Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry xxx (2008) xxx–xxx
Bioorganic & Medicinal Chemistry

Evaluation of α,β-unsaturated ketone-based probes for papain-family cysteine proteases

Zhimou Yang,[a] Marko Fonović,[a,c] Steven H. L. Verhelst,[a] Galia Blum[a] and Matthew Bogyo[a,b,*]

[a]*Department of Pathology, Stanford School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA*
[b]*Department of Microbiology and Immunology, Stanford School of Medicine, 300 Pasteur Drive, Stanford, CA 94305, USA*
[c]*Department of Biochemistry and Molecular Biology, Jozef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia*

Received 15 December 2007; revised 21 February 2008; accepted 27 February 2008

Abstract—The field of activity-based proteomics makes use of small molecule active site probes to monitor distinct subsets of enzymatic proteins. While a number of reactive functional groups have been applied to activity-based probes (ABPs) that target diverse families of proteases, there remains a continual need for further evaluation of new probe scaffolds and reactive functional groups for use in ABPs. In this study we evaluate the utility of the, α,β-unsaturated ketone reactive group for use in ABPs targeting the papain-family of cysteine proteases. We find that this reactive group shows highly selective labeling of cysteine cathepsins in both intact cells and total cell extracts. We observed a variable degree of background labeling that depended on the type of tag and linker used in the probe synthesis. The relative ease of synthesis of this class of compounds provides the potential for further derivatization to generate new families of cysteine protease ABPs with unique specificity and labeling properties.

1. Introduction

Most enzymes are regulated by a complex set of controls that prevent them from causing damage to a cell as the result of their uncontrolled activation. Proteases are no exception, as virtually all members of this enzyme class are regulated by initial activation of a zymogen form and then subsequent temporal control by endogenous inhibitors. Thus, methods such as transcript array profiling and standard proteomics are unable to provide information about the dynamic, post-translational regulatory mechanisms used to control the networks of proteolytic events involved in basic cell physiology or disease pathology. Therefore, it is necessary to develop tools that allow the activity levels of specific protease targets to be monitored within the context of a complex biological environment.

Activity-based probes (ABPs) are small molecules that form activity dependant covalent bonds to a target enzyme.[1–5] These probes contain three main elements: (1) a reactive functional group that facilitates the formation of the covalent bond with the active site catalytic residue of a target (2) a linker that can be used to control the specificity of binding interactions between the probe and target enzyme and (3) a tagging group that allows probe labeled targets to be isolated, biochemically characterized or imaged. While the past 10 years has seen a significant growth in both the diversity of ABPs and the types of enzymes that can be targeted by these probes,[6,7] there remains a need to develop new classes of ABPs in order to continue to expand the repertoire of proteins that can be studied using chemical proteomic methods. In particular, new reactive 'warheads' need to be evaluated for use with general probe scaffolds. These reactive functional groups often control the broad selectivity of the ABP and also have a dramatic impact on the overall cross-reactivity resulting from non-specific interactions with abundant background proteins.

While a number of effective reactive functional groups have been employed in ABPs that target cysteine proteases (for reviews see Refs. 2, 3, 7, and 8), there still remains room for improvement both at the level of

Keywords: Activity-based probes; Cysteine protease; Active site labeling; Irreversible inhibitor.
* Corresponding author. Tel.: +1 650 725 4132; fax: +1 650 725 7424; e-mail: mbogyo@stanford.edu

0968-0896/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2008.02.089

selectivity and also at the level of ease of synthesis. The peptide vinyl sulfones have found widespread use in probes that target papain-family cysteine proteases[9–11] as well as ubiquitin specific proteases[12–14] and the proteasome.[15,16] However, this reactive functional group is not effective for all classes of cysteine proteases. The highly related Michael acceptor, the α,β-unsaturated ketone has been extensively used in the development of inhibitors of specific classes of cysteine proteases.[17,18] Interestingly, while there is ample evidence to suggest that the α,β-unsaturated ketone has different specificity properties relative to the vinyl sulfone, this class of warhead has only been used in a single class of highly specific activity-based probes that target the deubiquitinating enzymes (DUBs).[19] This class of probe derives the majority of its selectivity from the large 76 amino-acid ubiquitin protein that is attached to the reactive warhead group. Thus, it remains unclear how valuable the α,β-unsaturated ketone is as a more general warhead for use with short peptide-based ABPs.

In this study we synthesized a number of simple peptide α,β-unsaturated ketones and tested them as general activity-based probes of the cysteine protease family. We find that biotinylated and fluorescent α,β-unsaturated ketones show very potent labeling of cathepsins with low overall background labeling in total tissue and cell extracts, respectively. Compared with the previously validated acyloxymethyl ketone (AOMK)-based probes, the α,β-unsaturated ketone showed enhanced labeling in tissue and cell extracts and similar levels of labeling in intact cells. Importantly, the α,β-unsaturated ketones were several orders of magnitude more potent toward the recombinant cathepsins B and L compared to a related AOMK, suggesting that they may provide higher levels of signal for in vivo imaging using the methods outlined for the AOMK-based ABPs.[20] Furthermore the α,β-unsaturated ketones can be readily synthesized from the corresponding peptide Weinreb amides and can therefore be used to make diverse peptide sequences for evaluations of new classes of specific ABPs.

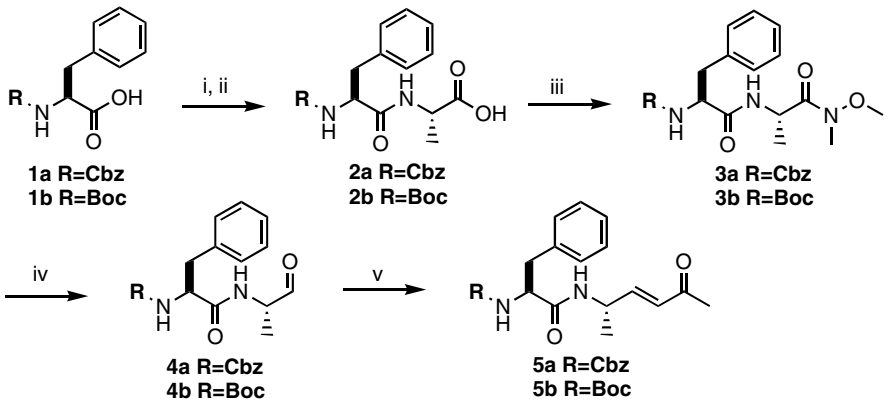
2. Results and discussion

2.1. Evaluation of the α,β-unsaturated ketone as a warhead for use in ABPs

We initially set out to evaluate the utility of the α,β-unsaturated ketone (αβUK) as a reactive group for use in ABPs that target cysteine proteases. We identified a related αβUK in a small molecule screen for compounds that block host cell invasion by the obligate intra-cellular parasite *Toxoplasma gondii*.[21] Based on the structure of this hit we designed and synthesized a series of analogs that contain a central P2 phenylalanine and P1 alanine peptide attached to the α,β-unsaturated methyl ketone. The simplest analog, **yzm09** (**5a**) contains a Cbz cap at its N-terminus. This capping group is replaced by biotin (**yzm13**, **5c**) or by a fluorescent tag (**yzm24**, **5d**) to yield the first generation ABPs. The synthesis of **yzm09** was accomplished by preparation of the dipeptide Z-Phe-Ala (**2a**) followed by conversion to the corresponding Weinreb amide (**3a**); the reduction of the Weinreb amide and further treatment with the Wittig reagent of 1-triphenyl-phosphoranylidene-2-propanone afforded a mixture of the desired product with trans-conformation and its cis isomer (Scheme 1). The ratio of trans/cis was determined to be approximately 4/1 based on HPLC analysis with the ratio being highly dependant on the steric properties of the P1 sidechain as has been shown for peptide vinyl sulfones.[22] The desired compound with trans-conformation (**yzm09**, **5a**) was separated by HPLC in reasonable yield (36% for five steps). The labeled analogs of **yzm09** were prepared from the Boc-protected compound **5b**. After deprotection of the Boc-group, biotin or BODIPY-TMR-X was used to generate the corresponding labeled probes (**yzm13** and **yzm24**, respectively, in Scheme 2).

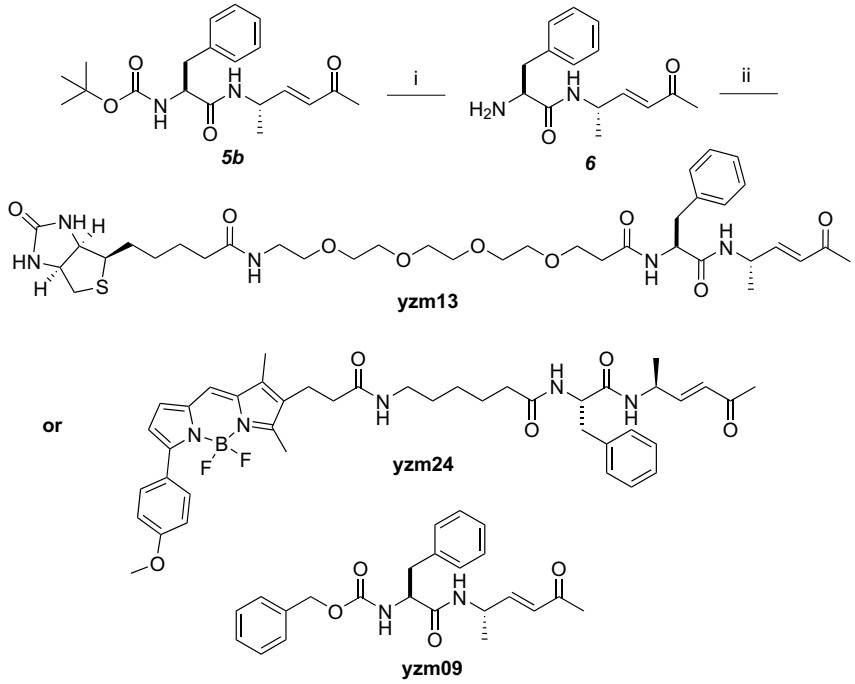
2.2. Evaluation of the biotinylated ABP yzm13

Having successfully synthesized two labeled analogs of the parent lead compound **yzm09**, we began testing the biotin probe **yzm13** for labeling of target proteases in



Scheme 1. Synthetic scheme for **yzm09** and **5b**. Reagents: (i) DCC, NHS, $CHCl_3$; (ii) L-alanine, $NaHCO_3$, acetone/H_2O; (iii) HBTU, DIPEA, *N*,*O*-dimethyl-hydroxylamine hydrochloride, DMF; (iv) $LiAlH_4$, THF; (v) 1-triphenyl-phosphoranylidene-2-propanone, DCM.

Please cite this article in press as: Yang, Z. et al., *Bioorg. Med. Chem.* (2008), doi:10.1016/j.bmc.2008.02.089



Scheme 2. Synthetic scheme for **yzm13** and **yzm24**. Reagents: (i) 4 M HCl in dioxane; (ii) DIPEA, acetone.

complex proteomes. We initially used **yzm13** in crude mouse tissue extracts derived from liver, spleen, heart, kidney, ovary, and uterus. Total protein extracts at pH 5.5 were pre-incubated with DMSO vehicle or the unlabeled parent compound **yzm09** followed by the addition of the biotin probe **yzm13** at a range of probe concentrations. Labeled proteins were analyzed by SDS–PAGE followed by detection using HRP-coupled streptavidin (Fig. 1). These initial labeling results confirmed that the biotin probe **yzm13** strongly labeled one predominant protein just below the 36 kDa molecular weight marker. The labeling of this predominant protein was completely blocked by pre-treatment of the samples with 10 μM concentrations of **yzm09** suggesting that it was the only selectively labeled protein in the extract. While a number of higher molecular weight species were observed at high concentrations of the probe, overall **yzm13** showed highly specific labeling of a single target that could be observed at probe concentrations as low as 10 nM. Thus, the αβUK is a potentially valuable warhead for use in ABPs as it shows effective labeling at very low concentrations of the probe.

We next wanted to identify the predominant labeled protease observed with **yzm13** in mouse tissue extracts. We therefore performed an affinity purification experiment in which mouse liver lysates with or without the pre-treatment of 10 μM of **yzm09** were labeled with 1.0 μM of **yzm13**. After enrichment of biotin labeled proteins on immobilized avidin, both samples were processed by 'on bead' trypsin digestion to release peptides for subsequent analysis by LC–MS/MS.[23] Analysis of all isolated peptides confirmed the presence of cathepsin B (Cat B) as well as several background proteins including endogenously biotinylated proteins such as propionyl-coenzyme A carboxylase and abundantly expressed proteins such as hemoglobin and transferases. Importantly, of all the nine proteins identified, only cathepsin B

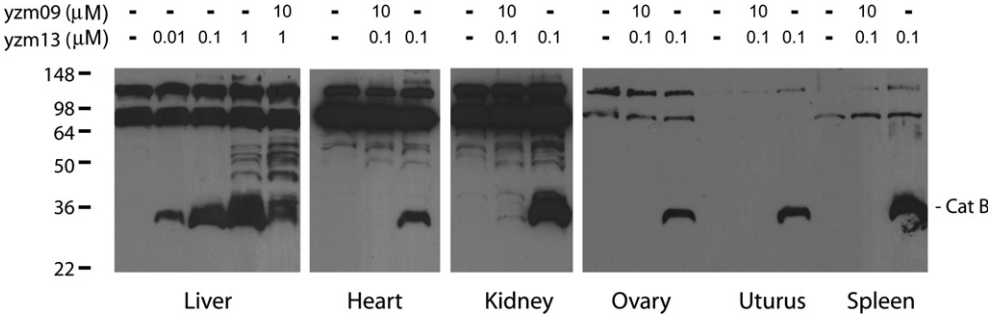


Figure 1. Labeling of mouse tissue lysates with the biotin-labeled probe **yzm13**. Total crude extracts (pH 5.5) from the indicated mouse tissues were normalized for total protein and labeled by addition of **yzm13** at the final concentrations indicated. In some samples the unlabeled control inhibitor **yzm09** was added prior to addition of the probe. Samples were resolved by SDS–PAGE and labeled proteins visualized by affinity blotting using streptavidin–HRP as outlined in Section 4.

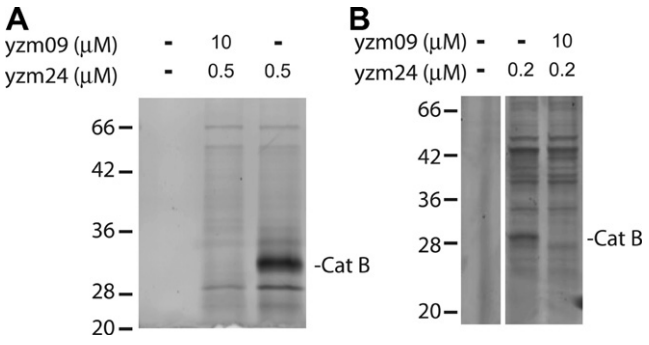


Figure 2. Labeling of total cell extracts with the fluorescently labeled probe **yzm24**. (A) Crude protein extracts from mouse spleen tissue (pH 5.5) was labeled by addition of **yzm24** at the indicated final concentration followed by SDS–PAGE and analysis of labeled proteins by scanning of the gel for fluorescence using a flatbed laser scanner. The unlabeled inhibitor **yzm09** was used for pre-treatment of the sample to identify specific labeled proteins. (B) MDA-MB231 cell lysates were labeled with **yzm24** at the indicated concentration and analyzed as outlined in (A).

showed a significant decrease in the number of peptides identified in the sample that had been pre-treated with **yzm09** to block specific labeling of targets. This result demonstrated that the only specific target of **yzm13** in total liver extracts is Cat B, consistent with the labeling observed by SDS–PAGE analysis.

2.3. Labeling complex proteomes with the fluorescent ABP yzm24

To confirm that the αβUK could also be used with fluorescent detection, we tested the ability of **yzm24** to label mouse tissue lysate and also cells lysate. The fluorescently labeled **yzm24** strongly labeled Cat B, as indicated by the intense band just above 28 kDa in mouse spleen lysate that was completely blocked by pre-treatment with **yzm09** (Fig. 2A). Again, we observed highly specific labeling of Cat B with very low overall background when the probe was used at 500 nM final concentrations, consistent with the observed labeling pattern of the biotin labeled probe **yzm13**. When **yzm24** was used to label total extracts from a human breast cancer cell lysate, again a band of labeled protein corresponding to cathepsin B was observed (Fig. 2B). While these cell extracts showed a significantly higher level of background labeling compared to the spleen lysates, only labeling of the specific cathepsin B band was abolished by pre-treatment of the sample with **yzm09**. Thus, the fluorescently labeled **yzm24** shows similar specific labeling patterns as the biotin probe **yzm13**. Both reagents are highly specific labels of cathepsin B and can be used at sub-micromolar concentrations.

2.4. Labeling in intact cells with the cell permeable fluorescent probe yzm24

As a final test of the αβUK probe, we wanted to test the utility of this probe class for labeling of endogenous cysteine proteases in intact cells. Since the biotin labeled probe **yzm13** was unlikely to penetrate the cells and gain access to intra-lysosomal pools of cathepsins we decided to focus our attention on the BODIPY labeled analog **yzm24**. Our group has previously shown that the BODIPY labeled version of the AOMK probe GB111 is useful for intact cell labeling experiments.[24] Thus, we compared the labeling of intact mouse fibroblast cells by the αβUK containing fluorescent probe **yzm24** to labeling by the related AOMK probe GB111. In wild type fibroblasts we observed the previously reported characteristic labeling pattern of GB111 that includes a single form of cathepsin B (32 kDa) and the heavy and light chain forms of cathepsin L (28 and 22 kDa).[24] We observed a highly similar labeling pattern for **yzm24** but with a higher level of background labeling (Fig. 3A). To confirm that **yzm24** was labeling both cathepsin B and cathepsin L as observed for GB111,

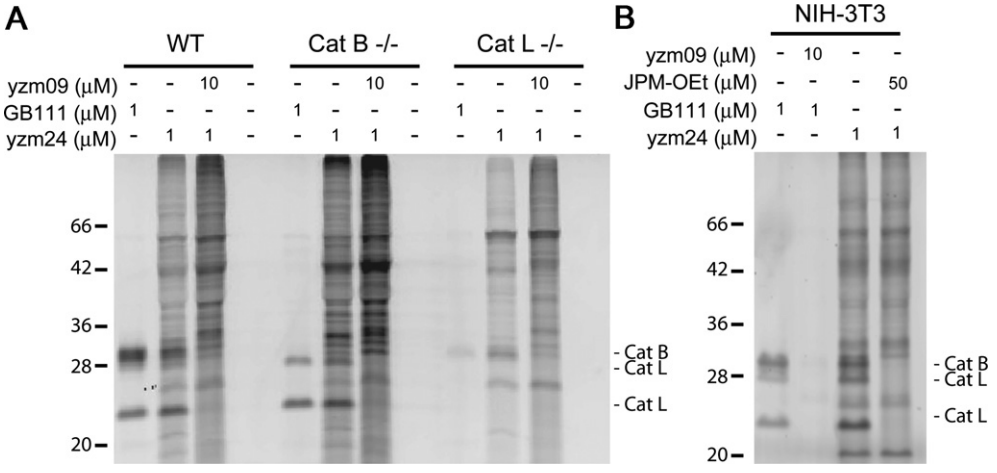


Figure 3. Labeling of intact cells with GB111 and **yzm24**. (A) Intact wildtype, cathepsin B deficient (Cat $B^{-/-}$) or cathepsin L deficient (Cat $L^{-/-}$) fibroblasts were labeled by addition of the fluorescent probes GB111 and **yzm24** to the culture media at the concentrations indicated. For some samples, the unlabeled inhibitor **yzm09** was used to pre-treat cells for 30 minutes prior to labeling with the probe. Cells were lysed by addition of SDS sample buffer and samples were analyzed by SDS–PAGE followed by scanning of the gels for fluorescence using a flatbed laser scanner. (B) Intact NIH-3T3 cells were pre-treated with DMSO vehicle, **yzm09** or the general papain-family protease inhibitor JPM-OEt and then labeled with GB111 or **yzm24** as in (A). The location of cathepsins B and L is indicated.

Please cite this article in press as: Yang, Z. et al., *Bioorg. Med. Chem.* (2008), doi:10.1016/j.bmc.2008.02.089

Table 1. Inhibition rate constants of **yzm09** for human cathepsin B and cathepsin L

Enzyme	Cat B (M^{-1} s^{-1})	Cat L (M^{-1} s^{-1})
yzm09	36,680 ± 3,952	1,634,400 ± 75,260

we labeled extracts for mouse fibroblasts derived from the cathepsin B and cathepsin L knock out mice (Cat $B^{-/-}$ and Cat $L^{-/-}$, respectively; Fig. 3A). As expected, the 32 kDa labeled protein disappeared in Cat $B^{-/-}$ cells and the lower two bands disappeared in Cat $L^{-/-}$ cells, thus confirming that **yzm24** efficiently labels both cathepsin B and cathepsin L in intact cells. This result was initially somewhat surprising considering the overall highly specific labeling of cathepsin B in both cell and tissue extracts. However, we and others have found that cathepsin L activity is lost once cells or tissues are lysed.[25] This may be due to the presence of endogenous inhibitors that are released upon lysis that quickly inactive the mature active forms of cathepsin L or it may be due to a relatively poor stability of cathepsin L in conditions used in our buffers.

As a final test of the αβUK containing probes, we wanted to confirm our findings that these compounds are efficient labels of both cathepsin B and L. We therefore measured the overall kinetic rate constants of inhibition of the parent **yzm09** for the recombinant forms of cathepsin B and cathepsin L. The results confirm that **yzm09** shows fast, irreversible inhibition of both cathepsins B and L (Table 1). The k_{ass} of **yzm09** for Cat B and Cat L are 36,680 and 1,634,400 M^{-1} s^{-1}, respectively, indicating that it is more than 10-fold more potent than the selective probe GB111.[24] Thus, these results suggest that the increased potency of the αβUK containing inhibitors relative to the AOMK containing probes may be due to the overall increased reactivity of this class of warhead.

3. Conclusions

Herein, we described the evaluation of the α,β-unsaturated ketone (αβUK) as a warhead for use in activity-based probes for papain-family cysteine proteases. We demonstrate that this class of reactive functional group allows rapid synthesis of ABPs that show excellent labeling properties in tissues and cell extracts as well as in intact cells. While the improved potency of this class of compounds toward the target cathepsins relative to previously reported AOMK-based probes may be beneficial in that it allows the use of lower probe concentrations, it comes with the price of potentially higher levels of background labeling. Thus, further studies may be required to determine if alternate peptide scaffolds can be used to enhance selectivity of the αβUK containing probes. Regardless we believe that this class of compound will find use for development of ABPs for diverse cysteine protease families, especially for those targets that are not efficiently labeled by currently used warheads such as the AOMK and vinyl sulfone. We are currently evaluating the properties of this probe class in vivo to determine if they represent useful tools for in vivo imaging studies.

4. Experimental

4.1. General

All chemicals and resin were purchased from commercial suppliers and used without further purification. All water sensitive reactions were conducted in anhydrous solvents under positive pressure of argon. Reactions were analyzed by LC/MS with an API 150EX single quadrupole mass spectrometer (Applied Biosystems). Reverse phase HPLC was used with an ÅKTA explorer 100 (Amersham Pharmacia Biotech) with C_{18} columns. High-resolution MS analyses were performed by the Stanford Proteomics and Integrative Research Facility with an Autoflex MALDA TOF/TOF mass spectrometer (Bruker). Fluorescent gels and plates were scanned with a Typhoon 9400 flatbed laser scanner (GE Healthcare).

4.2. Synthesis

*Cbz-FA-ketone (**yzm09**) and Boc-FA-ketone (**5b**).* Z-Phe (299 mg, 1 mmol) and NHS (115 mg, 1 mmol) were dissolved in 20 mL of chloroform, and DCC (230 mg, 1.1 mmol) was added to the above solution with stirring at room temperature. After 1 h the formed solid was filtered off and the solvent was removed under reduced pressure. The crude product was used in the next reaction without further purification.

L-Alanine (89 mg, 1 mmol) and $NaHCO_3$ (168 mg, 2 mmol) were dissolved in 20 mL of water with stirring. Next, a solution of crude product obtained above (dissolved in 40 mL acetone) was added. The resulting reaction mixture was stirred at room temperature overnight. The reaction mixture was concentrated under reduced pressure. Subsequently, 30 mL of water was added and the precipitate was removed by filtration. The filtrate was acidified to a pH of ~3, resulting in precipitation of the product. The solid was obtained by filtration, washed with water, and dried in vacuo (284 mg; 77% yield). The crude product was used without further purification.

About 185 mg of the crude material obtained in the previous step, *N*,*O*-dimethyl-hydroxylamine hydrochloride (47 mg, 0.5 mmol), and HBTU (190 mg, 0.5 mmol) were dissolved in 5 mL of anhydrous DMF. Next, 194 μL of DIEA was added. The resulting mixture was stirred at room temperature for 1 h, after which the solvent was concentrated under reduced pressure. 100 mL of ethyl acetate was added and the organic phase was washed with dilute acid (3× 40 mL) and brine (1× 40 mL), dried ($MgSO_4$), and concentrated under reduced pressure. The Weinreb amide **3a** was purified by flash column chromatography (17–50% ethyl acetate in hexanes) and obtained as a white solid (162 mg, 78%).

To a solution of Weinreb amide **3a** (162 mg) in 10 mL of anhydrous THF, 34 mg of $LiAlH_4$ was added at 0 °C.

After half an hour, 20 mL of 5% $KHSO_4$ and 50 mL of ethyl acetate were added. The organic layer was washed with $KHSO_4$ (2× 20 mL), brine, dried ($MgSO_4$), and concentrated under reduced pressure. The resulting aldehyde was obtained as a white solid.

To the solution of aldehyde (dried by coevaporation with toluene) in 15 mL DCM, 1-triphenyl-phosphoranylidene-2-propanone (160 mg) was added and stirred at room temperature overnight. The reaction mixture was concentrated and the resulting pale yellowish solid was purified by HPLC. 92 mg of **yzm09** was obtained (35.8% overall yield, based on Z-Phe-OH). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.16 (d, 3H, CH–CH_3), 2.194 (s, 3H, CH_3), 2.97–3.02 and 3.10–3.15 (m, 2H, –CH–CH_2–Ph), 4.33–4.35 (m, 1H, –CH–CH_2–Ph), 4.61–4.64 (m, 1H, CH–CH_3), 5.08 (s, 2H, –O–CH_2–Ph), 5.86–5.90 (d, 1H, –CH–CH=), 6.41–6.46 (d, 1H, =CH–CO–), and 7.18–7.34 (m, 10H, aromatic). MS (EI): $[M+H]^+$ Calculated: (394.2); found: (395.2).

A similar procedure was used to afford the compound **5b** (overall yield = 24.3%, based on Boc-Phe-OH). MS (EI): $[M+H]^+$ Calculated: (360.2); found: (361.2).

*Biotin-FA-ketone (**yzm13**) and BODIPY-TMRX-FA-Ketone (**yzm24**).* 100 mg of compound **5b** was dissolved in 6 mL of 4 M HCl in dioxane, the reaction mixture was stirred at room temperature for 4 h. The organic solvent was removed under reduced pressure. The obtained viscous liquid was coevaporated with toluene three times then dissolved in 10 mL of acetone and kept in the −20 °C refrigerator.

Compounds **yzm13** and **yzm24** were obtained by coupling the corresponding biotin or BODIPY NHS activated ester with a 2- to 3-fold excess of compound **6** (Scheme 2) in acetone with the assistance of 2 equiv of DIEA (with respect to the amine compound). They were obtained in 90% yield after purification by HPLC. High-resolution mass spectrometer (HRMS) data: $[M+H]^+$ Calculated for **yzm13**, $C_{36}H_{56}N_5O_9S^+$, 734.38; found 734.3797. Calculated for **yzm24**, $C_{42}H_{51}BF_2N_5O_5^+$, 754.40; found 754.3955.

4.3. Preparation of mouse tissue homogenates and protein labeling

Mouse tissue homogenates[23] were incubated in sodium acetate buffer (50 mM, pH 5.5) with **yzm13** (0.1 μM final concentration) in the presence of DTT (2 mM final concentration). Labeling was carried out for 1 h at room temperature. Samples were separated by SDS–PAGE and transferred to a nitrocellulose membrane (BioRad, USA). The membrane was blocked in casein (5% solution in PBS) and incubated with streptavidin–HRP (dilution 1:3500, Sigma, USA) for 1 h at room temperature. Biotinylated proteins were visualized using SuperSignal West Pico Chemiluminescent Substrate (Pierce, USA).

4.4. Affinity enrichment and elution of labeled proteins

After protein labeling, free probe was removed by a PD-10 gel filtration column (Amersham, USA) and sample was eluted in PBS buffer (pH 7.4). Streptavidin beads (Pierce, USA) were washed with PBS buffer and added to the PD-10 eluate. The sample was incubated at room temperature with shaking for 1 h. Streptavidin beads were separated from the unbound fraction by centrifugation. The supernatant was discarded and the beads were sequentially washed with a series of PBS buffers containing 0.05% SDS, 1 M NaCl, and 10% EtOH. Finally, the beads were washed with 100 mM ammonium hydrogen carbonate. Washing was performed three times with each buffer solution.

4.5. Sample preparation for LC–MS/MS analysis

For 'on bead' digestion, streptavidin beads with bound proteins were resuspended in 100 μL of denaturing buffer (50 mM sodium hydrogen carbonate, 6 M urea). Bound proteins were reduced in the presence of 10 mM DTT for 1 h. Samples were alkylated by addition of 200 mM iodoacetamide (20 μL) and incubated for 1 h in the dark. Unreacted iodoacetamide was neutralized by addition of 200 mM DTT (20 μL). The urea concentration was reduced by addition of dH_2O (800 μL). Samples were incubated with trypsin overnight at 37 °C and purified on a Vivapure C18 spin column according to the manufacturer's instructions (Sartorius, Germany) (see Table 2).

4.6. LC–MS/MS and database search

Samples were analyzed on a LCQ DecaXP Plus ion trap mass spectrometer (Thermo Fisher, USA) coupled to a nanoLC liquid chromatography unit (Eksigent, USA). Peptides were separated on a BioBasic Picofrit C18 capillary column (New Objective, USA). Elution was performed with acetonitrile gradient from 0% to 50% in the 0.1% solution of formic acid over 40 min with overall flowrate of 350 nL/min. The three most intense base

Table 2. MS identification of labeled proteins in liver lysate by 'on bead' digestion

Protein	Accession No.	MW (kDa)	Inhibitor	Probe
Cathepsin B	gi:74180941	37	3 (11%)	12 (40%)
Propionyl-coenzyme A carboxylase	gi:13905236	80	24 (44%)	28 (51%)
Betaine–homocysteine methyltransferase	gi:22477957	45	12 (55%)	12 (43%)
Chain D, chimeric human mouse carbonmonoxy hemoglobin	gi:18655689	16	8 (60%)	8 (60%)
Predicted: carbamoyl-phosphate synthetase 1 isoform 1	gi:82879179	165	7 (7%)	4 (5%)
Glycine *N*-methyltransferase	gi:15679953	33	3 (23%)	3 (20%)
S-Adenosylhomocysteine hydrolase	gi:26344433	36	3 (12%)	3 (12%)
Carbonic anhydrase 3	gi:31982861	29	2 (18%)	3 (18%)
Albumin 1	gi:26340966	69	2 (6%)	3 (10%)

Number of peptides identified for each protein is indicated with percentage of coverage in parentheses.

Please cite this article in press as: Yang, Z. et al., *Bioorg. Med. Chem.* (2008), doi:10.1016/j.bmc.2008.02.089

peaks in each scan were analyzed by MS/MS. Dynamic exclusion was set at a repeat count of 2 with exclusion duration of 2 min. Database searches were performed using the mouse NCBI protein database using the Sequest algorithm (Thermo Fisher, USA). Peptides with XCorr values over 1.5 (+1 charge), 2 (+2 charge) and 2.5 (+3 charge) and deltaCn values over 0.1 were considered for further evaluation. Protein and peptide hits were statistically reevaluated by Scaffold (Proteome Software, USA). Peptide identifications were accepted if they could be established at greater than 95% probability as specified by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm.[23]

4.7. Determination of kinetic rate constants of inhibition

The kinetics of inhibition was determined by the progress curve method under pseudo-first-order conditions with at least 10-fold molar excess of inhibitor. Recorded progress curves were analyzed by nonlinear regression according to the following equation:

$$[P] = v_z(1 - e^{-kt})/k$$

where [P] is the product, v_z is the velocity at time zero and k is the pseudo-first-order rate constant. Apparent rate constant (k_{app}) was determined from the slope of plot k versus [I]. Because of the irreversible and competitive mechanism of inhibition, k_{app} was converted to the association constant (k_{ass}) using the equation below:

$$k_{ass} = k_{app}(1 + [S]/K_m)$$

Activity of human cathepsin L was measured with the fluorogenic substrate Z-FR-AMC (Bachem; K_m = 7.1 μM) and cathepsin B was assayed against the fluorogenic substrate Z-RR-AMC (Bachem; K_m = 114 μM). Concentration of substrates during the measurement was 10 μM. Cathepsins B and L (1 nM final concentrations) were incubated with inhibitor concentrations ranging from 10 to 2000 nM in the presence of 10 μM of appropriate substrate. Total volume during the measurement was 100 μL. The increase in fluorescence (370-nm Ex, 460-nm Em) was continuously monitored for 30 min with a Spectramax M5 fluorescent plate reader (Molecular Devices), and inhibition curves were recorded. DMSO concentration during all measurements was 3.5%.

4.8. Cell cultures

WT mouse embryo fibroblasts (MEF), *Cat* $L^{-/-}$ MEF cells, and *Cat* $B^{-/-}$ MEF cells were cultured in DMEM supplemented with 10% FBS, 100 U ml^{-1} penicillin, 100 μg ml^{-1} streptomycin. All cells were cultured in a humidified atmosphere of 95% air and 5% CO_2 at 37 °C.

4.9. Labeling of intact cells with probes

WT and *Cat* $L^{-/-}$ (240,000 cells per well) and *Cat* $B^{-/-}$ (300,000 cells per well) were seeded in a six-well plate one day before treatment. Cells were pre-treated with **yzm09** (10 μM), or with control DMSO (0.1%) for 1 h and labeled for 1 h by addition of **yzm24** to culture medium. The final DMSO concentration was maintained at 0.2%. Cells were washed with PBS twice and lysed by addition of sample buffer (10% glycerol, 50 mM Tris/HCl, pH 6.8, 3% SDS, and 5% β-mercaptoethanol). Lysates were boiled for 10 min and cleared by centrifugation. Equal amounts of protein per lane were separated by 15% SDS–PAGE, and labeled proteases were visualized by scanning of the gel with a Typhoon flatbed laser scanner (Ex/Em 532/580 nm).

Acknowledgments

This work was supported by the National Institutes of Health National Technology Center for Networks and Pathways Grant U54 RR020843, R01 EB005011, P01 CA072006, and a Stanford Chemical and Systems Biology training grant (to S.V.).

References and notes

1. Campbell, D. A.; Szardenings, A. K. *Curr. Opin. Chem. Biol.* **2003**, *7*, 296.
2. Jeffery, D. A.; Bogyo, M. *Curr. Opin. Biotechnol.* **2003**, *14*, 87.
3. Kozarich, J. W. *Curr. Opin. Chem. Biol.* **2003**, *7*, 78.
4. Speers, A. E.; Cravatt, B. F. *Chembiochem* **2004**, *5*, 41.
5. Verhelst, S. H. L.; Bogyo, M. *QSAR Comb. Sci.* **2005**, *24*, 261.
6. Evans, M. J.; Cravatt, B. F. *Chem. Rev.* **2006**, *106*, 3279.
7. Powers, J. C.; Asgian, J. L.; Ekici, O. D.; James, K. E. *Chem. Rev.* **2002**, *102*, 4639.
8. Cravatt, B. F.; Sorensen, E. J. *Curr. Opin. Chem. Biol.* **2000**, *4*, 663.
9. Yuan, F.; Verhelst, S. H. L.; Blum, G.; Coussens, L. M.; Bogyo, M. *J. Am. Chem. Soc.* **2006**, *128*, 5616.
10. Palmer, J. T.; Rasnick, D.; Klaus, J. L.; Bromme, D. *J. Med. Chem.* **1995**, *38*, 3193.
11. Wang, G.; Yao, S. Q. *Org. Lett.* **2003**, *5*, 4437.
12. Borodovsky, A.; Ovaa, H.; Meester, W. J. N.; Venanzi, E. S.; Bogyo, M. S.; Hekking, B. G.; Ploegh, H. L.; Kessler, B. M.; Overkleeft, H. S. *Chembiochem* **2005**, *6*, 287.
13. Hemelaar, J.; Borodovsky, A.; Kessler, B. M.; Reverter, D.; Cook, J.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Gill, G.; Lima, C. D.; Ploegh, H. L.; Ovaa, H. *Mol. Cell. Biol.* **2004**, *24*, 84.
14. Overkleeft, H. S.; Bos, P. R.; Hekking, B. G.; Gordon, E. J.; Ploegh, H. L.; Kessler, B. M. *Tetrahedron Lett.* **2000**, *41*, 6005.
15. Bogyo, M.; Shin, S.; McMaster, J. S.; Ploegh, H. L. *Chem. Biol.* **1998**, *5*, 307.
16. Nazif, T.; Bogyo, M. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2967.
17. Santos, M. M. M.; Moreira, R. *Mini-Rev. Med. Chem.* **2007**, *7*, 1040.
18. Choe, Y.; Brinen, L. S.; Price, M. S.; Engel, J. C.; Lange, M.; Grisostomi, C.; Weston, S. G.; Pallai, P. V.; Cheng, H.; Hardy, L. W.; Hartsough, D. S.; McMakin, M.; Tilton, R. F.; Baldino, C. M.; Craik, C. S. *Bioorg. Med. Chem.* **2005**, *13*, 2141.

19. Borodovsky, A.; Ovaa, H.; Kolli, N.; Gan-Erdene, T.; Wilkinson, K. D.; Ploegh, H. L.; Kessler, B. M. *Chem. Biol.* **2002**, *9*, 1149.
20. Blum, G.; von Degenfeld, G.; Merchant, M. J.; Blau, H. M.; Bogyo, M. *Nat. Chem. Biol.* **2007**, *3*, 668.
21. Bogyo, M.; Phillips, C. unpublished observation.
22. Wang, G.; Mahesh, U.; Chen, G. Y. J.; Yao, S. Q. *Org. Lett.* **2003**, *5*, 737.
23. Fonovic, M.; Verhelst, S. H. L.; Sorum, M. T.; Bogyo, M. *Mol. Cell. Proteomic.* **2007**, *6*, 1761.
24. Blum, G.; Mullins, S. R.; Keren, K.; Fonovic, M.; Jedeszko, C.; Rice, M. J.; Sloane, B. F.; Bogyo, M. *Nat. Chem. Biol.* **2005**, *1*, 203.
25. Bogyo, M.; Verhelst, S.; Bellingard-Dubouchaud, V.; Toba, S.; Greenbaum, D. *Chem. Biol.* **2000**, *7*, 27.

Please cite this article in press as: Yang, Z. et al., *Bioorg. Med. Chem.* (2008), doi:10.1016/j.bmc.2008.02.089