

E-64 Analogs as Inhibitors of Cathepsin L and Cathepsin S: Importance of the S₂–P₂ Interactions for Potency and Selectivity¹

BARBARA J. GOUR-SALIN,² PAULE LACHANCE, PIERRE R. BONNEAU,³
ANDREW C. STORER, HEIDRUN KIRSCHKE,⁴ AND DIETER BROEMME⁵

Biotechnology Research Institute, 6100 Royalmount Avenue, Montreal, Quebec, H4P 2R2, Canada

Received January 17, 1994

A number of epoxysuccinyl amino acid benzyl esters (HO-Eps-AA-OBzl, **1**) in which the amino acid (AA) had been systematically varied were tested as inhibitors of cathepsins L and S. These E-64 analogs were designed to investigate whether selectivity for cathepsin L or cathepsin S could be attained by varying the amino acid bound to the essential epoxide ring which induces inhibition by alkylating the active site thiol of the cysteine proteases. The results indicate that the specificity of these analogs does not parallel that observed for substrates. This is possibly due to the fact that the direction of the peptide portion of these analogs, which is the reverse of that found for the substrate-like chloromethyl ketones (if the situation is analogous to papain), leads to differences in the orientation of the inhibitor side chain when bound in the S₂ subsite compared to substrates. The greatest selectivity was obtained with HO-Eps-Arg-OBzl which exhibited an 89-fold preference for cathepsin L over cathepsin S. A change from the L to the D stereochemistry for the phenylalanine analogs resulted in a 19-fold drop in k_2/K_i for cathepsin L and a 14-fold drop for cathepsin S. Both E-64 and Cbz-Phe-Ala-CH₂Cl form two hydrogen bonds with Gly 66 in the active site of papain. With the benzyl esters (HO-Eps-AA-OBzl) one of these hydrogen bonds is necessarily absent. In order to evaluate the importance of this hydrogen bond, three benzyl amide derivatives (HO-Eps-AA-NHBzl, **2**) were synthesized. In all cases the potency of the inhibitor was increased and indeed the HO-Eps-Phe-NHBzl analog is 64-fold more potent than the corresponding benzyl ester. For cathepsin L, there is also a 237-fold preference for L-Phe over D-Phe in the benzyl amide analog. In conclusion, although the information available from S₂–P₂ interactions with substrates cannot be used to enhance the selectivity of the E-64 analogs in a rational manner, the hydrogen-bonding interaction between the amide proton of the benzyl amid group in HO-Eps-AA-NHBzl and the S₂ subsite for both cathepsins L and S contributes to increase the potency of these types of inhibitors.

INTRODUCTION

The papain superfamily of proteolytic enzymes includes the plant cysteine proteinases papain and actinidin and the animal lysosomal cysteine proteinases ca-

¹ NRCC Publication No. 36168.

² To whom correspondence should be addressed at present address: Chemistry Dept., McGill University, 801 Sherbrooke St. W., Montreal, Quebec, Canada, H3A 2K6.

³ Bio-Mega Inc., 2100 Cunard Street, Laval, Quebec, H7S 2G5, Canada.

⁴ Institute of Biochemistry, Martin Luther University, Hollystr. 1, Halle (Saale), D-O-4020, Germany.

⁵ Present address: Kneip Pharmaceuticals Inc., 260 Littlefield, South San Francisco, CA 94080.

thepsins B, H, L, and S. Cathepsin L appears to play a physiologically major role not only in intracellular protein degradation and turnover (1) but also in intracellular events such as bone remodeling (2). In addition, cathepsin L is involved in a variety of diseases such as glomerulonephritis (3), arthritis (4), and cancer metastasis (5). Although less is known about cathepsin S it is believed to be involved in collagen (6) and elastin (7) degradation.

All these enzymes have closely related amino acid sequences and overall folding structures and utilize identical catalytic groups (8–12). In spite of these similarities, there is a great divergence in their proteolytic specificities, such as endopeptidase (papain (13), actinidin (14), cathepsin L (15), and cathepsin S (16)), aminopeptidase (cathepsin H (17)), and dipeptidylcarboxypeptidase (cathepsin B (18)). Therefore, it is of interest to investigate the role of amino acid substitutions around the active site evoking such a divergence in proteolytic activities. In addition, it is expected that an understanding of a set of closely related proteases would provide valuable information for the design of specific inhibitors as therapeutics and, as well, for the development of artificial proteinases having various specificities by protein engineering.

Schechter and Berger (19) have shown that the active site of papain can be considered to consist of seven subsites (S_1 – S_4 and S'_1 – S'_3), each able to accommodate one amino acid residue of a substrate. For papain, the nature of the residue at position P_2 has been shown to be the predominant factor in terms of determining specificity (20). Examination of the crystal structures of chloromethyl ketone derivatives of papain (21) allows identification of the residues in the S_2 subsite whose side chains make the most intimate contacts with the P_2 residue side chain of the substrate. These residues are Tyr 67, Pro 68, Val 133, Val 157, and Ala 160. The corresponding residues in cathepsins L and S can then be identified by sequence alignment (16, 22–28) (Table 1). Sequence comparisons indicate that Ser 205 of papain, located at the back of the hydrophobic pocket, is replaced by alanine in cathepsin L and by either Phe or Tyr in cathepsin S. Not surprisingly and consistent with the hydrophobic character of the S_2 subsite, studies indicate that cathepsin L, like papain, prefers bulky hydrophobic residues in P_2 (21, 29–31). The situation is not as well defined with cathepsin S. Recent work indicates (32) that a leucine residue is preferred in P_2 over a phenylalanine residue. Although one might expect the large aromatic side chain of either tyrosine 211 (bovine cathepsin S) or phenylalanine 211 (human cathepsin S) (Ser 205 in papain) to shorten the S_2 subsite and regulate its specificity, site-directed mutagenesis experiments with recombinant human cathepsin S indicate that the Gly 133 residue and not the large aromatic side chain of Phe 211 is responsible for the observed subsite specificity. Replacement of Gly 133 by an alanine results in a twofold better acceptance of phenylalanine over leucine. In contrast, replacement of Phe 211 by an alanine does not change the specificity of the subsite.

E-64 is a naturally occurring epoxide inhibitor of cysteine proteinases (33) which inhibits all the known cysteine proteases. It functions by alkylating the sulfur of the active-site cysteine residue. The challenge is to alter the structure of this molecule in such a manner as to make it specific to a particular member of the cysteine protease class. X-ray crystallographic studies of E-64 in the active site of

TABLE 1

Amino Acid Residues Whose Side Chains Are Involved in Substrate Binding in the S₂ Subsite of Papain and Corresponding Residues in Cathepsin L and Cathepsin S

Enzyme	Amino acid No.					
	67	68	133	157	160	205
Papain	Y	P	V	V	A	S
Human cathepsin L	L	M	A	M	G	A
Mouse cathepsin L	L	M	A	L	G	A
Rat cathepsin L	L	M	A	L	G	A
Chicken cathepsin L	L	M	A	L	G	A
Human cathepsin S	F	M	G	V	G	F
Bovine cathepsin S	F	M	G	V	G	Y

Note. Sequences are from Gal and Gottesman (24) for human cathepsin L, Troen *et al.* (28) for mouse cathepsin L, Wade *et al.* (27) for chicken cathepsin L, Ishidon *et al.* (25) for rat cathepsin L, Wiederanders *et al.* (26) for bovine cathepsin S, and Wiederanders *et al.* (16) for human cathepsin S. The amino acid numbering of papain starting at Ile 1 is used.

papain (34) and actinidin (35) indicate that its leucine side chain fits in the S₂ subsite of these enzymes. These interactions are schematically depicted in Fig. 1a. It can be seen that there are two hydrogen bonds between Gly 66 and the backbone of the peptide. In order to investigate the importance of the S₂-P₂ interactions in cathepsins L and S and to explore the possibility of developing specific inhibitors to these enzymes, E-64 analogs of the type HO-Eps-AA-OBzl, where AA has been varied systematically, were synthesized and tested for their inhibitory activity with these enzymes. In addition, a few analogs of the type HO-Eps-AA-NHBzl were also synthesized to evaluate the contribution of a hydrogen bond interaction between the amide proton of the benzyl amide group and the S₂ subsite (Fig. 1c).

RESULTS AND DISCUSSION

All of the epoxysuccinyl amino acid benzyl amides and esters were synthesized by coupling the amino acid benzyl amide or ester to the monoethylester of the epoxysuccinate as synthesized by the method of Mori *et al.* (36). The benzyl amides of the amino acids were synthesized by coupling the *t*-butoxycarbonyl (Boc)-protected amino acid with benzylamine followed by removal of the *t*-Boc-protecting group with trifluoroacetic acid. If the benzyl ester of the amino acid was not commercially available it was made by reacting the appropriately protected cesium salt of the required amino acid with benzyl bromide (37). The ethyl ester (EtO-Eps-AA-XBzl, X = O, NH) was then cleaved using pig liver esterase

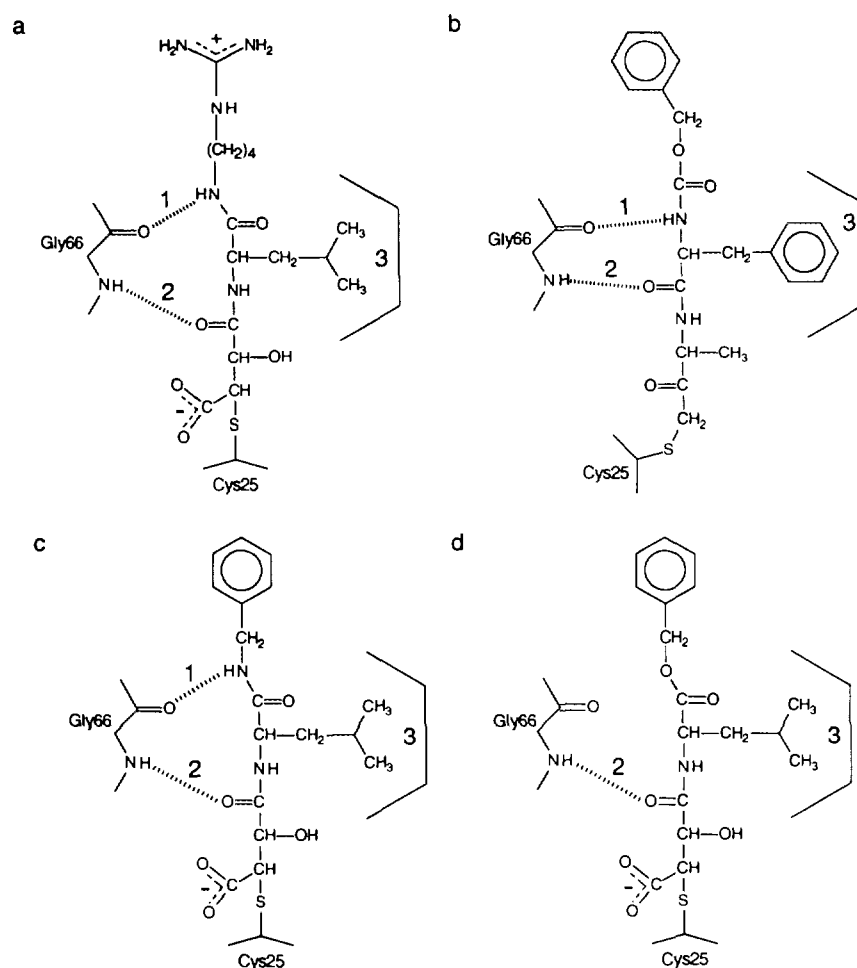
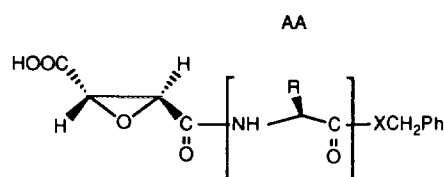


FIG. 1. Schematic representation of the interaction of various inhibitors in the S₂ subsite of papain. (a) E-64 (from the crystal structure (34)); (b) Cbz-Phe-Ala-CH₂Cl (from the crystal structure (21)); (c) proposed binding mode of HO-Eps-AA-NHBzl; (d) proposed binding mode of HO-Eps-AA-OBzl. Interactions in the S₂ subsite are represented by numbers (1 and 2, hydrogen bonds between main chain amide and carbonyl groups of the inhibitor and residue Gly 66 of papain; 3, interaction from the side chain of the inhibitor).

(38) to give the corresponding acid derivatives (HO-Eps-AA-XBzl, X = O, NH) of the epoxysuccinyl amino acid benzyl amides and esters (Scheme 1).

Table 2 summarizes the second-order rate constants k_2/K_i for inactivation of cathepsin L and cathepsin S by the E-64 analogs, HO-Eps-AA-OBzl. The inhibitory activity of HO-Eps-Leu-OBzl toward both cathepsin S and L is greater than E-64, indicating that for both of these enzymes the agmatine portion of E-64 is actually detrimental.

In order to compare the effect of varying P₂ in the inhibitors with the effect on



(1) X = O, AA = L-Phe, Leu, Arg, D-Phe, Ile, Val, Ala

(2) X = NH, AA = L-Phe, Leu, D-Phe

SCHEME 1

substrate hydrolysis, the specificity constant k_{cat}/K_M has been measured for several substrates of the type Cbz-AA-Arg-MCA under the same experimental conditions (Table 3). As mentioned earlier, cathepsin L prefers bulky hydrophobic residues in P_2 which was confirmed by our studies (Phe > Leu > Val > Arg). As predicted for cathepsin S, Leu is favored over Phe in P_2 . Somewhat surprisingly, Val is not favored, an indication that β -branching is not tolerated for cathepsin S. There is a 2440-fold drop in k_{cat}/K_M when one changes the leucine in P_2 to arginine, confirming the large influence interactions in the S_2 subsite have on hydrolysis rates.

It is clear from a comparison of Tables 2 and 3 that the variation of the residue in S_2 for the E-64 analogs does not parallel the variation found for substrates. For both enzymes the leucine analog of the epoxysuccinyl benzyl ester derivatives are the best inhibitors. It must be noted that for AA = Leu, Phe, D-Phe, Ile, the ratio

TABLE 2
Second-Order Rate Constants for Inactivation of Cathepsin L and Cathepsin S by HO-Eps-AA-OBzl^a

AA ^b	k_2/K_i ($10^3 \text{ M}^{-1} \text{ s}^{-1}$)			
	Cathepsin L		Cathepsin S	
Phe	434 ± 20	(1)	61 ± 10	(1)
Leu	791 ± 32	(0.55) ^c	171 ± 33	(0.36)
Arg	73 ± 11	(6)	0.82 ± 0.02	(74)
D-Phe	23 ± 2	(19)	4.4 ± 0.4	(14)
Ile	239 ± 6	(1.8)	60 ± 8	(1)
Val	8.1 ± 1.8	(53.6)	8.5 ± 0.1	(7.2)
Ala	3.7 ± 0.4	(116)	4.1 ± 0.01	(15)

^a The corresponding k_2/K_i values for E-64 are $218 \pm 20 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ and $99 \pm 5 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ with cathepsin L and cathepsin S, respectively.

^b All amino acids are "L" unless otherwise specified.

^c The number in parentheses represents the ratio of k_2/K_i values for Phe over AA.

TABLE 3
Second-Order Rate Constants for the Hydrolysis of
Cbz-AA-Arg-MCA with Cathepsin L and Cathepsin S

AA ^a	Substrate Cbz-AA-Arg-MCA k_{cat}/K_M ($10^3 \text{ M}^{-1} \text{ s}^{-1}$)			
	Cathepsin L		Cathepsin S	
Phe	5121 ± 255	(1) ^b	84.8 ± 2.7	(1)
Leu	2175 ± 187	(2.4)	237 ± 23	(0.36)
Val	498 ± 28	(10.3)	9.6 ± 2	(8.8)
Arg	6.6 ± 0.95	(776)	0.99 ± .23	(856)

^a All amino acids are "L" unless otherwise specified.

^b The number in parentheses represents the ratio of k_{cat}/K_M values for Phe over AA.

of the inactivation rate constants with cathepsin L over that with cathepsin S is 3.9–7.1 while for the smaller residues (AA = Ala, Val) the ratio goes down to 0.91–0.95. Increasing the size of the side chain, therefore, seems to increase the specificity toward cathepsin L. This possibly reflects the contribution of Ala 135 (Val 133 in papain) and Leu 161 (Val 157 in papain) to hydrophobic interactions with the amino acid side chain of the E-64 analogs. The inhibitor HO-Eps-Arg-OBzl displays selectivity for cathepsin L to a similar extent as the Cbz-Arg-Arg-MCA substrate. However, HO-Eps-Arg-OBzl displays a $k_2/K_i = 73 \times 10^3$ with cathepsin L, only 6-fold lower than that of HO-Eps-Phe-OBzl, whereas the substrate Cbz-Arg-Arg-MCA demonstrates a k_{cat}/K_M 78-fold lower than Cbz-Phe-Arg-MCA. In order to study the effect of stereochemistry L-Phe was replaced by D-Phe in the inhibitors, which led to a 19-fold drop in k_2/K_i for cathepsin L and a 14-fold drop for cathepsin S.

In order to understand the specificity differences between the E-64 analogs and the substrates, a closer examination of the binding modes of E-64 (Fig. 1a) (34) and Cbz-Phe-Ala-CH₂Cl (Fig. 1b) (21) in the active site of papain is necessary. The structure of the peptidyl chloromethyl ketone inhibitors is believed to bear relevance to enzyme–substrate interactions during the catalytic process (21). Contrary to most peptidyl inhibitors of cysteine proteases the alkylating epoxy-succinyl function of the molecule is attached to the amino group of leucine and not the C-terminal carbonyl group. Although both inhibitors bind in the S subsites of papain (34), the chloromethyl ketones bind from the N-terminal to C-terminal direction whereas E-64 binds in the reverse direction (C-terminal to N-terminal). This reverse arrangement results in differences in the nature of the hydrogen bonding in this site as well as a large difference in the orientation of the C α –C β bond in the S₂ subsite (39). Consequently, variations in P₂ of the inhibitors do not result in the same types of changes in activity as variations in P₂ of the substrates. However, both E-64 and Cbz-Phe-Ala-CH₂Cl form two hydrogen bonds with Gly 66 in the active site of papain. These two hydrogen bonds have been shown to be

TABLE 4
Second-Order Rate Constants for Inactivation of Cathepsin
L and Cathepsin S by HO-Eps-AA-NHBz^a

AA ^a	k_2/K_i ($10^3 \text{ M}^{-1} \text{ s}^{-1}$)	
	Cathepsin L	Cathepsin S
Phe	27,765 $\pm$ 1401 (1)	501 $\pm$ 26 (1)
Leu	4224 $\pm$ 398 ^b (6.5)	542 $\pm$ 136 (0.92)
D-Phe	117. $\pm$ 2 (237)	118 $\pm$ 2 (4.2)

^a All amino acids are "L" unless otherwise specified.

^b The number in parentheses represents the ratio of k_2/K_i values for Phe over AA.

of major importance for the substrate hydrolysis by papain (40). With the benzyl esters (HO-Eps-AA-OBzl), one of these hydrogen bonds is necessarily absent (Figs. 1c and 1d). In order to evaluate the importance of this hydrogen bond interaction in the S₂ subsite, three benzyl amides (HO-Eps-AA-NHBzl) were synthesized and their kinetic parameters measured (Table 4) (Fig. 1d). There is a significant preference for the benzyl amide inhibitors (HO-Eps-AA-NHBzl) for both enzymes. For cathepsin L, the benzyl amide inhibitor, HO-Eps-Phe-NHBzl, is 64-fold more potent than its benzyl ester analog and is also now 6.6-fold more potent than the leucine derivative. When L-Phe is converted to D-Phe there is a 237-fold drop in potency, much more than that observed for the benzyl esters. For cathepsin S, the leucine derivative (HO-Eps-Leu-NHBzl) and the phenylalanine analog (HO-Eps-Phe-NHBzl) are now almost of equal potency. For these enzymes, the potency of the inhibitor seems to be very dependent on the formation of the necessary hydrogen bonds between the inhibitor and an enzyme residue in the S₂ subsite (Gly 66 in papain).

From our results one cannot exclude the possibility that the epoxysuccinyl amino acid benzyl esters and benzyl amides used in this study bind to cathepsins L and S in a manner different from that of E-64 to papain or actinidin. However, based on the similarity between the cysteine proteinases and also between our epoxysuccinyl amino acid derivatives and E-64, we believe that their binding mode is essentially the same as that of E-64 in subsites S₁ and S₂ of the enzyme. The main difference between the E-64 type inhibitors and the substrate-like chloromethyl ketones is that the chemically reactive epoxysuccinyl group in the former is attached to the amino group of the amino acid residue, while in the latter the chloromethyl function is at the C-terminal part of the peptide. Thus, the amino acid residue in the epoxysuccinyl derivatives binds in the reverse direction compared to the chloromethyl ketones and substrates. Consequently, the interactions in the S₂ subsite are affected, and for this reason the specificity of the E-64 analogs cannot be altered in a rational manner by using the information available from S₂-P₂ interactions with substrates. However, modifications in the hydrogen-bonding interactions in the subsite lead to large perturbations in the potency of these types