Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs

Matthew Bogyo[1], Steven Verhelst[2], Valerie Bellingard-Dubouchaud[3], Sam Toba[4] and Doron Greenbaum[4]

Background: The lysosomal cysteine proteases of the papain family are some of the best studied proteolytic enzymes. Small-molecule inhibitors and fluorogenic substrate mimics have been used to probe the physiological roles of these proteases. A high degree of homology between family members and overlap in substrate specificity have made elucidating individual protease function, expression and activity difficult.

Results: Using peptide vinyl sulfones and epoxide as templates, we have generated probes that can be tagged with radioactive iodine. The resulting compounds covalently label various cathepsins and several unidentified polypeptides likely to be proteases. MB-074 was found to be a highly selective probe of cathepsin B activity. Probes that labeled several cathepsins were used to examine the specificity and cell permeability of the CA-074 family of inhibitors. Although CA-074 reportedly acts *in vivo*, we find it is unable to penetrate cells. Esterifying CA-074 resulted in a cell-permeable inhibitor with dramatically reduced activity and specificity for cathepsin B. The probes were also used to monitor protease activity in primary human tumor tissue and cells derived from human placenta.

Conclusions: We have generated a highly selective cathepsin B probe and several less specific reagents for the study of cathepsin biology. The reagents have several advantages over commonly used fluorogenic substrates, allowing inhibitor targets to be identified in a pool of total cellular enzymes. We have used the probes to show that cathepsin activity is regulated in tumor tissues and during differentiation of placental-derived cytotrophoblasts to invasive cells required for establishing blood circulation in a developing embryo.

Addresses: Departments of [1]Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94143, USA. [2]Leiden Institute of Chemistry, Gorleaus Laboratories, University of Leiden, 2300 RA Leiden, the Netherlands. [3]Stomatology and [4]Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143, USA.

Correspondence: Matthew Bogyo
E-mail: mbogyo@biochem.ucsf.edu

Key words: active-site labeling, inhibitors, cathepsins, epoxides, vinyl sulfones

Received: 10 September 1999
Revisions requested: 18 October 1999
Revisions received: 22 October 1999
Accepted: 25 October 1999

Published: 15 December 1999

Chemistry & Biology 2000, 7:27–38

1074-5521/00/$ – see front matter

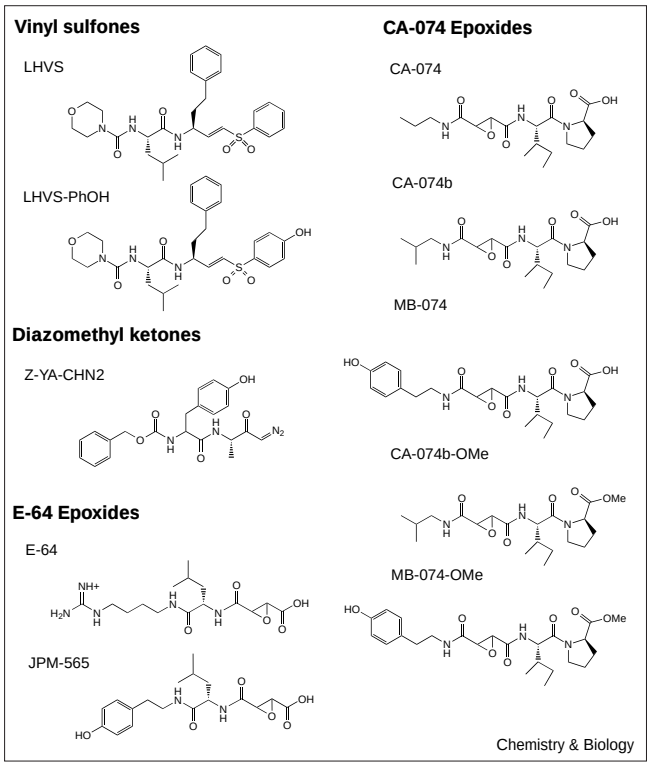
Introduction

The lysosomal cathepsins are proteases with diverse functions ranging from the generation of antigenic peptides to the turnover of misfolded proteins (for reviews see [1,2]). One of the largest subclasses of this family is a series of papain-like proteases that utilize a catalytic cysteine for proteolysis. Some of the most well known members of this cysteine protease family include cathepsins B, L, H, S, N and K. Much work has focused on these proteases and their involvement in several disease states (for a comprehensive review see relevant chapters in [3]).

Cathepsin B, in particular, has received much attention for its potential role in conditions such as cancer (reviewed in [4]), rheumatoid arthritis [5] and muscular dystrophy [6]. Most of these studies have focused on the development of diagnostic techniques for monitoring levels of cathepsin B expression using specific antibodies or by monitoring levels of mRNA [4]. These methods have been used in particular to examine the role of the lysosomal cathepsins in tumor progression and metastasis. Although such studies have established a connection between tumor invasiveness and cathepsin expression, they provide little information about actual levels of activity of individual proteases. For this reason, reagents that correlate more directly with activity could be a useful complement to the established expression methods.

Currently, one of the few sensitive methods available to monitor activity of proteases employs short peptides that have been modified with a fluorogenic moiety, which, when released by proteolysis, results in an easily detected fluorescent by-product. These types of substrates are often used *in vitro* to perform Michaelis–Menten kinetic analysis when purified enzyme is available. Such reporter substrates can also be used in zymograms to visualize protease bands within a native-type gel. The ease of synthesis of modified peptide fluorophores combined with information from cellular protein substrates has led to the development of substrates that are able to discriminate between often closely related proteases. This discrimination is far from absolute, however, and complications arise when substrate

Figure 1



Structures of various electrophilic substrate analogs. Compounds are all oriented as they are expected to bind in the active site of a target enzyme such that elements to the left of the electrophile occupy S-substrate binding sites and elements to the right of the electrophile occupy the S′-substrate binding sites.

hydrolysis is measured in complex mixtures of proteins. In addition, reactivity towards fluorogenic substrates can be accurately assessed only for known enzymes and fails to address other as-yet-unidentified proteases.

Some of the most commonly used tools to monitor cathepsin function are low molecular weight protease inhibitors (for reviews see [7,8]). Like substrate analogs, these reagents must be designed to accomplish the difficult task of discriminating between closely related enzyme family members. Selectivity has recently been achieved by exploiting cathepsin B's unique dipeptidyl peptidase activity. Compounds such as CA-074 (Figure 1) and CA-030 are peptide epoxides containing a short dipeptide free-acid that have been reported as specific cathepsin B inhibitors [9,10]. Structural studies show that these compounds bind with the peptide portion of the molecules, occupying the enzyme's S′ subsite, an orientation that is the inverse of that observed for other epoxide-containing inhibitors of the E-64 family [11]. Specificity for cathepsin B is achieved through binding interactions between the free carboxylate of proline and two highly conserved histidine residues from the so-called 'occluding loop' that directs the di-peptidyl peptidase activity unique to cathepsin B [12] facing the S′ substrate binding pocket (Figure 2).

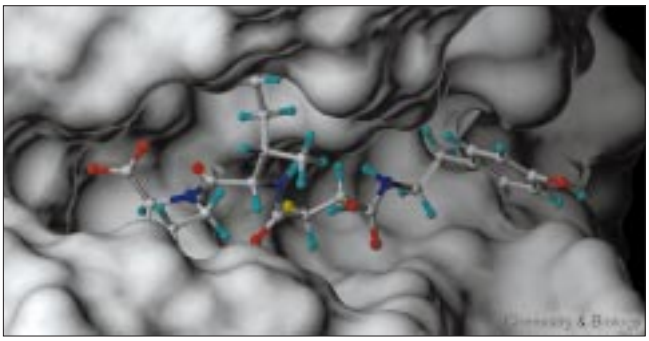
Several studies using CA-074 and CA-030 have been reported since their initial synthesis [11,13–16]. Specificity of these compounds for cathepsin B has only been addressed using fluorogenic substrates *in vitro* with purified enzymes or with crude cellular mixtures [9,10]. These types of studies, although informative, are plagued by the lack of synthetic substrates capable of absolute discrimination between cathepsin B and other cysteine proteases of the cathepsin and calpain family. In addition, no information regarding the reactivity of CA-074 and analogs towards unidentified proteases is available.

Effects of *in vivo* treatment of mice with CA-074 have been reported and these results were interpreted as a direct examination of the role of cathepsin B in antigen presentation [14–16]. The presence of the charged carboxylic acid moiety required for cathepsin B specificity, however, makes the cell permeability of this class of compounds questionable. To increase membrane permeability, an analog in which the proline free-acid group of CA-074 was converted to its methyl ester was reported [17]. This ester was shown to penetrate cells and to lead to a reduction in cathepsin B activity as measured by hydrolysis of a fluorogenic substrate. Buttle and colleagues [17] reason that although modification of the critical histidine-binding carboxylic acid results in reduced potency and specificity for cathepsin B, with time, nonspecific esterases convert the ester form of the inhibitor to the active acid form. The methyl ester CA-074-OMe has therefore been suggested to be a specific pro-inhibitor of cathepsin B [17].

The natural product E-64 is a peptide epoxide that inhibits most of the known lysosomal cysteine proteases by covalently modifying their active-site nucleophiles [18] (Figure 1). E-64 and related derivatives can be synthesized using a short and direct route from readily available diethyl tartrate. As a result, several analogs have been reported in which modifications to the pseudo-peptide portion of the molecule have been made [13,19,20]. In one such compound, JPM-565 (Figure 1), the guanidinium-containing arganine group has been replaced with tyramine. These and related compounds have been used for mechanistic studies of papain and other cysteine proteases of the papain family. Although the relative lack of specificity of these compounds makes *in vivo* inhibition studies difficult to interpret, in radiolabeled form they represent powerful tools to assess inhibitor binding to a wide range of protease targets simultaneously.

We report the development of a new peptide epoxide MB-074 that exploits the dipeptide Ile–Pro-OH peptide of CA-074 to create a highly selective affinity probe for cathepsin B. Replacing the isopropyl moiety of CA-074 with a phenol

Figure 2



Molecular modeling structure of MB-074 bound in the active site of cathepsin B. Note the orientation of the tyramine phenyl ring (right) and the free carboxylate of proline complexed with the occluding loop (left).

introduces a site for attachment of radioactive iodine. Covalent binding of the labeled inhibitor allows specificity to be assessed using simple sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE). The previously reported E-64 analog JPM-565, which already contains a phenol moiety, was used as a probe to monitor several identifiable lysosomal cathepsins as well as several yet to be identified polypeptides. The specificity and cell permeability of CA-074 and its derivatives were addressed using JPM-565 and other nonspecific cathepsin inhibitors. We also show that the modification of the free carboxylic acid of CA-074 and related compounds, although conferring membrane permeability, abrogates their specific binding to cathepsin B, making this a poor choice of sites for modification in the development of new cell-permeable cathepsin B inhibitors. Finally, the above reagents were used in two model systems to test their potential as tools to monitor the biological activity of individual cathepsins during the process of normal cell invasion as well as the unregulated progression of tumor metastasis.

Results

Inhibitor synthesis

On the basis of the finding that the dipeptide isoleucine–proline portion of CA-074 was responsible for specific cathepsin B binding [11], we synthesized analogs in which this specificity element is retained but the simple n-propyl moiety adjacent to the electrophilic epoxide is modified. Using a previously reported synthesis [13] for various peptide epoxides, we replaced the n-propylamino group with tyramine to make the compound MB-074 (Figure 1). We also synthesized the isobutyl analog of CA-074, named here CA-074b, to determine the effect of a less dramatic change in inhibitor structure. The compound JPM-565 was synthesized as reported elsewhere [20] and could be used directly for affinity labeling because of the tyramine moiety. Finally, we modified the dipeptide vinyl sulfone morpholinurea-leucyl–homophenyl–vinyl sulfone phenyl (LHVS; Figure 1) by replacing the phenyl group adjacent to the sulfone creating LHVS-PhOH (Figure 1), which contained a site for label attachment. Details of the synthesis and characterization of LHVS-PhOH will be the subject of a future publication.

Theoretical docking of MB-074 to the active site cathepsin B

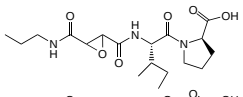
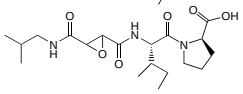
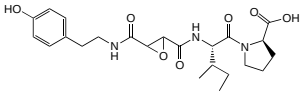
Theoretical energy-minimized structures were obtained for CA-074b, MB-074, I-MB-074 (2-iodo-tyramine analog of MB-074) and I_2-MB-074 (2,6-diodo-tyramine analog) bound to the active site of cathepsin B based on the reported crystal structure of cathepsin B inhibited by CA-030 [11]. Our modeling results suggest that CA-074b is a better binder than MB-074, which is in agreement with the experimental results described below. All the docked ligand structures retain nearly identical backbone atom positions to the crystal ligand (CA-030) structure (data not shown). MB-074's carboxy-terminal carboxylic group maintained interactions with the protein residues His110 and His111 in the prime side occluding loop and the isoleucine sidechain was buried deep into the S1′ pocket (Figure 2). The main structural differences among the various ligands arise from the interactions with the protein in the large S2 pocket. Turk *et al.* [11] have suggested that exploiting the interactions between the ligand and Glu245 in the S2 site should yield increased inhibitor selectivity. The i-butyl sidechain in CA-074b is extended into this S2 site, whereas the tyramine sidechain of MB-074 favors extension into the solvent (Figure 2). Similarly, the tyramine in I-MB-074 and I_2-MB-074 are also solvent exposed (data not shown). This orientation towards the solvent is mainly due to the phenol moiety of the MB-074 and its iodinated derivatives being too large to fit in the S2 site of cathepsin B. The favorable hydrogen-bond energy that would be gained by interaction between the phenol of MB-074 and Glu245 in the S2 pocket of cathepsin B is not enough to overcome the unfavorable van der Waals penalty caused by a clash between the phenyl ring of the MB-074 derivatives and other sidechains in the S2 pockets. Hence, this explains why addition of a large phenyl ring (and its iodination) has little effect on inhibitor potency compared with CA-074.

In vitro inhibition kinetics

Inhibition rate constants were obtained for MB-074, CA-074b and CA-074 binding to purified cathepsin B by standard Michaelis–Menten kinetic analysis as described in the Material and methods section (k_{obs}/[I]; Table 1). The cathepsin-L-like cysteine protease cruzain was also used to determine the reactivity of these same compounds against a homologous protease that does not have the occluding loop thought to be important for inhibitor binding to cathepsin B. The addition of a methyl group to the n-propyl moiety of CA-074 resulted in an inhibitor (CA-074b) with a potency for cathepsin B nearly identical to that of CA-074 (Table 1). Conversion to the tyramine analog MB-074 resulted in an active inhibitor with slightly

Table 1

Kinetic constants for inhibition of cathepsin B and cruzain by CA-074, CA-074b and MB-074.

Compound	Structure	$K_{obs}/[I]$ (1/M×1/s) Cathepsin B	IC_{50} (nM) Cruzain	IC_{50} (nM) Cathepsin B
CA-074		15,800±3750	>>10,000	40
CA-074b		29,000±1000	>>10,000	60
MB-074		2750±450	>>10,000	400

lower (less than tenfold reduced) activity than CA-074 and CA-074b towards cathepsin B. This reduced activity was presumably due to loss of binding interactions of the inhibitor in the S2-binding pocket of cathepsin B, consistent with the results from the modeling experiment described above. All three compounds showed only weak activity against the cathepsin-L-like cysteine protease cruzain as reflected in high IC_{50} values, which could not be accurately measured without inhibitor concentrations causing solubility problems. These data suggest that modifications to the n-propyl portion of CA-074 had little effect on specificity and binding *in vitro* and suggests this to be an appropriate site for chemical modification.

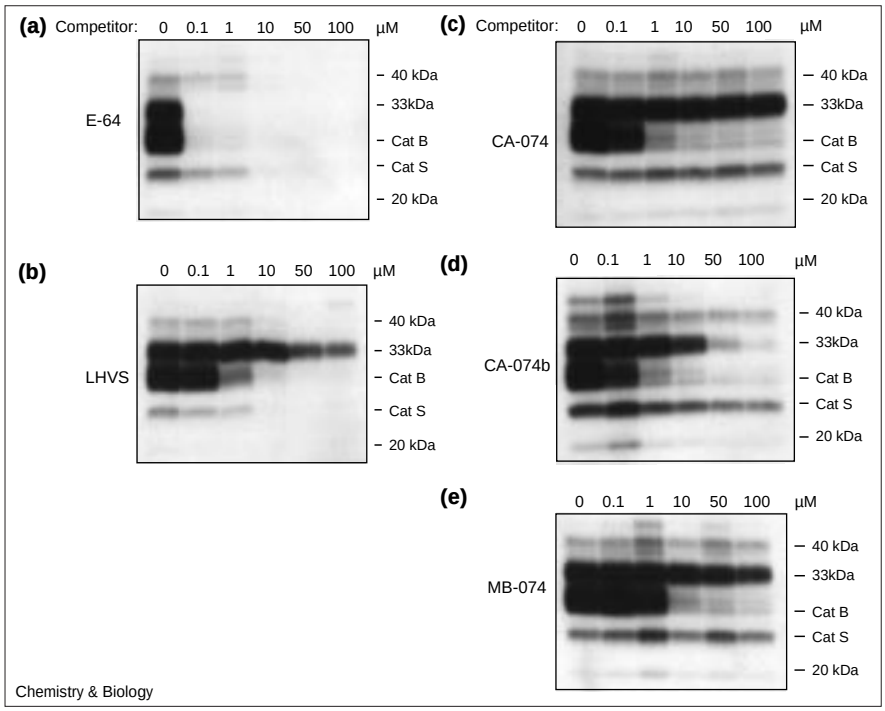
Specificity of inhibitors

Competition experiments were performed in a total cellular extract using the labeled, cysteine protease inhibitor ^{125}I-JPM-565 to better define the molecular targets of CA-074, its related family members and the reported cathepsin S specific compound LHVS [21–23]. Lysates from the dendritic cell line DC2.4 were prepared at acidic pH (pH 5.5) and incubated with increasing concentrations of E-64, LHVS, CA-074, CA-074b or MB-074. After preincubation with inhibitors, radiolabeled JPM-565 was added and the samples were analyzed by SDS–PAGE. Pretreatment of dendritic lysates with increasing concentrations of E-64, as expected, resulted in a complete loss of labeling, indicating that the two compounds share common targets and both are capable of acting as substrates for multiple cathepsins (Figure 3a). LHVS, a compound reported to be selective for cathepsin S at low concentrations, competed for binding to both cathepsin B and S at relatively similar concentrations. LHVS also modified two minor, unidentified polypeptides at ~20 and 40 kDa and at high concentrations showed partial modification of the major polypeptide at 33 kDa (Figure 3b). In contrast, CA-074, bound only cathepsin B at concentrations as high as 100 μM (Figure 3c). The closely related compound CA-074b showed identical inhibition kinetics to CA-074 for cathepsin B but at higher concentrations bound to an unidentified 33 kDa polypeptide, which was a major target of ^{125}I-JPM-565 (Figure 3d). In agreement with the observed *in vitro* inhibition kinetics, MB-074 showed an ~tenfold reduction in the concentration at which saturation of cathepsin B binding occurred when compared with CA-074 and CA-074b. Unlike the more closely structurally related CA-074b, however, MB-074 retained unique specificity for cathepsin B (Figure 3e). Thus, although the Ile–Pro-OH peptide portion of CA-074 is to a large extent responsible for the inhibitor's preference for cathepsin B, modifications at distal sites can interfere with this specificity.

MB-074 is a specific affinity label of cathepsin B

Although the competition data described above are pertinent to the selectivity of inhibitors for a given set of enzymes in a complex mixture of proteins, they fail to provide information about additional potential targets not visualized by ^{125}I-JPM-565 modification. A more comprehensive analysis of cellular targets of MB-074 was obtained by covalent attachment of radioactive iodine to the phenol ring of MB-074, followed by incubation with total cellular extracts from various cell types. Incubation of lysates from the dendritic cell line DC2.4 with labeled forms of the peptide diazomethyl ketone Z-YA-CHN_2, the peptide vinyl sulfone LHVS-PhOH and the peptide epoxide JPM-565 all resulted in modification of multiple polypeptides in the range 20–45 kDa (Figure 4a). The most predominantly labeled species was cathepsin B, which, because of its heterogeneous glycosylation, appeared as a diffuse band on SDS–PAGE at ~30 kDa. The radio-iodinated form of MB-074 showed modification of only the cathepsin B band. To determine whether this apparently unique specificity of MB-074 could be observed in other cell lines, we examined the labeling using the same four electrophiles in lysates prepared from the human pro-monocytic cell line U937 and primary cytotrophoblasts isolated from human placenta (Figure 4b,c). Labeling profiles in these cells were similar

Figure 3



CA-074 and MB-074 but not CA-074b are highly specific for cathepsin B in total cellular extracts. Lysates prepared from the dendritic cell line DC2.4 were treated with increasing concentrations (as indicated) of (a) E-64, (b) LHVS, (c) CA-074, (d) CA-074b or (e) MB-074 prior to addition of ^{125}I-JPM-565 and analysis by SDS–PAGE as described in the Materials and methods section. Lanes labeled 0 show the pattern of multiple polypeptides modified by the E-64 analog ^{125}I-JPM-565 in the absence of inhibitors. The position of inhibitor-labeled cathepsin B (Cat B) and cathepsin S (Cat S) are indicated. Predominant labeled polypeptides that could not be identified are indicated by their apparent molecular weight as judged by proximity to size standards.

to those observed for the dendritic cells except that the levels of cathepsin B activity were reduced, making it easier to distinguish bands around 30 kDa. In both of these cell types, ^{125}I-MB-074 labeled what appeared to be the single polypeptide band corresponding to cathepsin B.

To more carefully characterize the polypeptides modified by the above affinity reagents, lysates were incubated with various radio-iodinated inhibitors and separated using two-dimensional nonequilibrium pH gradient SDS–PAGE (NEPHGE; Figure 4d). As observed in the one-dimensional SDS–PAGE gels, ^{125}I-JPM-565 modified multiple polypeptides, some of which had identical molecular weights to cathepsin B and were therefore not visible in the one-dimensional SDS–PAGE profiles. As expected, ^{125}I-JPM-565 labeled polypeptides, including cathepsin B and cathepsin S (Figure 4d; identified by experiments described below). Yet, only the single 30 kDa diffuse cathepsin B spot was observed upon labeling with ^{125}I-MB-074, demonstrating the unique specificity of this compound for cathepsin B.

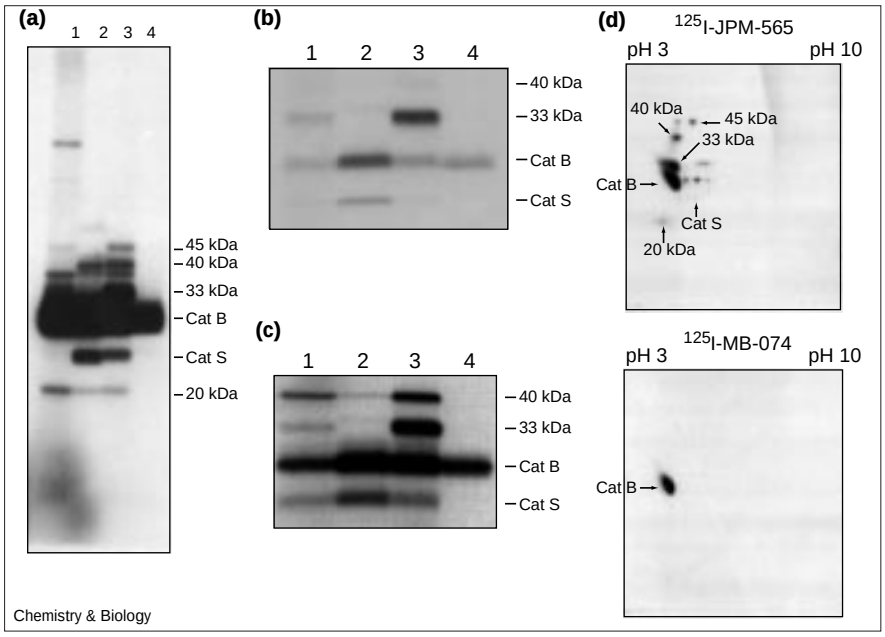
Cathepsin knock-outs reveal the identity of labeled polypeptides

To confirm the identity of the predominant polypeptides modified with either MB-074 or JPM-565, we examined the labeling profiles of primary spleenocyte lysates obtained from mice in which the cathepsin B, S or L genes had been deleted (Figure 5). As expected, ^{125}I-MB-074 modified the 30 kDa polypeptide (described above) in wild-type, cathepsin S and cathepsin L knock-out spleen cells. Labeling of the cathepsin B knock-out cells, however, resulted in complete loss of labeling of the 30 kDa polypeptide, identifying it as cathepsin B and as the sole target of ^{125}I-MB-074 (Figure 4a). The same knock-out cells could also be used to determine whether any of the additional polypeptide species modified by the nonspecific inhibitor JPM-565 could be identified as cathepsin L or S. As observed previously, labeling wild-type spleen lysates resulted in the modification of 20, 25, 30, 33 and 40 kDa polypeptides. Analysis of cathepsin B$^{-/-}$ lysates modified with ^{125}I-JPM-565 again identified the 30 kDa polypeptide as cathepsin B. Labeling of cathepsin S$^{-/-}$ lysates identified the 25 kDa polypeptide as cathepsin S, consistent with the known molecular weight of this protease. Labeling of cathepsin L$^{-/-}$ cells resulted in a pattern of polypeptides identical to that obtained for wild-type cells, indicating that none of the species modified by ^{125}I-JPM-565 represented cathepsin L.

CA-074b and MB-074 have no effect on cathepsin B *in vivo*

We next set out to address whether CA-074 and its derivatives were cell permeable and therefore could be used for *in vivo* studies as previously reported. Addition of either CA-074b or MB-074 to intact mouse dendritic cells (DC2.4) at concentrations as high as 100 μM resulted in no appreciable block of ^{125}I-LHVS labeling of any polypeptides, including cathepsin B (Figure 6a). As competition studies performed with both of these compounds showed dose-dependent activity in lysates, these data suggest that

Figure 4



The radio-iodinated form of MB-074 selectively labels cathepsin B in total extracts from multiple cell lines. Total lysates from the mouse dendritic cell line DC2.4 (a), the human pro-monocyte cell line U937 (b) or primary human cytotrophoblast isolated from placenta (c) were labeled by addition of ^{125}I-Z-YA-CHN$_2$ (lane 1), ^{125}I-LHVS-PhOH (lane 2), ^{125}I-JPM-565 (lane 3), or ^{125}I-MB-074 (lane 4) and analyzed by SDS–PAGE and autoradiography. (d) Samples of total DC2.4 lysates labeled with either ^{125}I-JPM-565 (top panel) or ^{125}I-MB-074 (bottom panel) described in (a) were quenched with NEPHGE sample buffer and then analyzed by two-dimensional NEPHGE and autoradiography. The positions of cathepsin B and S are noted based on known molecular weights, pI's and results described in Figure 5. Unidentified polypeptides are indicated by their approximate molecular weight as judged by standards.

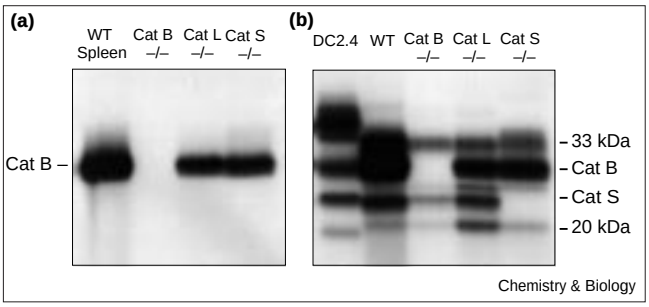
neither compound was capable of entering the lysosome and inhibiting their target. CA-074 treatment of mice has been reported to result in both resistance to Leishmania major infection and a switch from a Th2 to a Th1 immune response upon infection [15]. We wanted to determine whether these reported effects could be directly correlated with cathepsin B modification and therefore inhibition. Mice were treated for 2 days by interperitoneal injections every 12 hours with either CA-074b or a control vehicle every 12 hours as reported [15]. Spleens were then removed and intact spleenocytes were labeled with ^{125}I-LHVS (Figure 6b). Inhibitor treatment had no effect on cathepsin B labeling, consistent with the cell culture studies described above. Thus, although *in vivo* treatment of mice may result in effects on antigen presentation, these effects are not likely to be due to inhibition of cathepsin B as had been suggested previously [14–16].

Esterification of CA-074b and MB-074 leads to dramatically reduced potency and specificity for cathepsin B

The lack of cell permeability of CA-074 and its derivatives prompted us to examine the effects of modification to the charged carboxylic acid portion of this class of inhibitors. It has been reported that esterification of the proline acid moiety of CA-074 leads to a dramatic increase in cell permeability [17]. Still other reports focused on the effects of similar types of modification on specificity for cathepsin B [19]. The radio-labeled inhibitor ^{125}I-JPM-565 was used to determine the effects of methyl esterification on permeability and selectivity of CA-074b and MB-074. Pretreatment of intact dendritic cells with the methyl ester version of CA-074b and MB-074 followed by labeling of the cells with ^{125}I-LHVS showed that complete modification of cathepsin B required a > 50 times higher concentration of both compounds as compared with the parent compounds applied to DC2.4 cell lysates (compare Figure 7a with Figure 3d,e). In addition to the reduced potency towards cathepsin B, esterification of MB-074 and CA-074b led to a complete loss of specificity for cathepsin B. Both ester derivatives competed for labeling of cathepsin S and several unidentified polypeptides, including the same 20 kDa band observed above as a target for ^{125}I-LHVS. To further address the loss of specificity upon modification of the proline free acid of the CA-074 derivatives, MB-074 was iodinated and its targets identified by direct affinity labeling of intact dendritic cells (Figure 7b). As expected from the competition results, the ester derivative of MB-074 was able to penetrate intact cells and covalently modified multiple polypeptides. The profile of polypeptides modified by ^{125}I-MB-074-OMe was nearly identical to the profile of bands that were modified by the unlabeled MB-074-OMe in competition studies. These results suggest that if methyl ester derivatives are converted to free-acid forms once inside the cell, the process of conversion is slower than the rate of target modification. This class of pro-drugs is therefore unlikely to be of use as specific reagents for the study of cathepsin B function *in vivo*.

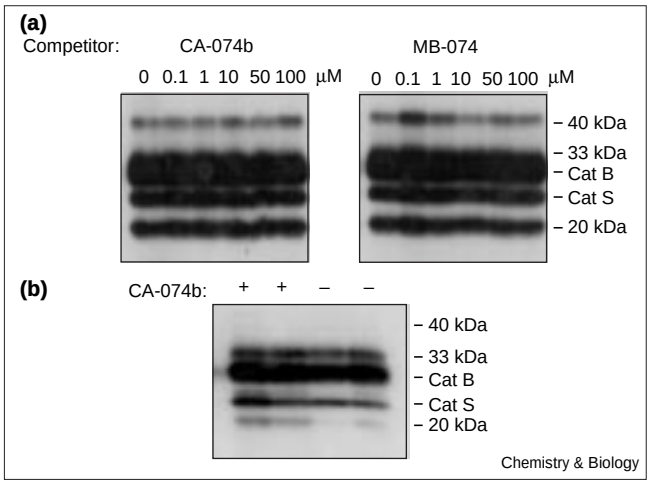
The inhibitor JPM-565, like the CA-074 derivatives, contains a charged free carboxylic acid moiety and is a relatively poor reagent for use in intact cells. The precursor ethyl ester form of JPM-565 (JPM-565-OEt) was iodinated and used to label intact dendritic cells (DC2.4). The ester form of JPM-565 showed a dramatic increase in labeling

Figure 5



Cathepsin knock-out cells identify cathepsin B as the sole target of ^{125}I-MB-074. Total lysates were obtained from spleenocytes of wild-type mice or mice in which the cathepsin B, S or L gene had been deleted. Lysates were labeled by addition of (a) ^{125}I-MB-074 or (b) ^{125}I-JPM-565 and analyzed using SDS–PAGE and autoradiography. The position of cathepsin B and S in the profiles was determined by the loss of labeling in the corresponding knock-out samples and is indicated along with several unidentified polypeptides (b).

Figure 6



The CA-074 analogs CA-074b and MB-074 show no activity *in vivo*. (a) Intact spleen cells were incubated with increasing concentrations of CA-074b and MB-074 as indicated and cathepsins visualized by labeling with ^{125}I-LHVS-PhOH followed by SDS–PAGE and autoradiography. (b) CA-074b (lanes labeled (+)) or control DMSO (lanes labeled (–)) was administered to mice by interperitoneal injections every 6 h for 48 h. Spleens were removed and isolated intact spleenocytes labeled with ^{125}I-LHVS-PhOH. Cathepsin B and S were identified based on migration in the gel as described above.

intensity of the same polypeptides modified by the parent JPM-565, indicating an increase in cell permeability (Figure 7b). Several additional polypeptides near 55 kDa that were not modified by JPM-565 were intensely labeled by JPM-565-OEt. Esterification of inhibitors such as JPM-565 can therefore render some otherwise ineffective compounds useful for *in vivo* studies.

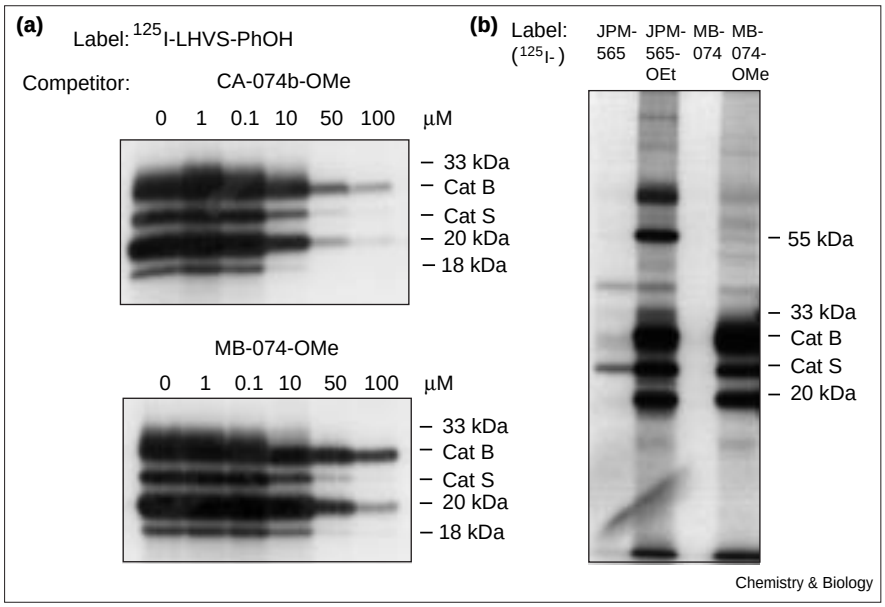
Affinity labels reveal cathepsin activity during invasion processes

The compound ^{125}I-JPM-565 was used to examine directly the levels of cathepsin activity in homogenates obtained from normal, primary tumor and metastatic tumor tissue from several human subjects. Homogenates were incubated with ^{125}I-JPM-565 and then analyzed using SDS–PAGE (Figure 8). In samples from all three patients an increase in cathepsin B labeling was observed in metastatic tissue when compared with normal or primary tumor samples. Furthermore, the extent of increase in cathepsin B labeling for metastatic samples relative to normal tissue was slightly different for each patient. These data fit well with the reported findings for cathepsin B's involvement in tumor progression and invasion [4]. Interestingly, cathepsin S labeling decreased in tumor samples, whereas an unknown polypeptide at 40 kDa was modified to a varying extent. These results suggest that although cathepsin B increases in most metastatic tissue, the levels of activity may differ amongst individual patients as well as tumors.

Levels of cathepsin activity were also monitored in an *in vitro* differentiation model for primary cytotrophoblasts derived from human placenta. During human placental development, specialized epithelial cells termed cytotrophoblasts invade the uterus, anchoring the conceptus to the decidua, tapping a supply of maternal blood. This unusual behavior requires cytotrophoblasts to assume highly specialized characteristics; some are commonly associated with tumor cells, whereas others are typical of endothelia. Unlike tumor invasion, however, trophoblast invasion is precisely regulated, confined spatially to the uterus and temporally to early pregnancy. Plating purified cytotrophoblast stem cells on the extracellular matrix (ECM) substrate Matrigel allows them to spontaneously differentiate along the invasive pathway, modulating their expression of stage-specific antigens in a manner analogous to that observed during uterine invasion *in situ* (reviewed in [24]). With regard to the cells' tumor-like properties, both adhesion molecule switching and matrix metalloproteinase-9 (MMP-9) activation have been shown to facilitate invasion [25–27]. However, little is known about the role of the lysosomal cathepsins in this pathway.

Cytotrophoblasts were incubated for various times after matrix attachment, collected and then lysed. Lysates were labeled with either ^{125}I-JPM-565 or the cathepsin-B-specific ^{125}I-MB-074 and labeled bands were visualized using SDS–PAGE followed by autoradiography (Figure 9). Labeling of samples with ^{125}I-MB-074 revealed only a single polypeptide corresponding to cathepsin B. Upon exposure of cells to the Matrigel, cathepsin B labeling increased in a time-dependent fashion, suggesting that activation of cathepsin B is regulated during the differentiation pathway towards an invasive phenotype in these cells. Labeling with the nonspecific ^{125}I-JPM-565 con-

Figure 7



The methyl esters CA-074b-OMe and MB-074-OMe target multiple polypeptides in DC2.4 cells. (a) DC2.4 cells were pre-incubated with increasing concentrations of CA-074b-OMe and MB-074-OMe as indicated, then labeled by addition of ^{125}I-LHVS-PhOH and samples analyzed using SDS–PAGE and autoradiography. (b) DC2.4 cells were directly labeled by addition of ^{125}I-JPM-565, ^{125}I-JPM-565-OEt, ^{125}I-MB-074, and ^{125}I-MB-074-OMe and labeled polypeptides visualized by SDS–PAGE and autoradiography.

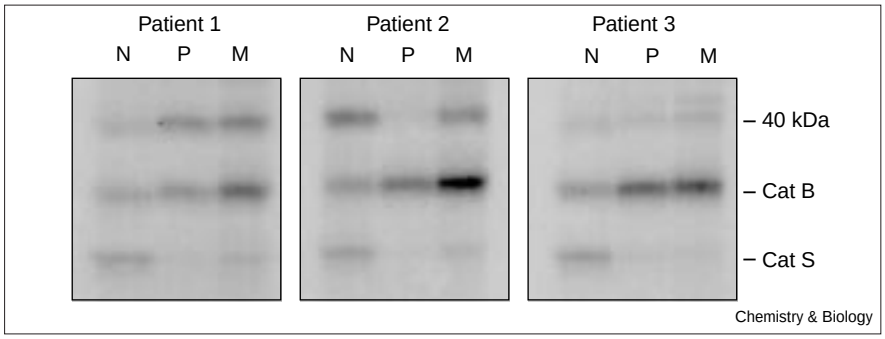
firmed the increase in cathepsin B activity after contact with the synthetic matrix. Cathepsin S, a target of JPM-565, was also activated in concert with cathepsin B activity, whereas another unidentified 33 kDa band showed no change in labeling throughout the 42 hour time course. These results suggest that the activities of a subset of cathepsins change upon differentiation of human cytotrophoblasts. Such modulation of cathepsin activity may be required for activation of this pathway.

Discussion

The studies reported here describe new reagents for studying cathepsin biology. Chemical modification of commonly used small-molecule protease inhibitors allows direct visualization of protease activity using SDS–PAGE. These reagents offer advantages over currently used techniques of in-gel zymograms and fluorogenic substrate hydrolysis assays because the target is modified within a native cellular environment. Furthermore, as individual protein species can be resolved using SDS–PAGE, highly specific reagents are not required in studies of closely related proteases. These affinity-labeling reagents are also particularly well suited to address the specificity of other commonly used protease inhibitors, both *in vitro* and *in vivo*.

Often small-molecule inhibitors are used to make conclusions about the role of a particular enzyme in a biological process without knowing all potential targets for inhibition by the reagent or to what extent inhibition of the desired target has been achieved. In the case of CA-074, a compound reported to be a selective inhibitor of cathepsin B, specificity has been addressed using *in vitro* kinetic studies with several well-characterized proteases [9,10]. Although this method provides some information about selectivity, it leaves open for question results obtained in crude extracts or intact cells where the composition of possible targets

Figure 8



Cathepsin activity in human tumor tissue extracts. Extracts were prepared from tissue samples of normal (N), primary tumor (P), and metastatic tumor (M) obtained for multiple patients (arbitrarily assigned 1, 2 and 3). The matched sets of samples were normalized with respect to total protein and cathepsins labeled by addition of ^{125}I-JPM-565 as described in the methods section. Cathepsin B was identified by labeling with the cathepsin B specific probe ^{125}I-MB-074 (data not shown). Cathepsin S was identified by its migration profile with respect to cathepsin B.

present is unknown. We show here, through both indirect and direct labeling experiments, that cathepsin B is the sole identifiable target of CA-074 and MB-074.

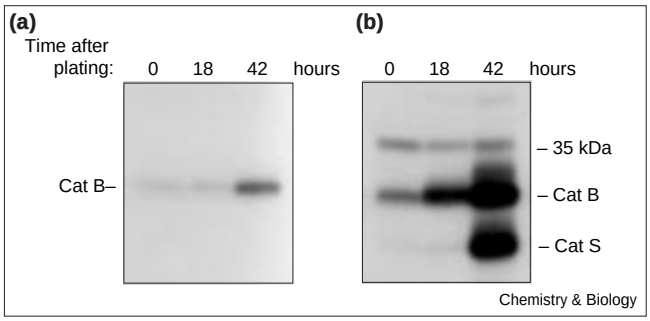
We also found that, although the important specificity element of CA-074 is thought to be the free-acid-containing Ile–Pro dipeptide, chemical modification at the distal side of the molecule can have a dramatic effect on specificity (Figure 3d). Addition of a phenol moiety to the pre-existing n-propyl group of CA-074 resulted in a compound (MB-074) that retained specificity for cathepsin B with only a slight loss in potency. Surprisingly, addition of a simple methyl group converting the n-propyl group of CA-074 to an i-butyl group (CA-074b) resulted in a compound that was equipotent to CA-074 towards cathepsin B but showed activity against an unidentified polypeptide at 33 kDa, potentially a member of the cathepsin protease family. These data suggest that although the Ile–Pro dipeptide contributes some degree of specificity to an inhibitor towards the dipeptidyl peptidase activity of cathepsin B, this is clearly not the only important specificity element. In fact, small changes in the structure of CA-074 have detrimental effects on specificity, whereas other larger changes to the structure have no effect whatsoever.

In vitro kinetic data for these inhibitors obtained for purified cathepsins show that the three closely related inhibitors MB-074, CA-074 and CA-074b are highly selective for cathepsin B over the cathepsin-L-like protease cruzain (Table 1). One would therefore conclude that all three would be of use in studying cathepsin B function *in vivo*. Yet without careful characterization and purification of the 33 kDa polypeptide found to be a target of CA-074b, the loss of specificity observed upon addition of the methyl group to CA-074 would certainly be overlooked.

In addition to incomplete information about target specificity for many commonly used inhibitors, further confusion can result from a lack of information about *in vivo* inhibition efficiency for a desired target. In the case of CA-074, *in vivo* effects of inhibitor treatment have been reported [9,14,15]. We show here, however, that treatment of either tissue culture cells or mice with CA-074 analogs has no effect on cathepsin B labeling by ^{125}I-LHVS-PhOH. This class of reagents is cell impermeable and therefore not likely to be of use for *in vivo* studies of cathepsin B.

To circumvent problems of permeability for the CA-074 family of inhibitors, studies have focused on the effects of modification of the charged carboxy-terminal proline carboxylate [17,19]. Simple methyl esterification of this acid group results in cell-permeable 'pro-inhibitors' that, once inside the cell, are thought to be processed to the highly selective free-acid form, CA-074 [17]. Our initial concerns about the use of a pro-drug form of CA-074 were that these esters would be reactive towards other cathepsins,

Figure 9



Activity of cathepsin B and cathepsin S is up-regulated during initiation of the invasive phenotype in primary human cytotrophoblasts. Primary human cytotrophoblasts were isolated from placenta tissue and plated on the synthetic matrix, Matrigel. Cells were harvested from the matrix after various times of incubation as indicted. Total cellular extract obtained from these time points were incubated with either (a) ^{125}I-MB-074 or (b) ^{125}I-JPM-565. Labeled polypeptides were visualized by SDS–PAGE and autoradiography. The identity of cathepsin B was determined by ^{125}I-MB-074 labeling. The identity of cathepsin S was determined by its migration relative to cathepsin B as observed in previous experiments.

and that the rate of inhibition would be faster than the rate of ester hydrolysis. Treatment of cells with the pro-inhibitor, over time, would therefore lead to a steady state in which multiple cathepsin activities had been blocked. In fact, we found that this was the case, as treatment of cells with methyl ester derivatives of both CA-074b and MB-074, for times reported to be sufficient for ester hydrolysis, lead to a block in labeling of multiple cathepsin species. This suggests that to make cell-permeable, specific inhibitors it will be necessary to modify the n-propyl portion of the molecule such that it imparts cell permeability to the compound without disrupting the dipeptide free-acid portion of the molecule. Perhaps the use of short membrane-targeting sequences in place of the i-propyl group will render the modified peptide cell permeable without loss of specificity for cathepsin B. Several studies have focused on the specific delivery of protease inhibitors to the lysosome by conjugation with protein ligands such as monoclonal antibodies or transferrin [28,29]. Such targeting strategies might be well suited to effective delivery of the otherwise impermeable CA-074.

Radiolabeled forms of the cysteine protease inhibitors Z-Tyr–Ala–CHN_2 and Fmoc-Tyr–Ala–CHN_2 have previously been used to study cathepsins in living cells. Shaw, Mason and coworkers [30,31] were able to approximate concentrations of cathepsin B and L in various cell lines, including several tumor-derived cell lines. They found that levels of cathepsin B and L protease activity correlate with invasiveness of tumors. These studies have shown, however, that it is difficult to quantitate levels of cathepsin B and L activity, mainly because cathepsin B and L comigrate in SDS–PAGE gels and because of the inconsistency of

protein recovery using immunoprecipitation. Our results described here for inhibitor labeling of primary human tissues agree with these reported findings, as well as with studies reporting cathepsin expression in tumors. In addition, introducing ^{125}I-MB-074 as a cathepsin-B-specific label will allow cathepsin B activity to be directly monitored without concern for contributions from overlapping proteases, a problem associated with the less specific labels.

The process by which placental cytotrophoblasts invade the uterine wall and establish connections with maternal blood has been considered to be similar to invasion mechanisms used by tumor cells during metastasis [24]. Studies have focused on the expression of cell-adhesion markers and matrix metalloproteases in *in vitro* invasion models established using purified primary cells [25,26,32]. Our results indicate that cathepsin activity is stimulated when primary cytotrophoblasts are induced to undergo the differentiation cascade leading to the invasive phenotype (Figure 9). Cathepsin B and S activity, in particular, are dramatically up-regulated after exposure of cells to Matrigel. Although these findings are only preliminary, they demonstrate the use of labeled inhibitors to follow protease activity in total extracts. Work is already underway to establish the role these enzyme may play in the invasion pathway. In particular, it is not clear whether activation of cathepsins is a cause or effect of invasion. Answers to such questions are likely to require specific inhibitors that can be used *in vivo* coupled with the labeling techniques described here.

Significance

The lysosomal cysteine proteases of the papain family known as cathepsins perform many tasks within the cell. Reduced or elevated levels of activity of these enzymes are associated with many disease states. Techniques used to specifically monitor protein levels of individual cathepsins often fail to correlate with enzyme activity, whereas methods for determining activity are usually unable to provide absolute discrimination between closely related proteases. For this reason new reagents that allow individual protease activities to be simply and accurately detected by chemically modifying targets are likely to be of value for studying a wide range of proteolytic enzymes. We show here that specific affinity labels such as MB-074 and several other less specific reagents that covalently tag their targets in the process of inhibition provide information about the lysosomal cysteine proteases that would otherwise be difficult or impossible to obtain. Specifically, these reagents show that cathepsins may be differentially expressed in different tumor types and in individual patients. Moreover, using these types of compounds as activity-dependent diagnostics could provide a simple method for profiling of cathepsin activity upon biopsy of tumor tissue, providing information that may be of use for determining an effective route for chemotherapeutic intervention. These studies also show that activity of specific cathepsins is regulated during the process by which cells of the placenta begin to invade the maternal host tissue, establishing blood flow to a developing fetus. Although these studies provide solid examples of the uses for active-site-directed affinity labels, diversification of these reagents is likely to generate additional tools that have great potential for use in studying virtually any process or disease state where proteases are suspected to play an important role.

Materials and methods

Reagents

Morpholineurea-leucinyl-homophenalanyl-vinylsulfone-phenyl (LHVS) and JPM-565 were kind gifts of Hidde Ploegh (Harvard Medical School, Boston, MA). Z-YA-CHN_2 was purchased from Enzyme Systems Products (Dublin, CA). E-64 was purchased from Peptides International (Louisville, KY). CA-074 was a kind gift of James McKerrow (University of California, San Francisco, San Francisco, CA).

Inhibitor synthesis

Synthesis of inhibitors that have not been previously reported are described below. All inhibitors were dissolved in dimethylsulfoxide (DMSO) unless otherwise noted. Typically DMSO stock solutions were prepared such that after dilution, the final working concentration of DMSO did not exceed 1%.

Synthesis of phenolic containing peptide vinyl sulfones

The phenolic vinyl sulfone LHVS-PhOH was synthesized using protocols previously described for several vinyl sulfone inhibitors of the proteasome [33]. Briefly, Boc-homophenylalanine was converted to the corresponding Weinreb Amide, reduced with lithium aluminum hydride to yield Boc-homophenylalanal, which was then converted to the corresponding phenolic vinyl sulfone by reaction with a phosphorous ylide using standard Wittig chemistry. Subsequent removal of the Boc group and coupling to the morpholine urea capped leucine yielded the target compound LHVS-PhOH in good yield. The synthetic details and further applications of this compound will be the focus of a future publication.

Synthesis of CA-074 analogs

Peptide epoxides related to published cathepsin B inhibitor CA-074 were synthesized essentially as described [13,19] (Figure 10). Synthetic details are included in the Supplementary material.

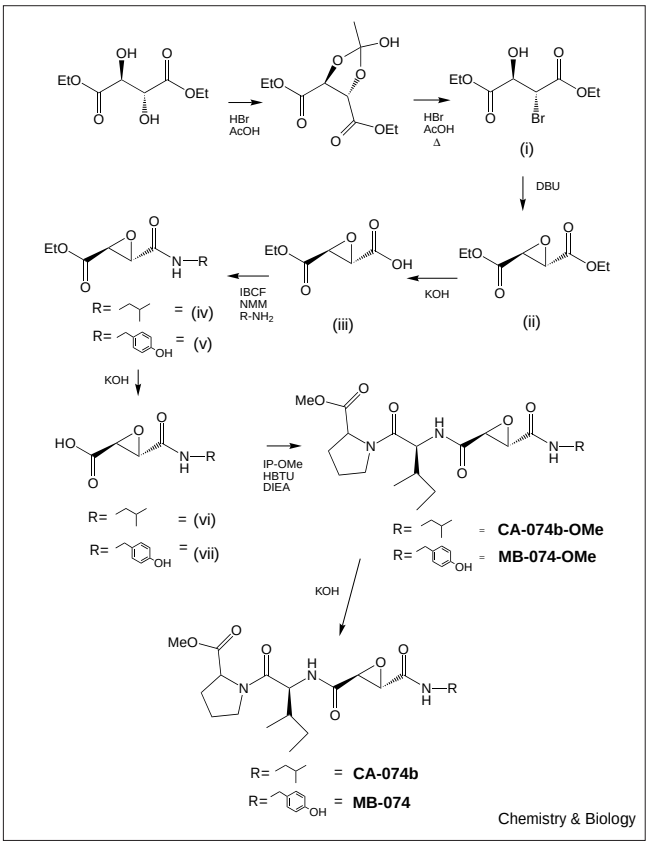
Covalent docking of ligands to the active site of cathepsin B

The 2.0 Å resolution of human cathepsin B crystal structure complexed with CA-030 (pdb identifier: 1CSB) [11] was used in the docking study. Structures of CA-074b, MB-074, I-MB-074, and I_2-MB-074 were generated from the crystal template of CA-030 using the SYBYL program from Tripos [34]. Amber 95 [35] charges were applied to the protein and Gasteiger–Marsilli charges [36] were calculated for the ligands. Using a modified DOCK4.0 [37] algorithm, a covalent bond was placed between the sulfur atom of Cys29 of the protein and the C2 carbon atom (numbering according to Turk *et al.* [11]) of the ligands. The harmonic bond and angle parameters obtained from parm95 [35] of AMBER were applied to the S-C2 bond and CA-S-C2 angle. The ligands were allowed to grow flexibly from the anchor site of 'S-C2'. The resultant top non-bonded force field docked ligand structures were then re-scored using a generalized-Born model to account for the solvation [38].

In vivo *treatment with CA-074b*

Female Balb/c mice were injected interperitoneal with CA-074b (0.25 mg in 500 µl saline) or vehicle (DMSO, 5 µl in 500 µl saline) every 12 h for 2 days as described in [15]. After completion of the injection time-course spleens were removed and spleenocytes isolated as described below.

Figure 10



Synthesis of the peptide epoxides CA-074b-OMe, MB-074-OMe, CA-074b and MB-074.

Cell isolation and culture

Cytotrophoblasts were isolated from pools of first trimester placentas by published methods [25]. In all cases, remaining leukocytes were removed using an antibody to CD-45, a protein tyrosine phosphatase found on bone-marrow-derived cells [39] coupled to magnetic beads. The resulting cells were plated on Matrigel-coated substrates (Collaborative Biomedical Research, Bedford, MA) and cultured in DME H21 minimal essential medium containing 2% Nutridoma (Boehringer Mannheim, Indianapolis, IN) and 50 μg/ml gentamicin [25] under standard tissue culture conditions (5% CO_2/95% air).

Primary mouse spleenocytes were obtained from normal, cathepsin B, cathepsin L and cathepsin S deficient mice (kind gift of Hidde Ploegh, Harvard Medical School, Boston, MA) by physical disruption of an intact spleen submerged in RPMI media supplemented with 10% fetal calf serum. Large tissue matter was removed by filtering through a sterile mesh and cells isolated by centrifugation. Red blood cells were removed by lysis in a hypotonic solution (0.17 M ammonium chloride) followed by centrifugation and washing of the remaining spleenocytes with PBS.

The dendritic cell line DC2.4 was generated by Kenneth Rock and was obtained from the Dana-Farber Cancer Institute (Boston, MA) under transfer agreement. All cell lines (U937, DC2.4) were maintained in RPMI medium supplemented with 10% fetal calf serum using standard tissue culture protocols.

Kinetic assays

Simple IC_{50} values were determined by pre-treatment of recombinant cruzain lacking the carboxy-terminal domain (4 nM; kind gift of James McKerrow) or cathepsin B (4 nM; Sigma, St. Louis, MO) in 100 assay buffer (100 mM sodium acetate pH 5.5, 10 mM DTT) with each of the three inhibitors CA-074, CA-074b, and MB-074 over a 10–10000 nM range for 5 min at room temperature. The substrate Z-FR-MCA (100 μM; Bachem, King of Prussia, PA) was then added and progress curves monitored for 10 min. Slopes of each linear curve were determined by linear least squares fitting and recorded as a ratio of uninhibited versus inhibited reaction rates. IC_{50} values were approximated by extrapolation from plots of concentration vs % residual activity to determine inhibitor concentrations that resulted in 50% inhibition of hydrolysis activity.

Values of $k_{obs}/[I]$ were obtained as follows. Inhibitor dilutions that resulted in simple progress curves over a wide range of k_{obs} were used to determine kinetic parameters. Progress curves of Z-FR-MCA (20 μM for cruzain; 100 μM for cathepsin B) hydrolysis in the presence of inhibitor were measured for 10 min. Values for k_{obs}, the pseudo-first order rate constant for inactivation of enzyme, were obtained for the equation:

$$\text{Fluorescence} = v_s t + [(v_o - v_s)/k_{obs}][1 - \exp(-k_{obs} t)]$$

for pseudo-first order dynamics using UltraFit (Biosoft). Kinetic data is displayed as averages of multiple inhibitor concentrations performed in duplicate for $k_{obs}/[I]$ where [I] is the concentration of inhibitor used to obtain each progress curve.

Radiolabeling of inhibitors

All compounds were iodinated and isolated using the reported protocol [31,40]. Briefly, a glass test tube was coated with 100 μg of Iodo-gen® (Pierce, Rockford, IL). The tube was placed on ice and inhibitor (25 μl of a 5 mM solution in ethanol/phosphate buffer pH 7.5 (v:v)) was added to the tube. Finally, 1 mCi of $Na^{125}I$ (10 μl) was added to the tube and incubation continued on ice for 15 min. The labeled inhibitor was purified by application to a sep-pak® (Waters, Milford, MA) column containing a C18 stationary phase. After the sample was applied to the column, the column was washed with 30 ml of phosphate buffer pH 7.5. Labeled inhibitor was then eluted using 100% acetonitrile. Fractions of 1 ml were collected and the fractions with the greatest number of counts were pooled.

Preparation of total cellular lysates from DC2.4, U937, primary spleen cells, cytotrophoblasts and primary human tissue

Cells were lysed using a bead smashing technique described previously [33]. Briefly, cells were washed 3× with cold PBS and then collected by centrifugation. A volume of glass beads (< 106 microns, acid washed; Sigma Chem. Co., St. Louis, MO) equivalent to the volume of the pellet were added followed by a similar volume of homogenization buffer (50 mM Tris pH 5.5, 1 mM DTT, 5 mM $MgCl_2$, 250 mM sucrose). Cells were vortexed at high speed for 1 min. The beads and broken cell debris were removed by centrifugation at 10,000 g for 5 min. The resultant homogenate was centrifuged at 10,000 g for 20 min to remove unbroken cells and nuclei. Protein concentration was determined using the BCA protocol described by the manufacturer (Pierce Chem. Co., Rockford, IL).

Frozen tissue pieces from time-matched colon cancer sets, or individual tumors (tumor, nearby normal tissue, and liver metastasizes stored at –70°C), were homogenized in 250 μl/ mg ice cold PBS with 0.1% Triton × 100 using a Tenbrock homogenizer on ice. The homogenate was centrifuged in a 4°C microfuge at 16 K g for 10 min. High salt extracts were prepared by resuspending these pellets at 10 μl/mg in 50 mM Na acetate pH 6.1, 2 M NaCl , vortexing for 3 min and collecting the supernatants as above. Pellets and supernatants were stored at –70°C. Protein levels were determined for each sample by quantitation using the BCA method. Equal amounts of total protein were used for all labeling experiments.

Labeling of lysates and intact cells with iodinated inhibitors

Equivalent amounts of radioactive inhibitor stock solutions (~ 10^6 cpm per sample) were used for all labeling experiments. It should be noted, however, that not all inhibitors may have been modified to the same

extent during iodination [40]. It is difficult to determine the amount of unlabeled inhibitor present in a given labeling stock so these reagents should not be used as quantitative measures of inhibitory or protease activity. Relative inhibition of radiolabeling of specific proteases by cold competition can be used to make quantitative observations. Samples of lysates (50 μg total protein in 50 μl buffer; 50 mM tris pH 5.5, 5 mM $MgCl_2$, 2 mM DTT) were labeled for 2 h on ice unless noted otherwise and were quenched by dilution of 4× SDS sample buffer to 1× (for one-dimensional SDS–PAGE) or by dissolving urea to a final concentration of 9.5 M (for two-dimensional SDS–PAGE).

Competition labeling experiments

Nonradioactive inhibitors (CA-074, CA-074b, MB-074, MB-074-OMe, CA-074b-OMe, E-64 and LHVS) were added to lysates on ice (prepared as described above) at the concentrations indicated. Samples were pre-incubated for 1 h and then labeled by addition of the inhibitor (^{125}I-JPM-565, ^{125}I-LHVS-PhOH; ~10^6 cpm/sample). Labeling was quenched after 2 h on ice by addition of 4× SDS sample buffer to a final concentration of 1× followed by boiling and analysis by SDS–PAGE.

Gel electrophoresis

One-dimensional SDS–PAGE, two-dimensional NEPHGE were performed as described previously [41].

Supplementary material

Supplementary material including the synthesis of MB-074 and CA-074b is available at http://current-biology.com/supmat/supmatin.htm.

Acknowledgements

We would like to thank Jim McKerrow, Juan Engel and Elizabeth Hansell (UCSF) for the preparation of tissue extracts from human tumors and for assistance with enzyme kinetics. We thank Hidde Ploegh and Rebecca Bryant (Harvard Medical School) for preparation of spleenocytes from cathepsin knock-out mice and for critical evaluation of the manuscript. We thank Erwin Kuntz, Susan Fisher, Richard Locksley and Charles Craik (UCSF) for use of reagents and resources and for critical discussions of this work. Finally, we thank Ben Kelly (UCSF) for assistance with our *in vivo* mouse study. This work was supported by funding from the Sandler Foundation (M.B. and D.G.).

References

1. Chapman, H.A., Reese, R.J. & Shi, G.-P. (1997). Emerging roles for cysteine proteases in human biology. *Annu. Rev. Physiol.* 59, 63-88.
2. Chapman, H.A. (1998). Endosomal proteolysis and MHC class II fundtion. *Curr. Opin. Immunol.* 10, 93-102.
3. Barrett, A.J., Rawlings, N.D. & Woessner, J.F. (1998). *Handbook of Proteolytic Enzymes.* Academic Press, London.
4. Sloane, B.F., Moin, K., Krepela, E. & Rozhin, J. (1990). Cathepsin B and its endogenous inhibitors: the role in tumor malignancy. *Cancer Metastasis Rev.* 9, 333-352.
5. Werb, Z. (1989). Proteinases and matrix degredation. In *Textbook of Rheumatology.* (W.N. Keller, E.D. Harris, S. Ruddy & C.S. Sledge, eds), pp. 300-321. W.B. Saunder Co., Philadelphia.
6. Katunuma, N.,& Kominami, E. (1987). Abnormal expression of lysosomal cysteine proteinases in muscle wasting diseses. *Rev. Physiol. Biochem. Pharmacol.* 108, 1-20.
7. Otto, H.-H. & Schirmeister, T. (1997). Cysteine proteases and their inhibitors. *Chem. Rev.* 97, 133-171.
8. Shaw, E. (1990). Cysteinyl proteinases and their selective inactivation. *Adv. Enzymol. Relat. Areas Mol. Biol.* 63, 271-347.
9. Towatari, T., Nikawa, T., Murata, M., Yokoo, C., Tamai, M., Hanada, K. & Katunuma, N. (1991). Novel epoxysuccinyl peptides; a selective inhibitor of cathepsin B, *in vivo*. *FEBS Lett.* 280, 311-315.
10. Murata, M., *et al.,* & Katunuma, N. (1991). Novel epoxysuccinyl peptides; a selective inhibitor of cathepsin B, *in vitro*. *FEBS Lett.* 280, 307-310.
11. Turk, D., *et al.,* & Turk, V. (1995). Crystal structure of cathepsin B inhibited with CA-030 at 2.0 Å resolution: a basis for the design of specific epoxysuccinyl inhibitors. *Biochemistry* 34, 4791-4797.
12. Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T. & Mort, J.S. (1997). Role of the occluding loop in cathepsin activity. *J. Biol. Chem.* 272, 1197-1202.
13. Meara, J.P. & Rich, D.H. (1996). Mechanistic studies on the inactivation of papain by epoxysuccinyl inhibitors. *J. Med. Chem.* 39, 3357-3366.
14. Matsunaga, Y., Saibara, T., Kido, H. & Katunuma, N. (1993). Participation of cathepsin B in processing of antigen presentation to MHC class II. *FEBS Lett.* 324, 325-330.
15. Maekawa, Y., *et al.,* & Katunuma, N. (1998). Switch of CD4+ T cell differentiation from Th2 to Th1 by treatment with cathepsin B inhibitor in experiemental Leishmaniasis. *J. Immunol.* 161, 2120-2127.
16. Katunuma, N., *et al.,* & Kakiuchi, T. (1998). Novel physiological functions of cathepsins B and L on antigen processing and osteoclastic bone resorption. *Adv. Enzyme Reg.* 235-251.
17. Buttle, D.J., Murata, M., Knight, C.G. & Barrett, A.J. (1992). CA074 methyl ester: a proinhibitor for intracellular cathepsin B. *Arch. Biochem. Biophys.* 299, 377-380.
18. Barrett, A.J. (1986). Introduction to proteinases. In *Protease Inhibitors.* (A.J. Barrett & G. Salvesen, eds), pp. 3-22. Elsevier, Amsterdam.
19. Gour-Salin, B.J., Lachance, P., Plouffe, C., Storer, A.C. & Ménard, R. (1993). Epoxysuccinyl dipeptides as selective inhibitors of cathepsin B. *J. Med. Chem.* 36, 720-725.
20. Shi, G-P., Munger, J.S., Meara, J.P., Rich, D.H. & Chapman, H.A. (1992). Molecular cloning and expression of human aveolar macrophage cathepsin S, an elastinolytic cysteine protease. *J. Biol. Chem.* 267, 7258-7262.
21. Palmer, J.T., Rasnick, D., Klaus, J.L. & Bromme, D. (1995). Vinyl sulfones as mechanism-based cysteine protease inhibitors. *J. Med. Chem.* 38, 3193-3196.
22. Brömme, D., Klaus, J.L., Okamoto, K., Rasnick, D. & Palmer, J.T. (1996). Peptidyl vinyl sulfones: a new class of potent and selective cysteine protease inhibitors. *Biochem J.* 85-89.
23. Riese, R.J., *et al.,* & Chapman, H.A. (1998). Cathepsin S activity regulates antigen presentation and immunity. *J. Clin. Invest.* 101, 2351-2363.
24. Damsky, C.H. & Fisher, S.J. (1998). Trophoblast pseudo-vasculogenesis: faking it with endothelial adhesion receptors. *Curr. Opin. Cell Biol.* 10, 660-666.
25. Librach, C.L., *et al.,* & Fisher, S.J. (1991). 92-kD type IV collagenase mediates invasion of human cytotrophoblasts. *J. Cell Biol.* 113, 437-449.
26. Damsky, C.H., *et al.,* & Fisher, S.J. (1994). Integrin switching regulates normal trophoblast invasion. *Development* 120, 3657-3666.
27. Zhou, Y., *et al.,* & Damsky, C. H. (1997). Human cytotrophoblasts adopt a vascular phenotype as they differentiate. *J. Clin. Invest.* 99, 2139-2151.
28. Xing, R., Wu, F. & Mason, R. (1998). Control of breast tumor cell growth using a targeted cysteine protease inhibitor. *Cancer Res.* 58, 904-909.
29. Xing, R., & Mason, R. (1998). Design of a transferrin-proteinase inhibitor conjugate to probe for active cysteine proteinases in endosomes. *Biochem. J.* 336, 667-673.
30. Mason, R.W., Wilcox, D., Wilkstrom, P. & Shaw, E.N. (1989). The identification of active forms of cysteine proteinases in Kirsten-virus-transformed mouse fibroblasts by use of a specific radiolabeled inhibitor. *Biochem. J.* 257, 125-129.
31. Xing, R., Addington, A.K. & Mason, R.W. (1998). Quantification of cathepsin B and L in cells. *Biochem. J.* 332, 499-505.
32. Fisher, S.J., *et al.,* & Damsky, C.H. (1989). Adhesive and degradative properties of human placental cytotrophoblast cells in vitro. *J. Cell Biol.* 109, 891-902.
33. Bogyo, M., Shin, S., McMaster, J.S. & Ploegh, H. (1998). Substrate binding and sequence preference of the proteasome revealed by active-site-directed affinity probes. *Chem. Biol.* 5, 307-320.
34. SYBYL (1999). Tripos Associates, St. Louis, MO.
35. Cornell, W.D., *et al.,* & Kollman, P.A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117, 5179-5197.
36. Gasteiger, J. & Marsili, M. (1980). Iterative partial equilization of orbital electronegativity – a raid access to atomic charges. *Tetrahedron* 36, 3219-3228.
37. Ewing, T. & Kuntz, I.D. (1998). *DOCK*. UCSF, San Francisco.
38. Zou, X., Sun, Y. & Kuntz, I.D. (1999). Inclusion of solvation in ligand binding free energy calculations using the generalized-born model. *J. Am. Chem. Soc.,* 121, 8033-8043.
39. Charbonneau, H., *et al.,* & Walsh, K.A. (1989). Human placenta protein-tyrosine-phosphatase: amino acid sequence and relationship to a family of receptor-like proteins. *Proc. Natl Acad. Sci. USA* 86, 5252-5256.
40. Bogyo, M., McMaster, J.S., Gaczynska, M., Tortorella, D., Goldberg, A.L. & Ploegh, H.L. (1997). Covalent modification of the active site Thr of proteasomal β-subunits and the *E. coli* homolog HslV by a new class of inhibitors. *Proc. Natl Acad. Sci. USA* 94, 6629-6634.
41. Monaco, J.J., & McDevitt, H.O. (1986). The LMP antigens: a stable MHC-controlled protein complex. *Hum. Immunol.* 15, 416-426.

Supplementary material

Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs

Matthew Bogyo, Steven Verhelst, Valerie Bellingard-Dubouchaud, Sam Toba and Doron Greenbaum

Chemistry & Biology 2000, 7:27–38

*Synthesis of diethyl (2*R,*3*S*)-2-bromo-3-hydroxy-butanedioate (**I**).* A 30% solution of HBr in AcOH (35 ml) was added to diethyl-D-tartrate (7.2 g, 6 ml, 35 mmol) and stirred for 1 h at 0°C. The solution was allowed to warm up to room temperature and stirred overnight. Then it was poured into ice-water, and extracted with ether (3×). The combined ether layers were washed with water and brine, dried (Na_2SO_4), and concentrated. The remaining oil was dissolved in EtOH (35 ml). A 30% solution of HBr in AcOH (21 ml) was added, and the resulting solution was refluxed for 4 h, and then concentrated *in vacuo*, removing most of the acetic acid by co-evaporation with toluene. The yellow oil was dissolved in EtOAc and washed with water, sat. $NaHCO_3$, brine and dried ($MgSO_4$). Silica column chromatography (16–25% EtOAc in hexanes) yielded I (5.67 g, 21.1 mmol, 60%). Diethyl (2*R*,3*S*)-2-bromo-3-hydroxy-butanedioate: 1H NMR (300 MHz, $CDCL_3$) δ 1.33 (m, 6H), 2.1 (d, 1H), 4.19–4.36 (m, 4H), 4.7 (dd, 1H).

*Synthesis of diethyl (2*S,*3*S*)-oxirane-2,3-dicarboxylate (**II**).* To I (2.67 g, 9.9 mmol) in ether (25 ml) DBU (1.63 ml, 10.9 mmol, 1.1 eq) was added dropwise. After 0.5 h TLC analysis revealed complete conversion of the starting material. The solid was filtered off, and the filtrate was washed with citric acid buffer (pH = 3), brine, and dried ($MgSO_4$). Silica column chromatography (9–13% EtOAc in hexanes) provided II (1.42 g, 7.53 mmol, 76%). Diethyl (2S,3S)-oxirane-2,3-dicarboxylate: ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.33 (t, 6H), 3.68 (s, 2H), 4.19–4.36 (m, 4H).

*Synthesis of ethyl (2*S,*3*S*)-oxirane-2,3-dicarboxylate (**III**).* To a solution of II (1.35 g, 7.18 mmol) in EtOH (35 ml) at 4°C was slowly added a solution of KOH in ethanol (7.2 ml, 1.00 M). After 3 h the reaction was allowed to warm to room temperature and stirred for an additional hour. The solvent was removed under reduced pressure. The residue was dissolved in water, and extracted with EtOAc (2×). The resulting water layer was acidified to pH 1–2 with 3 M hydrochloric acid, and extracted with EtOAc (4×). The organic layers were washed with a 1:1 mixture of brine and 1 M hydrochloric acid, dried ($MgSO_4$), and concentrated *in vacuo*, yielding a colorless oil (872 mg, 5.45 mmol, 76%). The compound was used without further purification. Ethyl (2S,3S)-oxirane-2,3-dicarboxylate: ^{1}H NMR (300 MHz, DMSO-d_6) δ 1.22 (t, 3H), 3.60 (d, 1H), 3.69 (d, 1H), 4.17 (q, 2H).

*Synthesis of N-[[*L*-*trans*-3(ethoxycarbonyl)oxiran-2-yl]- 3-methylbutaneamide (**IV**).* Epoxy acid III (400 mg, 2.5 mmol) was dissolved in anhydrous THF under inert atmosphere and N-methyl morpholine (NMM; 303 µl, 2.75 mmol) added dropwise while stirring. The reaction was cooled to –10°C in an ice/sodium chloride slurry and isobutylchloroformate (IBCF; 358 µl, 2.75 mmol) was added slowly. To the stirring solution was added i-butylamine (273 µl, 2.75 mmol). The reaction was stirred for 1h at reduced temperature then allowed to warm to room temperature for 2h. The reaction mixture was diluted with ethyl acetate and the organic layer washed with 3× volumes of brine, water and 0.1 N HCl. The remaining organic layer was dried, concentrated by rotary evaporation resulting in a crude oil. Flash chromatography yielded the desired product as a pale yellow oil (430 mg, 2.0 mmol, 80%). N-[[L-*trans*-3(ethoxycarbonyl)oxiran-2-yl]-3-methylbutaneamide (IV): ^{1}H NMR (300 MHz, $CDCl_3$) δ 0.83 (d, 6H), 1.32 (t, 3H), 1.92 (m, 1H), 3.24 (t, 2H), 3.60 (d, 1H), 3.72 (d, 1H), 4.32 (m, 2H), 6.19 (d, 1H).

*Synthesis of N-[[*L*-trans-3(ethoxycarbonyl)oxiran-2-yl]-tyramide (**V**).* Compound V was synthesized exactly as described for compound IV except tyramine was used instead of i-butylamine. N-[[L-*trans*-3(ethoxycarbonyl)oxiran-2-yl]-tyramine amide (V): ^{1}H NMR (300 MHz, $CDCl_3$) δ 1.35 (t, 3H), 2.77 (m, 2H), 3.39 (d, 1H), 3.48 (m, 2H), 3.64 (d, 1H), 4.27 (m, (2H), 6. 33 (t, 1H), 6.81 (d, 2H), 6.99 (d, 2H).

*Synthesis of N-[(*L*-trans-3-carboxyoxirane-2-yl)carbonyl]-tyramide (**VI**) and N-[(*L*-trans-3-carboxyoxirane-2-yl)carbonyl]-3-methyl-butaneamide (**VII**).* Compounds VI and VII were synthesized by removal of the ethyl ester of compounds V and IV, respectively. Ester hydrolysis was achieved by dissolving each compound in a minimal volume of ethanol followed by addition of 1 equivalent of sodium hydroxide as a 1 N aqueous solution. The resulting reaction was stirred for 2 h at room temperature at which time 1 equivalent of acetic acid was added and water added to the mixture. The water solution was washed with 3 volumes of ethyl acetate. The organic layers were combined, dried and concentrated by rotary evaporation. The crude colorless oils VI and VII were used without further purification.

Synthesis of Boc-isoleucine-proline methyl ester (Boc-IP-OMe). Boc-isoleucine (3.17 g, 13.2 mmol) and hydroxybenzatriazole (HOBt; 1.95 g, 14.4 mmol) were dissolved in a minimal amount of a 1:1 (v:v) solution of DMF:CH_2Cl_2. Dicyclohexylcarbodiimide (DCC; 2.72 g, 13.2 mmol) was added as a solid in one portion to the stirring reaction mixture. The resulting reaction was stirred for 30 min at room temperature, the solid removed by filtration and the supernatant added to a flask containing proline methyl ester hydrochloride salt (2.0 g, 12 mmol). Diisopropylethylamine (DIEA, 4.25 ml, 24 mmol) was added and the reaction stirred at room temperature for 1 h. Solvent was removed by rotary evaporation resulting in a yellow oil, which was dissolved in ethyl acetate, washed with 3× volumes of 0.1 N hydrochloric acid, 3× volumes saturated sodium carbonate, dried and re-concentrated by rotary evaporation. The crude oil was purified by flash chromatography (1:3; v:v ethyl acetate:hexanes) yielding Boc-IP-OMe as a colorless oil (3.6 g, 10.3 mmol, 84%). Boc-IP-OMe: ^{1}H NMR (300 MHz, $CDCl_3$) δ 0.82 (t, 3H), 1.04 (d, 3H), 1.24 (m, 2H), 1.44 (m, 9H), 1.76 (m, 1H), 2.08 (m, 2H), 2.36 (q, 2H), 3.67 (m, 2H), 3.65 (s, 3H), 4.35 (t, 1H), 4.55 (m, 1H), 5.24 (d, 1H).

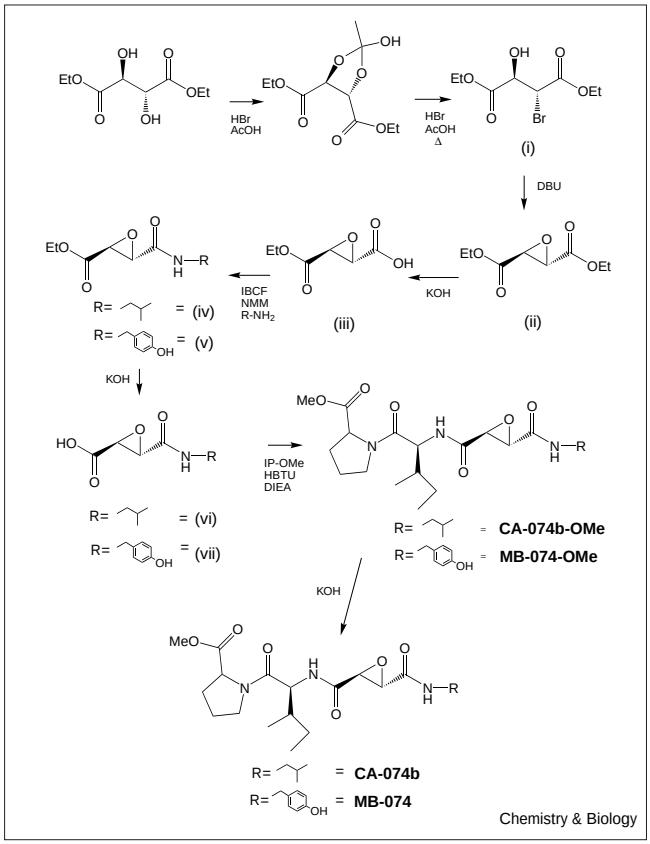
Synthesis of NH_2-isoleucine–proline-OMe (IP-OMe). The Boc group was removed by dissolving IP-OMe in a minimal amount of methylene chloride and adding an equal volume of trifluoroacetic acid (TFA). The reaction was stirred for 30 min at room temperature, a 2× volume of toluene added and the product concentrated by rotary evaporation. The resulting crude clear oil was used without further purification.

Synthesis of MB-074-OMe. Epoxy acid VI (71 mg, 308 µmol) was dissolved in a minimal volume of a 1:1 (v:v) mixture of methylene chloride:dimethylformamide. To the solution was added HBTU (140 mg, 369 µmol) followed by NH_2-IP-OMe (169 mg, 462 µmol) and then DIEA (136 µl, 770 µmol). The reaction was quenched by dilution into ethyl acetate followed by extraction with 3× volumes of brine and 0.1 N HCl. The resulting organic phase was dried and concentrated by rotary evaporation to yield a crude oil which upon purification by flash chromatography (4:1 ethyl acetate:hexanes) yielded MB-074-OMe as a white solid (60.6 mg, 125 µmol, 55%). MB-074-OMe ^{1}H NMR (300 MHz, $CDCl_3$) δ 0.81 (t, 3H), 1.09 (d, 3H), 1.18 (m, 2H), 1.56 (m, 2H), 1.92 (m, 2H), 2.05 (m, 2H), 2.31 (m, 2H), 2.65-2.82 (dm, 2H), 2.83 (s, 2H), 3. 19 (s, 1H), 3.42 (m, 1H), 3.58 (m, 1H), 3. 77 (m, 2H), 3.78 (s, 3H), 4. 42 (m, 1H), 4.62 (t, 1H), 6. 36 (t, 1H), 6.79 (d, 2H), 6.99 (d, 2H).

Synthesis of CA-074b-OMe. CA-074b-OMe was synthesized exactly as described for MB-074-OMe except epoxy acid VII was used instead of epoxy acid VI. CA-074b-OMe 1H NMR (300 MHz, $CDCl_3/CD_3OD$) δ 0.75 (m, 6H), 1.12 (m, 3H), 1.19 (m, 2H), 1.58-1.70 (dm, 1H), 1.85 (m, 1H), 2.16 (m, 2H), 2.34 (m, 2H), 2.82-2.95 (dd, 2H), 3.16 (m, 2H), 3.38 (s, 1H), 3.58 (s, 1H), 3.63–3.92 (dm, 2H), 3.76 (s, 3H), 4. 42 (m, 1H), 4.62 (t, 1H).

Synthesis of MB-074 and CA-074b. MB-074 and CA-074b were synthesized from MB-074-OMe and CA-074b by hydrolysis of the corresponding ester. Ester hydrolysis was achieved by dissolving each compound in a minimal volume of ethanol followed by addition of 1 equivalent of sodium hydroxide as a 1N aqueous solution. The resulting reaction was stirred for 2 h at room temperature or until the reaction was judged complete by TLC analysis. The reaction mixture was concentrated to dryness by rotary evaporation and the residue re-dissolved in a minimal amount of methanol. Ice cold diethyl ether was added and the product sodium salt was collected by centrifugation and washed with two portions of diethyl ether. The inhibitors were deemed pure by HPLC analysis and used directly for all further experiments. MB-074: ^{1}H NMR (300 MHz, $CDCl_3$) δ 0.81 (t, 3H), 1.09 (d, 3H), 1.18 (m, 2H), 1.56 (m, 2H), 1.92 (m, 2H), 2.05 (m, 2H), 2.31 (m, 2H), 2.75-2.92 (dm, 2H), 2.83 (s, 2H), 3. 19 (s, 1H), 3.42 (m, 1H), 3.58 (m, 1H), 3. 77 (m, 2H), 4.32 (t, 1H)4. 42 (m, 1H), , 6. 36 (t, 1H), 6.79 (d, 2H), 6.99 (d, 2H). CA-074b 1H NMR (300 MHz, $CDCl_3/CD_3OD$) δ 0.75 (m, 6H), 1.12 (m, 3H), 1.19 (m, 2H), 1.58-1.70 (dm, 1H), 1.85 (m, 1H), 2.16 (m, 2H), 2.34 (m, 2H), 2.82–2.95 (dd, 2H), 3.16 (m, 2H), 3.38 (s, 1H), 3.58 (s, 1H), 3.63–3.92 (dm, 2H), 4. 42 (m, 1H), 4.62 (t, 1H).

Figure S1



Synthesis of the peptide epoxides CA-074b-OMe, MB-074-OMe, CA-074b and MB-074.