Published in final edited form as:
Comb Chem High Throughput Screen. 2010 May ; 13(4): 337–351.

Design, Synthesis and Biological Evaluation of a Library of Thiocarbazates and their Activity as Cysteine Protease Inhibitors

Zhuqing Liu[a,b], **Michael C. Myers**[a,b], **Parag P. Shah**[a,c], **Mary Pat Beavers**[a,c], **Phillip A. Benedetti**[a,b], **Scott L. Diamond**[a,c], **Amos B. Smith III**[a,b], and **Donna M. Huryn**[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

Abstract

Recently, we identified a novel class of potent cathepsin L inhibitors, characterized by a thiocarbazate warhead. Given the potential of these compounds to inhibit other cysteine proteases, we designed and synthesized a library of thiocarbazates containing diversity elements at three positions. Biological characterization of this library for activity against a panel proteases indicated a significant preference for members of the papain family of cysteine proteases over serine, metallo-, and certain classes of cysteine proteases, such as caspases. Several very potent inhibitors of Cathepsin L and S were identified. The SAR data was employed in docking studies in an effort to understand the structural elements required for Cathepsin S inhibition. This study provides the basis for the design of highly potent and selective inhibitors of the papain family of cysteine proteases.

Keywords

thiocarbazates; Cathepsin B; Cathepsin S; Cathepsin L; cysteine protease inhibitor; library

Introduction

Cysteine proteases, ubiquitous in Nature, are frequent targets of drug discovery efforts due to role they play in a number of physiological and pathophysiological processes. In mammals, three classes of cysteine proteases have been characterized: papain-like (such as the cysteinyl cathepsins), calpains, and caspases. i›ii Papain-like cysteine proteases play a role in protein turnover, and their overexpression or disregulation has been implicated in certain inflammatory diseases, in cancer, and in osteoporosis, arthritis, among other diseases. Inappropriate activity of calpains has also been associated with a number of disease conditions including neurodegeneration, muscular dystrophy and diabetes. Caspases play a role in the inflammatory process and in apoptosis; their inhibition has been suggested as an approach to degenerative diseases such as arthritis and stroke. A number of infectious agents (bacteria, viruses, and protozoa) also utilize either host or their own cysteine proteases for infectivity, virulence and/or replication processes, and targeting these proteases has been a popular strategy for anti-infective drug discovery efforts.iii

Correspondence to: Amos B. Smith, III; Donna M. Huryn.

Inhibitors of cysteine proteases typically rely on the presence of a "warhead" to provide a site for nucleophilic attack by the active site cysteine thiolate.[iv] Examples of warheads include peptidyl halomethyl ketones,[v] peptidyl diazomethanes,[vi] peptide aldehydes,[vii] acyloxymethyl ketones,[viii] epoxysuccinyl derivatives,[ix] epoxyketones,[x] α-aminoalkyl epoxides,[xi] α¯keto-aldehydes,[xii] azepanones,[xiii] aziridines and azodicarboxamides,[xiv,xv] vinyl sulfones,[xvi] and aza peptides.[xvii]

Recently, we disclosed a novel series of Cathepsin L inhibitors characterized by a unique thiocarbazate warhead.xviii Through further optimization, potent and selective inhibitors of Cathepsin L (e.g. **1**) were developed and characterized.[xix,xx,xxi]Since this warhead had not been previously described, we explored its potential to inhibit other proteases. Our study entailed the design and synthesis of a library of thiocarbazates and their biological characterization in a series of protease assays to define the potential of this chemotype to exhibit broad protease inhibition. Based on the results of these studies, the library was tested against specific cathepsins; docking studies to rationalize our findings were also carried out. The results from these efforts are reported herein.

Design of the Thiocarbazate Library

Our strategy involved the design of a library containing a thiocarbazate scaffold incorporating a variety of functional groups at three different positions A, B and C (Figure 1). From the outset, we sought to optimize diversity of the final products in terms of size, shape and functionality. At the same time, we focused on maintaining optimal physical properties to ensure solubility, permeability and other "drug-like" properties. Finally, we maintained a strict requirement for an expedient synthesis that would produce a minimum of 10mg of final product in purities of at least 95% as determined by LC/MS analysis. Modifications at the A position involved changes in size, as well as replacement of the *t*-butyloxycarbonyl group. Previous work in our laboratories identified an issue of instability with the presence of the free amine at this position, therefore all modifications took that possibility into consideration. Position B underwent the most extensive modifications, where changes in size, polarity, acidity, and functionality were incorporated. Thiocarbazates derived from natural amino acids, such as methionine, valine, alanine, glutamic acid, leucine, proline, phenylalanine, tyrosine, threnonine, serine, glutamic acid, lysine, arginine and histine, along with unnatural amino acids were prepared. Modifications at C involved incorporation of ring constraints, removal of the amide bond and exploration of size requirements; a variety of acetamides derived from aniline, primary amines, and methyl esters were included. Examples of substituents at position C include differentially substituted anilines, quinolines and isoquinolines, non-aromatic amines, morpholines, indoline, and pyridinone.

To ensure that the library contained compounds with a range of acceptable physical properties such as logP, molecular weight, polar surface area etc, we relied on the cheminformatics tool, Leadscope.[xxii] Table 1 below details this analysis. The aggregated library exhibited excellent properties according to Lipinski-like parameters:xxiii average LogP = 3.5, average H-bond acceptors = 4.8, average H-bond donors = 3.9. The average molecular weight (526) is typical of small molecule protease inhibitors, but slightly above the Lipinski target of <500. Consequently, the number of Lipinski violations on average was 1, which is reflective of the molecular weight characteristics. Polar surface area and rotable bonds were on average 134 A and 16, respectively. While the PSA was within the range suggested by Veber[xxiv] for orally available drugs, the rotatable bond count was somewhat higher than suggested as optimal.

Synthesis of the Thiocarbazate Library

Based on chemistry developed in our laboratory[18,19] we designed a versatile synthetic strategy to prepare the thiocarbazate library as illustrated in Scheme 1. A variety of acids were treated

with ethylchloroformate to form the corresponding mixed anhydride, which were not isolated. Treatment with hydrazine monohydrate then furnished the hydrazides which were isolated after aqueous workup. Reaction with carbonyl sulfide gas in ethanol[xxv] formed the thiosemicarbazide intermediates that were directly treated with an alkylating agent to form the thiocarbazate products. The thiocarbazates were isolated and purified by HPLC to at least 95% purity. All compounds were characterized by high resolution mass spectroscopy, and a subset further characterized by ^{1}H NMR, ^{13}C NMR and IR. The general procedure for the preparation of thiocarbazates was amenable to all of the thiocarbazates produced in the library (Table 2).

While many of the amino acid and acid starting materials were commercially available, several required preparation. Hydrazides derived from commercially available *R* and *S*-2-methyl-3-hydroxypropionate were generated directly from the ester as shown in Scheme 2, then converted to the corresponding thiocarbazates (e.g., **67–70**). Subsequent etherification of **69** afforded the butyldimethylsilyl (TBS),[xxvi] and *para*-methoxyl benzyl (PMB)xxvii ether analogs **71** and **72**. Beta amino acids, such as those incorporated into thiocarbazates **73–81**, were prepared using a modified Arndt-Eistert protocol to furnish the desired β-amino acid in yields of 70–95%.xxviii

Characterization of Thiocarbazate Library

As neither thiocarbazates nor their activity as protease inhibitors have been described previously, we profiled a subset of twenty-two compounds at a concentration of 10 μM for inhibitory activity against 75 different proteases.[xxix] (Figure 2). The proteases chosen covered a broad spectrum of classes including serine proteases, metallo- proteases, aspartyl proteases and cysteine proteases. The aim of this study was to determine quickly whether thiocarbazates as a class displayed selectivity towards different families of proteases, or displayed broad protease inhibition properties.

The heatmap illustrated in Figure 1 allows us to draw several broad conclusions. First, the thiocarbazates as a class exhibit selectivity towards cysteine proteases. No significant activity was detected against any other protease family member (e.g. serine, aspartyl, metallo-). Second, even among the cysteine protease family, a preference for the papain family is observed, as only modest activity against representatives of the calpain or caspase families is observed. Finally, several thiocarbazates exhibit potent activity (>95% inhibition at 10 μM) against Cathepsins L, S, V, K and papain.

Based on this study, we focused further attention on a more thorough biological characterization of this library against Cathepsins B, L and S. This choice was based on the potential for potent inhibition, as well as selectivity. In addition, all three cathepsins represent important targets for drug discovery efforts. Towards this end, IC_{50}'s were generated for all 82 members of the thiocarbazate library (Table 1) against Cathepsin B, L and S.

Based on this data, we offer several generalizations for broad cathepsin inhibition. First, the substitution patterns displayed at positions A and B are key determinants for inhibitory activity. Alpha amino acid derived thiocarbazates are the preferred substituents, while those prepared from beta amino acid derivatives (**74–81**) and acid precursors (**67–73**) are devoid of activity. Among the thiocarbazates that incorporate alpha amino acids, the size of substituent B had a profound effect on potency. Thiocarbazates containing large groups (such as 3-indolemethylene, benzyl, 4-benzyloxybenzyl) or medium sized substituents (such as isopropyl, methyl thioethyl) were more potent than those incorporating smaller groups [e.g., hydrogen (**64**), methyl (**25**).] Incorporation of the specific amino acids proline (**32, 33, 80, 81**), histadine (**65, 66**) or glutamic acid (**27, 28, 59, 60, 61**) proved detrimental to activity. Stereochemistry preferences at position B were also explored through the preparation a key enantiomeric pairs such as **4** and **16**; **46** and **48**; and **47** and **49**. In those examples, the thiocarbazate derived from

the L-amino acid was more potent than that derived from the D isomer. Structural requirements for activity at position C involved a strong preference for a carbonyl group (**6** vs. **4**; **30** vs. **29**; **35** vs. **34**). In all cases where the direct comparison could be made, removal of the carbonyl group significantly diminishes the inhibitory activity against all three cathepsins. In contrast, the specific nature of the amine side chain influenced activity far less.

Cathepsin B

Inhibitors of Cathepsin B have been proposed for the treatment of cancer, osteoporosis, arthritis, and viruses.[2] As such, there is an interest in identifying potent and selective inhibitors of this member of the papain family. As a class, a number of thiocarbazates effectively inhibit Cathepsin B with potencies in the range of 500nM to 10 μM. The most potent thiocarbazates, with IC_{50}'s in the sub-micromolar range, incorporate leucine in the A/B area (e.g. **21, 22, 23**) and a variety of substituents at C. Within this series, however, Cathepsin B activity often paralleled that of Cathepsin L, with some evidence of modest selectivity against Cathepsin S.

Cathepsin L

The role of Cathepsin L in bone and cartilage remodeling, as well as the infectivity of certain infectious organisms, has focused interest on identifying inhibitors of this protease. A number of very potent (<100nM) inhibitors of Cathepsin L were identified among the thiocarbazate library. The most potent examples (e.g. **4, 5, 7, 8, 9, 12, 15, 39**) contained an aromatic group at position B, with a preference for 3-indolemethyl, over benzyl and substituted benzyl analogs. In contrast, there is a tolerance for a broad array of C substituents with anilides (**4, 8, 9**), constrained anilides (**5**), alkyl amides (**7, 12, 39**) and esters (**15**) exhibiting potent inhibition. In many cases, at least 10-fold selectivity was observed over the Cathepsin S, and even higher selectivity vs. Cathepsin B.

Cathepsin S

Inhibitors of Cathepsin S have garnered significant attention as potential treatment for several autoimmune diseases such as psoriasis, rheumatoid arthritis, multiple sclerosis and asthma. Among the thiocarbazate library, a number of sub-micromolar inhibitors were identified. Unlike Cathepsin B and L, Cathepsin S shows a broad tolerance for the nature of the substituent at B. Potent inhibition (<100nM) was observed among thiocarbazates containing aromatic (**7, 12, 34**), alkyl (**24**), thio ethers (**17, 18, 19, 20**), and benzyl ethers (**51, 52, 53**). Similar to the Cathepsin S SAR, tolerances for diverse size and functionalities at position C were observed. The most potent Cathepsin S inhibitors typically were even more potent Cathepsin L inhibitors (e.g. **7, 9, 12**). However certain trends suggest that selectivity over Cathepsin L is possible. For example, several alkyl thioether analogs (e.g. **17, 18**) exhibit approximately 2-fold selectivity over Cathepsin L, and even greater selectivity over Cathepsin B. Benzyl ether analogs such as **44, 45, 48** and **49**, albeit active in the low micromolar range, exhibit selectivity for Cathepsin S.

Cathepsin S Docking Studies

The publicly available x-ray crystal structure of Cathepsin S complexed to a small molecule inhibitor (2hh5.pdb) was selected to predict the protein/ligand binding interactions of thiocarbazates **69** and **20**. Both inhibitors are presumed to bind within the catalytic triad binding site of Cathepsin S. These two compounds were selected for docking studies given their differences in inhibitory activity: **69** has an IC_{50} of only 25μM, whereas **20** has an IC_{50} of 310nM. Each compound was docked independently in the binding site of Cathepsin S (Figure 2).[xxx] As expected, a significant difference between the total energy scores for **69** and **20** was calculated. The total energy score for the potent inhibitor (**20**) was −6.13 kcal/mol, while the

score for the weakly active **69** was only −1.54 kcal/mol. It is presumed that each of the two compounds binds covalently to the active site Cys25 thiolate in the protein with the electrophilic carbon adjacent to the sulfur in the thiocarbazates. The distance between the Cys25 sulfur and the electrophilic carbon of the inhibitors is in the range of 2.9–3.3 angstrom for both compounds, further supporting covalent binding between the inhibitor and the protein.

Examination of the non-covalent interactions between **20** and the catalytic binding site residues of cathepsin S reveals strong hydrophobic interactions within the S2 subsite, specifically between Phe 211 of Cathepsin S and the indole of **20**. In the S1' subsite, there is also a strong hydrophobic pocket formed by the interaction of the O-*tert*-butyl group of **20** and Trp 186. Non-covalent binding interactions of the weakly active inhibitor were also noted. However, docking studies reveal that **69** does not provide anyside-chain hydrophobic group for interaction within the S2 subsite (Figure 2), as seen with **20**. This observation could explain the lack of potency of **69**. The overlay of **69** and **20** in the binding site of cathepsin S illustrates this void in the S1' pocket around Trp 186 for **69**.

Conclusions

Based on our earlier finding that thiocarbazates can act as inhibitors of proteases, we designed and synthesized a library with structural modifications at three different positions. The library was designed to incorporate diverse size, shapes and functionalities in order to profile thiocarbazates as a class of inhibitors in a series of representative proteases. Preliminary profiling indicated that thiocarbazates are selective for the papain-like cystein protease; exhibiting limited or no activity against serine, aspartyl, metallo-proteases, nor members of the calpain or caspase families of cysteine proteases. Complete IC_{50} characterization against Cathepsins B, L and S identified a number of potent and selective inhibitors of Cathepsin L. In addition, trends to direct the design of new thiocarbazates exhibiting potent and selective inhibitory activity against Cathepsins B and S were also developed. Further characterization of the Cathepsin L inhibitors disclosed, as well as the development of improved inhibitors of Cathepsin B and L will be reported in due course.

Experimental Section

General procedure for the synthesis of thiocarbazates

To a solution of acid (10.0 mmol) in THF (15 mL) was added NEt_3 (1.45 mL). After stirring at room temperature for 5 min, the mixture was cooled to −10 °C followed by the addition of the solution of ethyl chloroformate (1.0 mL) in THF (15 ml). After stirring 10 min, the mixture was filtrated and washed with THF (2 × 5 mL). The combined filtrated was slowly added to a solution of hydrazine (2 mL) in methanol (26 mL) at 0 °C. The mixture was allowed to warm to room temperature and was concentrated after stirring 1 h. The residue was diluted with ethyl acetate (80 mL) and washed successfully with saturated $NaHCO_3$ (12 mL) and brine (12 mL). The organic layer was dried over Na_2SO_4 and filtrated. After the removal of volatile solvents, the residue was used without further purification.

The hydrazide (1.0 mmol, 1.0 equiv.) made above was dissolved into a solution of KOH in 95% EtOH (0.25 M, 4.4 mL, 1.1 equiv.) in a 25 mL one-necked round bottom flask at room temperature. After stirring for 5 min, a balloon of carbonyl sulfide gas was attached to the flask. The flask was purged with the gas (5 s) and a full balloon was reattached. The reaction mixture was stirred for 15 h at 23 °C, followed by the addition of a bromide (1.1 mmol, 1.1 equiv.). The reaction was monitored by LC-MS, and the alkylating agent was typically consumed within 20 to 60 min. The reaction mixture was filtered and the filtrate was concentrated *in vacuo*. The residue was purified by preparative reverse phase HPLC and identified by HRMS.

Using this procedures compounds **4–15** and **73** were prepared from L-N-Boc trytophan; **16** from D-N-Boc- tryptophan; **17–20** and **76–77** from N-Boc methionine; **21–24** from N-Boc valine; **25–26** and **78–79** from N-Boc alanine; **27–28** from N-Boc- Asp(OBzl); **29–31** from N-Boc-leucine;**32–33** and **80–81**from N-Boc-proline; **34** from N-Boc-phenylalanine; **36–43** from N-Boc-Try(Bzl); **44–45** from N-Boc-Thr(Bzl); **46–47**, **50–56** and **74–75** from N-Boc-serine; **48–49** from D-N-Boc-serine; **57–58** from N-Boc-Lys(Z); **59–61** from N-Boc-Glu(Xan); **62–63** from N-Boc-Arg(NO2); **64** from N-Boc-glycine; **65–66** from N-Boc-histine; **62–72** from 2-methyl-3-hydroxypropionate.

Synthesis of thiocarbazates derived from 2-methyl-3-hydroxypropionate

Synthesis of 67–70—To a solution of 2-methyl-3-hydroxypropionate (1 mL, 7.94 mmol) in THF (10 mL) was added hydrazine hydate (1.16 mL, 11.91 mmol). The mixture was stirred at room temperature overnight and concentrated to give hydrazide (820 mg, 88%). The hydroxyl thiocarbazates **67–70** were obtained following the procedure described above.

Synthesis of 71—To the solution of hydroxyl thiocarbazate (34 mg, 0.1 mmol) in CH_2Cl_2 (8 mL) was added imidazole (13.6 mg, 0.20 mmol) and TBSCl (17 mg, 0.11 mmol) at 0° C. The resulting mixture was warm up and stirred at room temperature for 48h. The mixture was washed with 1N HCl, saturated $NaHCO_3$ and brine solution and extracted with CH_2Cl_2. The combined organic layer was dried and concentrated. The residue was purified by HPLC to give TBS protected thiocarbazate (22mg, 50%).

Synthesis of 72—To the solution of hydroxyl thiocarbazate (34 mg, 0.1 mmol) in CH_2Cl_2 (7 mL) was added 2-(4-methoxylbenzyoxy)-3-nitropyridine (39 mg, 0.15 mmol) and CSA (5 mg). The resulting mixture was stirred at room temperature for 3 h. The mixture was washed with saturated $NaHCO_3$ and extracted with CH_2Cl_2. The combined organic layer was dried and concentrated. The residue was purified by HPLC to give PMB protected thiocarbazate (7 mg, 16%).

NMR of representative thiocarbazates

Compound 1: ^{1}H NMR (500 MHz, DMSO-d_6, VT-350K) δ10.61 (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).

Compound 5: ^{1}H NMR (500 MHz, DMSO-d_6,) δ 9.79–10.80 (m, 3H), 7.67 (m, 1H), 7.46 (m, 1H), 7.33 (m, 1H), 6.93–7.18 (m, 5H), 6.78 & 6.31–6.40 (br s, 1H), 4.23 (br s, 1H), 4.04 (br, s, 1 H), 3.87 (br s, 1H), 3.70 (br s, 2H), 3.12 (m, 1H), 2.90 (m, 1H), 2.69 (m, 2H), 1.88 (m, 2H), 1.29 & 1.13 (br s, 9H).

Compound 14: ^{1}H NMR (500 MHz, DMSO-d_6,) δ 9.93–10.87 (m, 3H), 7.65 (m, 2H), 7.31 (m, 2H), 7.13 (m, 2H), 7.04 (m, 2H), 6.87(m, 1H), 4.28 (m, 1H), 3.86 (m, 2H), 3.12 (m, 1H), 2.90 (m, 1H), 2.69 (m, 2H), 2.10 (s, 3 H), 2.15-1.81 (m, 2H), 1.28 & 1.14 (br s, 9H).

Compound 19: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.94 (br s, 1H), 7.21-7.11 (m, 5H), 5.33 (br s, 1H), 4.72 (m, 2H), 4.42 (d. *J* = 5.4 Hz, 1H), 3.93 (br s, 1H), 3.91 (br s, 1H), 3.82 (t, *J* = 5.8 Hz, 1H), 3.75 (t, *J* = 5.9 Hz, 1H), 2.94 (t. *J* = 5.8 Hz, 1H), 2.85 (t, *J* = 5.9 Hz, 1H), 2.63-2.58 (m, 2H), 2.12 (br s, 1H), 1.96 (ddd, *J* = 14.2, 7.1, 7.1 Hz, 1H), 1.43 (s, 9H).

Compound 24: ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.82 (br s, 1H), 8.35 (br s, 1H), 7.21(s, 1H), 7.18(d, *J* = 8.3 Hz, 2H), 6.96 (d, *J* 8.3 Hz, 2H), 5.15 (br s, 1H), 4.07-4.01 (m, 1H), 3.62 (br s,

2H), 2.70 (s, 2H), 2.69 (s, 2H), 2.05 (br s, 1H), 1.75 (t, *J* = 3.0 Hz, 4H), 1.41 (s, 9H), 1.00 (d, *J* = 6.7 Hz, 3H), 0.97 (d. *J* = 6.7 Hz, 3H).

Compound 37: **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.41 (d, *J* = 6.9 Hz, 2H), 7.38-7.35 (m, 2H), 7.32-7.30 (m, 1H), 7.20-7.17 (m, 2H), 7.16-7.11 (m, 4H), 6.89 (dd, *J* = 8.6, 3.2 Hz, 2H), 5.06 (br s, 1H), 5.01 (s, 2H), 4.70 (d, *J* = 14.5 Hz, 2H), 4.44 (br s, 1H), 3.90 (d, *J* = 9.5 Hz, 2H), 3.81 (t, *J* = 6.0 Hz, 1H), 3.73 (t, *J* = 6.0 Hz, 1H), 3.10 (d, *J* = 14.1, 6.3 Hz, 1H), 2.97 (t, *J* = 6.9 Hz, 1H), 2.93 (t, *J* = 5.8 Hz, 1H), 2.84 (t, *J* = 5.8 Hz, 1H), 1.37 & 1.37 (s, 9H).

Compound 38: **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.67 (d, *J* 8.6 Hz, 2H), 8.61 (br s, 1H), 8.15 (d, *J* = 8.4 Hz, 2H), 7.61 (t, *J* = 8.1 Hz, 1H), 7.40-7.36 (m, 4H), 7.32-7.29(m, 1H), 7.16 (t, *J* = 7.9 Hz, 1H), 7.12 (d, *J* = 8.4 Hz, 2H), 6.88 (d, *J* = 8.4 Hz, 2H), 5.15 (br s, 1H), 4.99 (s, 2H), 4.45 (br, d, *J* = 6.0 Hz, 1H), 3.77 (br s, 2H), 3.11-3.08 (m, 1H), 2.97 (br s, 1H), 1.37 (s, 9H).

Compound 72: **^{1}H NMR** (500 MHz, DMSO-d_6,) δ 9.37–10.16 (m, 3H), 7.40 (d, *J* = 7.5 Hz, 1H), 7.11 (m, 5H), 6.89 (d, *J* = 8.0 Hz, 2H), 4.48 & 4.38 (br s, 2H), 3.80 (m, 1H), 3.73 (s, 3H), 3.69 (s, 1H), 3.56 (m, 2H), 2.58 (m, 3H), 1.11 (t, *J* = 7.5 Hz, 3H), 1.03 (m, 3H).

Compound 85: **^{1}H NMR** (500 MHz, $CDCl_3$) δ 7.47-7.27 (m, 5H), 7.20-7.05 (m, 4H), 5.37 (br s, 1H), 4.52 (s, 2H), 4.14-4.11 (m, 1H), 3.92 (br s, 2H), 3.80 (t, *J* = 6.4 Hz, 2H), 3.63-3.56 (m, 1H), 3.54-3.51(m, 1H), 2.74-2.58 (m, 2H), 2.66-2.53(m, 2H), 1.99-1.91 (m ,2H), 1.42 (s, 9H).

IC_{50} Determinations

IC_{50} determinations were conducted with the following assay buffer: 20 mM sodium acetate, 1 mM EDTA, and 5 mM cysteine, pH 5.5. Compounds were serially diluted in DMSO and transferred into a 96-well Corning 3686 assay microplate to give a 16-point two-fold serial dilution dose response ranging from 25 microM to 760 pM. Human liver cathepsin L (Calbiochem 219402) was activated by incubating with assay buffer for 30 min. Upon activation, cathepsin L (300 pM, 8.7 ng/mL) was incubated with 1 microM Z-Phe-Arg-AMC substrate (Sigma C9521) and test compound in 100 mL 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 PerkinElmer 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.

Cathespin Assays

Human spleen cathepsin S (Calbiochem 219344, 40 ng/mL) was assayed using 15 microM Z-Phe-Arg-AMC substrate. Human liver cathepsin B (Calbiochem 219362, 65 ng/mL) was assayed using 15 microM Z-Arg-Arg-AMC substrate (Bachem I-1135). All reactions were performed in 20 mM sodium acetate buffer containing 1 mM EDTA and 5 mM cysteine, pH 5.5.

Acknowledgments

Financial support for this work was provided by the NIH (5U54HG003915). 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

i. Schirmeister T, Kaeppler U. Mini. Rev. Med. Chem 2003;3:361. [PubMed: 12678829]

ii. Vasiljevera O, Reinheckel T, Peters C, Turk D, Turk V, Turk B. Curr. Pharm. Des 2007;13:385.

iii. McKerrow JH, Engel JC, Caffrey CR. Bioorg Med. Chem 1999;7:639. [PubMed: 10353643]

iv. Hernandez AH, Roush WR. Curr. Opin. Chem. Biol 2002;6:459. [PubMed: 12133721]

v. a) Schoellmann G, Shaw E. Biochemistry 1963;2:252. [PubMed: 13992248] b) Rasnick D. Anal. Biochem 1985;149:461. [PubMed: 4073501] c) Rauber P, Angliker H, Walker B, Shaw E. Biochem. J 1986;239:633. [PubMed: 3827817]

vi. Crawford C, Mason RW, Wikstrom P, Crawford ES. Biochem. J 1988;253:751. [PubMed: 2845932]

vii. Yasuma T, Oi S, Choh N, Nomura T, Furuyama N, Nishimura A, Fujisawa Y, Sohda T. J. Med. Chem 1998;41:4301. [PubMed: 9784105]

viii. Smith RA, Copp LJ, Coles PJ, Pauls HW, Robinson VJ, Spencer RW, Heard SB, Krantz A. J. Am. Chem. Soc 1988;110:4429.

ix. Fujishima A, Imaia Y, Nomura T, Fujisawa Y, Yamamoto Y, Sugawara T. FEBS Lett 1997;407:47. [PubMed: 9141479]

x. Spaltenstein A, Leban JJ, Huang JJ, Reinhardt KR, Viveros OH, Sigafoos J, Crouch R. Tetrahedron Lett 1996;37:1343.

xi. Albeck A, Persky R, Kliper S. Bioorg. Med. Chem. Lett 1995;5:1767.

xii. Lynas JF, Hawthorne SJ, Walker B. Bioorg. Med. Chem. Lett 2000;10:1771. [PubMed: 10937745]

xiii. Marquis RW, James I, Zeng J, Trout REL, Thompson S, Rahman A, Yamashita D, Cie R, Ru Y, Gress CJ, Blake S, Lark MA, Hwang S, Tomaszek T, Offen P, Head MS, Cummings MD, Veber DV. J. Med. Chem 2005;48:6870. [PubMed: 16250645]

xiv. Schirmeister T. J. Med. Chem 1999;42:560. [PubMed: 10052963]

xv. Radim V, Matthias B, Christoph G, Nikolaus S, Josef S, Werner S, Franziska S, Mladenovic M, Bernd E, Matthias L, Knut B, Tanja S. Chem. Med. Chem 2006;1:1126. [PubMed: 16933358]

xvi. Palmer JT, Rasnick D, Klaus JL, Brömme D. J. Med. Chem 1995;38:3193. [PubMed: 7650671]

xvii. Xing R, Hanzlik RP. J. Med. Chem 1998;41:1344. [PubMed: 9548822]

xviii. Myers MC, Shah PP, Diamond SL, Huryn DM, Smith AB III. Bioorg. Med. Chem. Lett 2008;18:210. [PubMed: 18060772]

xix. Myers MC, Shah PP, Beavers MP, Napper AD, Diamond SL, Smith AB III, Huryn DM. Bioorg. Med. Chem. Lett 2008;18:3646. [PubMed: 18499453]

xx. Shah PP, Myers MC, Beavers MP, Purvis JE, Jing H, Grieser HJ, Sharlow ER, Napper AD, Huryn DM, Cooperman BS, Smith AB III, Diamond SL. Mol. Pharmacol 2008;74:3441.

xxi. Beavers MP, Myers MC, Shah PP, Purvis JE, Diamond SL, Copperman BS, Huryn DM, Smith AB III. J. Chem. Info & Mod 2008;48:1464.

xxii. Leadscope, Inc.

xxiii. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Adv. Drug Del Rev 2001;46:3.

xxiv. Veber DF, Johnson SR, Cheng H-Y, Smith BR, Ward KW, Kopple KD. J. Med. Chem 2002;45:2615. [PubMed: 12036371]

xxv. Chande MS, Singh-Jathar K. Indian J. Chem 1998;37B:352.

xxvi. Patrick P, Walailak S, Filisaty V. Org. Biomol. Chem 2004;2:2220. [PubMed: 15280959]

xxvii. Hayashi S, Hirano K, Yorimitsu H, Oshima K. J. Am. Chem. Soc 2007;129:12650. [PubMed: 17902673]

xxviii. Linder MR, Steurer S, Podlech J. Org. Syn 2002;79:154.

xxix. www.reactionbiology.com

xxx. Tully DC, Liu H, Chatterjee AK, Alper PB, Epple R, Williams JA, Roberts MJ, Woodmansee DH, Masick BT, Tumanut C, Li J, Spraggon G, Hornsby M, Lesley SA, Harris JL, Karnewsky DS. Bioorg. Med. Chem. Lett 2006;16:5112. [PubMed: 16876402]

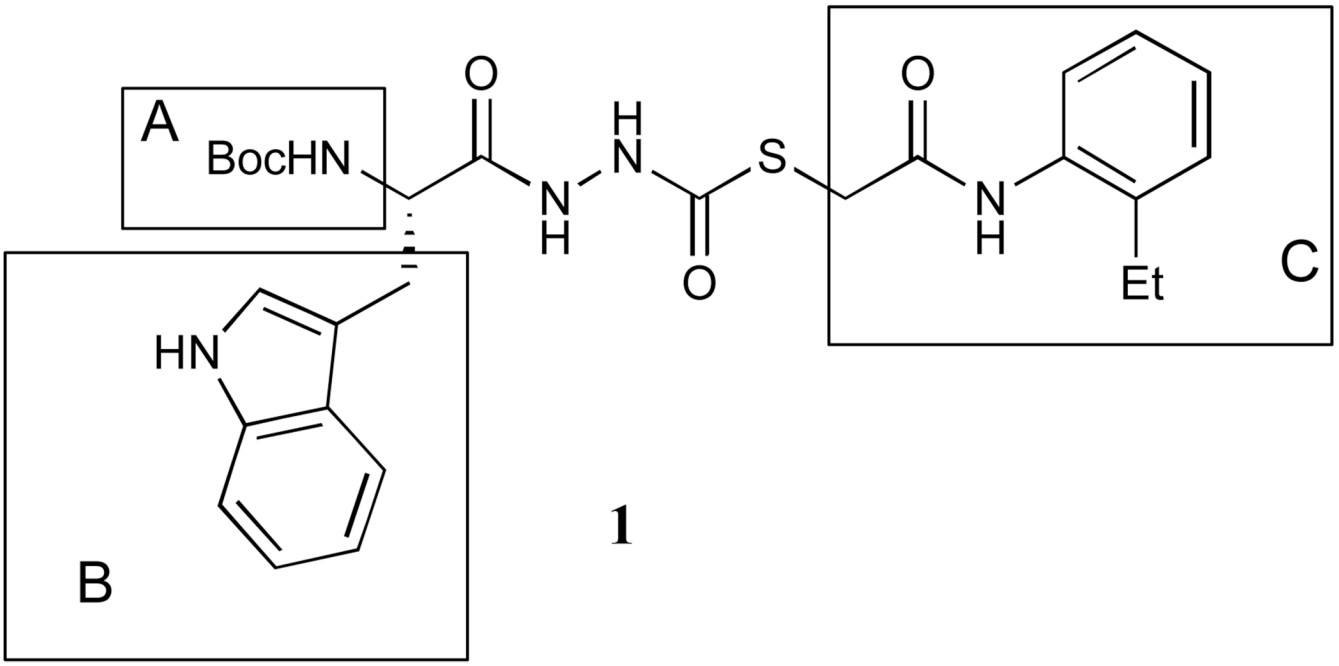


Figure 1.
Design of thiocarbazate library

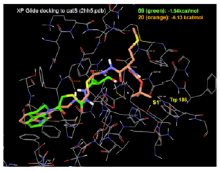


Figure 2.
Docking poses of 20 and 69 into Cathepsin S Model.

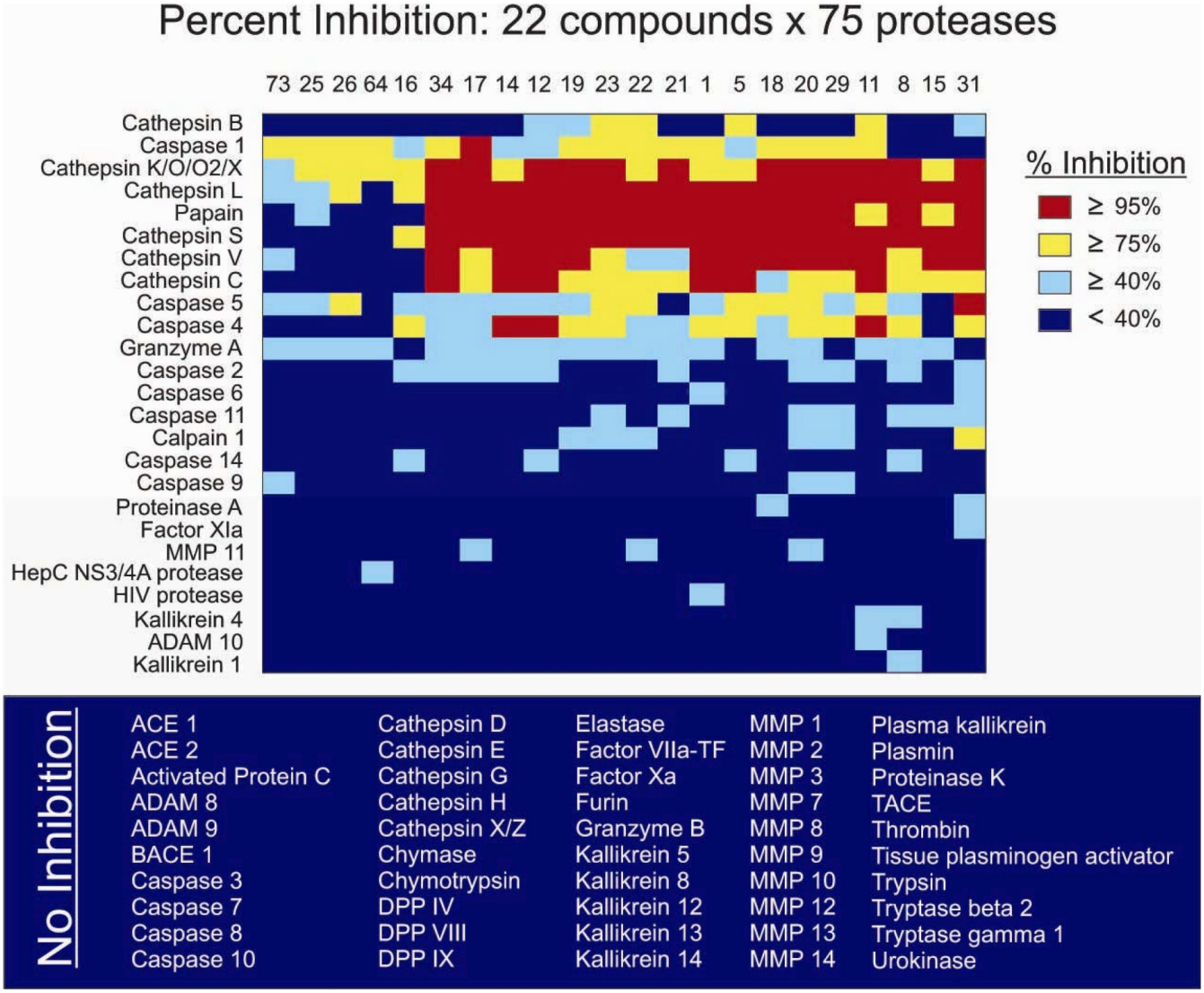
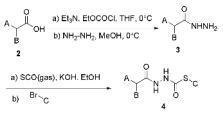
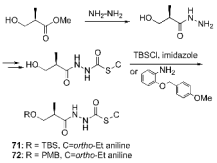


Figure 1.
Protease profiling heatmap of 22 thiocarbazates at 10μM against 75 proteases.



Scheme 1.
General synthesis of thiocarbazates



Scheme 2.
Preparation of thiocarbazates **68–72**

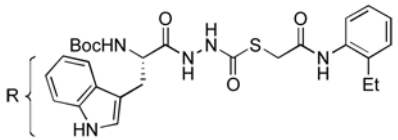
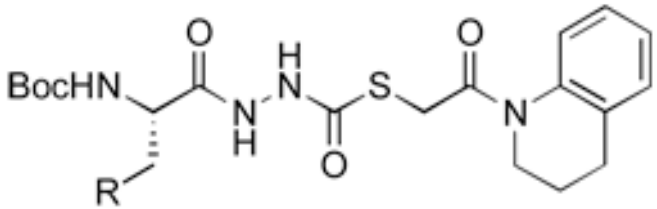
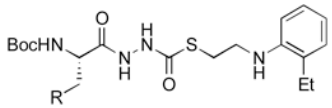
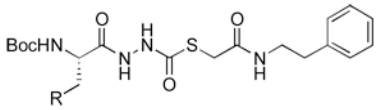
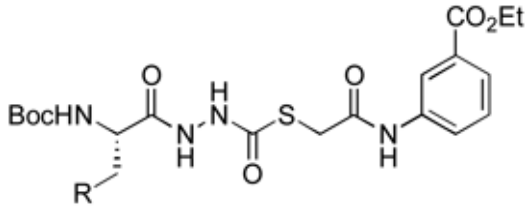
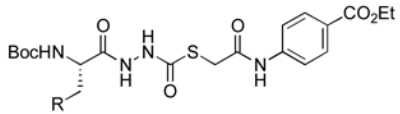
Table 1

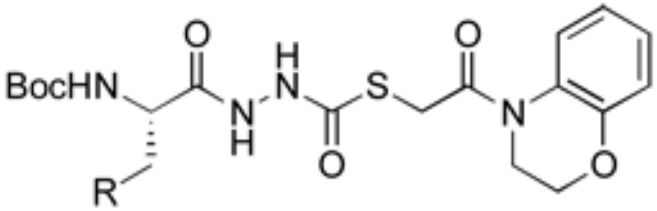
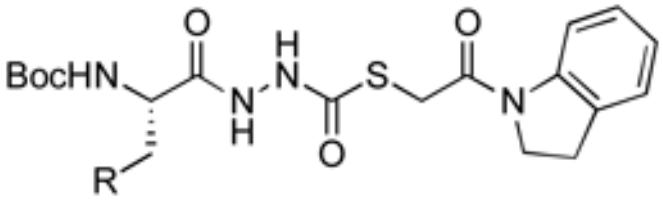
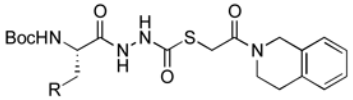
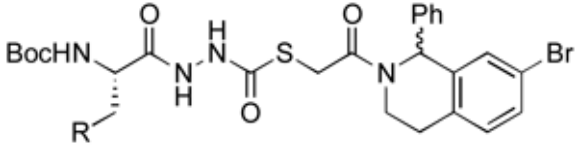
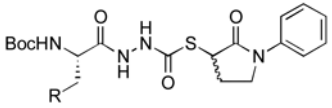
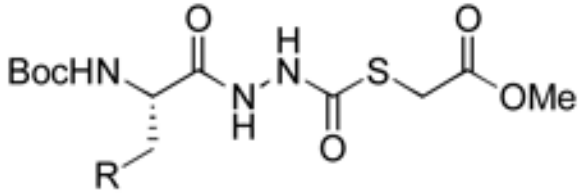
Physical properties of thiocarbazate library.

	Average	Range	Median	Mode
AlogP	3.5	1.0–6.0	3.5	4.1
H-Bond Acceptors	4.8	3–7	5.0	5.0
H-Bond Donors	3.9	2–7	4.0	4.0
Lipinski Violations	1.0	0–3	1.0	1.0
Molecular Weight	526.1	367.2–706.6	536.6	539.6
Polar Surface Area	134.3	95.5–210.8	132.6	125.6
Rotatable Bonds	16	9–23	16	16

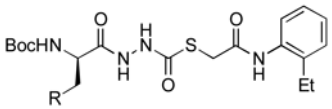
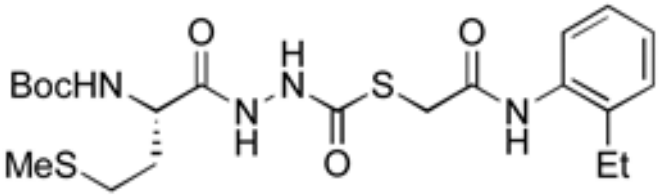
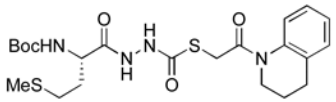
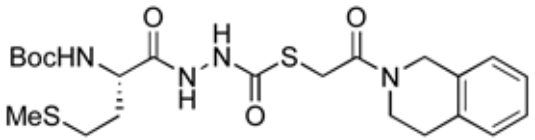
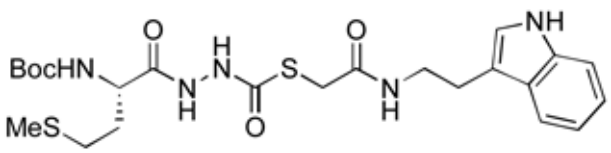
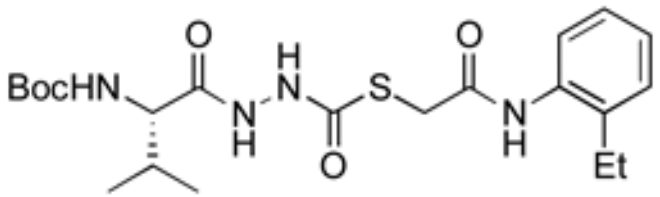
Table 2

Thiocarbazates synthesized and assayed against cathepsins B, L, and S

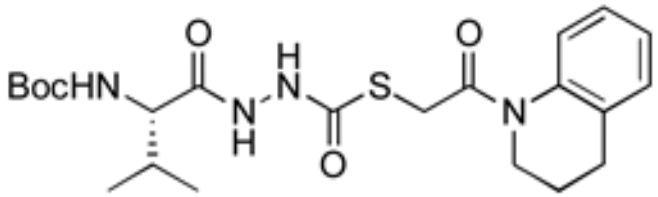
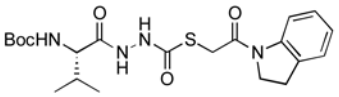
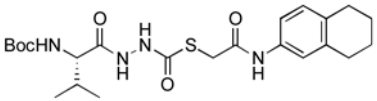
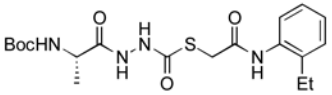
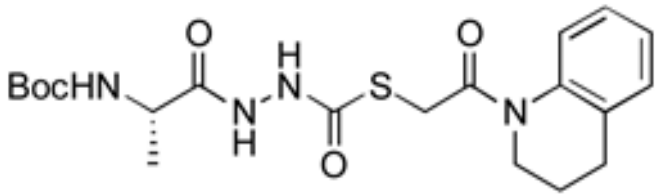
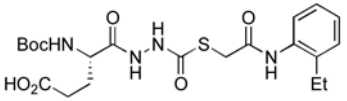
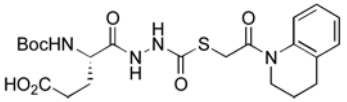
Thiocarbazate	Structure	HRMS	IC_{50} (mM) Cat B	Cat L	Cat S
1		[M+Na]+562.2126	4.45	0.056	0.96
5		[M+Na]+574.2102	4.46	0.096	0.93
6		[M+H]+487.2383	>25.0	>25.0	>25.0
7		[M+Na]+562.2094	6.59	0.029	0.32
8		[M+Na]+606.1995	>25.0	0.062	0.68
9		[M+Na]+606.1995	>25.0	0.057	0.58

Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
10		[M+Na]+576.1874	3.23	0.22	0.95
11		[M+Na]+590.1843	2.60	0.11	1.65
12		[M+Na]+574.2109	3.25	0.019	0.53
13		[M+Na]+728.1520	>25.0	1.03	2.90
14		[M+Na]+560.1924	14.04	1.05	2.07
15		[M+Na]+473.1458	6.33	0.098	0.57

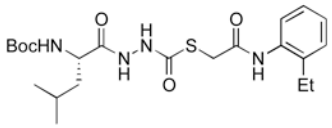
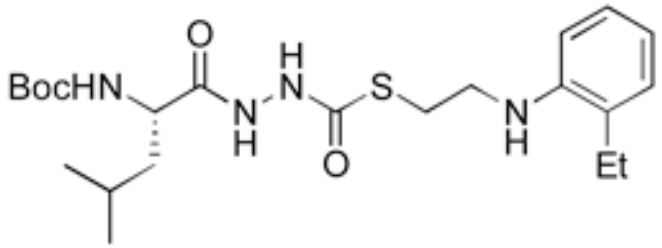
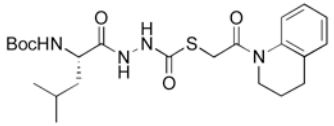
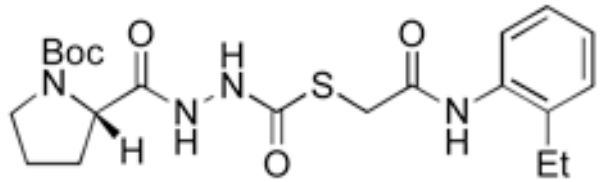
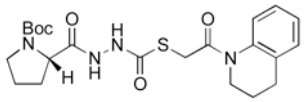
Comb Chem High Throughput Screen. Author manuscript; available in PMC 2011 May 1.

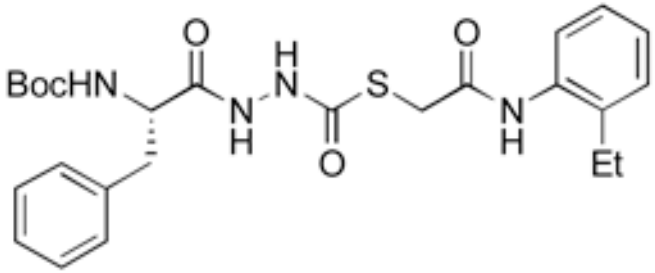
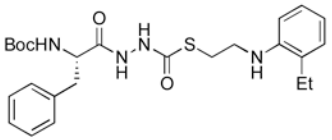
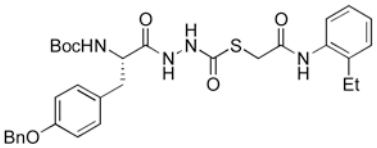
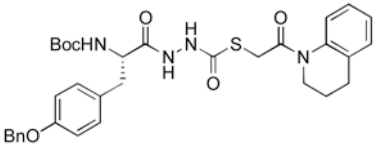
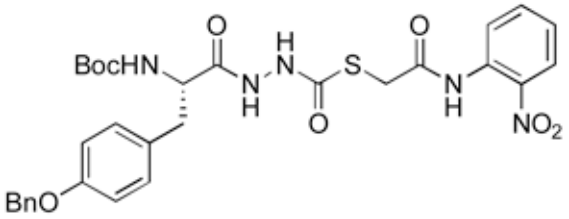
Thiocarbazate	Structure	HRMS	IC_{50} (mM) Cat B	Cat L	Cat S
16		[M+Na]+562.2126	>25.0	>25.0	>25.0
17		[M+Na]+507.1713	2.03	1.29	0.69
18		[M+Na]+519.1699	2.21	1.34	0.63
19		[M+Na]+519.1709	2.31	0.52	0.41
20		[M+Na]+546.1807	2.42	0.33	0.31
21		[M+Na]+475.2003	0.87	0.66	3.74

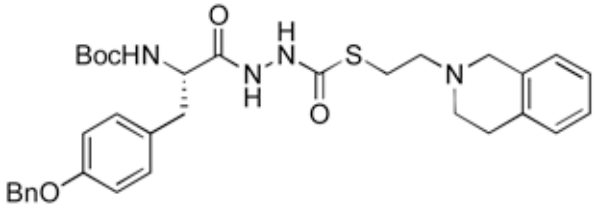
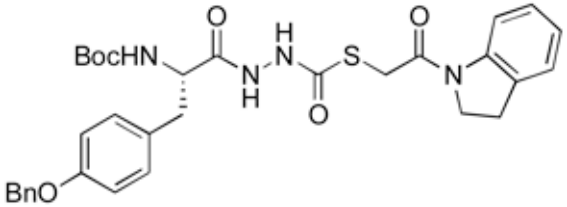
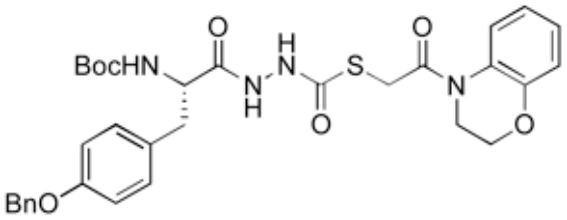
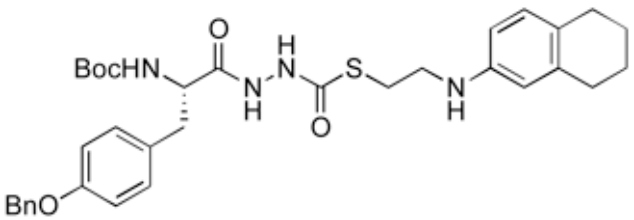
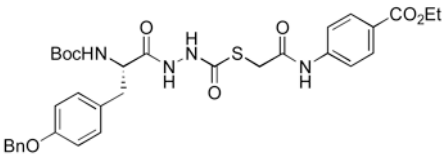
Comb Chem High Throughput Screen. Author manuscript; available in PMC 2011 May 1.

Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
22		[M+Na]+487.1994	0.63	0.68	4.24
23		[M+Na]+473.1838	0.44	0.45	7.24
24		[M+Na]+501.2155	>25.0	0.11	0.62
25		[M+Na]+447.1685	>25.0	>25.0	>25.0
26		[M+Na]+459.1674	>25.0	>25.0	>25.0
27		[M+Na]+505.1733	>25.0	>25.0	>25.0
28		[M+Na]+517.1725	>25.0	>25.0	>25.0

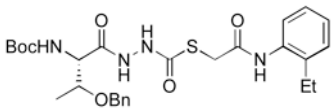
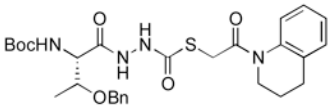
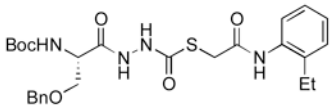
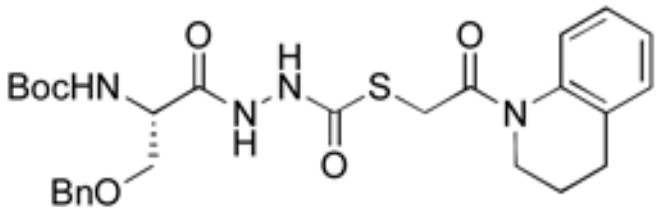
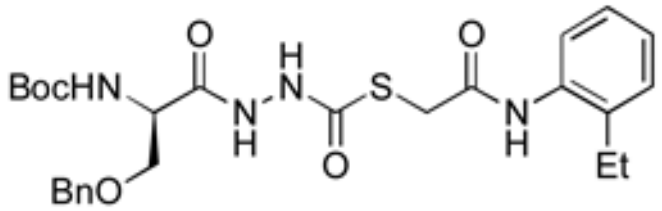
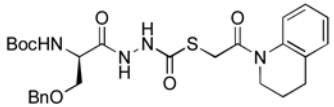
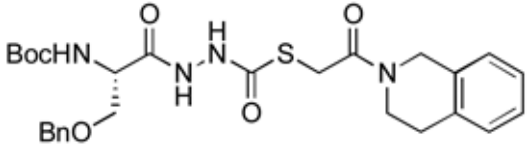
Comb Chem High Throughput Screen. Author manuscript; available in PMC 2011 May 1.

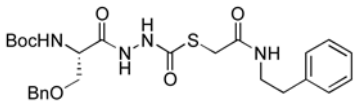
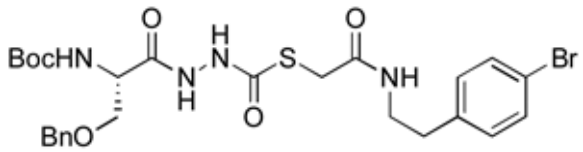
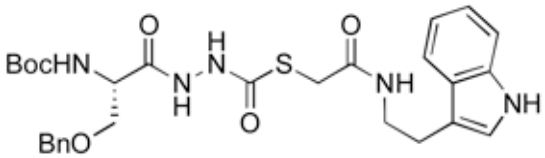
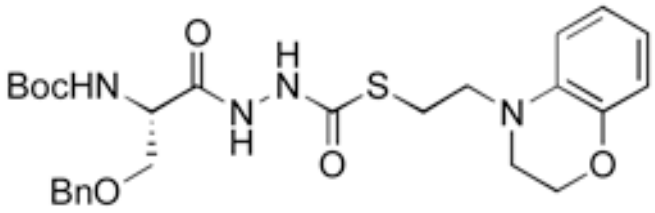
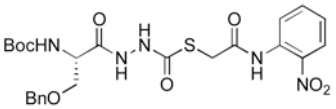
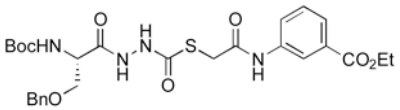
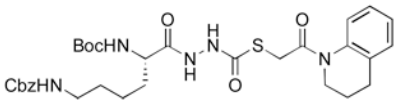
Thiocarbazate	Structure	HRMS	IC_{50} (mM) Cat B	Cat L	Cat S
29		[M+H]+467.2318	2.63	1.09	1.65
30		[M+H]+453.2516	>25.0	>25.0	>25.0
31		[M+H]+479.2327	2.72	1.64	1.64
32		[M+Na]+473.1819	>25.0	>25.0	>25.0
33		[M+H]+463.2014	>25.0	>25.0	>25.0

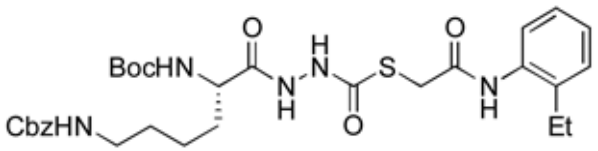
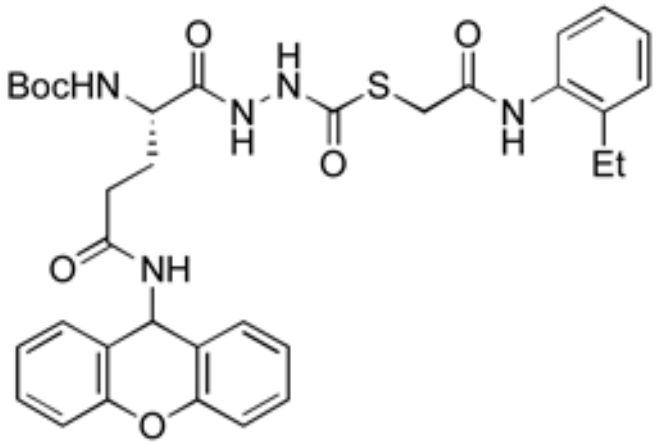
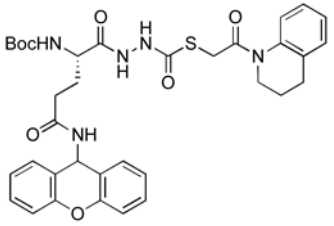
Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
34		[M+Na]+523.1977	3.06	0.12	0.93
35		[M+H]+487.2383	>25.0	>25.0	>25.0
36		[M+Na]+629.2419	5.08	0.83	2.83
37		[M+Na]+619.2906	3.94	0.63	3.34
38		[M+Na]+646.1946	2.77	0.59	2.08

Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
39		[M+Na]+641.2401	3.19	0.086	1.31
40		[M+Na]+627.2245	3.12	0.70	8.00
41		[M+H]+643.2224	3.03	1.56	3.98
42		[M+H]+655.2554	>25.0	0.50	1.90
43		[M+H]+673.2322	>25.0	0.14	1.08

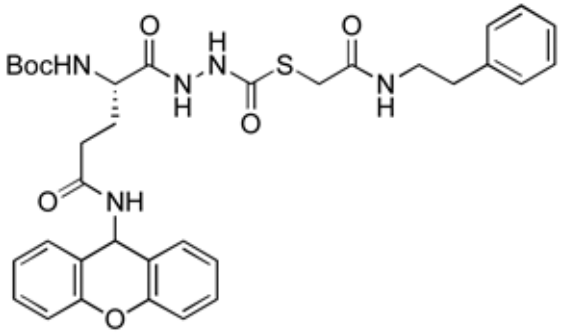
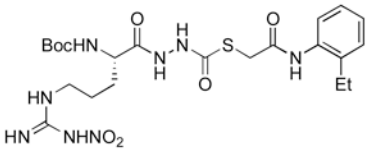
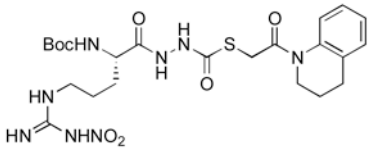
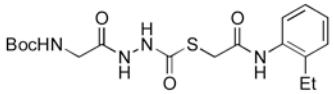
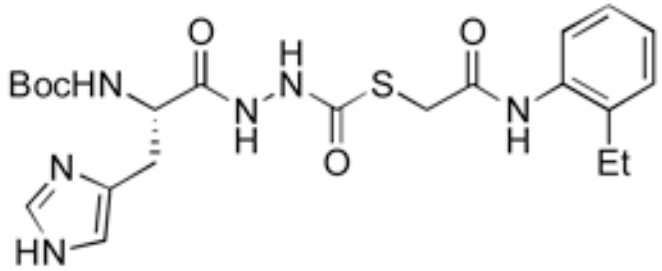
Comb Chem High Throughput Screen. Author manuscript; available in PMC 2011 May 1.

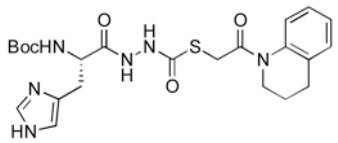
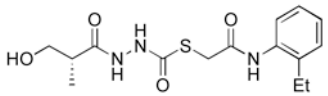
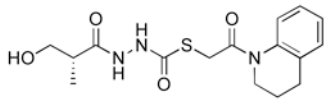
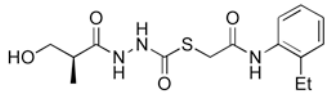
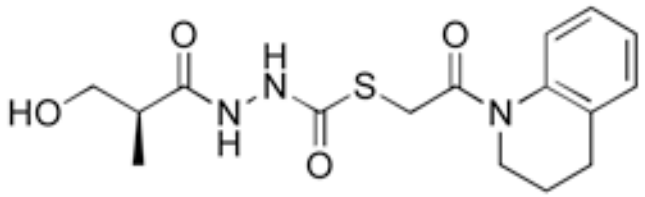
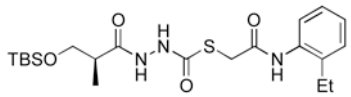
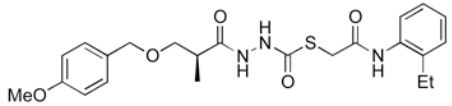
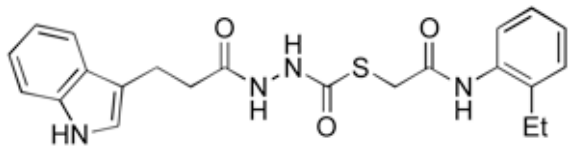
Thiocarbazate	Structure	HRMS	IC_{50} (mM) Cat B	Cat L	Cat S
44		[M+Na]+567.2239	3.67	13.7	3.32
45		[M+H]+579.2236	4.35	>25.0	5.15
46		[M+Na]+553.2101	14.3	1.27	2.15
47		[M+H]+565.2097	13.2	1.51	2.64
48		[M+Na]+553.2087	>25.0	15.4	4.42
49		[M+Na]+565.2088	>25.0	17.9	5.89
50		[M+Na]+565.2083	9.28	0.37	0.95

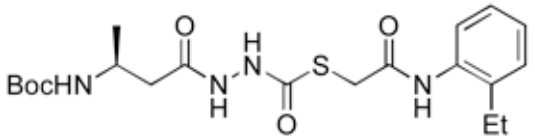
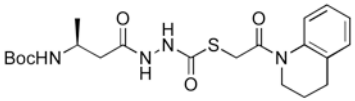
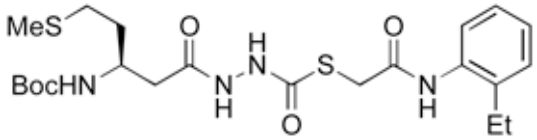
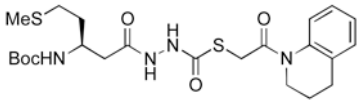
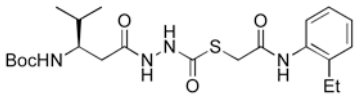
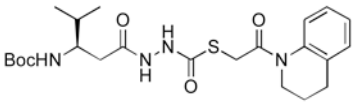
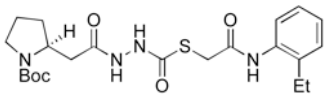
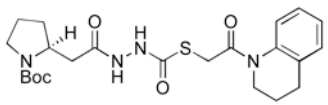
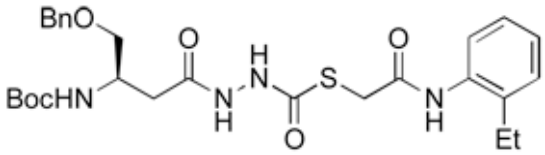
Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
51		[M+Na]+553.2082	>25.0	0.51	0.65
52		[M+Na]+609.1385	>25.0	0.58	0.91
53		[M+Na]+592.2198	>25.0	0.38	0.92
54		[M+Na]+567.1900	13.7	1.85	3.00
55		[M+Na]+570.1649	5.07	0.70	1.44
56		[M+Na]+597.1998	>25.0	0.43	1.64
57		[M+Na]+68.2802	4.35	>25.0	5.15

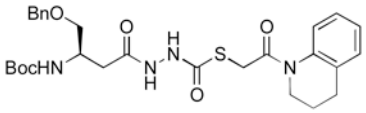
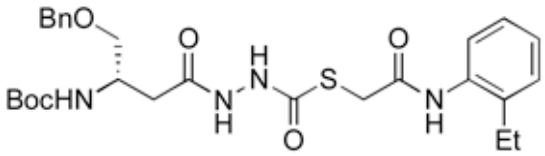
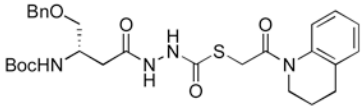
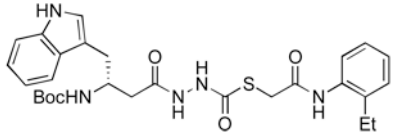
Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
58		[M+Na]+616.286	14.3	1.27	2.15
59		[M+H]+662.2664	>25.0	>25.0	>25.0
60		[M+Na]+674.2639	>25.0	>25.0	>25.0

Comb Chem High Throughput Screen. Author manuscript; available in PMC 2011 May 1.

Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
61		[M+Na]+662.2669	>25.0	>25.0	>25.0
62		[M+Na]+554.2271	>25.0	3.67	2.15
63		[M+Na]+565.2214	19.7	>25.0	3.07
64		[M+Na]+433.1513	>25.0	>25.0	>25.0
65		[M+H]+491.2057	13.7	>25.0	>25.0

Thiocarbazate	Structure	HRMS	IC$_{50}$ (mM) Cat B	Cat L	Cat S
66		[M+H]+503.2071	11.6	>25.0	>25.0
67		[M+Na]+362.1144	>25.0	>25.0	>25.0
68		[M+H]+374.1152	>25.0	>25.0	>25.0
69		[M+Na]+362.1144	>25.0	>25.0	>25.0
70		[M+Na]+374.1152	>25.0	>25.0	>25.0
71		[M+Na]+476.2003	>25.0	>25.0	>25.0
72		[M+Na]+482.1727	>25.0	>25.0	>25.0
73		[M+H]+447.1485	>25.0	>25.0	>25.0

Thiocarbazate	Structure	HRMS	IC_{50} (mM) Cat B	Cat L	Cat S
74		[M+H]+439.2004	>25.0	>25.0	>25.0
75		[M+H]+451.2012	>25.0.	>25.0	>25.0
76		[M+H]+499.2047	>25.0	5.52	2.41
77		[M+H]+511.2044	>25.0	6.95	3.01
78		[M+Na]+489.2125	11.5	7.52	>25.0
79		[M+H]+479.2319	15.8	>25.0	>25.0
80		[M+H]+465.2172	>25.0	>25.0	>25.0
81		[M+H]+477.2177	>25.0	>25.0	>25.0
82		[M+Na]+567.2276	>25.0	>25.0	4.47

Thiocarbazate	Structure	HRMS	IC_{50} (mM) Cat B	Cat L	Cat S
83		[M+H]+557.2435	>25.0	2.12	5.95
84		[M+H]+545.2438	>25.0	>25.0	15.2
85		[M+H]+557.2454	>25.0	>25.0	>25.0
86		[M+Na]+576.2249	>25.0	>25.0	>25.0

Comb Chem High Throughput Screen. Author manuscript; available in PMC 2011 May 1.