Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

A small molecule oxocarbazate inhibitor of human cathepsin L blocks SARS and Ebola pseudotype virus infection into HEK 293T cells

Parag P. Shah[1], Tianhua Wang[1], Rachel L. Kaletsky, Michael C. Myers, Jeremy E. Purvis, Huiyan Jing, Donna M. Huryn, Doron C. Greenbaum, Amos B. Smith III, Paul. Bates, and Scott L. Diamond

Department of Chemical and Biomolecular Engineering (P.P.S, T.W, J.E.P, H.J, S.L.D.), Department of Microbiology (R.L.K, P.B), Department of Pharmacology (D.C.G.), Department of Chemistry (M.C.M., D.M.H, A.B.S), Penn Center for Molecular Discovery, Institute for Medicine and Engineering, University of Pennsylvania, Philadelphia Pennsylvania.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Running title: A Cat-L inhibitor blocks SARS and Ebola pseudotype virus entry

Corresponding author: Scott L. Diamond, Penn Center for Molecular Discovery, University of Pennsylvania, 1024 Vagelos Laboratories, Philadelphia, PA 19104-6383. Phone: 215-573-5702. Fax: 215-573-7227. E-mail: sld@seas.upenn.edu

Text pages: 13 (Abstract through Discussion)

Tables: 1

Figures: 5

References: 26

Abstract: 240

Introduction: 499

Discussion: 718

Nonstandard abbreviations: CID, PubChem compound identifier; AMC, 7-amido-4-methylcoumarin; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; SARS, severe acute respiratory syndrome;

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

A tetrahydroquinoline oxocarbazate (PubChem CID 23631927) was tested as an inhibitor of human cathepsin L (EC 3.4.22.15) and as an entry blocker of SARS coronavirus and Ebola pseudotype virus. In the cathepsin L inhibition assay, the oxocarbazate caused a time-dependent 17-fold drop in IC_{50} from 6.9 nM (no pre-incubation) to 0.4 nM (4 hr pre-incubation). Slowly reversible inhibition was demonstrated in a dilution assay. A transient kinetic analysis using a single step, competitive inhibition model provided rate constants of k_{on} = 153,000 $M^{-1}s^{-1}$ and k_{off} = 4.40 x 10^{-5} s^{-1} (K_i = 0.29 nM). The compound also displayed cathepsin L/B selectivity of > 700-fold and was nontoxic to human aortic endothelial cells at 100 μM. The oxocarbazate and a related thiocarbazate (CID 16725315) were tested in a SARS-CoV and Ebola virus-pseudotype infection assay with the oxocarbozate, but not the thiocarbazate, demonstrating activity in blocking both SARS-CoV (IC_{50} = 273±49 nM) and Ebola virus (IC_{50} = 193±39 nM) entry into HEK 293T cells. In order to trace the intracellular action of the inhibitors with intracellular cathepsin L, the activity based probe DCG-04 was used to label the active site of cysteine proteases in 293T lysates. The reduction in active cathepsin L in inhibitor treated cells correlated well with the observed potency of inhibitors observed in the virus pseudotype infection assay. Overall, the oxocarbazate CID 23631927 was a sub-nanomolar slow-binding, reversible inhibitor of human cathepsin L that blocked SARS-CoV and Ebola pseudotype virus entry in human cells.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

Cathepsin L is one of the eleven members (cf. B, C, F, H, K, L, O, S, V, W and X) of human lysosomal cysteine proteases which fall in the C1 family (papain family) of the CA clan (Rossi, et al., 2004). These enzymes were traditionally linked to non-specific proteolytic activity within lysosomes. More recently, cathepsin L has been implicated in regulatory events relating to cancer, diabetes, immunological responses, degradation of the articular cartilage matrix and other pathological processes (Chapman, et al., 1997; Turk and Guncar, 2003; Maehr, et al., 2005; Vasiljeva, et al., 2007) including osteoporosis, rheumatoid arthritis, and tumor metastasis (McGrath, 1999; Turk, et al., 2001; Potts, et al., 2004; Schedel, et al., 2004; Palermo and Joyce, 2008). Inhibitors of cathepsin L block viral entry of Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) and Ebola virus and impair conversion of Hendra virus glycoprotein into the mature, active form (Simmons, et al., 2005; Chandran, et al., 2005; Pager and Dutch, 2005). With respect to the development of antiviral agents, inhibitors of human cathepsin L are not subject to resistance due to rapid mutations of the viral genome. Cathepsin L is therefore an attractive target for drug development.

In 2002, SARS-CoV infection resulted in 8,096 cases of infection and 774 deaths (Enserink, 2003). Following virus uptake after angiotensin-converting enzyme 2 (ACE2) receptor binding, cathepsin L-mediated proteolysis induces conformational changes in the SARS-CoV S glycoprotein to trigger the endosomal membrane fusion process (Simmons, et al., 2005). Similarly, cleavage of the Ebola glycoprotein by cathepsin L reveals the receptor interaction domain within GP and cathepsin L activity is required for entry into the host cell (Chandran, et al., 2005; Schornberg, et al., 2006; Lee, et al., 2008). The Ebola virus causes fatal

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

hemorrhagic fever. The first outbreaks were reported in Sudan and Zaire in 1976 (Feldmann, et al., 2003). Inhibition of cathepsin L thus holds the promise for therapeutic intervention for both SARS-CoV and Ebola virus infection.

High throughput screening for cathepsin L inhibitors first identified CID 16725315 (**Fig.1A**) as a novel thiocarbazate compound exhibiting potent inhibition against cathepsin L (Shah, et al., 2008; Myers, et al., 2008a). Molecular docking simulation of the thiocarbazate against papain structure revealed key contacts for this ligand-protein interaction. (Beavers, et al., 2008) Addition medicinal chemistry efforts resulted in the synthesis of an even more potent tetrahydroquinoline oxocarbazate compound CID 23631927 (**Fig. 1B**) (Myers, et al., 2008a; Beavers, et al., 2008; Myers, et al., 2008b).

In this paper, a full kinetic characterization of the oxocarbazate inhibitor is reported along with a measure of the potency in blocking entry of SARS-CoV and Ebola virus pseudotypes. This compound displayed an IC_{50} of ~200 and 300 nM in preventing 293T cell infection by the Ebola and SARS-CoV pseudotypes respectively. Labeling of intracellular active cathepsin L using an activity based probe, DCG-04, permitted quantitative estimation of the interaction of cathepsin L inhibitors with the target enzyme in HEK 293T cells.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Material and Methods

Cathepsin L Assay. As described in prior work (Shah, et al., 2008)**,** the cathepsin L assay was carried out using 1 µM *Z*-Phe-Arg-7-amido-4-methylcoumarin (*Z*-Phe-Arg-AMC; Sigma, St. Louis, MO) and 8.7 ng/ml human liver Cathepsin L (Calbiochem, San Diego, CA). Assay buffer consisted of 20 mM sodium acetate, 1 mM EDTA, and 5 mM cysteine, pH 5.5. The enzyme was incubated in the buffer for 30 min to ensure the catalytic cysteine was in the reduced form. The enzymatic reaction (25°C) was read on an Envision florescent microplate reader (PerkinElmer, Waltham, MA) and the fluorescence signal was measured at Ex 355nm and Em 460nm. CID 23631927 was stored as a 10 mM DMSO aliquots (-80°C) and used fresh in each experiment. Samples of CID 23631927 can be obtained from the NIH Molecular Libraries Small Molecule Repository (BioFocus, South San Francisco, CA).

Inhibitor Potency Determination. A 16-point, 2-fold serial dilution dose response assay was performed in triplicate. Each well of a 96-well assay plate (Corning Life Sciences, Acton, MA) contained 38 µl of water and 2 µl of inhibitor in dimethyl sulfoxide (DMSO). Positive and negative controls were present to serve as internal controls into which 2 µl of DMSO was transferred in place of the inhibitor. Ten microliters of 10 µM Z-Phe-Arg-AMC in 5x concentrated assay buffer and 50 µl of 17.4 ng/ml Cathepsin L in assay buffer were added sequentially to initiate the protease reaction. A total of 50 µl of assay buffer was dispensed in place of enzyme into negative control wells. This resulted in a final dose response concentration range from 50 µM to 1.5 nM of inhibitor (2% DMSO) in a 100-µl final reaction volume. A four-

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

parameter logistic model (IDBS XLfit eq. 205; Guildford, Surrey, UK) was used to fit the measured data and find IC_{50} values.

Time dependent inhibition and reversibility test. Pre-incubation assays were performed to further establish the time dependency of inhibition. Enzyme and inhibitor were pre-incubated for various time points in a 96-well microplate before the addition of substrate to initiate the enzymatic reaction. A total of 47.5 µl of cathepsin L (18.3 ng/ml), and 47.5 µl of CID 23631927 at various concentrations in assay buffer were incubated up to 4 hr. A volume of 5 µl of Z-Phe-Arg-AMC substrate was then added and the plate was monitored for AMC hydrolysis. A dilution assay protocol was deployed as previously described (Shah, et al., 2008) to monitor reversibility. The concentrated mixture of the enzyme (870 ng/ml) and the inhibitor CID 23631927 (25 nM) were incubated for 1 hr before being diluted 100-fold in a Corning 96-well plate with assay buffer containing 1 µM Z-Phe-Arg-AMC to a final volume of 200 µl.

Kinetic Data Fitting. In the kinetic simulations, the concentrations of chemical species ([E], [S], [I], [P], [ES], [EI]) over time were calculated using a system of ordinary differential equations for each reaction step as adopted for the parameter fitting in prior work (Shah, et al., 2008). Progress curves at each inhibitor concentration were fit into a five-parameter (k_1, k_{-1}, k_{on}, k_{off}, k_{cat}) kinetic inhibition model using MATLAB 7.8.0 (The MathWorks, Inc). The objective function used in the minimization search was computed as the sum of the square of the relative errors between the simulated values and experimental data.

Selectivity against Cathepsin B. CID 23631927 was assayed for inhibition against cathepsin B. Human liver cathepsin B (65 ng/ml; Calbiochem) was used and the corresponding substrate was

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

15 µM of Z-Arg-Arg-AMC. The reactions were performed in 20 mM sodium acetate buffer containing 5 mM cysteine and 1 mM EDTA, pH 5.5.

Cytotoxicity. CID 23631927 was tested for cytotoxicity against human aortic endothelial cells (HAECs) as previously described (Shah, et al., 2008). Cells were seeded at 1000 cells in 25 µl per well in a 384 well plate. The plate was centrifuged and incubated at 37 °C for 24 hr. CID 23631927 and doxorubicin positive control were then serially diluted in EGM-2 endothelial cell media (Lonza CC4176). A total of 5 µl of these serial dilutions was added to the wells in triplicate, resulting in final concentrations of compound from 100 µM to 156 nM (0.17% DMSO). The plate was centrifuged and incubated at 37 °C for 24 hr. A total of 30 µl CellTiter-Glo™ (Promega G7570) was added to each well and centrifuged. After 10 min, luminescence was measured using the Envision microplate reader.

Virus Pseudotype Preparation and Virus Infection Assay. Pseudovirions were prepared essentially as described previously(Kaletsky, et al., 2007). In brief, 293T cells were cultured at 37°C and 5% CO_2 in Dulbecco's Modified Eagle's Medium (DMEM) (Invitrogen) supplemented with 10% Fetal Bovine Serum. Pseudotypes were produced by transfecting 293T cells with 10 µg of an HIV-luciferase vector, pNL-luc, and 30 µg of a plasmid encoding the Ebola glycoprotein or SARS Spike proteins (Simmons, et al., 2005; Kaletsky, et al., 2007). Vesicular stomatitis virus glycoprotein (VSV-G) pseudovirions were prepared as above using 6 µg of a VSV-G-encoding plasmid. Virions were concentrated by ultracentrifugation at 40,000 rpm in a SW41 rotor (Beckman) through a 20% sucrose cushion for 1 hr at 4°C. The pellets were resuspended in PBS overnight at 4°C. To measure infection, 293T cells were pretreated for 1 hr

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

with inhibitors or MDL28170 (Sigma) as a positive control (Simmons, et al., 2005). The media was removed and replaced with the same inhibitors at double the final concentration. An equal volume of pseudotypes was then added, and cells were spin-infected at 1,200g for 2 hr at 4°C. Following spin-infection, the cells were incubated for 6 hr, and the medium was replaced with fresh media without drug. Cell lysates were assayed for luciferase activity (Promega) at 40 hr post-infection.

Labeling active cathepsin L in HEK 293T using DCG-04. Oxocarbazate (CID 23631927), thiocarbazate (CID 16725315), MDL28170 (*Z-Val-Phe-CHO*) (Calbiochem) and Cathepsin-L Inhibitor III (*Z-Phe-Tyr(t-Bu)-diazomethylketone*) (Calbiochem) were dissolved in DMSO to give 10mM stock solution. DCG-04 (biotin-Lys-C5 alkyl linker-Tyr-Leu-epoxide) was prepared as previously described (Greenbaum, et al., 2000). HEK 293T cells were seeded into 6-well tissue culture plates (BD) at a concentration of 10^6 cells / well. Inhibitors were added into the well in a 1:1000 dilution to give final concentration at 10 µM and incubated with the cells for 3 hr. The adherent cells were then scrapped, pelleted, lysed in 100 mM sodium acetate and 0.25% triton-X (pH 5.5), and subjected to 3X freeze-thaw cycles, followed by sonification (30 sec) (Branson Ultraxonic Systems, Branson). Following centrifugation (13200 rpm 10 min at 4°C), the supernatant was incubated with 5 µM cysteine protease probe DCG-04 (Greenbaum, et al., 2000) at 4°C overnight. The incubated sample (120 µg total protein per lane) was mixed with Nupage 4X LDS sample buffer (Invitrogen) and 10X sample reducing agent (Invitrogen) and denatured at 95°C for 3 min. SDS-PAGE gel separation was performed with Xcell Sure Lock system (Invitrogen) on 10% Bis-Tris gel (Invitrogen) at 125V for 1.5 hr following transfer to a PVDF membrane at 30V (1.5 hr), the membrane was blocked in casein solution (Vector

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Laboratories) for 1 hr and then incubated with avidin and biotinylated horse-radish peroxidase, followed by washing (4X) and detection with West Pico chemiluminescent Substrate (Thermo Scientific) in LAS-3000 image reader (FujiFilm). Band intensity was quantified using the software Multi Gauge 3.0 (FujiFilm). To confirm the position of the Cat-L band, the membrane was later incubated with western blot striping buffer (Thermo Scientific) for 10 min and re-detected with anti-Cat-L rabbit pAb (Calbiochem) in 1:2000 dilution for 1hr. Goat anti-rabbit IgG peroxidase conjugate (Calbiochem) was used as the secondary antibody (1:8000 dilution). Protein loading was again checked using Mouse anti-β-Actin mAb (Sigma) with 1:2000 dilution followed by 1:5000 diluted goat anti-mouse IgG peroxidase conjugate (Pierce).

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Results

Kinetics of human cathepsin L inhibition.

CID 23631927 was found to inhibit cathepsin L with an IC_{50} of 6.9 ± 1.0 nM when assayed immediately after mixing with the enzyme. Upon pre-incubation with cathepsin L for a period of 1, 2, or 4 hr, the tetrahydroquinoline oxocarbazate demonstrated a time-dependent reduction of IC_{50} from 2.3 ± 0.1 nM (1-hr) to 1.2 ± 0.1 nM (2-hr) to 0.4 ± 0.1 nM (4-hr pre-incubation) (**Fig. 1C**), demonstrating a slow-on rate of inhibition.

Inhibitor reversibility was evaluated using a pre-incubation/dilution assay (Copeland, 2005). After pre-incubation of the enzyme with inhibitor at 10 x IC_{50}, most of the enzyme (>90%) is associated with the inhibitor as shown in **Fig. 2A**. Upon 100-fold dilution with buffer when the reaction was initiated with substrate, a total of 28% of enzymatic activity returned after 6000 s for the 1-hr pre-incubated reaction. For the case of 4-hr pre-incubation, the substrate cleavage rate after 8820 s was 8.9 times greater than the initial hydrolysis rate (**Fig. 2C**), indicating that the enzyme-inhibitor complex was very slowly reversible.

Transient kinetic data analysis was performed on the inhibition data for the oxocarbazate, CID 23631927. Given the close structural resemblance to the thiocarbazate, CID 16725315, the same inhibition model (**Fig. 3A and 3B**) was used for both inhibitors as reported previously (Shah, et al., 2008). MATLAB code was developed for the optimization and the fitted parameters for both inhibitors are presented in **Table 1** for comparison. The slow-on rate of inhibition is clearly apparent in **Fig. 3C,** as product formation rate slowly dropped in the assay. The best fit parameters for CID 23631927 were k_1 = 2.27 x 10^6 $M^{-1}s^{-1}$, k_{-1} = 0.29 s^{-1}, k_{cat} = 3.78 s^{-1}, k_{on} =

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

153,000 $M^{-1}s^{-1}$, and k_{off} = 4.40 x 10^{-5} s^{-1}. The five parameters, k_1, k_{-1}, k_{cat} and k_{off} do not very substantially between the thiocarbazate (CID 16725315) and the oxocarbazate (CID 23631927) tests. For the two slow on-rate inhibitors, the fitted kinetic parameter k_{on} for CID 23631927 was rate controlling displaying a rate almost 5X greater than that for CID 16725315, resulting in a calculated inhibition constant for the oxocarbazate of K_i = 0.29 nM. This calculated K_i value corresponds well with the IC_{50} value (0.4 nM) displayed after 4-hr preincubation of cathespin L with CID 23631927.

Cathespsin L/B Selectivity and Cytotoxicity

The IC_{50} of CID 23631927 against human liver cathepsin B was measured to be 5.072 μM ± 0.883 μM (*not shown*) with no pre-incubation, resulting in a cathepsin L/B selectivity ratio of 735, considerably higher than the 40-50 fold selectivity previously reported for the thiocarbazate CID 16725315 (Shah, et al., 2008). Prior work with SID26681509 indicated that the selectivity was not a function of time (Shah, et al., 2008). Importantly CID 23631927 proved to be non-toxic to human aortic endothelial cells up to 100 μM (*not shown*). Thus, the oxocarbazate inhibitor SID46493575 is one of the most potent and selective inhibitors of human cathepsin L reported to date.

Inhibition of virus pseudotype infection

CID 23631927 inhibited entry of both SARS-CoV (IC_{50} = 273 ± 49 nM) and Ebola virus (IC_{50} = 193 ± 39 nM) pseudotypes in cellular assays. Vesicular stomatitis virus glycoprotein

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

(VSV-G), which does not rely upon cathepsin L activity to infect host cells, was used as a control and was not inhibited by CID 23631927 (**Fig. 4**). While CID 23631927 revealed potency in blocking both SARS-CoV and Ebola virus pseudotype entry into 293T cells, CID 16725315 was found to have little inhibition for either SARS-CoV entry (IC_{50} > 10 μM) or Ebola viral entry (IC_{50} > 10 μM) (*not shown*). This difference was postulated to be the result of differing cell permeability properties of the compounds, thus motivating the assay of intracellular cathepsin L inhibition exploiting a labeling probe.

Inhibition of intracellular Cathepsin L

Cells were incubated with 4 different of cathepsin L inhibitors at 10 μM for 3 hr to permit time for each compound to permeate through cell membranes and interact with cathepsin L via inhibition of the active site of the enzyme. The chemical probe DCG-04, an epoxide electrophile with biotin tag (Greenbaum, et al., 2002) labels the active cathepsin L at the catalytic site. Any intracellular cathepsin L that had been inhibited during the 3 hr incubation would thus be unable to bind the probe DCG-04, especially given the slow reversibility of the carbazate inhibitors. In the western blot (**Fig. 5A**), band intensity was quantified with Multi Gauge 3.0 (FujiFilm) and correlated with the active cathepsin L activity remaining in the cell. The cathepsin L band position was verified in **Fig. 5B** and protein loading in each lane was checked using β-actin (**Fig. 5C**). The quantified results were normalized against Lane 1 (cells treated with DMSO only), and presented in **Fig. 5D**. The oxocarbazate inhibitor (CID 23631927) caused a 38 % reduction in the active intracellular cathepsin L level compared to Lane 1. The commercially available cysteine protease inhibitor MDL28170 revealed an IC_{50} against cathepsin L of 2.5nM and IC_{50} against SARS in the range of 100 nM (Simmons, et al., 2005). This positive control employing

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

MDL28170 in Lane 3 led to a 53% reduction of cathepsin L activity compared to Lane 1. The thiocarbazate inhibitor (CID 16725315) however did not affect cathepsin L level in the cell as illustrated in Lane 4, potentially indicating the poor cellular uptake of this inhibitor , that in turn correlates with the poor activity in the viral entry assay. As a postitive control, the cathepsin L inhibitor III (Calbiochem), a diazomethyl ketone (Lane 5) led to a 29% reduction of cathepsin-L activity compared to Lane 1.

Discussion

The oxocarbazate CID 23631927, designed based on the knowledge of the previously reported thiocarbazate chemotype (Myers, et al., 2008b), has been demonstrated to be a sub-nanomolar (K_i = 0.29 nM), slow-on, slow-off inhibitor of human cathepsin L that blocks SARS-CoV and Ebola pseudotype virus entry in human cells (at 273 nM and 193 nM IC50s, respectively) with minimal in vitro toxicity. Given the >700-fold cathepsin L/B selectivity of the oxocarbazate inhibitor, we suggest that inhibition of cathepsin L alone is sufficient to block Ebola pseudotype virus entry in HEK 293T cells without the need to simultaneously block cathepsin B.

Individual kinetic parameters were estimated through transient kinetic analyses on the reaction progress curves at various inhibitor concentrations. Compared to CID 16725315, CID 23631927 had substantially higher k_{on} values, resulting in a lower K_i. This improved rate of binding between CID 23631927 and cathepsin L may be attributed to the optimized tetrahydroquinoline structure which has both stronger hydrophobic interactions in the P1' subsite and a better hydrogen bond network with the receptor enzyme, as observed and predicted in the

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

docked pose for the ligand-enzyme complex (Beavers, et al., 2008). It is to be noted that the oxocarbazate and thiocarbazate inhibitors are likely covalent reversible inhibitors since replacement of the thiocarbazate sulfur in CID 16725315 results in a loss of activity (Myers, et al., 2008a). It is quite possible that the mechanism of inhibition involves a tetrahydral intermediate by attack of the active site Cys25 residue on the thiocarbazate carbonyl or oxocarbazate carbonyl. As it is difficult to predict temperature dependency of the IC_{50}, we directly measured it and found a small shift from IC_{50} = 5.9 ± 0.6 nM (at 25°C, n=3) to IC_{50} = 10.6 ± 0.9 nM (at 37°C, n=3) (data not shown).

Endosomal cathepsins play important roles in mediating virus entry into host cells for SARS-CoV (Huang, et al., 2006) and Ebola virus (Schornberg, et al., 2006). In particular, Cathepsin L has been singled out as possibly playing a crucial role in triggering membrane fusion within endosomes via proteolysis (Simmons, et al., 2005; Kaletsky, et al., 2007). In the present study, cathepsin L inhibitor CID 23631927 was demonstrated to be effective in blocking the virus entry in both SARS and Ebola pseudotype infection. This result validates the possibility of preventing the invasion and spread of such diseases using cathepsin L inhibitors and as such establishes this oxacarbazate as a viable candidate for further optimization.

The design of CID 23631927 entailed substituting a tetrahydroquinoline anilide moiety for the 2-ethylphenyl anilide group in CID 16725315. This structural modification improved ligand-protein interaction in the S1' pocket as demonstrated in molecular docking studies (Beavers, et al., 2008). In addition, we hypothesize that entropic considerations contributed to increased potency. Using the ADME property prediction software Qikprop 3.1 (Schrodinger LLC, New York, NY), we found CID 23631927 to have a predicted log value for the partition

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

coefficient, cLogP(octanol/water) to be 5.296. This cLogP is higher than that of CID 16725315 (4.724). This estimate of cLogP is based upon the condition that the inhibitor is uncharged, a condition that is likely upheld since the pKa = -3.6 (Hur and Guven, 2002) for the indole group of the oxocarbazate and thiocarbazate inhibitors. The predicted increase in hydrophobicity may result in higher permeability of CID 23631927 into cells, hence resulting in better potency in both the virus infection assay (Fig. 4) and intracellular inhibition of cathepsin L (Fig. 5). The poor activity of the thiocarbazate in the cell based viral entry assay may be partially explained by the labeling assay in which thiocarbazate treatment had little effect in blocking intracellular cathepsin L activity. The t-BOC group on the inhibitors is a common amino protecting group that we have found prevent self-cleavage of the thiocarbazate (Myers, et al., 2008b). We note that this protecting group may be unstable under acidic conditions of the GI tract and the oral availability of the compound, with or without the t-BOC protecting group, is unknown.

Taken together the results of this study demonstrating the efficacy of CID 23631927 in blocking virus entry into human cells in conjunction with the structure of CID 23631927 offers both a new starting point for intervention in SARS and Ebola virus infection and a novel probe to explore the mechanism of virus entry.

Acknowledgment

We thank Dr. Rajesh Chandramohanadas for help in DCG04 activity based probe labeling work.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

References

Beavers MP, Myers MC, Shah PP, Purvis JE, Diamond SL, Cooperman BS, Huryn DM and Smith,Amos B.,,III (2008) Molecular docking of cathepsin L inhibitors in the binding site of papain. *Journal of Chemical Information and Modeling* 48:1464-1472.

Chandran K, Sullivan NJ, Felbor U, Whelan SP and Cunningham JM (2005) Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* 308:1643-1645.

Chapman HA, Riese RJ and Shi GP (1997) Emerging roles for cysteine proteases in human biology. *Annu Rev Physiol* 59:63-88.

Copeland RA (2005) *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*. J. Wiley, Hoboken, NJ.

Enserink M (2003) SARS: A pandemic prevented. *Science* 302:2045-2045.

Feldmann H, Jones S, Klenk HD and Schnittler HJ (2003) Ebola virus: from discovery to vaccine. *Nature Reviews Immunology* 3:677-685.

Greenbaum D, Medzihradszky KF, Burlingame A and Bogyo M (2000) Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem Biol* 7:569-581.

Greenbaum DC, Baruch A, Grainger M, Bozdech Z, Medzihradszky KF, Engel J, DeRisi J, Holder AA and Bogyo M (2002) A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* 298:2002-2006.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Huang IC, Bosch BJ, Li F, Li WH, Lee KH, Ghiran S, Vasilieva N, Dermody TS, Harrison SC, Dormitzer PR, Farzan M, Rottier PJM and Choe H (2006) SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. *J Biol Chem* 281:3198-3203.

Hur D and Guven A (2002) The acidities of some indoles. Journal of Molecular Structure-Theochem 583:1-18.

Kaletsky RL, Simmons G and Bates P (2007) Proteolysis of the Ebola virus glycoproteins enhances virus binding and infectivity. *J Virol* 81:13378-13384.

Lee JE, Fusco ML, Hessell AJ, Oswald WB, Burton DR and Saphire EO (2008) Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 454:177-U27.

Maehr R, Mintern JD, Herman E, Lennon-Dumenil AM, Mathis D, Benoist C and Ploegh HL (2005) Cathepsin L is essential for onset of autoimmune diabetes in NOD mice. *J Clin Invest* 115:2934-2943.

McGrath ME (1999) The lysosomal cysteine proteases. *Annu Rev Biophys Biomol Struct* 28:181-204.

Myers MC, Shah PP, Beavers MP, Napper AD, Diamond SL, Smith AB and Huryn DM (2008a) Design, synthesis, and evaluation of inhibitors of cathepsin L: Exploiting a unique thiocarbazate chemotype. *Bioorg Med Chem Lett* 18:3646-3651.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Myers MC, Shah PP, Diamond SL, Huryn DM and Smith AB (2008b) Identification and synthesis of a unique thiocarbazate cathepsin L inhibitor. *Bioorg Med Chem Lett* 18:210-214.

Pager CT and Dutch RE (2005) Cathepsin L is involved in proteolytic processing of the Hendra virus fusion protein. *J Virol* 79:12714-12720.

Palermo C and Joyce JA (2008) Cysteine cathepsin proteases as pharmacological targets in cancer. *Trends Pharmacol Sci* 29:22-28.

Potts W, Bowyer J, Jones H, Tucker D, Freemont AJ, Millest A, Martin C, Vernon W, Neerunjun D, Slynn G, Harper F and Maciewicz R (2004) Cathepsin L-deficient mice exhibit abnormal skin and bone development and show increased resistance to osteoporosis following ovariectomy. *Int J Exp Pathol* 85:85-96.

Rossi A, Deveraux Q, Turk B and Sali A (2004) Comprehensive search for cysteine cathepsins in the human genome. *Biol Chem* 385:363-372.

Schedel J, Seemayer CA, Pap T, Neidhart M, Kuchen S, Michel BA, Gay RE, Muller-Ladner U, Gay S and Zacharias W (2004) Targeting cathepsin L (CL) by specific ribozymes decreases CL protein synthesis and cartilage destruction in rheumatoid arthritis. *Gene Ther* 11:1040-1047.

Schornberg K, Matsuyama S, Kabsch K, Delos S, Bouton A and White J (2006) Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J Virol* 80:4174-4178.

Shah PP, Myers MC, Beavers MP, Purvis JE, Jing H, Grieser HJ, Sharlow ER, Napper AD, Huryn DM, Cooperman BS, Smith AB and Diamond SL (2008) Kinetic characterization and

molecular docking of a novel, potent, and selective slow-binding inhibitor of human cathepsin L. *Mol Pharmacol* 74:34-41.

Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL and Bates P (2005) Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc Natl Acad Sci U S A* 102:11876-11881.

Turk D and Guncar G (2003) Lysosomal cysteine proteases (cathepsins): promising drug targets. *Acta Crystallogr Sect D-Biol Crystallogr* 59:203-213.

Turk V, Turk B and Turk D (2001) Lysosomal cysteine proteases: facts and opportunities. *EMBO J* 20:4629-4633.

Vasiljeva O, Reinheckel T, Peters C, Turk D, Turk V and Turk B (2007) Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Curr Pharm Des* 13:387-403.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Footnotes

This work was supported by National Institute of Health [U54-HG003915 (S.L.D.), T32-AI55400 (R.L.K), U01-AI070369 (P.B), U54-AI057168 (P.B)].

Please submit requests for reprints to:

Scott L. Diamond

Penn Center for Molecular Discovery

University of Pennsylvania

1024 Vagelos Laboratories

Philadelphia, PA 19104-6383

USA

E-mail: sld@seas.upenn.edu

[1]These authors contributed equally to this work.

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Legends for figures

Figure 1. ***A***, Thiocarbazate CID 16725315. ***B***, Oxocarbazate CID 23631927. ***C***, Activity of CID 23631927 against human cathepsin L after preincubation with the enzyme target for 0 h (○), 1 h(Δ), 2 h (□), and 4 h (▲).

Figure 2. ***A,*** dilution protocol for determination of reversibility: 870 ng/ml of cathesin L and 25 nM of the inhibitor were combined and incubated at room temperature. ***B.*** Reversibility data for CID 23631927 after 0 min (○), 15 min (Δ), 1 hr (□), and 4 hr (▲) pre-incubation with cathepsin L and upon 100-fold dilution into assay buffer containing 1 μM Z-Phe-Arg-AMC as substrate and the final inhibitor concentration was at 0.25 nM. A full enzyme-substrate reaction without inhibitor (●) served as a positive control. ***C.*** Enlarged reaction progress curve with 4 hr pre-incubation of cathepsin L and CID 23631927 to clearly demonstrate the recovering rate of enzymatic cleavage of the substrate.

Figure 3. ***A,*** Single-step mechanism for simple, reversible, slow binding inhibition. ***B,*** ordinary differential equations governing the single-step mechanism of inhibition. ***C,*** Nonlinear regression for transient dynamics of CID 23631927 using code written in MATLAB to fit in a five-parameter inhibition kinetic model. Reaction progress curves (●) are shown for 8.7 ng/mL human cathepsin L enzyme and 1 μM Z-Phe-Arg-AMC substrate with varying concentrations of CID 23631927 inhibitor.

Figure 4. CID 23631927 inhibition of SARS-CoV and Ebola virus pseudotype infections. CID 23631927 inhibits SARS-CoV pseudotype entry in to cells at IC_{50} = 273 ± 49 nM and Ebola

Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

virus pseudotype at IC_{50} = 193 ± 39 nM; Vesicular stomatitis virus glycoprotein pseudovirions were used as a control and was not inhibited

Figure 5. Western blot to detect amount of intracellular active Cat-L after treatment with 10 μM inhibitors ***A.*** Intracellular cysteine protease labeled by DCG-04 was detected at about 30 kDa in western blot. Lane 0, papain (21kDa), a purified cysteine protease served as a positive control for DCG-04 probe. Lane 1, treated with DMSO only without inhibitor; Lane 2, Oxocarbazate, CID 23631927; Lane 3, MDL28170; Lane 4, Thiocarbazate, CID 16725315; Lane 5, Cat-L inhibitor III. ***B.*** This band was subsequently confirmed to be cathepsin L by anti-cathepsin L rabbit polyclonal antibody. ***C.*** Protein Loading in each lane was checked using β-Actin (MW=42kDa) as a control. ***D.*** Statistical Analysis of repeated Western blots was done using Multi Gauge 3.0 (FujiFilm). Data are presented as mean±Standard Error (n=3)**.**

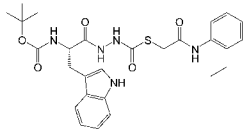
Molecular Pharmacology Fast Forward. Published on May 13, 2010 as DOI: 10.1124/mol.110.064261
This article has not been copyedited and formatted. The final version may differ from this version.

Table 1. Comparison of kinetic parameters and inhibition potency data between CID 16725315 and CID 23631927

	Thiocarbazate CID 16725315	**Oxocabazate CID 23631927**
k_1	3.09 x 10^6 $M^{-1}s^{-1}$	2.27 x 10^6 $M^{-1}s^{-1}$
k_{-1}	0.32 s^{-1}	0.29 s^{-1}
k_{cat}	4.31 s^{-1}	3.78 s^{-1}
k_{on}	3.27 x 10^4 $M^{-1}s^{-1}$	1.53 x 10^5 $M^{-1}s^{-1}$
k_{off}	3.75 x 10^{-5} s^{-1}	4.40 x 10^{-5} s^{-1}
$K_i = k_{off}/k_{on}$	1.15 nM	0.29 nM
Human cathepsin L IC50 (0 hr)	56 ± 4 nM	6.9 ± 1.0 nM
Human cathepsin L IC50 (4 hr)	1.0 ± 0.5 nM	0.4 ± 0.1 nM
SARS IC50	-	273 ± 49 nM
Ebola IC50	-	193 ± 39 nM
%Inhibition of intracellular Cat L.	2	38
Cat L/B Selectivity	40-50	735

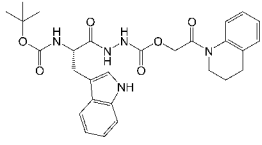
Figure 1

A



CID 16725315

B



CID 23631927

C

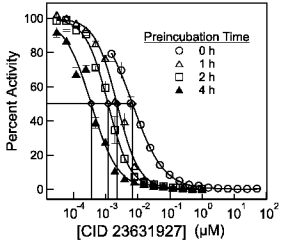


Figure 2

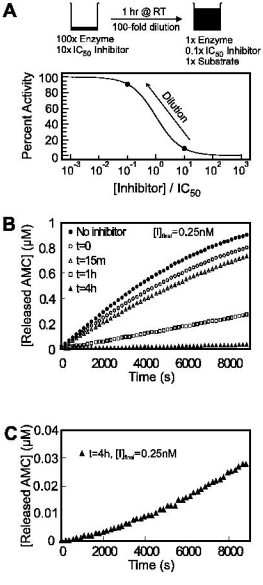
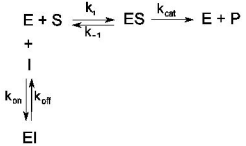

A
1 hr @ RT
100-fold dilution
100x Enzyme
10x IC50 Inhibitor
1x Enzyme
0.1x IC50 Inhibitor
1x Substrate
Percent Activity
Dilution
[Inhibitor] / IC50
B
No inhibitor
[I]final=0.25nM
t=0
t=15m
t=1h
t=4h
[Released AMC] (μM)
Time (s)
C
t=4h, [I]final=0.25nM
[Released AMC] (μM)
Time (s)

Figure 3

A



B

$$\frac{d[E]}{dt} = k_{off}[EI] + k_{-1}[ES] - k_{on}[E][I] - k_1[E][S] + k_{cat}[ES]$$

$$\frac{d[S]}{dt} = k_{-1}[ES] - k_1[E][S]$$

$$\frac{d[I]}{dt} = k_{off}[EI] - k_{on}[E][I]$$

$$\frac{d[P]}{dt} = k_{cat}[ES]$$

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] - k_{cat}[ES]$$

$$\frac{d[EI]}{dt} = k_{on}[E][I] - k_{off}[ES]$$

C

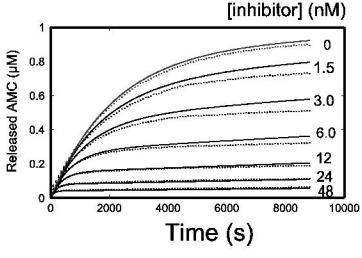


Figure 4

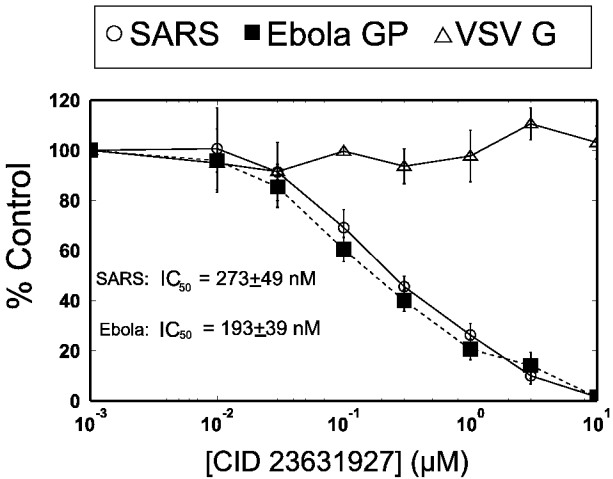
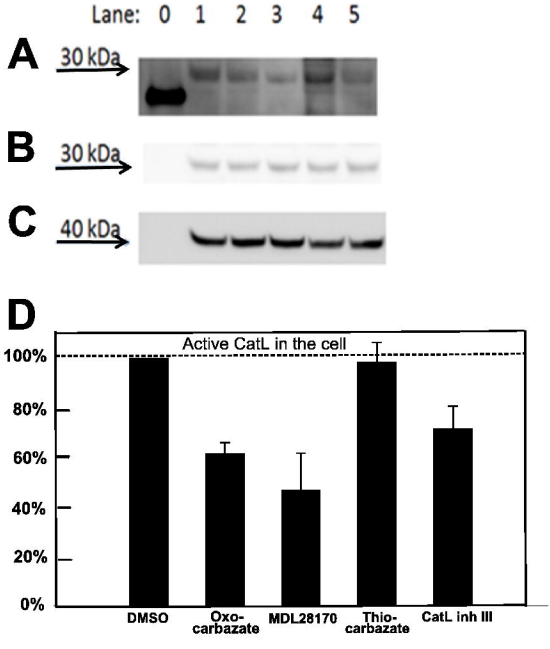

○ SARS ■ Ebola GP △ VSV G
% Control
120
100
80
60
40
20
0
SARS: IC50 = 273±49 nM
Ebola: IC50 = 193±39 nM
10-3
10-2
10-1
100
101
[CID 23631927] (μM)

Figure 5


Lane: 0 1 2 3 4 5
A
30 kDa
B
30 kDa
C
40 kDa
D
Active CatL in the cell
100%
80%
60%
40%
20%
0%
DMSO
Oxo-carbazate
MDL28170
Thio-carbazate
CatL inh III