

Evaluation of an Orthogonal Pooling Strategy for Rapid High-Throughput Screening of Proteases

Nuzhat Motlekar, Scott L. Diamond, and Andrew D. Napper

Abstract: Orthogonal pooling was evaluated as a strategy for the rapid screening of multiple cysteine and serine proteases against large compound libraries. To validate the method the human cysteine protease cathepsin B was screened against a library of 64,000 individual compounds and also against the same library mixed 10 compounds per well. The orthogonal pooling method used resulted in each compound being present in two wells, mixed with a different set of nine other compounds in each location. Thus hits were identified based on activity in both locations, avoiding the need for retesting of each component of active mixtures. Hits were tested in dose–response both in the dithiothreitol (DTT)-containing buffer used in the primary HTS and in buffer containing cysteine in place of DTT to rule out artifacts due to oxidative inactivation of the enzyme. Comparison of the confirmed actives from single-compound and mixture screening showed that mixture screening identified all of the actives from single-compound HTS. Based on these results the orthogonal pooling strategy has been used successfully to rapidly screen several cysteine and serine proteases.

Introduction

ONE OF OUR GOALS at the Penn Center for Molecular Discovery,¹ Philadelphia, PA, is to develop capabilities for screening multiple members of target classes, for example, cysteine and serine proteases. As part of the Molecular Libraries Screening Center Network (MLSCN) we have access to a diverse library of more than 200,000 compounds. Our aim is to test the entire MLSCN compound library against multiple cysteine and serine proteases to obtain a profile of activity against these enzyme classes. This profile may then be used to immediately identify selective compounds during subsequent screening of novel enzyme targets. It may also be possible to identify a subset of the library with an enhanced hit rate towards these enzyme families that might provide the basis for gene family screening. One strategy that enables the rapid screening of large compound libraries against multiple targets is to screen compound mixtures instead of individual compounds; reagent costs and screening times can thereby be reduced significantly. We chose to evaluate an orthogonal pooling strategy that

gives 10 compounds per well, as described in detail below. An advantage of this method is that each compound is in two plates, mixed with a different set of nine other compounds at each location. This duplication improves the validity of statistical interpretation of the HTS and greatly simplifies deconvolution of the results as activity in both wells containing a given compound immediately identifies that compound as a hit. This method has been used with some success in several pharmaceutical companies.^{2–4}

Human liver cathepsin B (EC 3.4.22.1) was chosen as a prototype from the cysteine protease family to validate the compound pooling strategy. Cathepsin B is a lysosomal cysteine protease that is involved in many physiological processes, such as remodeling of the extracellular matrix during wound healing, apoptosis, and activation of thyroxin and renin.^{5–7} In addition to its physiological roles, cathepsin B is important in many pathological processes, such as cancer, inflammation, and infection. Overexpression and secretion of cathepsin B occur in many types of tumors and correlate positively with invasion and metastasis.⁸ Cathepsin B facilitates tu-

Penn Center for Molecular Discovery, Institute for Medicine and Engineering, and Department of Chemical and Biomolecular Engineering, University of Pennsylvania, Philadelphia, Pennsylvania.

ABBREVIATIONS: AMC, aminomethyl coumarin; CV, coefficient of variation; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; IC₅₀, 50% inhibitory concentration; LC-MS, liquid chromatography-mass spectrometry; MLSCN, Molecular Libraries Screening Center Network; QC, quality control.

mor invasion by dissolving extracellular matrix, and inhibitors of the enzyme have been shown to reduce both tumor cell motility and invasiveness *in vitro*.^{9–11} Inhibitors of cathepsin B are therefore of interest as potential anticancer agents. There has also been a recent resurgence of interest in cathepsin B due to research showing that proteolysis by this enzyme is required for the entry and replication of the Ebola and severe acute respiratory syndrome viruses in human cells. Studies have confirmed roles for cathepsin B and cathepsin L in Ebola virus glycoprotein-mediated infection.^{12,13} Cathepsin B inhibitors were shown to diminish the multiplication of infectious Ebola virus-Zaire in cultured cells and may merit investigation as anti-Ebola virus drugs.¹⁴

Validation of the compound pooling strategy is provided here by the results of screening cathepsin B using a simple end-point assay monitoring the release of the fluorophore aminomethyl coumarin (AMC) upon enzymatic hydrolysis of an AMC-labeled dipeptide. Pooling allowed us to screen 64,000 compounds in a single 3-h HTS run. We also screened cathepsin B by single-compound HTS, allowing us to compare the hit profile obtained by single and mixture screening. The most important consideration was to ensure that all of the active compounds discovered by single-compound screening were also identified by mixture screening.

Materials and Methods

Materials

Human liver cathepsin B (catalog number 219362) was purchased from Calbiochem (San Diego, CA). Substrate Z-Arg-Arg-AMC was from Bachem (King of Prussia, PA). Assay buffer components and AMC standard were purchased from Sigma-Aldrich (St. Louis, MO). Low-volume non-binding surface 384-well black plates used for the fluorescent assay were from Corning (Lowell, MA). Polypropylene V-bottom plates from Greiner Bio-One (Monroe, NC) were used for compound storage and dilution. Heat seals were applied using a PlateLoc[®] heat sealer (Velocity 11, Menlo Park, CA), and polystyrene lids were from Nunc[®] (Rochester, NY). The compounds tested here were supplied by BioFocus DPI (South San Francisco, CA) as part of the MLSCN.¹⁵

Compound library

A library of 64,000 compounds from the Molecular Libraries Small Molecule repository was shipped on dry ice from BioFocus DPI as frozen 10 mM solutions in dimethyl sulfoxide (DMSO) in heat-sealed 384-well polypropylene plates. Prior to compound plating and shipping, quality control (QC) testing at BioFocus DPI confirmed compound solubility in DMSO, the presence

of the expected molecular ion in liquid chromatography-mass spectrometry (LC-MS), and >90% purity by evaporative light scattering or ultraviolet at 214 nm. On arrival, plates were stored in a desiccator (<20% relative humidity) at -25°C . After storage for approximately 3 months, plates were allowed to warm to room temperature in a desiccator. The thawed plates were used to prepare dilution plates for single-compound and mixture HTS as described below and then refrozen.

Preparation of plates for single-compound and mixture HTS

For single-compound HTS, a set of 0.5 mM dilution plates were prepared on a Perkin Elmer (Boston, MA) EP3 workstation using a 384-tip pipetting head. Two microliters from each BioFocus DPI plate was pipetted into a polypropylene V-bottom plate containing 18 μl per well of DMSO and mixed by repeated aspiration and dispense (15 μl). Plates were heat-sealed and stored at room temperature for 1–2 weeks in a desiccator prior to use. Plates for mixture HTS were prepared as shown in Fig. 1. Two hundred plates containing the 64,000 compounds from BioFocus DPI were arranged in two 10 \times 10 grids of 100 plates each. Sets of 10 plates were pooled to give a single mixture plate as shown, resulting in 20 mixture plates per 100 single-compound plates. An automated plate pooling protocol was set up on a Thermo-CRS (Burlington, ON, Canada) screening deck, controlled by Polara operating software. The heat seals were removed by hand from each set of 10 single-compound plates and replaced with loose-fitting polystyrene plate lids. Each plate in turn was delidded and placed on a Perkin Elmer EP3 workstation. Two microliters of compound in DMSO from each plate was transferred by pipetting head (384 disposable tips) into a single mixture plate, giving a final mixture volume of 20 μl . The concentration of each mixture was 10 mM in DMSO (1 mM per compound). After pipetting was complete each plate was immediately heat-sealed to minimize exposure to atmospheric moisture. In a separate protocol the mixture plates were then diluted fourfold by transfer of 5 μl of each mixture into 15 μl of DMSO to give mixtures 2.5 mM in DMSO (250 μM per compound). Plates were heat-sealed and stored at room temperature for 2–3 days in a desiccator prior to use.

Preparation of plates for 50% inhibitory concentration (IC_{50}) testing

Hits selected from single-compound and mixture HTS were cherry-picked from the set of 10 mM compound plates supplied by BioFocus DPI, and each compound was transferred to a single well in row A of a 384-well V-bottom polypropylene plate. Frozen compound plates were allowed to warm to room temperature in a desic-

AU1

AU2

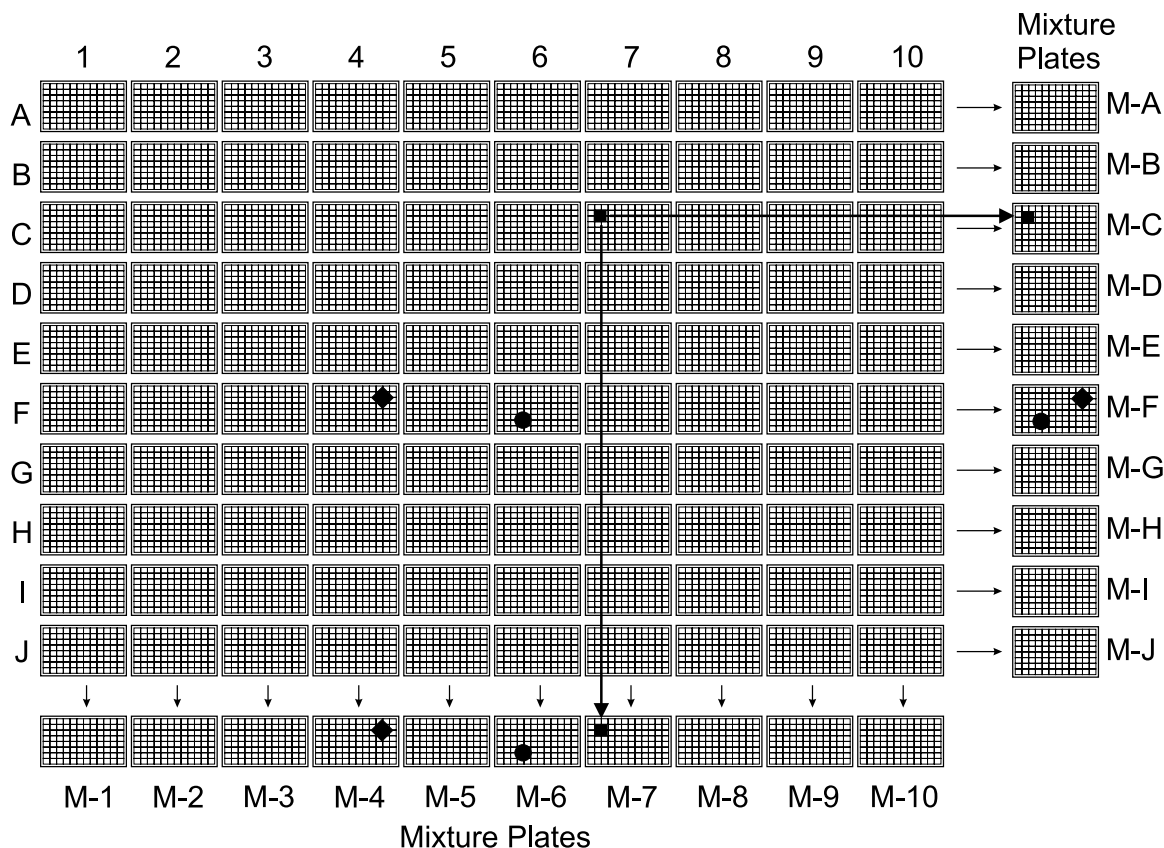


FIG. 1. Orthogonal pooling of 100 compound plates to give 20 mixture plates. Each compound is in two locations in the mixture plates, mixed with a different set of nine other compounds in each location. For example, the single compound shown in plate C7 (■, plate in row C, column 7) is mixed with nine compounds in mixture plate M-C, and also with nine different compounds in mixture plate M-7.

ator. Compound pipetting was carried out using disposable liquid level sensing tips on a Perkin Elmer Janus four-tip workstation. Ten microliters of DMSO was added to row A of 384-well V-bottom polypropylene plates, and 10 μl of each selected compound was transferred to one well out of A3 to A22. An additional 10 μl of DMSO was added to wells A1, A2, A23, and A24, followed by 20 μl per well of DMSO added to the entire plate using a Thermo Fisher Scientific (Waltham, MA) Multidrop. The compounds were twofold serially diluted by transfer of 20 μl row by row from row A to row P using a single row of disposable tips on a Perkin Elmer EP3 384-tip pipetting head, after which 20 μl was discarded from row P. The resulting dose-response plates contained 16 twofold dilutions (2.5 mM–50 nM) of each compound, arranged one compound per column in columns 3–22. Plates were heat-sealed and stored at room temperature for 1–2 days in a desiccator prior to use.

LC-MS analysis of hits

Purity and integrity of compounds identified as hits in the mixture and single-compound HTS were analyzed by

LC-MS on a Waters (Milford, MA) ZQ system. Forty microliters of a 100 μM solution of each compound in DMSO was eluted from a SunFire™ (Waters) C18 column (4.6 \times 50 mm) with a 5-min gradient of 90:10 to 10:90 water (0.05% formic acid):acetonitrile (0.05% formic acid). Compound integrity and purity were determined by identification of the expected molecular ion and peak integration by evaporative light scattering and ultraviolet absorbance at 214 nm.

Determination of substrate K_m

The assay buffer consisted of 100 mM sodium-potassium phosphate (pH 6.8) (86 mM potassium phosphate, monobasic; 7 mM sodium phosphate, monobasic; and 7 mM sodium phosphate, tribasic), 1 mM EDTA, and 2 mM dithiothreitol (DTT). Z-Arg-Arg-AMC substrate was serially diluted from 400 to 6 μM in assay buffer containing 4% DMSO, and 5 μl per well of each dilution was transferred by multichannel pipette into four columns of a low-volume 384-well assay plate. Cathepsin B was activated prior to addition to the assay plate by incubating in assay buffer for 15 min. The assay was started by ad-

AU3

dition of 5 μl of activated enzyme to two columns to give final concentrations of enzyme, substrate, and DMSO of 2.36 nM, 3–200 μM , and 2%, respectively. The remaining two columns consisted of control wells at each substrate concentration containing 5 μl of assay buffer without enzyme. The mixture was allowed to incubate at room temperature, and fluorescence was read in a Perkin Elmer EnVision™ microplate fluorimeter (excitation 355 nm, emission 460 nm) at 1-min intervals for 1 h. Fluorescence values were corrected for background readings from the control wells containing substrate in buffer but no enzyme. Rate of change of fluorescence was calculated from the initial linear timecourse over the first 15 min and converted to rate of AMC release in $\mu\text{M s}^{-1}$ using a standard curve of AMC fluorescence. Rate of AMC release versus substrate concentration was plotted and fit in XLfit® (IDBS, Guildford, UK) using fit model 350 (Michaelis-Menten equation).

HTS assay

Compound mixtures were added by pintool transfer on a Perkin Elmer EP3 workstation. A 384-pin pintool (V&P Scientific, San Diego, CA) was used to transfer 200 nl of a 2.5 mM solution in DMSO into assay plates containing 4 μl of water per well, giving a final concentration of each compound in the assay of 5 μM in 2% DMSO. Single compounds were added likewise from a 0.5 mM solution in DMSO giving a final concentration in the assay of 10 μM in 2% DMSO. QC plates contained the cysteine protease inhibitor E-64 at a concentration of 50 nM, obtained by addition of 200 nl of a 2.5 μM solution in DMSO. The assay buffer consisted of 100 mM sodium-potassium phosphate (pH 6.8) (86 mM potassium phosphate, monobasic; 7 mM sodium phosphate, monobasic; and 7 mM sodium phosphate, tribasic), 1 mM EDTA, and 2 mM DTT. Cathepsin B was activated prior to addition to the assay plate by incubating in assay buffer for 15 min. The assay was started by addition of substrate (1 μl) and activated enzyme (5 μl) using Multidrop reagent dispensers to give final concentrations of 2.36 nM cathepsin B, 15 μM Z-Arg-Arg-AMC substrate, and 2% DMSO in 10 μl of assay buffer. After 1 h at room temperature fluorescence was read in an EnVision microplate fluorimeter (excitation 355 nm, emission 460 nm).

IC₅₀ assay

This assay was identical to the HTS assay, except that 16 twofold dilutions of each compound were tested. Compounds were twofold serially diluted in DMSO from 2.5 mM to 50 nM and transferred by pintool into assay plates to give a final concentration range of 50 μM –1.5 nM in 2% DMSO. For the IC₅₀ assay in the presence of cysteine, DTT in the cathepsin B assay buffer was replaced

with 2 mM cysteine. Enzyme was activated in the cysteine-containing buffer for 1 h and then assayed as described above.

HTS data analysis

Data were analyzed using ActivityBase (IDBS). Each HTS plate contained individual compounds (10 μM) or mixtures (10 compounds at 5 μM each) in columns 3–22, controls (enzyme, no compound) in columns 2 and 24, and blanks (no enzyme) in columns 1 and 23. Percentage inhibition was calculated for each compound from the signal in fluorescence units and the means of the plate controls and plate blanks using the following equation:

$$\% \text{ Inhibition} = 100 \times (1 - [(\text{signal} - \text{blank mean}) / (\text{control mean} - \text{blank mean})])$$

Compounds from the single-compound HTS that gave >33% inhibition were selected as hits and retested in dose–response. Percentage inhibition results from the mixture HTS were retrieved in SARgen (IDBS) together with the identity of the individual compounds within each mixture. As each compound was present in two mixtures the data were rearranged using a custom Excel (Microsoft, Redmond, WA) macro (developed and generously contributed by Dr. Mandar Ghatnekar and Rajaram Gurumurthi, Infosys Technologies Ltd., Bangalore, India) to align both percentage inhibition values associated with each compound. Compounds that gave >20% inhibition in both mixture locations were selected and retested individually in dose–response.

IC₅₀ data analysis

IC₅₀ plates contained one compound per column in columns 3–22, controls (enzyme, no compound) in columns 2 and 24, and blanks (no enzyme) in columns 1 and 23. Each column from 3 to 22 contained 16 twofold dilutions of a single compound, ranging in concentration from 50 μM to 1.5 nM. Percentage activity was calculated for each dilution of each compound from the signal in fluorescence units and the means of the plate controls and plate blanks using the following equation:

$$\% \text{ Activity} = 100 \times [(\text{signal} - \text{blank mean}) / (\text{control mean} - \text{blank mean})]$$

Dose–response curves of percentage activity were fit in XLfit, using a four-parameter logistic fit (equation 205 with maximum percentage activity and minimum percentage activity fixed at 100% and 0%, respectively).



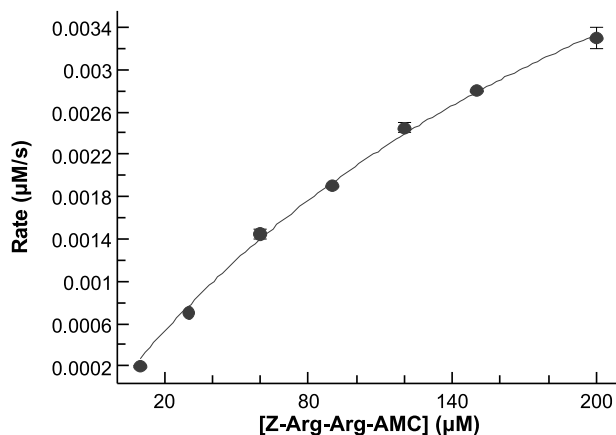



FIG. 2. Determination of K_m of Z-Arg-Arg-AMC substrate with cathepsin B. The assay protocol and data analysis were as described in Materials and Methods. Plot symbols and error bars represent mean of two determinations \pm SD. $K_m = 288 \mu M$; $k_{cat} = 3.45 s^{-1}$.

Results and Discussion

Assay development and HTS validation

The fluorescence assay used for cathepsin B screening and IC_{50} confirmation was based on a previously published protocol.¹⁶ The cathepsin B purchased from Calbiochem was supplied at a concentration of 1,600 $\mu g/ml$. The appropriate concentration of enzyme for a 1-h assay at room temperature in low-volume non-binding surface 384-well plates was determined by assaying in the pres-

ence of 15 μM Z-Arg-Arg-AMC substrate and 2% DMSO. Based on a serial dilution of the enzyme from the Calbiochem stock a final dilution of 1:25,000, corresponding to 2.36 nM enzyme, was found to give a robust signal and linear time course over a 60-min interval.

The K_m of the Z-Arg-Arg-AMC substrate with cathepsin B in 2% DMSO in assay buffer was determined using 2.36 nM enzyme and substrate concentrations ranging from 3 to 200 μM . A nonlinear fit of the data (Fig. 2) gave a K_m of 288 μM and a k_{cat} of 3.45 s^{-1} . These kinetic parameters are consistent with the values of 184 μM and 158 s^{-1} , respectively, previously reported for affinity-purified human cathepsin B¹⁶ when allowance is made for the significant differences in assay conditions between the two studies (pH of 6.0 and temperature of 37°C in Baricos *et al.*¹⁶ as compared with a pH of 6.8 and room temperature in the present study).

The cathepsin B assay was validated for enzyme inhibition HTS using the well-known cysteine protease inhibitor E-64. This has been shown to be an active site-directed irreversible inhibitor of cathepsin B.¹⁷ Dose-response testing of E-64 in the IC_{50} assay used in this study gave an IC_{50} of 12 nM (Fig. 3A). Thus at concentrations of 12 nM E-64 and 2.36 nM cathepsin B we observed an inhibition half-life of 60 min, which is consistent with the previously published half-life of 155 s obtained at higher reagent concentrations (50 nM E-64 and 10 nM cathepsin B) and higher temperature (40°C vs. room temperature) than in the current study.¹⁷ Based on these IC_{50} results both mixture and single-compound screening included analysis of the performance of QC

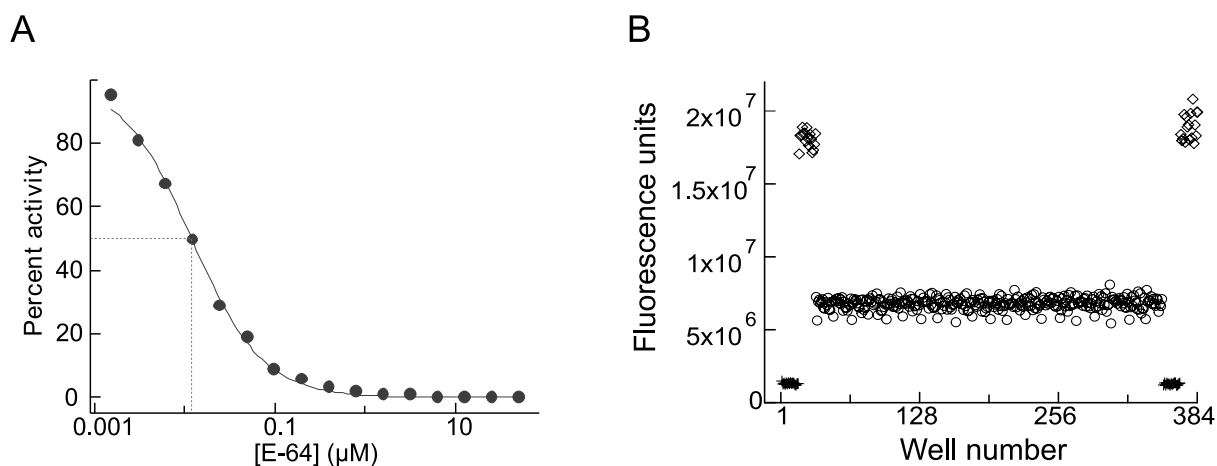


FIG. 3. Validation of cathepsin B assay using E-64. **(A)** Dose-response testing of E-64 against cathepsin B. The assay protocol and data analysis were as described in Materials and Methods (IC_{50} assay). Each plot symbol represents a single determination. $IC_{50} = 12 nM$. **(B)** Graphical representation of results of testing QC plate in cathepsin B HTS assay. Blanks (+) were located in wells 1–16 and 353–368 (columns 1 and 23). Controls (\diamond) were located in wells 17–32 and 369–384 (columns 2 and 24). E-64 (\circ) was added to wells 33–352 (columns 3–22) at a final concentration of 50 nM in 2% DMSO in assay buffer. Control and blank wells contained 2% DMSO in assay buffer. Substrate was added to all wells to a final concentration of 15 μM . Cathepsin B was added to all wells except blanks to a final concentration of 1 nM. Blank wells contained no enzyme. The assay protocol was as described in Materials and Methods.

plates containing E-64 at 50 nM. These plates showed consistent inhibition by E-64 and very robust coefficient of variations (CVs), signal-background ratios, and Z' -factors. The representative example shown in Fig. 3B gave 61% inhibition by E-64 and control CV of 4.5%, blank CV of 5.1%, signal-background ratio (mean control/mean blank) of 14.2, and Z' -factor of 0.84.

Compound handling and analysis

Compound plates were stored frozen at -25°C after delivery from BioFocus DPI. After storage for approximately 3 months, plates were thawed, used to prepare dilution plates for single-compound and mixture HTS, and refrozen. After an additional 1–2 months HTS hits were cherry-picked from the original compound plates; thus these compounds had undergone one freeze-thaw cycle since HTS. The cherry-picked hits were serially diluted in DMSO, after which the resulting dose–response plates were immediately heat-sealed and stored at room temperature in a desiccator for no more than 1–2 days prior to IC_{50} testing. Overall, the plate handling was designed to minimize exposure of the compounds to water and to reduce storage times as much as practically possible. To rule out compound degradation as a cause of data inconsistencies, hits were analyzed for purity and integrity by LC-MS and found to match the compounds originally analyzed and shipped from BioFocus DPI. In addition, the compound series **3a–i** (Fig. 4) selected for hit-to-lead follow-up was reordered from BioFocus DPI, and the newly delivered compounds were compared with the in-house cherry-picks by LC-MS and IC_{50} testing.¹⁸

Screening and hit confirmation

A library of 64,000 compounds was screened at 5 μM in mixtures and individually at 10 μM . Each HTS run included two QC plates containing E-64 at 50 nM as a reference inhibitor; 3,840 wells gave mean inhibition of $69 \pm 3.2\%$. E-64 was also present in the screening library, and, consistent with previous studies,¹⁶ gave 100% inhibition both as a single compound at 10 μM and as a mixture component at 5 μM . Dose–response confirmation gave IC_{50} values of 12 nM and 20 nM in the presence of DTT and cysteine, respectively (Table 1, compound **1b**). To complete single-compound screening 209 plates were assayed over 5 days, and mixture HTS was completed in a single 40-plate run. All plates gave a Z' -factor of at least 0.8. As compounds in the mixture HTS were each present in two mixtures, data were rearranged to identify individual compounds that showed $>20\%$ inhibition in both locations, and these compounds were selected as hits. This hit threshold was more than fivefold larger than the mean CV of the plate controls (3.9%). For our analysis, hits in the single-compound HTS were defined as those compounds that exceeded 33% inhibition;

this threshold was more than eightfold above the mean CV of the plate controls (3.7%).

The higher hit threshold used in the analysis of the single-compound HTS as compared to the mixture HTS (33% compared with 20%) was designed to compensate for the difference in compound concentration (10 μM in the single-compound HTS and 5 μM in the mixture HTS). For a compound showing inhibition that saturates at 100%, the following equation may be used to estimate its IC_{50} from the percentage inhibition observed at a single screening concentration:

$$\text{Percentage inhibition} = 100[I]/([I] + \text{IC}_{50}) \quad (1)$$

where $[I]$ = inhibitor concentration. Rearranging:

$$\text{IC}_{50} = [I](100 - \text{percentage inhibition}) / \text{percentage inhibition} \quad (2)$$

Applying Eq. 2 to the mixture HTS, we may predict an IC_{50} of 20 μM for a compound that shows 20% inhibition at a concentration of 5 μM :

$$\text{IC}_{50} = 5(100 - 20)/20 (\mu\text{M}) = 20 \mu\text{M}$$

Equation 1 may then be used to estimate the percentage inhibition that the same compound with IC_{50} of 20 μM would give when screened in the single-compound HTS at 10 μM :

$$\text{Percentage inhibition} = 100(10)/(10 + 20) (\%) = 33\%$$

Hence a hit threshold of 33% in the single compound HTS at 10 μM would be expected to generate the same hits as a threshold of 20% in the mixture HTS at 5 μM .

Hits from mixture or single-compound HTS (or both) were cherry-picked from 10 mM master plates and tested in dose–response from 50 μM to 1.5 nM. Full HTS and dose–response results from the mixture screening are available in PubChem under assay IDs 488 and 830, and single-compound results are available under assay IDs 453 and 820.^{19,20} All compounds that were active in dose–response in the presence of DTT were also tested with cysteine in the assay buffer in place of DTT. It has been reported previously that redox-sensitive compounds can be reduced by DTT to produce reactive oxygen species such as hydrogen peroxide. Under these conditions enzymes that contain an active-site cysteine are inactivated by thiol oxidation. There are numerous examples of such compounds that react with DTT and thus cause enzyme inactivation but show no activity in the presence of cysteine.^{21,22} Thus compounds that inhibited cathepsin B in the presence of DTT but not cysteine were judged to be artifacts. Table 1 lists the compounds con-

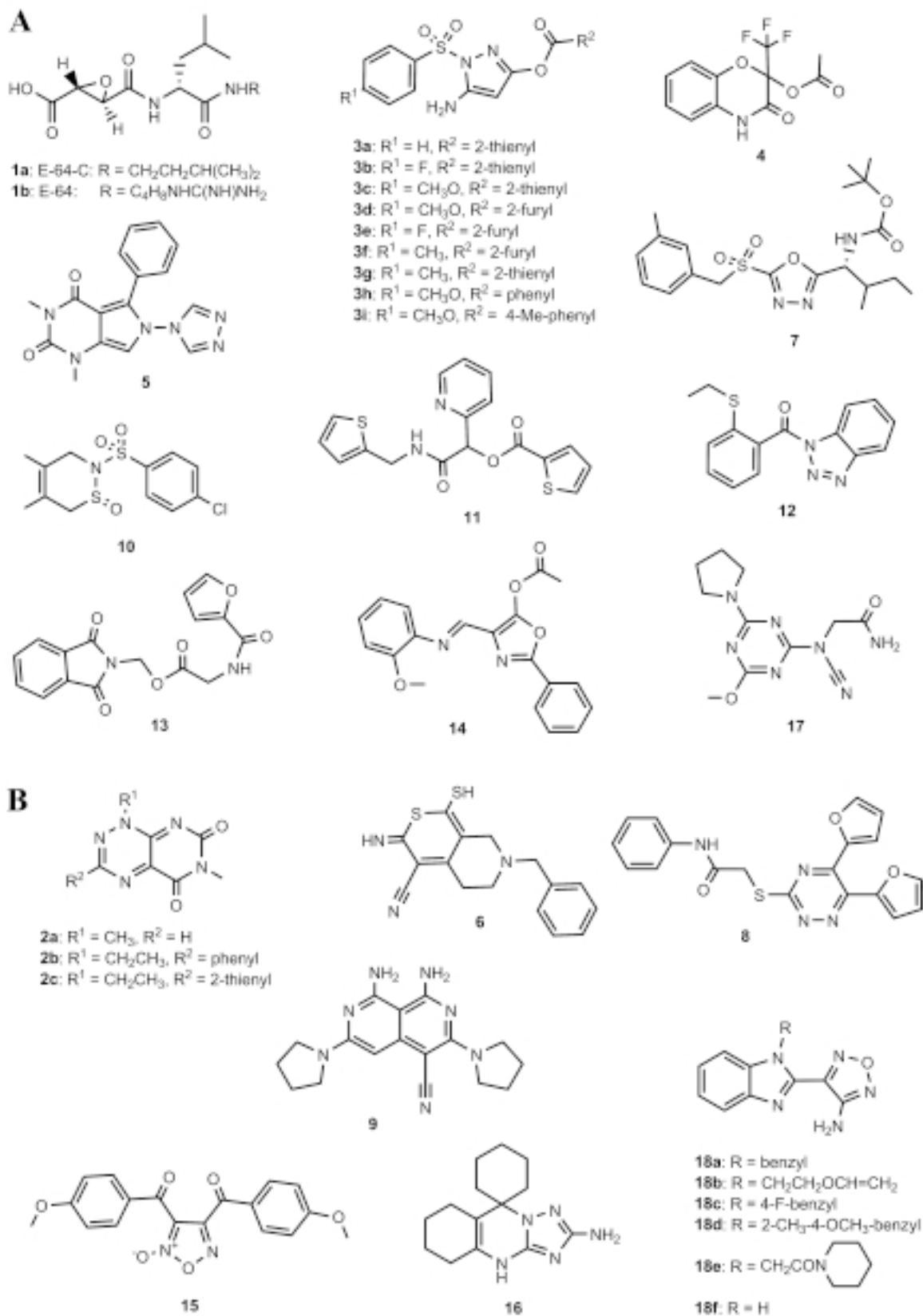


FIG. 4. Cathepsin B inhibitors confirmed by dose-response testing. (A) Compounds active in the presence of DTT or cysteine; not DTT-reactive artifacts. (B) DTT-reactive artifacts active only in the presence of DTT.

TABLE 1. CONFIRMED ACTIVE COMPOUNDS FROM CATHEPSIN B MIXTURE AND SINGLE-COMPOUND HTS

Compound	Pubchem SID	Single- compound HTS	Mixture HTS percentage inhibition		Single- compound HTS hits	Mixture HTS hits	DTT- reactive artifacts	IC ₅₀ (μM)	
		percentage inhibition	Location 1	Location 2				DTT-containing buffer	Cys-containing buffer
1a	855916	100.0	100.3	100.4	Hit	Hit		<0.0015	<0.0015
1b (E-64)	855920	100.0	98.6	100.9	Hit	Hit		0.0123 ± 0.0003	0.0204 ± 0.002
2a	4251194	83.0	75.3	83.9	Hit	Hit	Artifact	0.0461 ± 0.0026	>50
2b	845167	76.2	8.4	85.7	Hit		Artifact	0.071 ± 0.007	>50
2c	850758	81.9	86.9	85.0	Hit	Hit	Artifact	0.0722 ± 0.0033	>50
3a	4249135	75.7	65.4	72.2	Hit	Hit		0.247 ± 0.024	0.365 ± 0.047
3b	4247730	55.6	46.0	64.7	Hit	Hit		0.435 ± 0.089	0.908 ± 0.241
3c	4245669	76.0	70.4	66.5	Hit	Hit		0.692 ± 0.079	0.69 ± 0.231
3d	851299	78.2	61.8	73.1	Hit	Hit		0.924 ± 0.034	3.29 ± 0.19
3e	845259	61.2	45.1	74.1	Hit	Hit		1.26 ± 0.03	8.56 ± 0.37
3f	849441	76.1	49.9	68.4	Hit	Hit		1.75 ± 0.05	8.07 ± 0.45
3g	844213	63.6	36.6	57.0	Hit	Hit		1.99 ± 0.14	3.04 ± 1.05
3h	852843	39.1	26.3	32.1	Hit	Hit		9.56 ± 0.58	21.4 ± 4.5
3i	844423	48.7	33.1	30.4	Hit	Hit		12.3 ± 5.7	16.9 ± 1.4
4	4249019	74.7	51.8	69.6	Hit	Hit		0.845 ± 0.016	1.21 ± 0.11
5	4241883	59.2	49.0	51.5	Hit	Hit		1.19 ± 0.07	5.95 ± 0.76
6	845964	80.5	15.5	57.2	Hit		Artifact	1.64 ± 0.04	>50
7	864233	33.7	20.4	25.2	Hit	Hit		2.09 ± 0.13	41.5 ± 10.8
8	860234	57.1	8.3	1.7	Hit		Artifact	2.25 ± 0.18	>50
9	3712249	45.4	77.6	76.0	Hit	Hit	Artifact	3.17 ± 0.55	>50
10	3713060	23.7	32.5	30.9		Hit		4.17 ± 0.66	7.52 ± 1.24
11	845947	35.0	25.6	31.3	Hit	Hit		6.36 ± 0.22	6.44 ± 0.33
12	4255516	32.2	26.5	32.8		Hit		7.11 ± 0.81	14.5 ± 2.1
13	4249181	21.5	25.9	24.5		Hit		8.56 ± 0.6	7.91 ± 0.74
14	857378	40.6	23.6	41.9	Hit	Hit		8.93 ± 0.75	17.7 ± 0.7
15	856149	31.0	40.6	33.4		Hit	Artifact	11.5 ± 2.2	>50
16	7973426	39.1	58.8	79.0	Hit	Hit	Artifact	12.9 ± 0.4	>50
17	4250069	36.7	25.8	20.8	Hit	Hit		14.2 ± 1.3	13.4 ± 0.5
18a	4258256	4.7	21.1	64.5		Hit	Artifact	22.4 ± 2.4	ND
18b	7977171	15.0	47.5	62.7		Hit	Artifact	29.7 ± 2	ND
18c	852689	73.1	76.8	82.5	Hit	Hit	Artifact	33.9 ± 0.9	>50
18d	4261352	45.2	22.8	16.2	Hit		Artifact	37.2 ± 0.4	>50
18e	850777	81.2	86.0	86.2	Hit	Hit	Artifact	44.6 ± 0.2	>50
18f	4259392	18.1	57.6	46.5		Hit	Artifact	46.1 ± 1.8	ND

For compound, numbers refer to structures shown in Fig. 4. For Pubchem SID, numbers are the substance ID used to retrieve biological data and chemical properties from PubChem. Single-compound HTS hits are defined as compounds that gave single-compound HTS percentage inhibition >33%. Mixture HTS hits are defined as compounds that gave mixture HTS percentage inhibition >20% in both mixture wells where they were present (locations 1 and 2). DTT-reactive artifacts are compounds that were active against cathepsin B in the presence of DTT but inactive in the presence of cysteine. IC₅₀ values are expressed as the mean of three independent determinations ±SD.

firmed active in dose–response and shows the corresponding HTS percentage inhibition and IC₅₀ values in the presence of DTT or cysteine. Structures of the compounds listed in Table 1 are shown in Fig. 4. It is noteworthy that the structures of the DTT-reactive compounds (Fig. 4B) all contain conjugated nitrogen-carbon bonds. This functionality has previously been correlated with redox activity in the presence of DTT.²²

Comparison between mixture and single-compound HTS results

Hit confirmation results from the mixture and single-compound screening were analyzed and compared (Table

2). The mixture screening gave a higher retest rate than the single-compound screening (79% vs. 55%). Furthermore, of the compounds selected as hits from the mixture HTS that were found to be inactive on retest, four were found to originate from a mixture that contained a confirmed active compound (Table 2, row 3). For example, compound SID 4264967 was identified as a hit based on mixture HTS percentage inhibition values of 98.6% and 26.9%. However, the mixture with percentage inhibition of 98.6% also contained E-64, which gave an IC₅₀ of 0.012 μM on retest (Table 1, compound **1b**). Thus E-64 and not SID 4264967 most likely gave the observed mixture percentage inhibition of 98.6%. Thus the true

TABLE 2. SUMMARY OF MIXTURE AND SINGLE-COMPOUND HTS RESULTS

Hit classification ^a	Number of hits	
	Mixture HTS	Single-compound HTS
Primary HTS hits	38	49
Active in dose-response	30	27
Marginal in dose-response	3	9
Inactive in dose-response (another active present)	4	N/A
Inactive in dose-response (no other active present)	1	13
Retest rate	79%	55%
False-positive rate	2.6%	27%

^aPrimary HTS hits were defined as >20% inhibition in both locations in mixture HTS, or >33% inhibition in the one location in single-compound HTS. Active in dose-response was defined as $IC_{50} > 50 \mu M$. Marginal in dose-response was defined as $IC_{50} > 50 \mu M$ with percentage inhibition 30–50% at $50 \mu M$. Inactive in dose-response (another active present) was defined as dose-response testing in which individual compounds gave no activity, defined as $IC_{50} > 50 \mu M$ and percentage inhibition at $50 \mu M < 30\%$. These compounds were selected for dose-response testing as the two mixtures from which they originated gave percentage inhibition >20% in the mixture HTS. However, the mixture that gave the highest percentage inhibition in HTS was found also to contain one of the 30 compounds confirmed active in dose-response testing. Inactive in dose-response (no other active present) was defined as dose-response testing in which individual compounds gave no activity, defined as $IC_{50} > 50 \mu M$ and percentage inhibition at $50 \mu M < 30\%$. In the case of single-compound HTS, these compounds were selected for dose-response testing as they were identified as hits that gave percentage inhibition >33%. In the case of mixture HTS, this compound was selected for dose-response testing as the two mixtures from which it originated gave percentage inhibition >20%. Despite the discrepancy between activity in the mixture HTS and lack of activity in dose-response, none of the 30 confirmed active compounds was present in the original mixtures. Retest rate was defined as active in dose-response as a percentage of primary HTS hits. False-positive rate was defined as inactive in dose-response (no other active present) as a percentage of primary HTS hits. N/A, not applicable.

← AU8

false-positive rate in the mixture HTS was >10-fold lower than that for the single-compound HTS (2.6% vs. 27%). These percentages are based on one HTS campaign under each set of conditions so we have not determined the degree to which the differences are statistically significant. Nevertheless, the mixture HTS does appear to be the more reliable predictor of compound activity on subsequent dose-response confirmation. This presu-

ably reflects the statistical value of screening each mixture component in duplicate rather than in one well in the case of the single compound HTS.

The most important factor to consider in validating the mixture screening is the false-negative rate. Thus we addressed the following question: are there any confirmed active compounds from the single-compound screening that were missed in the mixture HTS? Inspection of Table

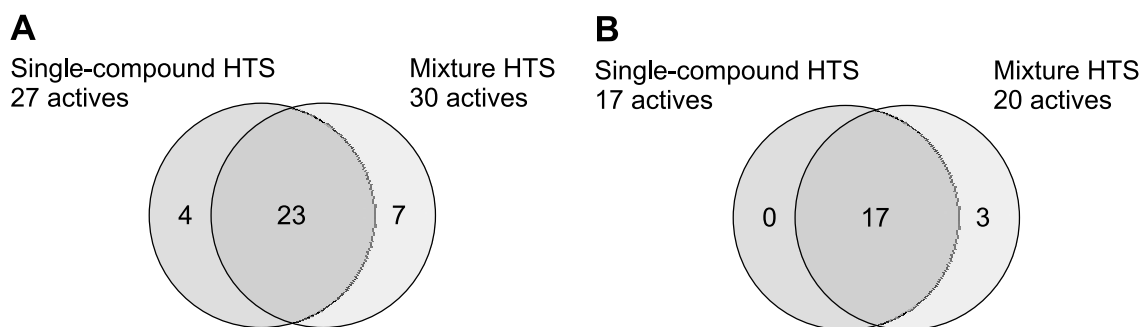


FIG. 5. Overlap of single compound and mixture actives. **(A)** All compounds included. Total single-compound HTS hits active in IC_{50} testing = 27; number of single-compound IC_{50} actives that were hits in mixture HTS = 23 (85%); number of single-compound IC_{50} actives missed in mixture HTS = 4 (15%). **(B)** DTT-reactive artifacts excluded. Total single-compound HTS hits active in IC_{50} testing = 17; number of single-compound IC_{50} actives that were hits in mixture HTS = 17 (100%); number of single-compound IC_{50} actives missed in mixture HTS = 0.

1 reveals that the majority of confirmed active compounds were hits in both screens. However, some compounds were hits in either single-compound or mixture HTS but not both. The extent of the overlap of the HTS hits is illustrated by the Venn diagram in Fig. 5. Of 27 confirmed actives identified from single-compound HTS, 23 were also identified from mixture HTS. Thus the mixture screening missed four confirmed actives from the single-compound screening. However, this analysis looks significantly different if the DTT-reactive artifacts are excluded. As discussed above, numerous compounds were found to cause enzyme inactivation by reacting with DTT in the assay buffer. After retesting in dose-response using a cysteine-containing buffer there remained 17 confirmed actives from single-compound screening, all of which were hits in the mixture HTS. Thus after excluding artifacts mixture screening identified all of the actives from the single-compound screening. Interestingly, mixture HTS identified an additional three compounds that were not classified as hits in the single-compound HTS. However, these compounds did show 20–30% inhibition in the single-compound HTS, below the 33% threshold used for hit selection. The fact that the activity of the DTT-reactive artifacts did not correlate well between mixture and single-compound HTS suggests that other compounds in the mixture may modulate the reactivity of these artifacts, by affecting either initial compound reduction by DTT or inactivation of the cathepsin B enzyme by the reactive oxygen species generated by the DTT-compound redox chemistry.

Hit-to-lead follow-up

Inspection of the structures of the confirmed actives in Table 1 allowed us to group the hits into 20 structural classes (Fig. 4). After excluding analogs of the known cysteine protease inhibitor E-64 (compounds **1a** and **1b**) and DTT-reactive artifacts 11 structural classes remained. The series of aminopyrazoles represented by **3a–i** presented a promising structure–activity relationship, with IC_{50} values ranging from 0.247 to 12.3 μM . Furthermore, these compounds appeared not to be DTT-mediated redox artifacts and were accessible to medicinal chemistry optimization. Results from synthesis and testing of a series of analogs are available in Pubchem under assay ID 523.¹⁸ A full report of this hit-to-lead effort has been published.²³

Conclusions

Screening of 64,000 compounds from the MLSCN library as mixtures of 10 compounds per well against cathepsin B identified 20 actives. These actives included all those identified by single-compound screening against cathepsin B. Thus mixture screening represents a viable

strategy for HTS of proteases. Based on this validation we have used this mixture HTS method to profile the MLSCN library against the cysteine protease cathepsin S²⁴ and the serine proteases cathepsin G,²⁵ complement factor C1s,^{26,27} and coagulation factors XIa^{28,29} and XIIa.^{30,31}

Acknowledgments

We are grateful to Dr. Michael Myers for performing LC-MS analyses. This work was supported by the NIH Molecular Libraries Screening Center Network (grant U54HG003915-02).

References

1. Penn Center for Molecular Discovery: [online]. <http://www.seas.upenn.edu/~pcmd/> (last accessed January 15, 2008).
2. Devlin JJ, Liang A, Trinh L, Polokoff MA, Senator D, Zheng W, *et al.*: High capacity screening of pooled compounds: identification of the active compound without re-assay of pool members. *Drug Dev Res* 1996;37:80–85.
3. Hann M, Hudson B, Lewell X, Lifely R, Miller L, Ramsden N: Strategic pooling of compounds for high-throughput screening. *J Chem Inform Comput Sci* 1999;39:897–902.
4. Ferrand S, Schmid A, Engeloch C, Glickman JF: Statistical evaluation of a self-deconvoluting matrix strategy for high-throughput screening of the CXCR3 receptor. *Assay Drug Dev Technol* 2005;3:413–424.
5. Frlan R, Gobec S: Inhibitors of cathepsin B. *Curr Med Chem* 2006;13:2309–2327.
6. Stoka V, Turk B, Turk V: Lysosomal cysteine proteases: structural features and their role in apoptosis. *IUBMB Life* 2005;57:347–353.
7. Turk B, Stoka V, Rozman-Pungercar J, Cirman T, Droga-Mazovec G, Oreic K, *et al.*: Apoptotic pathways: involvement of lysosomal proteases. *Biol Chem* 2002;383:1035–1044.
8. Jedeszko C, Sloane BF: Cysteine cathepsins in human cancer. *Biol Chem* 2004;385:1017–1027.
9. Michaud S, Gour BJ: Cathepsin B inhibitors as potential anti-metastatic agents. *Expert Opin Ther Patents* 1998;8:645–672.
10. Schmitt M, Janicke F, Graeff H: Tumor-associated proteases. *Fibrinolysis* 1992;6:3–26.
11. Vasiljeva O, Reinheckel T, Peters C, Turk D, Turk V, Turk B: Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Curr Pharm Design* 2007;13:387–403.
12. Schornberg K, Matsuyama S, Kabsch K, Delos S, Bouton A, White J: Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol* 2006;80:4174–4178.
13. Haagmans BL, Osterhaus ADME: Coronaviruses and their therapy. *Antiviral Res* 2006;71:397–403.
14. Chandran K, Sullivan NJ, Felbor U, Whelan SP, Cunningham JM: Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* 2005;308:1643–1645.

15. BioFocus DPI: [online]. http://mlsmr.glp.com/MLSMR_HomePage/ (last accessed January 15, 2008).
16. Baricos W, Zhou Y, Mason R, Barrett A: Human kidney cathepsins B and L. *Biochem J* 1998;252:301–304.
17. Barrett A, Kembhavi A, Brown M, Kirschke H, Knight G, Tamai M, Hanada K: L-trans-Epoxy succinyl-leucyl-amido(4-guanidini)butane (E64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochem J* 1982;201:189–198.
18. Penn Center for Molecular Discovery: Cathepsin B Inhibitor Series SAR Studies [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=523> (last accessed January 15, 2008).
19. Penn Center for Molecular Discovery: Cathepsin B [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=453> (last accessed January 15, 2008).
20. Penn Center for Molecular Discovery: Cathepsin B compound mixture screening [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=488> (last accessed January 15, 2008).
21. Crowe A, Ballatore C, Hyde E, Trojanowski JQ, Lee VM: High throughput screening for small molecule inhibitors of heparin-induced tau fibril formation. *Biochem Biophys Res Commun* 2007;358:1–6.
22. University of Pittsburgh Molecular Library Screening Center: Cdc25B Catalytic Domain Protein Tyrosine Phosphatase Probe Assessment: Compound Dose Dependent Redox Cycling H₂O₂ Generation in 0.5 mM DTT [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=682> (last accessed January 15, 2008).
23. Myers MC, Napper A, Motlekar N, Shah P, Chiu CH, Beavers MP, *et al.*: Identification and characterization of 3-substituted pyrazolyl esters as alternate substrates for cathepsin B: the confounding effects of DTT and cysteine in biological assays. *Bioorg Med Chem Lett* 2007;17:4761–4766.
24. Penn Center for Molecular Discovery: Cathepsin S [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=501> (last accessed January 15, 2008).
25. Penn Center for Molecular Discovery: Cathepsin G [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=581> (last accessed January 15, 2008).
26. Penn Center for Molecular Discovery: Complement Factor C1 [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=538> (last accessed January 15, 2008).
27. Penn Center for Molecular Discovery: Complement factor C1s IC50 from mixture screen [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=787> (last accessed January 15, 2008).
28. Penn Center for Molecular Discovery: Factor XIa Mixture HTS [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=680> (last accessed January 15, 2008).
29. Penn Center for Molecular Discovery: Factor XIa Dose Response Confirmation [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=679> (last accessed January 15, 2008).
30. Penn Center for Molecular Discovery: Factor XIIa Mixture HTS [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=684> (last accessed January 15, 2008).
31. Penn Center for Molecular Discovery: Factor XIIa Dose Response Confirmation [online]. <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=716> (last accessed January 15, 2008).

Address reprint requests to:

Andrew D. Napper, Ph.D.

Penn Center for Molecular Discovery

1160 Vagelos Research Laboratories

University of Pennsylvania

3340 Smith Walk

Philadelphia, PA 19104

E-mail: napper@seas.upenn.edu

AU7

MOTLEKAR

AU1

Verify supplier's location.

AU2

Verify supplier's location.

AU3

Verify supplier's location.

AU4

Material from original reference 16 incorporated into text.

AU5

Verify or correct location.

AU6

Original reference 23 cited out of order; check all renumbering is done correctly.

AU7

Verify degree.

AU8

N/A, not applicable as meant?