, 83–111.