University of Pennsylvania
ScholarlyCommons

Departmental Papers (CBE)

Department of Chemical & Biomolecular Engineering

1-1-2008

Identification and synthesis of a unique thiocarbazate cathepsin L inhibitor

Michael C. Myers
University of Pennsylvania, myersmic@sas.upenn.edu

Parag P. Shah
University of Pennsylvania

Scott L. Diamond
University of Pennsylvania, sld@seas.upenn.edu

Donna M. Huryn
University of Pennsylvania

Amos B. Smith
University of Pennsylvania, smithab@sas.upenn.edu

Postprint version. Published in *Bioorganic and Medicinal Chemistry Letters*, Volume 18, Issue 1, January 2008, pages 210-214.
Publisher URL: http://dx.doi.org/10.1016/j.bmcl.2007.10.107

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cbe_papers/114
For more information, please contact repository@pobox.upenn.edu.

Identification and synthesis of a unique thiocarbazate cathepsin L inhibitor

Abstract
Library samples containing 2,5-disubstituted oxadiazoles were identified as potent hits in a high throughput screen (HTS) of the NIH Molecular Libraries Small Molecule Repository (MLSMR) directed at discovering inhibitors of cathepsin L. However, when synthesized in pure form, the putative actives were found to be devoid of biological activity. Analyses by LC–MS of original library samples indicated the presence of a number of impurities, in addition to the oxadiazoles. Synthesis and bioassay of the probable impurities led to the identification of a thiocarbazate that likely originated via ring opening of the oxadiazole. Previously unknown, thiocarbazates (-)-11 and (-)-12 were independently synthesized as single enantiomers and found to inhibit cathepsin 20 L in the low nanomolar range.

Keywords
thiocarbazate; cathepsin L inhibitor; cysteine protease inhibitor; MLSCN

Comments
Postprint version. Published in *Bioorganic and Medicinal Chemistry Letters*, Volume 18, Issue 1, January 2008, pages 210-214.
Publisher URL: http://dx.doi.org/10.1016/j.bmcl.2007.10.107

This journal article is available at ScholarlyCommons: http://repository.upenn.edu/cbe_papers/114

Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters xxx (2007) xxx–xxx
Bioorganic & Medicinal Chemistry Letters

Identification and synthesis of a unique thiocarbazate cathepsin L inhibitor

Michael C. Myers,[a,b] Parag P. Shah,[a,c] Scott L. Diamond,[a,c,*] Donna M. Huryn[a,b,*] and Amos B. Smith, III[a,b,*]

[a]*Penn Center for Molecular Discovery, University of Pennsylvania, 1024 Vagelos Research Laboratories, Philadelphia, PA 19104-6383, USA*
[b]*Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, PA 19104-6323, USA*
[c]*Institute for Medicine and Engineering, University of Pennsylvania, 1024 Vagelos Research Laboratories, Philadelphia, PA 19104-6383, USA*

Received 6 October 2007; revised 22 October 2007; accepted 25 October 2007

Abstract—Library samples containing 2,5-disubstituted oxadiazoles were identified as potent hits in a high throughput screen (HTS) of the NIH Molecular Libraries Small Molecule Repository (MLSMR) directed at discovering inhibitors of cathepsin L. However, when synthesized in pure form, the putative actives were found to be devoid of biological activity. Analyses by LC–MS of original library samples indicated the presence of a number of impurities, in addition to the oxadiazoles. Synthesis and bioassay of the probable impurities led to the identification of a thiocarbazate that likely originated via ring opening of the oxadiazole. Previously unknown, thiocarbazates (−)-**11** and (−)-**12** were independently synthesized as single enantiomers and found to inhibit cathepsin L in the low nanomolar range.
© 2007 Elsevier Ltd. All rights reserved.

The cathepsins comprise a family of lysosomal protease enzymes whose primary function (i.e., protein degradation) plays a critical role in normal cellular homeostasis.[1] Overexpression of cathepsin L and/or abnormal activity has been implicated in a number of disease states.[2] For example, cathepsin L is responsible for bone resorption through degradation of collagen type I; this disregulation is believed to lead to osteo- and rheumatoid arthritis.[3] In addition, several infective organisms, such as SARS and Ebola viruses, utilize cathepsin L-like proteins for replication in human cells.[4] The large number of disease states associated with cathepsin L calls for an understanding of the biological function.[2]

Recently, the Penn Center for Molecular Discovery (PCMD),[5] completed a high throughput screening (HTS) campaign of the NIH Molecular Libraries Small Molecule Repository (MLSMR) to identify inhibitors of members of the papain-like cysteine protease family, including cathepsins B, L, and S.[6] In this letter, we detail our continuing efforts to create a comprehensive, publicly available profile of small-molecule inhibitors of the cysteine protease class, and herein describe the identification of a novel class of potent cathepsin L inhibitors.

Previously reported inhibitors of cathepsin L include the peptides, leupeptin and aprotinin, and the fluoromethyl ketone, *Z*-LLL-FMK.[3,7] The few known, potent small-molecule inhibitors are either peptidic and therefore likely to suffer from physiological instability and poor permeability, or are non-selective for cathepsin L.[3,8,9] The identification of potent, selective, stable, and cell permeable small-molecule inhibitors would therefore be a valuable tool to interrogate cathepsin L and cathepsin L-like function, as well as to provide potential starting points for drug discovery and development.[10–15]

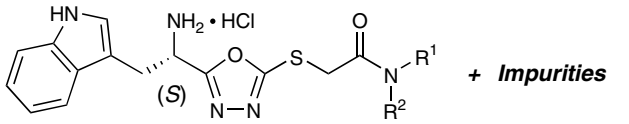
Initial HTS results of our cathepsin L screen indicated that several structurally related oxadiazoles exhibited potent inhibitory activity (Table 1).[16–18]

The most potent hit, cataloged in the MLSMR as disubstituted oxadiazole **1**, exhibited an IC_{50} of 0.13 μM.[17] Wells containing related analogs (**2**–**5**), differing only

Keywords: Thiocarbazate; Cathepsin L inhibitor; Cysteine protease inhibitor; MLSCN.
Q1 * Corresponding authors. E-mail: huryn@sas.upenn.edu

0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.10.107

Table 1. Cathepsin L inhibitory activity of oxadiazole-containing library samples[a]



PubChem SID	R^1	R^2	IC_{50}[b] (μM)
861540 (**1**)	2-Ethylbenzene	H	0.13 ± 0.01
861087 (**2**)	2,4-Dimethylbenzene	H	0.16 ± 0.05
861840 (**3**)	Me	Me	0.17 ± 0.02
861542 (**4**)	2,3-Dimethylbenzene	H	0.30 ± 0.04
861992 (**5**)	Et	Et	0.51 ± 0.02

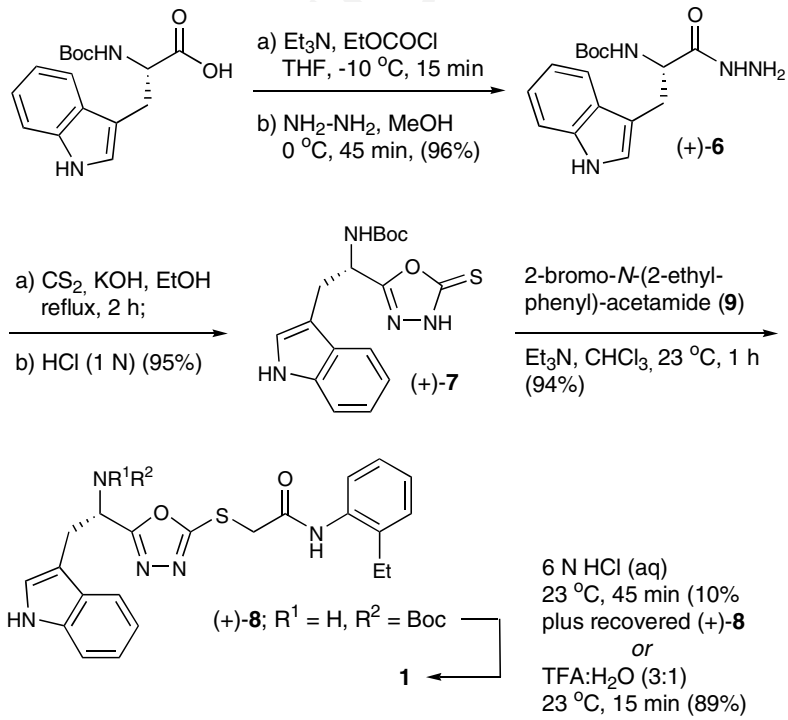
[a] Library samples containing the parent oxadiazole with impurities.
[b] IC_{50} values are reported as means ± standard deviation (number of determinations = 3).

in the amide nitrogen substitution, displayed a range of activities (IC_{50} = 0.16–0.51 μM), suggesting a nascent structure activity relationship. As is our practice, library samples containing the putative active oxadiazoles were evaluated for both purity and integrity by LC–MS analysis. This analysis indicated that the primary constituent of each sample was indeed the expected oxadiazole, in up to 60% purity, however numerous impurities were also present. To confirm the biological activity attributed to **1**, a synthetic sequence was developed to generate the oxadiazoles in pure form.

Although little literature precedence exists for the construction of compounds such as **1**, we realized that modification of the Woodward–Confalone[19] approach to α-amino acid substituted oxadiazolethiones could provide an entry to this class of sulfur-based 2,5-disubstituted oxadiazoles. Toward this end, commercially available L-Boc-Trp-OH was converted in high yield to the corresponding hydrazide (+)-**6**, on a 10-g scale (Scheme 1).[20]

Cyclization with carbon disulfide in ethanol at reflux afforded the expected substituted thione (+)-**7** as a stable solid in excellent yield, which upon chemoselective alkylation with 2-bromo-*N*-(2-ethyl-phenyl)-acetamide[21] (**9**) efficiently generated (+)-**8** (94% yield). Further study indicated that isolation of (+)-**7** was unnecessary, and (+)-**8** could be prepared directly from (+)-**6**, in a single flask. Both sequences furnished (+)-**8** in high yield. In similar fashion, the antipode, (−)-**8**, was prepared starting from D-Boc-Trp-OH. High enantiomeric purities (>99%) were demonstrated via chiral supercritical fluid chromatography.

Initial attempts to remove the Boc-group utilizing HCl in dioxane (4 N) or HCl in water (6 N) were compromised by the poor solubility of (+)-**8**. However, small quantities of the HCl salt of **1** could be obtained by treatment with 6 N HCl over 45 min. Upon LC–MS analysis of synthetic **1**, we recognized that the impurities formed during the acid-promoted Boc-deprotection step were identical to those found in the library screening samples (vide infra). Optimal conditions to remove the Boc-group were eventually developed. Specifically, treatment of (+)-**8** with 25% water in TFA for 15 min, followed in turn by adjustment of the pH to 8.0 with $NaHCO_3$ and aqueous extraction with methylene chloride-furnished the free-base of **1** in 89% yield and 99% after column chromatography. In the end, oxadiazole **1** was prepared on gram scale from L-Boc-Trp-OH; the overall yield was 76%.



Scheme 1. Synthesis of 2,5-disubstituted oxadiazoles.

Please cite this article in press as: Myers, M. C. et al., *Bioorg. Med. Chem. Lett.* (2007), doi:10.1016/j.bmcl.2007.10.107

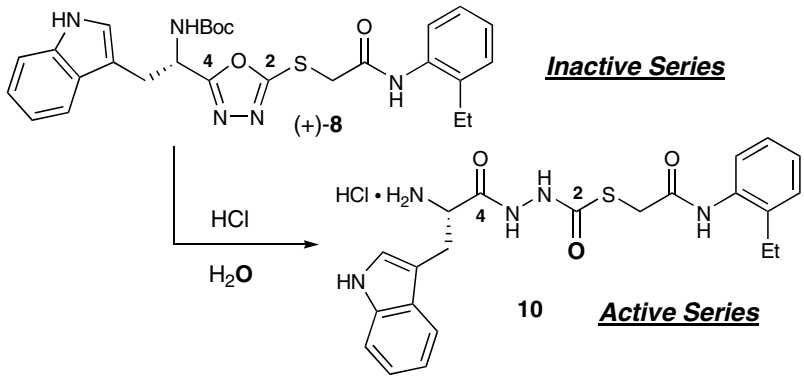


Figure 1. Conversion of oxadiazole (+)-**8** to thiocarbazate **10**.

To our surprise, the free-base of **1** was found to be completely devoid of activity when assayed against cathepsin L.[18] This result suggested that an impurity present in the original library sample was responsible for the observed activity. Based on LC–MS analyses of the biologically active samples, we hypothesized that the active component was likely the ring-opened product **10**, formed via acid-promoted addition of H_2O to (+)-**8** (Fig. 1).

The presence of a molecular [M+1] ion equal to 440 amu in the LC–MS analysis of impure samples was consistent with the presence of **10**, thereby providing strong support for the hypothesis.

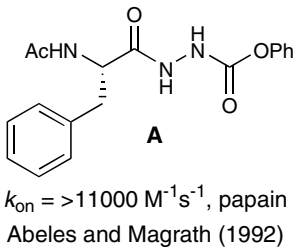
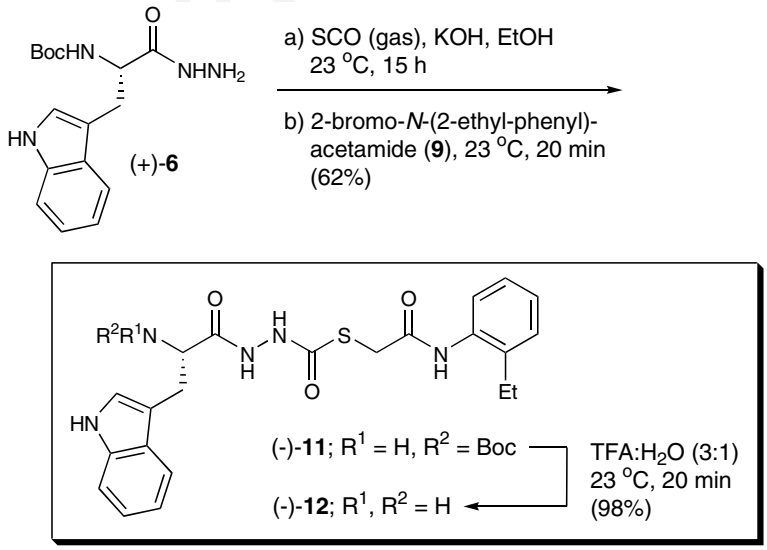


Figure 2. Known aza-peptide cathepsin inhibitor **A**.

Thiocarbazates such as **10** have not been described previously in the literature. However, these compounds do bear structural resemblance to aza-peptides[22–24] (e.g., **A**; Fig. 2),[25] examples, of which have been reported to exhibit cysteine protease inhibitory activity through a mechanism involving attack by the active site cysteine on the carbamate carbonyl.[26]

To test the hypothesis that **10** was indeed the active species, we devised an expedient synthesis of this structural class. We envisioned that introduction of the C2 carbonyl unit of **10** could be achieved via chemistry in parallel to that used to prepare **1**, beginning with hydrazide (+)-**6**. We began the synthesis by converting tryptophan hydrazide (+)-**6** to an intermediate thiosemicarbazide (not shown) employing carbonyl sulfide (S=C=O) gas dissolved in ethanol (Scheme 2).[27]

The intermediate thiosemicarbazide did not precipitate from solution, therefore 2-bromo-*N*-(2-ethyl-phenyl)-acetamide[21] (**9**) was added to the reaction flask to generate thiocarbazate (−)-**11**.[28] Pleasingly, the yield for the two-step, one-flask operation was 62%. Deprotection of (−)-**11**, employing our previously developed condi-



Scheme 2. Synthetic procedure to prepare thiocarbazates.

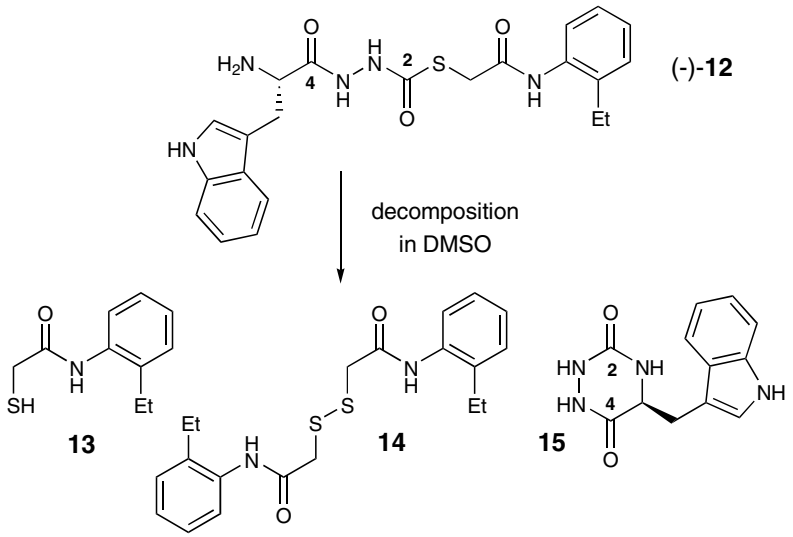


Figure 3. Decomposition products of (−)-**12** in DMSO.

tions (25% water in TFA), generated the free-base (−)-**12** in 98% yield with >95% purity. This three-step sequence permitted construction of (−)-**12** on gram scale, starting from L-Boc-Trp-OH; the overall yield was 58%.

As anticipated, both (−)-**11**(*S*) and (−)-**12**(*S*) exhibited potent inhibitory activity against cathepsin L with IC_{50} values of 56 nM and 133 nM, respectively.[29] The *R*-enantiomer, (+)-**11**(*R*),[30] was only modestly active against cathepsin L (IC_{50} = 34 μM).[29] Some variability in IC_{50} values determined for (−)-**12** prompted us to explore the stability of both (−)-**11** and (−)-**12** in solvents relevant to the bioassay. While (−)-**11** was found to be completely stable in DMSO, as well as in the buffer employed in the Cathepsin L assay [e.g., NaOAc (20 mM), pH 5.5; EDTA (1 mM); cysteine (5 mM)], the free-base (−)-**12** proved unstable in DMSO, generating decomposition products **13**, **14**, and **15** after only 1 h at room temperature (Fig. 3).

Presumably **13**–**15** are formed from (−)-**12** via intramolecular attack of the primary amine on the C2 thiocarbazate moiety, and release of **13**. These products, prepared either via synthesis (**13** and **14**) or by HPLC purification of decomposed material (**15**), were assayed for cathepsin L activity and found to be inactive. Due to the instability of (−)-**12** under the assay conditions, IC_{50} values vary somewhat, rendering interpretation of the bioassay data difficult. We are, however, confident in the accuracy of the bioassay results for (−)-**11**, as this compound is stable under all conditions evaluated. Thus, future efforts will be focused on the more stable thiocarbazate (−)-**11** and related congeners.

As most cysteine protease inhibitors contain an electrophilic 'warhead,' and work through a mechanism involving reaction with the active site cysteine,[7,26] we strongly believe these novel thiocarbazates behave similarly, and are active by virtue of their electrophilic carbonyl moiety. The formation of **15** supports this mechanism and is indicative of the electrophilicity of the C2 carbonyl moiety. Reactivity at the anilide carbonyl is also a possibility and cannot be ruled out. However, we observe no evidence of side-products resulting from reaction at this position (cf. Fig. 3). A complete biochemical characterization of this novel chemotype is underway.[31]

In summary, samples from the NIH MLSMR revealed promising cathepsin L inhibitory activity attributed to a series of 2,5-disubstituted oxadiazoles (**1**–**5**). Analytical analyses (LC–MS) of the library samples, however, indicated numerous impurities, thus requiring the development of an efficient synthesis for the putative active 2,5-disubstituted oxadiazoles. Synthetic samples of pure 2,5-disubstituted oxadiazoles were found to be completely devoid of cathepsin L inhibitory activity. Careful LC–MS investigation of both the library and synthetic samples revealed a thiocarbazate to be the active component. Bioassay of the synthetic thiocarbazates confirmed the hypothesis: (−)-**11** and (−)-**12** display potent inhibitory activity against cathepsin L, with IC_{50} values of 56 nM and 133 nM, respectively, however instability of (−)-**12** was also noted. The design, syntheses, and biological evaluation of analogs of the highly potent lead thiocarbazate class are now ongoing in our laboratories, and will be reported in due course.

Acknowledgments

Financial support for this work was provided by the NIH (5U54HG003915-02 and 5U54HG003915-03). We thank Professor Barry S. Cooperman (University of Pennsylvania) and Dr. Mary Pat Beavers (University of Pennsylvania) for helpful discussions. We also thank Christina Kraml (Lotus Separations) for chiral SFC analysis. Finally we thank Dr. G.T. Furst (NMR Facility) and Dr. R. Kohli (Mass Spectrometry Facility) at the University of Pennsylvania for assistance in obtaining NMR and high-resolution mass spectra.

References and notes

1. McGrath, M. E. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 181.
2. Vasiljeva, O.; Reinheckel, T.; Peters, C.; Turk, D.; Turk, V.; Turk, B. *Curr. Pharm. Des.* **2007**, *13*, 385.
3. Esser, R. E.; Angelo, R. A.; Murphey, M. D.; Watts, L. M.; Thornburg, L. P.; Palmer, J. T.; Talhouk, J. W.; Smith, R. E. *Arthritis Rheum.* **1994**, *37*, 236.
4. Simmons, G.; Gosalia, D. N.; Rennekamp, A. J.; Reeves, J. D.; Diamond, S. L.; Bates, P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 11876.
5. Penn Center for Molecular Discovery: http://www.seas.upenn.edu/~pcmd/ Molecular Library Screening Center Network: http://nihroadmap.nih.gov/molecularlibraries/.
6. Myers, M. C.; Napper, A. D.; Motlekar, N.; Shah, P. P.; Chiu, C.-H.; Beavers, M. P.; Diamond, S. L.; Huryn, D. M., ; Smith, A. B., III *Bioorg. Med. Chem. Lett.* **2007**, *17*, 4761.
7. Otto, H.-H.; Schirmeister, T. *Chem. Rev.* **1997**, *97*, 133.
8. Montaser, M.; Lalmanach, G.; Mach, L. *Biol. Chem.* **2002**, *383*, 1305.
9. Fujishima, A.; Imai, Y.; Nomura, T.; Fujisawa, Y.; Yamamoto, Y.; Sugawara, T. *FEBS Lett.* **1997**, *407*, 47.
10. Schirmeister, T.; Kaeppler, U. *Mini-Rev. Med. Chem.* **2003**, *3*, 361.
11. Hernandez, A. A.; Roush, W. R. *Curr. Opin. Chem. Biol.* **2002**, *6*, 459.
12. Marquis, R. W.; James, I.; Zeng, J.; Trout, R. E. L.; Thompson, S.; Rahman, A.; Yamashita, D. S.; Xie, R.; Gress, C. J.; Blake, S.; Lark, M. A.; Hwang, S.-M.; Tomaszek, T.; Offen, P.; Head, M. S.; Cummings, M. D.; Veber, D. F. *J. Med. Chem.* **2005**, *48*, 6870.
13. Zhou, N. E.; Guo, D.; Thomas, G.; Reddy, A. V. N.; Kaleta, J.; Purisima, E.; Menard, R.; Micetich, R. G.; Singh, R. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 139.
14. Chiva, C.; Barthe, P.; Codina, A.; Gairí, M.; Molina, F.; Granier, C.; Pugnière, M.; Inui, T.; Nishio, H.; Nishiuchi, Y.; Kimura, T.; Sakakibara, S.; Albericio, F.; Giralt, E. *J. Am. Chem. Soc.* **2003**, *125*, 1508.
15. Huang, L.; Ellman, J. A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2993.
16. PubChem: http://pubchem.ncbi.nlm.nih.gov/.
17. PubChem web address for PCMD cathepsin L hits: http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=460.
18. HTS and initial followup (cherry picks) studies were conducted with the following assay buffer: 20 mM sodium acetate, 1 mM EDTA, and 5 mM DTT, pH 5.5. Confirmatory results were obtained utilizing the following assay conditions, replacing DTT with cysteine in the assay buffer: compounds were serially diluted in DMSO and transferred into a 96-well Corning 3686 assay microplate to give 16 dilutions ranging from 50 μM to 1.5 nM. Human liver cathepsin L (Calbiochem 219402) was activated by incubating with assay buffer for 30 min. Assay buffer consisted of 20 mM sodium acetate, 1 mM EDTA, and 5 mM cysteine, pH 5.5. Upon activation, cathepsin L (300 pM) was incubated with 1 μM *Z*-Phe-Arg-AMC substrate and test compound in 100 μL of assay buffer for 1 h at room temperature. Fluorescence of AMC released by enzyme-catalyzed hydrolysis of *Z*-Phe-Arg-AMC was read on a Perkin-Elmer Envision microplate reader (excitation 355 nm, emission 460 nm). Data were scaled using internal controls and fitted to a four-parameter logistic model (IDBS XLfit equation 205) to obtain IC_{50} values in triplicate.
19. Confalone, P. N.; Woodward, R. B. *J. Am. Chem. Soc.* **1983**, *105*, 902.
20. Verardo, G.; Geatti, P.; Lesa, B. *Synthesis* **2005**, 559.
21. For a general procedure for the preparation of α-bromo anilides, see von Geldern, T. W.; Tasker, A. S.; Sorensen, B. K.; Winn, M.; Szczepankiewicz, B. G.; Dixon, D. B.; Chiou, W. J.; Wang, L.; Wessale, J. L.; Adler, A.; Marsh, K. C.; Nguyen, B.; Opgenorth, T. J. *J. Med. Chem.* **1999**, *42*, 3668.
22. Baggio, R.; Shi, Y.-Q.; Wu, Y.-Q.; Abeles, R. H. *Biochemistry* **1996**, *35*, 3351.
23. Thompson, S. K.; Halbert, S. M.; Bossard, M. J.; Tomaszek, T. A.; Levy, M. A.; Zhao, B.; Smith, W. W.; Abdel-Meguid, S. S.; Janson, C. A.; D'Alessio, K. J.; McQueney, M. S.; Amegadzie, B. Y.; Hanning, C. R.; DesJarlais, R. L.; Briand, J.; Sarkar, S. K.; Huddleston, M. J.; Ijames, C. F.; Carr, S. A.; Garnes, K. T.; Shu, A.; Heys, J. R.; Bradbeer, J.; Zembryki, D.; Lee-Rykaczewski, L.; James, I. E.; Lark, M. W.; Drake, F. H.; Gowen, M.; Gleason, J. G.; Veber, D. F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14249.
24. Xing, R.; Hanzlik, R. P. *J. Med. Chem.* **1998**, *41*, 1344.
25. Magrath, J.; Abeles, R. H. *J. Med. Chem.* **1992**, *35*, 4279.
26. For a review on irreversible inhibitors of serine, cysteine, and threonine proteases, see Powers, J. C.; Asglan, J. L.; Ekici, O. D.; James, K. E. *Chem. Rev.* **2002**, *102*, 4639.
27. Chande, M. S.; Singh-Jathar, K. *Indian J. Chem.* **1998**, *37B*, 352.
28. Characterization data for (−)-**11** (PubChem SID 26681509) mp = 131 °C; $[\alpha]_D^{23}$ −14.8 (*c* 0.5, AcOH); IR (KBr) 3412, 3343, 2971, 1680, 1529, 1167 cm^{-1}; ^1H NMR (500 MHz, DMSO-d_6, VT-350K) δ 10.61 (br s, 1H), 10.04, (br s, 1H), 9.88 (br s, 1H), 9.14 (br s, 1H), 7.61 (d, *J* = 7.6 Hz, 1H), 7.43 (d, *J* = 7.4 Hz, 1H), 7.32 (d, *J* = 8.2 Hz, 1H), 7.21 (d, *J* = 7.3 Hz, 1H), 7.16–7.11 (m, 3H), 7.05 (t, *J* = 7.1 Hz, 1H), 6.97 (t, *J* = 7.4 Hz, 1H), 6.47 (br s, 1H), 4.31 (br s, 1H), 3.73 (br s, 2H), 3.17 (dd, *J* = 14.7, 4.1 Hz, 2H), 2.96 (m, 2H), 2.59 (q, *J* = 7.5 Hz, 2H), 1.29 (br s, 9H), 1.13 (t, *J* = 7.5 Hz, 3H); ^{13}C NMR (125 MHz, DMSO-d_6, major rotamer) δ 172.7, 171.6, 166.9, 155.2, 137.3, 136.0, 135.4, 128.5, 127.1, 125.7, 125.5, 125.2, 124.0, 120.9, 118.5, 118.1, 111.3, 109.9, 78.0, 53.6, 33.1, 28.1, 27.7, 23.6, 14.2; high-resolution mass spectrum (ES+) *m/z* 562.2126 [$(M+Na)^+$; Calcd for $C_{27}H_{33}N_5O_5S$-Na: 562.2100]. $[\alpha]_D^{23}$ −14.8.
29. IC_{50} values as means ± standard deviations: (−)-**11**(*S*) 56 nM ± 4 nM; (−)-**12**(*S*) 133 nM ± 2 nM; (+)-**11**(*R*) 34 μM ± 2 μM.
30. Optical rotation for (+)-**11**: $[\alpha]_D^{23}$ +12.8. The enantiomeric purity of both (−)-**11** and (+)-**11** was assayed using an OD-RH chiral column with the following LC parameters: 1.0 mL/min with a linear gradient of 90% water in acetonitrile to 10% water in acetonitrile over 15 min. Using this method, baseline separation was obtained for the enantiomers and retention times for (−)-**11** and (+)-**11** were 14.01 min and 13.02 min, respectively. The synthesis, outlined in Scheme 2, produced both enantiomers in >99% enantiomeric purity.
31. Shah, P. P.; Myers, M. C.; Beavers, M. P.; Purvis, J. E.; Huryn, D. M.; Smith, A. B., III.; Diamond, S. L., in preparation. Q2