Center Probe Report

Probe reports are generated for each compound that meets or surpasses the MLSCN requirements for a new probe.

Please provide the following specific information:

Project Title: Multiplexed HTS of Serine and Cysteine Proteases
Grant Number: MH076406-01

Screening Center Name: Penn Center for Molecular Discovery
Principal Investigator of Screening Center: Scott L. Diamond, Ph.D.

Assay Submitter & Institution: Scott L. Diamond, University of Pennsylvania

PubChem Bioassay Identifier (AID): 460

Probe PubChem Identifier: SID 46493575, CID 23631927

Please provide the following descriptions:

Specific Aim:

To perform a high throughput screen of the NIH Molecular Libraries Small Molecule Repository (MLSMR) to identify novel, potent, and selective inhibitors of human cathepsin L.

Significance/Rationale [1]:

"Cathepsins are a family of ubiquitous lysosomal proteases that have demonstrated important physiopathological roles in inflammation, osteo- and rheumatoid arthritis, osteoporosis, Alzheimer's disease, multiple sclerosis, pancreatitis, apoptosis, liver disorders, lung disorders, myocardial disorders, diabetes, muscular dystrophy, tumor invasion and metastasis [2-5]. Cathepsin L is an endopeptidase that is catalytically the most active known lysosomal cysteine protease [6]. It is a papain-like cysteine protease of interest due to involvement in many biological processes including epidermal homeostasis, hair follicle morphogenesis and cycling, extracellular matrix degradation, and viral entry of Severe Acute Respiratory Syndrome (SARS) coronavirus, Ebola virus, and Hendra virus [5, 7-10]. Due to the roles of many cathepsin L-like proteases in diseases such as malaria (falcipain), leishmaniasis (*Leishmania major* cathepsin L), Chagas disease (cruzipain), African trypanosomiasis (congopain), toxoplasmosis (*Toxoplasma gondii* cathepsin L), amoebiasis (histolysain), and sleeping sickness (rhodesain), inhibitors of human cathepsin L may be highly valuable as therapeutic treatments against these infectious diseases." [1]

Assay Implementation and Screening

PubChem Bioassay Name: Cathepsin L

PubChem Primary Assay Description [11]:

"Human liver cathepsin L (EC 3.4.22.15) is a lysosomal cysteine protease. Recent interest in cathepsin L has been generated by research showing that proteolysis by this enzyme is required for the entry and replication of the SARS and Ebola viruses in human cells. Thus cathepsin L inhibitors have potential as novel anti-viral agents.

Cathepsin L inhibitors may also be active against Plasmodium falciparum, the parasite responsible for human malaria. Plasmodium contains cathepsin L-like cysteine proteases known as falcipains that appear to promote virulence of the parasite through haemoglobin digestion and erythrocyte invasion.

The high-throughput screen for cathepsin L inhibitors reported here consisted of an end-point assay monitoring the release of the fluorophore aminomethyl coumarin (AMC) upon enzymatic hydrolysis of an AMC-labeled dipeptide.

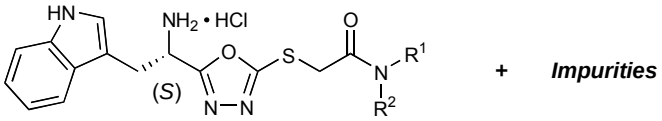
This assay is a part of the Molecular Libraries Screening Center Network." [11]

Center Summary of the Primary Screen:

57,821 compounds from the NIH Molecular Libraries Small Molecule Repository (MLSMR) were screened against human liver cathepsin L. This screen was performed in 384-well Corning 3676 black, low-volume, non-binding surface (NBS)-coated polystyrene plates using a total reaction volume of 10 µL per well. Final well concentrations were: 10 µM compound (2% DMSO), 300 pM (8.7 ng/mL) human liver cathepsin L, 1 µM Z-FR-AMC substrate, and 20 mM sodium acetate buffer containing 1 mM EDTA and 5 mM DTT, pH 5.5. Cathepsin L was activated in the assay buffer for 30 minutes prior to dispensing into wells. Assay plates were incubated at room temperature for one hour and the fluorescence intensity of each well was read with a PerkinElmer Envision plate reader (Ex: 355 nm, Em: 460 nm) to measure hydrolysis of the AMC substrate. The screen correctly identified E-64 and E-64c members of the library as potent inhibitors of cathepsin L. A Z' factor of 0.73 was calculated for this screen, indicating good plate uniformity throughout the run.

From the 57,821 compound primary HTS, 102 compounds (0.18%) showed >45% inhibition against human cathepsin L. Upon a confirmatory dose response assay, 49 (48%) of these small molecules inhibited human cathepsin L activity with an IC_{50}<50 µM. "Library samples containing 2,5-disubstituted oxadiazoles were identified as potent hits in a high throughput screen of the NIH Molecular Libraries Small Molecule Repository directed at discovering inhibitors of cathepsin L (**Table I**). 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 (**Figure 1**). Previously unknown, thiocarbazates were independently synthesized as single enantiomers and found to inhibit cathepsin L in the low nanomolar range." [12]

Table I. Cathepsin L inhibitory activity of oxadiazole-containing library samples[a] [12].



PubChem SID	R^1	R^2	IC_{50} (μM)[b]
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 mean ± standard deviation (number of determinations = 3).

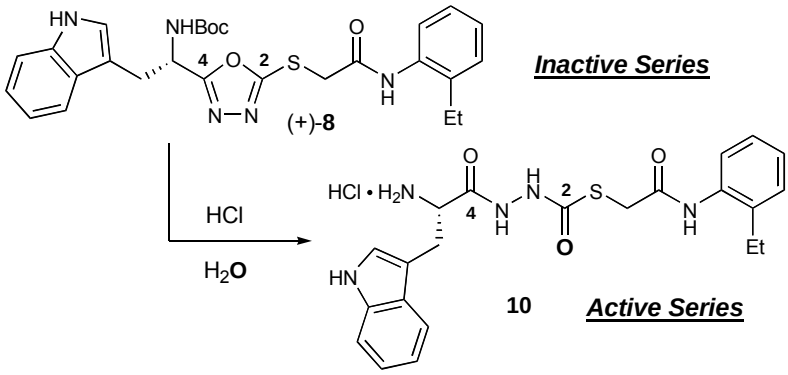


Figure 1. Conversion of oxadiazole (+)-**8** to thiocarbazate **10**. [12]

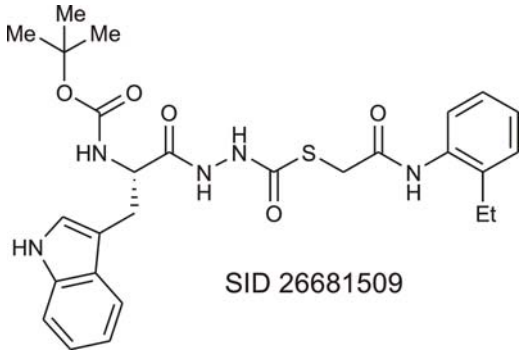


Figure 2. Previously reported thiocarbazate probe SID 26681509. [12]

Probe Optimization

Center comments on chemistry strategy leading to probe identification:

Building upon knowledge of a previously reported thiocarbazate probe (**Figure 2**) [12] for cathepsin L, we sought out ways to make this inhibitor even more active against the biological target by focusing on molecular interactions between the inhibitor and the crystal structure of papain. Researchers have previously used papain as a model to study the interactions of compounds with cathepsin L [1, 13]. Upon observing the rather nice fit of the 2-ethylanilide moiety of SID 26681509 into the hydrophobic pocket of the papain binding site, we designed a new compound by completing a ring structure connecting the ethyl group of the 2-ethylanilide to the nearby nitrogen atom. The resulting molecule would become more hydrophobic, and would thereby bind stronger to the cathepsin L active site and more rapidly inhibit the cysteine protease [14].

In addition, synthesis of analogs to this molecule replacing the sulfur atom with carbon or oxygen revealed that oxygen is preferred in that position (**Table II**). This resulted in the development of SID 46493575 (**Figure 3**), a tetrahydroquinoline probe for cathepsin L inhibition. The synthetic route for preparation of SID 46393575 is shown in **Figure 4**.

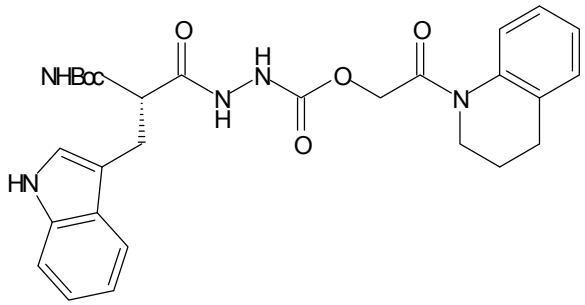


Figure 3. Newly developed tetrahydroquinoline probe SID 46493575.

Table II. Cathepsin L inhibitory activity (IC_{50} data, n=3) of synthesized compounds replacing the sulfur atom of SID 26681509 with either carbon or oxygen. Inhibitory data is also shown for the tetrahydroquinoline compounds, replacing the same sulfur atom with either carbon or oxygen. Compounds do not show inhibition against cathepsin L when carbon is present in that position.

	C	**S**	**O**
2-ethylanilide	> 50 μM	0.056 ± 0.004 μM	0.028 ± 0.004 μM
Tetrahydroquinoline	> 50 μM	0.041 ± 0.002 μM	0.007 ± 0.001 μM

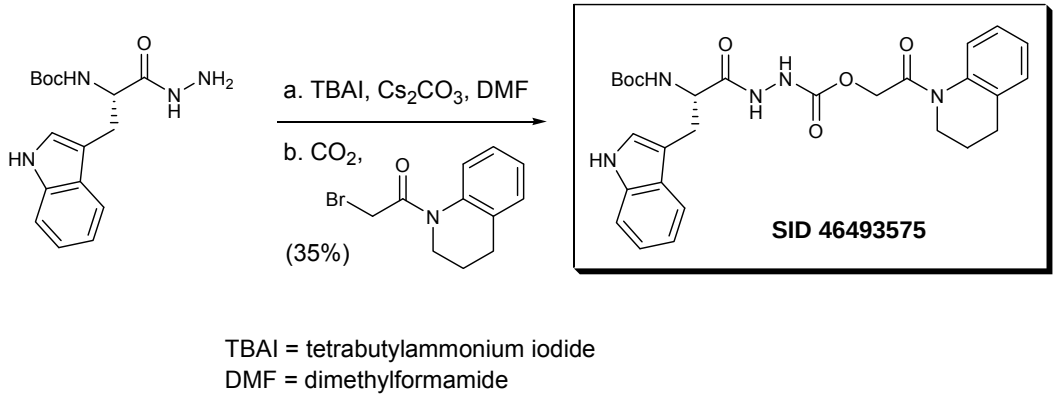


Figure 4. Synthetic procedure for preparation of tetrahydroquinoline probe SID 46493575 [15].

We also observed an additional hydrogen bonding interaction between the conserved His159 and the tetrahydroquinoline probe SID 46493575 (**Figure 5**). The docking score for tetrahydroquinoline SID 46493575 in papain was correspondingly better at -10.00 kcal/mol (as compared to -9.03 kcal/mol for thiocarbazate SID 26681509 in papain) [14].

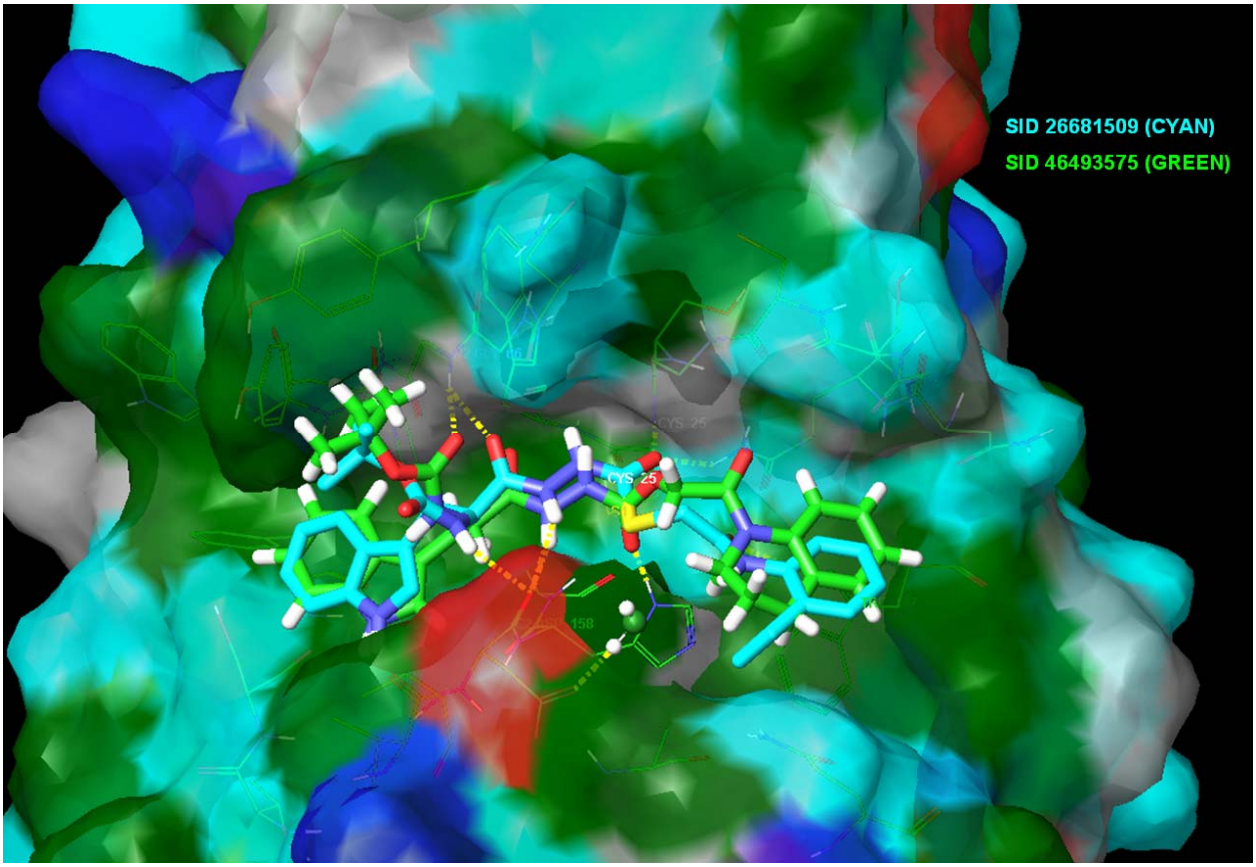


Figure 5. Theoretical model of tetrahydroquinoline probe SID 46493575 (green) overlaid with thiocarbazate probe SID 26681509 (cyan) in the binding site of papain. An additional H-bonding interaction can be observed with SID 46493575 and His159.

Severe acute respiratory syndrome coronavirus (SARS-CoV) and Ebola virus are hypothesized to function by trafficking to an intracellular compartment, wherein the contents of the viral package are released due to proteolytic cleavage by cathepsin L [10]. By inhibiting human cathepsin L, we hoped that SARS-CoV and Ebola viral entry could be inhibited as well. Therefore, SID 46493575 was tested in vitro in a SARS coronavirus pseudotype entry assay and a Ebola virus pseudotype entry assay using 293T cells. Preliminary data is reported in **Figures 6** and **7**. SID 46493575 is shown to inhibit entry of both SARS-CoV (IC_{50} = ~200 nM) and Ebola virus (IC_{50} = ~200 nM). Vesicular stomatitis virus (VSV), which is not presumed to rely upon cathepsin L, was used as a control [16].

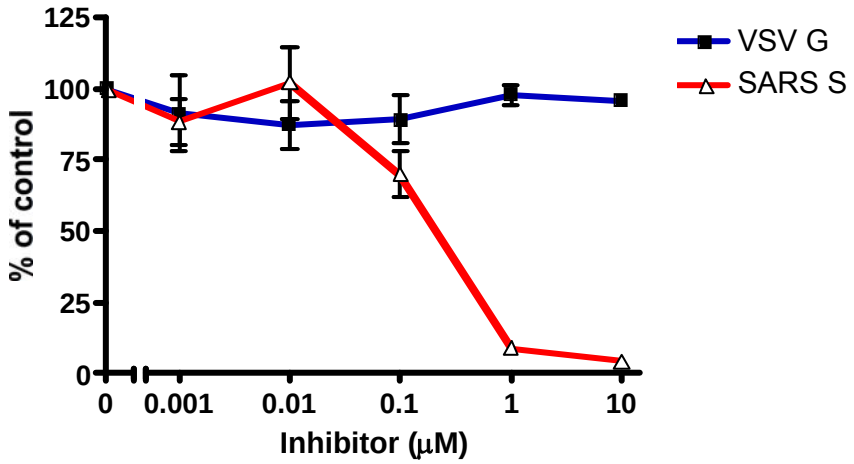


Figure 6. Activity of SID 46493575 against SARS-CoV pseudotype infection.

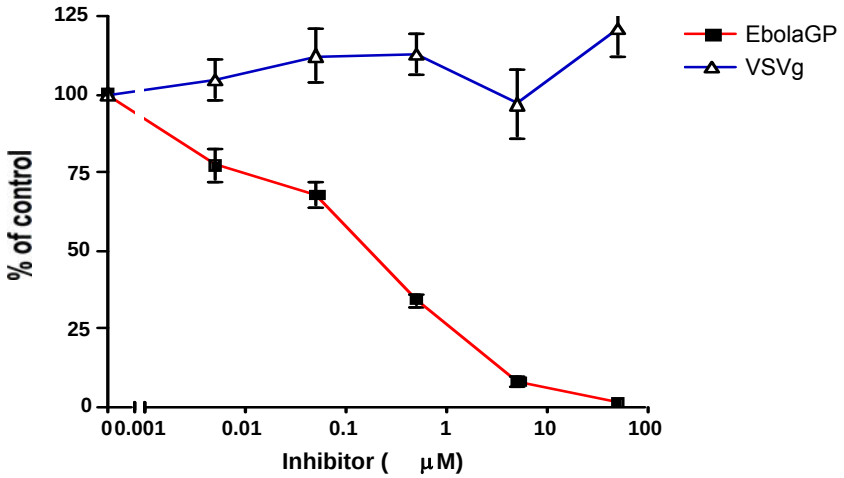


Figure 7. Activity of SID 46493575 against Ebola virus pseudotype infection.

Center summary of probe properties and recommendations for the scientific use of probe as research tool [16]:

SID 46493575, a small molecule tetrahydroquinoline, was determined to be an extremely potent slowly reversible inhibitor of human cathepsin L with an IC_{50} of 6.9 nM. After a 4 hr preincubation with cathepsin L, this compound became even more potent, demonstrating an IC_{50} of 0.4 nM. Molecular docking studies involving the papain/SID 26681509 computationally derived crystal structure [1] were used in designing SID 46493575. These studies revealed key hydrophobic interactions that could be exploited in the synthesis of a structurally-related compound [14]. SID 46493575 displays 725-fold greater selectivity toward cathepsin L than cathepsin B. SID 46493575 is non-toxic at 100 μM to zebrafish in a live organism assay.

Most interesting to note, however, is the fact that SID 46493575 inhibits both SARS coronavirus and Ebola virus pseudotype entry into cells in vitro. SID 26681509 was shown to have little inhibition of SARS-CoV entry (IC_{50} > 10 μM) and Ebola viral entry (IC_{50} > 10 μM). These results suggest that crossing the hydrophobic barriers of the many membranes required to reach the endosomes is perhaps eased by the closing of the tetrahydroquinoline ring structure in SID 46493575.

A summary of probe properties of SID 46493575 is shown in **Table III**. Further kinetic analyses as well as structure-activity relationships are currently under investigation to develop even more potent and selective compounds to inhibit processes involving cathepsin L and cathepsin L-like proteases [15, 16].

Table III. Summary of properties of tetrahydroquinoline probe SID 46493575.

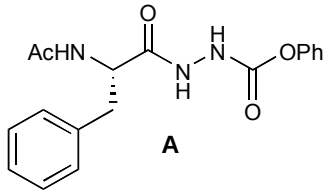
IC_{50} (no preincubation)	6.9 nM
IC_{50} (4 hr preincubation)	0.4 nM
(IC_{50} Cat B) / (IC_{50} Cat L)	724.6
Toxicity Viral entry assays	Non-toxic to zebrafish at 100 μM. Inhibits SARS-CoV pseudotype infection with IC_{50} = ~200 nM. Inhibits Ebola virus pseudotype infection with IC_{50} = ~200 nM.

Appendices

Comparative data on probe, similar compound structures and prior probes [12]:

"Previously reported inhibitors of cathepsin L include the peptides, leupeptin and aprotinin, and the fluoromethyl ketone, Z-LLL-FMK [17, 18]. 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, 17, 19]. 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 [18, 20-24].

Thiocarbazates such as **10** have not been described previously in the literature. However, these compounds do bear structural resemblance to aza-peptides [25-27] (e.g. **A**; **Figure 8** [28]), 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 [29]." [12]



k_{on} = >11000 $M^{-1}s^{-1}$, papain
Abeles and Magrath (1992)

Figure 8. Known aza-peptide cathepsin inhibitor **A**.

Quinoline structures have previously been reported as inhibitors of malaria. However, their mechanism of action is not fully understood. Foley and Tilley state that quinolines may interfere with hemoglobin digestion in the various blood stages of the parasitic life cycle [30]. The cysteine protease falcipains are regarded as cathepsin L-like proteases and could be inhibited by the quinolines, thus inhibiting disease progression. We plan to test this tetrahydroquinoline probe SID 46493575 in a luciferase assay against infected red blood cells.

Comparative data showing probe specificity for target [16]:

SID 46493575 was tested for selectivity against human cathepsin B (**Table IV**) with no preincubation of enzyme and inhibitor. The selectivity ratio ($IC_{50,CatB}/IC_{50,CatL}$) of SID 46493575 showed a 725-fold preference for inhibition of human cathepsin L over human cathepsin B.

Table IV. Selectivity of SID 46493575 against human cathepsin B.

Enzyme	IC_{50} (μM)	Selectivity Index
Cathepsin L, human liver	0.007 ± 0.001	1.0
Cathepsin B, human liver	5.072 ± 0.883	724.6

References

1. Shah, P.P., et al., *Kinetic Characterization and Molecular Docking of a Novel, Potent, and Selective Slow-binding Inhibitor of Human Cathepsin L.* 2007. Submitted.
2. Turk, D. and G. Guncar, *Lysosomal cysteine proteases (cathepsins): promising drug targets.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 2): p. 203-13.
3. Fujishima, A., et al., *The crystal structure of human cathepsin L complexed with E-64.* FEBS Lett, 1997. **407**(1): p. 47-50.
4. Leist, M. and M. Jaattela, *Triggering of apoptosis by cathepsins.* Cell Death Differ, 2001. **8**(4): p. 324-6.
5. Nomura, T. and N. Katunuma, *Involvement of cathepsins in the invasion, metastasis and proliferation of cancer cells.* J Med Invest, 2005. **52**(1-2): p. 1-9.
6. Otto, H.H. and T. Schirmeister, *Cysteine Proteases and Their Inhibitors.* Chem Rev, 1997. **97**(1): p. 133-172.
7. Turk, V., B. Turk, and D. Turk, *Lysosomal cysteine proteases: facts and opportunities.* Embo J, 2001. **20**(17): p. 4629-33.
8. Chandran, K., et al., *Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection.* Science, 2005. **308**(5728): p. 1643-5.

9. Pager, C.T. and R.E. Dutch, *Cathepsin L is involved in proteolytic processing of the Hendra virus fusion protein.* J Virol, 2005. **79**(20): p. 12714-20.
10. Simmons, G., et al., *Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry.* Proc Natl Acad Sci U S A, 2005. **102**(33): p. 11876-81.
11. PubChem web address for PCMD cathepsin L hits: http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=460.
12. Myers, M.C., et al., *Identification and Synthesis of a Unique Thiocarbazate Cathepsin L Inhibitor.* Bioorg Med Chem Lett, 2008. **18**(1): p. 210-214.
13. Tsuge, H., et al., *Inhibition mechanism of cathepsin L-specific inhibitors based on the crystal structure of papain-CLIK148 complex.* Biochem Biophys Res Commun, 1999. **266**(2): p. 411-6.
14. Beavers, M.P., et al., *Molecular Docking of Cathepsin L Inhibitors in the Binding Site of Papain.* 2008. In preparation.
15. Myers, M.C., et al., *Structure-Activity Relationships of a Thiocarbazate Probe for Human Cathepsin L Inhibition.* 2008. In preparation.
16. Shah, P.P., et al., *Tetrahydroquinoline Resistance to Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Ebola Virus Pseudotype Infections.* 2008. In preparation.
17. Esser, R.E., et al., *Cysteine proteinase inhibitors decrease articular cartilage and bone destruction in chronic inflammatory arthritis.* Arthritis Rheum, 1994. **37**(2): p. 236-47.
18. Marquis, R.W., et al., *Azepanone-based inhibitors of human cathepsin L.* J Med Chem, 2005. **48**(22): p. 6870-8.
19. Montaser, M., G. Lalmanach, and L. Mach, *CA-074, but not its methyl ester CA-074Me, is a selective inhibitor of cathepsin B within living cells.* Biol Chem, 2002. **383**(7-8): p. 1305-8.
20. Schirmeister, T. and U. Kaeppler, *Non-peptidic inhibitors of cysteine proteases.* Mini Rev Med Chem, 2003. **3**(4): p. 361-73.
21. Hernandez, A.A. and W.R. Roush, *Recent advances in the synthesis, design and selection of cysteine protease inhibitors.* Curr Opin Chem Biol, 2002. **6**(4): p. 459-65.
22. Zhou, N.E., et al., *3-Acylamino-azetidin-2-one as a novel class of cysteine proteases inhibitors.* Bioorg Med Chem Lett, 2003. **13**(1): p. 139-41.
23. Huang, L. and J.A. Ellman, *General solid-phase method to prepare novel cyclic ketone inhibitors of the cysteine protease cruzain.* Bioorg Med Chem Lett, 2002. **12**(20): p. 2993-6.
24. Chiva, C., et al., *Synthesis and NMR structure of p41icf, a potent inhibitor of human cathepsin L.* J Am Chem Soc, 2003. **125**(6): p. 1508-17.
25. Baggio, R., et al., *From good substrates to good inhibitors: design of inhibitors for serine and thiol proteases.* Biochemistry, 1996. **35**(11): p. 3351-3.
26. Thompson, S.K., et al., *Design of potent and selective human cathepsin K inhibitors that span the active site.* Proc Natl Acad Sci U S A, 1997. **94**(26): p. 14249-54.
27. Xing, R. and R.P. Hanzlik, *Azapeptides as inhibitors and active site titrants for cysteine proteinases.* J Med Chem, 1998. **41**(8): p. 1344-51.
28. Magrath, J. and R.H. Abeles, *Cysteine protease inhibition by azapeptide esters.* J Med Chem, 1992. **35**(23): p. 4279-83.
29. Powers, J.C., et al., *Irreversible inhibitors of serine, cysteine, and threonine proteases.* Chem Rev, 2002. **102**(12): p. 4639-750.
30. Foley, M. and L. Tilley, *Quinoline antimalarials: mechanisms of action and resistance and prospects for new agents.* Pharmacol Ther, 1998. **79**(1): p. 55-87.