Bioorganic & Medicinal Chemistry Letters 23 (2013) 2968–2973

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

Identification of new peptide amides as selective cathepsin L inhibitors: The first step towards selective irreversible inhibitors?

Ana Torkar [a], Brigita Lenarčič [b], Tamara Lah [a,b], Vincent Dive [c], Laurent Devel [c,*]

[a] Department of Genetic Toxicology and Cancer Biology, National Institute of Biology, Večna pot 111, 1000 Ljubljana, Slovenia
[b] Faculty of Chemistry and Chemical Technology, University of Ljubljana, Cesta v Mestni log 88a, 1000 Ljubljana, Slovenia
[c] Service d'Ingénierie Moléculaire des Protéines (SIMOPRO), Institut de Biologie et de Technologies de Saclay (iBiTec-S), Commissariat à l'Énergie Atomique (CEA), CE-Saclay, Bât. 152, F-91191 Gif-sur-Yvette, France

ARTICLE INFO

Article history:
Received 17 January 2013
Revised 8 March 2013
Accepted 11 March 2013
Available online 20 March 2013

Keywords:
Cysteine protease
Cathepsin L
Cathepsin B
Reversible inhibitors
Acyloxyethyl ketone inhibitors

ABSTRACT

A small library of peptide amides was designed to profile the cathepsin L active site. Within the cathepsin family of cysteine proteases, the first round of selection was on cathepsin L and cathepsin B, and then selected hits were further evaluated for binding to cathepsin K and cathepsin S. Five highly selective sequences with submicromolar affinities towards cathepsin L were identified. An acyloxymethyl ketone warhead was then attached to these sequences. Although these original irreversible inhibitors inactivate cathepsin L, it appears that the nature of the warhead drastically impact the selectivity profile of the resulting covalent inhibitors.

The endopeptidase cathepsin L (CatL) belongs to the papain-like cysteine proteases, which comprise 11 human cysteine cathepsins: cathepsins B, C, H, F, K, L, O, S, V, W and X/Z. In spite of the great similarities in the primary and tertiary structures among the cathepsins, these proteases have distinct enzymatic properties, as reflected in their different substrate specificities. Consequently the cathepsins have different roles in normal and pathological states,[1,2] including cancers.[3–5] Human[6] and animal[7] studies have clearly shown over-expression of CatL in various pathological processes, along with some of these other proteases.[8–10] CatL and CatB have been implicated in several steps of tumorigenesis and cancer progression.[5,11,12] Simultaneously increased protein levels of CatL and CatB have been reported in breast, gastrointestinal, head and neck, and lung cancers, and in glioblastoma and other cancers.[6]

In glioblastoma multiformae, which represent the most invasive stage of glial tumour astrocytomas, CatB has been shown to be relevant for tumour-cell invasion,[13,14] apoptosis[15] and angiogenesis. Multivariate analyses have shown that CatB also has a strong impact on prognosis.[16] In contrast, the role of elevated CatL[17,18] is less clear, although it appears to be related to tumour-cell invasion, and more likely to tumour-cell apoptosis[9,19] and resistance to therapy.[20] A tool for the detection and selective inhibition of the active forms of CatL would help to define its role in cancer progression, as well as in other pathologies.

The papain-like cysteine proteases share highly conserved active sites, with a catalytic dyad that is made up of a key nucleophile (Cys-25) and a general acid (His-159). The superposition of various X-ray structures of these cysteine cathepsins has provided support for only three well-defined substrate-binding sites: S_2, S_1 and $S_{1'}$. Outside of these regions, further sub-sites are not considered to be actual binding sites, but are more appropriately considered to be areas in which the substrate residues find their most favourable binding positions.[1]

Most of the synthetic cysteine protease inhibitors reported to date are covalent inhibitors that exploit this extremely well-defined structural context.[21] Indeed, these inhibitors incorporate a peptide or pseudo-peptide segment in their structure that corresponds to the sequences of good substrates, along with an additional reactive electrophilic moiety, or 'warhead', that can covalently modify the catalytic cysteine[21,22] (e.g., Fig. 1, compound A). Irreversible[23–26] and reversible[27–31] inhibitors are obtained according to the chemical nature of this electrophile. Although a number of potent cysteine protease inhibitors have been described in the literature, achieving selectivity towards only one member of the cathepsin family remains a challenge. In any series of covalent inhibitors, addressing the selectivity for a given target requires the optimisation of both the initial affinity binding step and the subsequent inactivation step. A direct evaluation of libraries of covalent

Abbreviations: AOMK, acyloxymethyl ketone; Cat, cathepsin.

* Corresponding author. Tel.: +33 01 69089565; fax: +33 01 69089071.
E-mail address: laurent.devel@cea.fr (L. Devel).

0960-894X/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.bmcl.2013.03.041

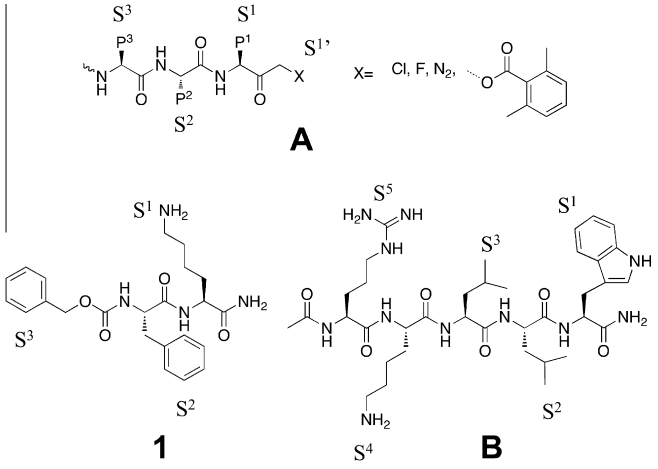
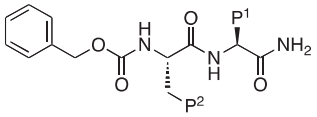


Figure 1. Structures of the covalent inhibitors of the cysteine proteases (A), the peptide amide as a potent CatL inhibitor (B), and the starting structure of the reversible inhibitor, compound **1**.

inhibitors for these targeted enzymes can be envisaged, but access to a large set of compounds often remains difficult to implement, mainly due to the sophisticated chemistry that is used to incorporate the reactive moieties.[32] In contrast, an indirect approach would imply initial profiling of the targeted enzyme active site with peptide or pseudo-peptide sequences, and then providing the selected segments with the necessary warhead. Such an approach represents a major advantage for easy and rapid screening of a large set of compounds with high chemical diversity using standard peptide chemistry. However, this strategy does not allow the simultaneous profiling of primed and non-primed regions. Such an approach was recently reported for CatK, which led to the identification of a relatively selective irreversible CatK inhibitor.[33] In this context, we have here evaluated the feasibility of such a strategy to access selective covalent inhibitors of CatL.

Compared to the approach described by Choe et al.[33] that relied on the screening of fluorogenic substrates, we evaluated a library of simple di-peptide and tri-peptide amides. This approach is supported by the work of Brinker et al.[34] that led to the identification of a peptide amide as a potent CatL inhibitor (Fig. 1, compound B; K_i 130 nM) that can interact within the non-primed region of the active site without being cleaved at its C-terminal position. As a starting point, we re-explored and optimised the affinity and selectivity profile of the CbzPheLys sequence, a motif that is classically encountered in irreversible inhibitors and series of activity-based probes.[21,24] Thus, from the pseudo-peptide amide **1** (Fig. 1), chemical diversity was introduced alternatively at each position (P^1, P^2 and P^3). Several substrates and inhibitors of cysteine cathepsins have a basic residue in the P^1 position.[24,33,35,36] To evaluate the precise impact of a P^1 lysine on the affinity and selectivity profile towards CatL, basic residues that varied in length and/or flexibility, and also uncharged residues, were tested in the P^1 position (Table 1, compounds **2**–**12**). The substrate specificity profiles for CatL show a preference for bulky hydrophobic and aromatic residues in the P^2 position.[24,33–35,37] On the assumption that a substituted phenyl moiety would be better tolerated within the S_2 sub-site of CatL compared to other cysteine proteases, several aromatic side chains that varied in length and substitution were introduced in the P^2 position (Table 1, compounds **13**–**23**). Finally, as already mentioned, the S^3 and S^4 substrate-binding sites of CatL are not considered to be actual binding sites, which suggests that they are susceptible to accommodate a large panel of various residues that might be either hydrophobic or basic.[35,38] The carboxybenzoyl moiety was therefore replaced by various acyl moieties, and by randomly selected amino acids (Table 2, compounds **24**–**45**). This small library

Table 1
Effects of the P^1 and P^2 substituent on affinity towards CatL and CatB



	P_1	P_2	K_i[a] (μM) CatL	K_i[a] (μM) CatB	F_B[b]
1	$-(CH_2)_4-NH_2$	$-(C_6H_5)$	4.2	3406	**810**
2	$-(CH_2)_4-NH-CH_3$	$-(C_6H_5)$	2.6	1547	**595**
3	$-(CH_2)_3-CH(OH)-NH_2$	$-(C_6H_5)$	27	1734	64
4	$-(CH_2)_3-NH_2$	$-(C_6H_5)$	3.5	2460	**703**
5	$-NH_2$	$-(C_6H_5)$	28	ND	ND[c]
6	$-(CH_2)-(4\text{-}NH_2-C_6H_4)$	$-(C_6H_5)$	26	2787	**107**
7	$-(CH_2)_3-NH-(C{=}NH)-NH_2$	$-(C_6H_5)$	28	349	12
8	$-(CH_2)$-1*H*-Imidazol	$-(C_6H_5)$	14	ND	ND[c]
9	$-(CH_2)_2-S-CH_3$	$-(C_6H_5)$	64	7096	**111**
10	$-(CH_2)-(CH)-(CH_3)_2$	$-(C_6H_5)$	14	4341	**310**
11	$-(CH)-(CH_3)_2$	$-(C_6H_5)$	17	ND	ND[c]
12	$-CH_3$	$-(C_6H_5)$	16	ND	ND[c]
13	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(C_6H_5)$	2.4	77	32
14	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(3\text{-}NH_2-C_6H_4)$	1.8	38	21
15	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(3\text{-}Cl-C_6H_4)$	1.7	58	34
16	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(3\text{-}CH_3-C_6H_4)$	2.5	389	**156**
17	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(2\text{-}CH_3-C_6H_4)$	23	245	11
18	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(4\text{-}CH_3-C_6H_4)$	1.7	113	66
19	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p\text{-}(C_6H_4)-p\text{-}(C_6H_5)$	2.6	74	28
20	$-(CH_2)_4-NH_2$	$-(C_6H_4)-p$-(2-Thiophen)	0.12	39	**325**
21	$-(CH_2)_4-NH_2$	1*H*-Imidazol	3.1	56	18
22	$-(CH_2)_4-NH_2$	1*H*-Indol	4.2	100	24
23	$-(CH_2)_4-NH_2$	$-(4\text{-}OH-C_6H_4)$	1.4	135	96

Bold values refer to selectivity factors of 3 orders of magnitude by comparison with those of only 2 orders or less.
[a] K_i values were determined at a minimum of three concentrations of inhibitors (measured in duplicates).
[b] $F_B = K_i(\text{CatB})/K_i(\text{CatL})$.
[c] ND, not determined.