HHS Public Access

Author manuscript

Bioorg Med Chem Lett. Author manuscript; available in PMC 2020 January 01.

Published in final edited form as:

Bioorg Med Chem Lett. 2019 January 01; 29(1): 36–39. doi:10.1016/j.bmcl.2018.11.019.

Discovery of Covalent Enzyme Inhibitors Using Virtual Docking of Covalent Fragments

Sandipan Roy Chowdhury[a], **Steven Kennedy**[a], **Kai Zhu**[b], **Rama Mishra**[c], **Patrick Chuong**[a], **Alyssa-uyen Nguyen**[a], **Stefan G. Kathman**[d], and **Alexander V. Statsyuk**[a]

[a]Department of Pharmacological and Pharmaceutical Sciences, College of Pharmacy, University of Houston, Houston, Texas 77204-5037, USA

[b]Schrödinger Inc., 120 West 45th Street, New York, New York 10036, USA

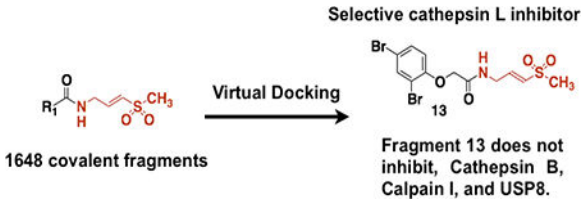
[c]Center for Molecular Innovation and Drug Discovery, Northwestern University, Evanston, IL 60208, USA

[d]Department of Chemistry, Northwestern University, 2170 Campus Drive, Evanston, Illinois 60208, USA.

Abstract

Here we present a virtual docking screen of 1648 commercially available covalent fragments, and identified covalent inhibitors of cysteine protease cathepsin L. These inhibitors did not inhibit closely related protease cathepsin B. Thus, we have established virtual docking of covalent fragments as an approach to discover covalent enzyme inhibitors.

Graphical Abstract



Covalent fragments is a technology useful to discover covalent probes for proteins *in vitro* and *in cellulo.*[1,2] There are some advantages of covalent fragments when compared to reversible fragments. First, it is easier to obtain a crystal structure of the protein with covalently bound fragment for structural optimization.[3] Second, intracellular targets of covalent fragments can be easily identified using modern click chemistry methods.[2] In recent years computational algorithms have been used as screening methods in drug discovery, but they face unique challenges when applied to fragments.[4] Reversible fragments are often promiscuous due to their small size and weak binding affinity to their protein target

(K_i μM-mM range) and display dynamic binding modes to the protein targets. This makes it challenging to screen fragments using computational approaches and to accurately predict fragment binding modes. Thus, ligand docking algorithms that have been developed to describe the binding of drug-like compounds to protein targets, may not perform well for fragments, and have to be tested separately.[4] Covalent fragments on the other hand have to adopt a specific geometric orientation at the ligand binding site for the covalent reaction, which leads to reduced promiscuity and more defined binding modes. Additional $n\rightarrow\pi^*$ interactions between the covalent fragment and the nucleophile at the protein surface may further enhance binding interactions between the covalent fragment and the protein.[5] However it is currently not clear if covalent docking algorithms designed for covalent drug-like inhibitors can be used to predict the binding mode of covalent fragments and their growth vectors. Attempts to address this challenge have been done for covalent reversible cyanoacrylamide fragments, but not for irreversible covalent fragments that require stringent design rules to avoid non-specific covalent labeling.[6] This paper addresses this challenge by showing that covalent docking algorithms can be used to identify covalent fragments that inhibit drug target of interest. In doing so we designed a small 1648-member library of vinylsulfones, conducted virtual docking, and discovered covalent fragments that inhibit cysteine protease cathepsin L but not related protease cathepsin B, calpain I, or deubiquitinating enzyme USP8.

First, we asked if the virtual docking of covalent fragments can correlate with the experiment. In our earlier studies we tested a 100-member library of methyl acrylates with balanced reactivity against cysteine protease papain and identified three covalent fragments that inhibited papain.[1] Here we used Shrodinger CovDock program to dock a library of the same 100 fragments to the papain active site cysteine (PDB:1KHQ). In addition, we also screened three previously reported peptide papain inhibitors (Figure 1A-B).[1] The docking procedure used CovDoc (Schrodinger 2018–2 version) with the distance cutoffs (8Å Cβ to the ligand, 5Å sulfur to the ligand bonding carbon) to decide if the covalent reaction can happen or a candidate pose should be kept for further consideration. From six known papain inhibitors **1–6**, four papain inhibitors **1**,**2**,**4**, and **6** scored favorably (Figure 1A, top 20 hits). Overall, we observed a correlation between ranking and reported k_{inact}/K_i values: two most potent inhibitors **1** and **4** (highest k_{inact}/K_i values) received highest rankings in CovDock. Docking poses of compounds **1**, **2**, **4** and **6** showed that they form a hydrogen bond with the backbone amide carbonyl of Asp^{158}. Similar hydrogen bond between Asp^{158} and the glycine of ZLFG-DAM (DAM – diazomethyl) is observed in the X-Ray crystal structure of the covalent complex of papain and its covalent inhibitor.[7] Compounds **1**, **4**, and **6** also formed hydrogen bonds with Gly^{66}, however these hydrogen bonds seem to be not essential in the context of CovDock. This is because compound **2** did not form hydrogen bonds with Gly^{66} (because it is an enantiomer of **1**), yet received high score and inhibited papain under experimental conditions. Nevertheless, a hydrogen bond between Gly^{66} and bound covalent papain inhibitors has been observed in X-ray crystal structures, and perhaps could be exploited in virtual docking. Two other compounds **3** and **5** did not receive favorable docking scores and were ranked 96^{th} and 48^{th} respectively (Figure S1).

Next, we performed covalent docking of 1648 vinylsulfone-based covalent fragments against cysteine protease cathepsin L. Cathepsin L is well validated drug target to treat cancer, osteoporosis and autoimmune disorders, and selective covalent inhibitors of cathepsin L are known.[8] Initially we faced standard challenges of a) selecting electrophile which is not hyper reactive, and b) figuring out how to link fragments and electrophile in the library. To solve these challenges, we used the design rules (or guidelines) for covalent fragment libraries that we outlined earlier.[9]

We turned our attention to vinylsulfone class of covalent cysteine protease inhibitors (Figure 2A). Vinyl sulfones are frequently used to inhibit cysteine proteases, they are ten-fold more reactive toward thiols than methyl acrylates, and have different geometry when compared to acrylates (tetragonal sulfur atom vs sp^2-hybridized carbon).[1] Thus, by conducting virtual docking on vinylsulfones we could explore different covalent fragment chemotypes, and reactivity. Specifically, we analyzed previously reported vinyl sulfones **7** and **8**, which inhibit cathepsin B and L.[10] These compounds have a general formula R^1-C(=O)-NH-CH_2-CH=CH-SO_2-Ph, and would be considered prototypes for covalent fragments of this type. Compound **7** is more selective at inhibiting cathepsin L as judged by K_i values. Structural modification of R^1 in **7**, led to compound **8**, which showed reversed selectivity (K_i values) and increased potency (k_{inact}/K_i values) for cathepsin B. Importantly, for both compounds screened against cathepsin B and L k_{inact} stayed within 0.17–0.36 s^{-1} range, suggesting that structural modifications of R^1 group do not significantly increase vinylsulfone electrophile reactivity. This indicates that structural modifications of R^1 group can change inhibitor selectivity and potency. Therefore, it is rational to assume that a library of covalent fragments of general formula R^1-C(=O)-NH-CH_2-CH=CH-SO_2-Ph, where R^1- is a variable fragment, will display similar variability in selectivity and potency at least when screened against cathepsins or related proteases. Given that the active site cysteine in cathepsin L is highly reactive (pK_a ~3.5) most likely covalent fragment libraries such as R^1-C(=O)-NH-CH_2-CH=CH-SO_2-Ph will display similar variations in potency and selectivity when screened against less reactive cysteines with pK_a >3.5. Our own experience with acrylate based covalent fragment libraries confirm these predictions.[3]

One of the key features of fragments is their low molecular weight. Covalent fragments R^1-C(=O)-NH-CH_2-CH=CH-SO_2-Ph contain phenyl group attached to sulfone, and this group adds extra molecular weight. We decided to reduce the molecular weight by replacing the phenyl group with the methyl group, which led to the general formula R^1-C(=O)-NH-CH_2-CH=CH-SO_2-CH_3. Inhibitors of this type display different selectivity/potency toward cathepsins, and therefore it is safe to design the library of covalent fragments of this type.[10] We also expected that the intrinsic reactivity of covalent fragments R^1-C(=O)-NH-CH_2-CH=CH-SO_2-CH_3 will not be significantly affected by R_1 group based on our previous kinetic studies (Figure 2B).[1] All these considerations suggest that a library of covalent fragments R^1-C(=O)-NH-CH_2-CH=CH-SO_2-CH_3 will not have hyper reactive fragments, and R_1- group will be determining the potency and selectivity of these fragments against drug targets. Because every member of this library can be rapidly synthesized from the carboxylic acid R^1-COOH and corresponding amine NH_2-CH_2-CH=CH-SO_2-CH_3, we

constructed a small 1648-member virtual library of covalent fragments for which carboxylic acids were commercially available.

To dock 1648 vinylsulfone ligands, we used human cathepsin L apo form crystal structure (4AXL) with 1.92Å resolution.[11] Cys^{25} is the active site residue of cathepsin L. We first carried the non-covalent docking flowed by minimizing the covalent complex in Optimized Potential for Liquid Simulation (OPLS) force field. The complex was minimized and the average RMSD between the minimized and non-minimized ligand was found to be in between 1.5 to 2Å. This step was required to rank ligands and prioritize the compounds for synthesis. Ligand docked poses and their energetics were analyzed and out of 1648 ligands, 33 were showed good docking poses, lower complex energies, and which had the highest docking scores ≥ 6.0. Out of 33 ligands, we selected five compounds for synthesis based on pricing and structural diversity. Upon further experimental testing compound **11** (racemic mixture) showed time and dose dependent inhibition of cathepsin L (k_{inact} = 0.0006 s^{-1}, K_i = 146 ± 17 μM) (Figure S2). Other compounds had solubility issues. Thus, we proceeded to investigate compound **11** and its analogues.

The docking pose of compound **11** resembled those of many other covalent inhibitors of cysteine proteases (Figure 3A-B). Methylvinyl sulfone moiety forms covalent bond with the active site cysteine of cathepsin L at the S1 pocket, and sulfone moiety interacts with Q19, H163, and W189 of cathepsin L. The amide bond functionality in **11** formed two hydrogen bonds with D162 and G68 respectively (shown with dashed lines). Glycolic acid moiety and dibromophenyl group extended in the direction of S2 pocket of cathepsin L. There 2,4-dibromophenyl group formed hydrophobic contacts with L69 and M70. The key feature of fragments is the ability to have growth vectors. We overlaid crystal structures of known cathepsin L inhibitor **12** (IC_{50} 22 nM) and compound **11** bound to cathepsin L.[12] Compound **12** is a reversible covalent nitrile-inhibitor with two chlorophenyl moieties projecting toward S2 and S3 pockets respectively, where they form a combination of hydrophobic and halogen bond contacts. This structural analysis suggests that compound **11** may have potentially two growth vectors: one extending from dibromophenyl moiety toward S2 pocket, and one from the methyl group extending toward S3 pocket of cathepsin. Both of these vectors are relatively far from the vinylsulfone electrophile and chemical modifications at those sites should not inhibit the covalent labeling of the catalytic cysteine.

Compound **11** has a methyl group at the chiral center that we suggested can be used as a potential growth vector. To test if this methyl group provides critical contribution to binding and covalent labeling of cathepsin L, we prepared compound **13**, which lacks this methyl group. This step also allowed us to eliminate the chiral center, since the corresponding acid to make **11** is available only in its racemic form. Compound **13** inhibited cathepsin L with k_{inact}/K_i values of 5 $M^{-1}s^{-1}$ (k_{inact} = 0.0003 s^{-1}, K_i =59 ± 10 μM, Figure 4). Thus, the methyl group is not critical for fragment binding, and chemical modification at that position may be used to grow the fragment. Importantly, compound **13** shifts the K_m of the cathepsin L substrate (Figure 4A) indicating that compound **13** is a substrate competitive inhibitor under these reaction conditions.[13] Compound **13** did not inhibit other classes of cysteine proteases such as Cathepsin B, Calpain I, and USP8 with and without preincubation of the inhibitor (Figure S3). Finally, gel filtration experiments have shown that compound **13**

inhibits cathepsin L irreversibly (Figure 4D). Taken together, our experiments show that virtual docking of covalent fragment libraries can be used to discover covalent fragments that can act as selective enzyme inhibitors. Judicious choice of electrophiles and their proper connection to fragments is needed to avoid hyper reactive hits.

Our initial proof of concept studies validates virtual docking methods for covalent fragment screening, outlines the path forward to construct larger libraries of covalent fragments and to discover covalent fragment leads for a wider range of nucleophilic drug targets. The discovered covalent inhibitor **13** of Cathepsin L have weak binding to Cathepsin L (K_i =59 ± 10 μM) and slow k_{inact} (0.0003 s^{-1}), k_{inact}/K_i 5 $M^{-1}s^{-1}$. How far are these values from those of pharmacologically useful probes? It is already encouraging that compound **13** displays selectivity and does not inhibit related cysteine proteases Cathepsin B, Calpain I, and USP8. Recently covalent K-Ras^{G12C} inhibitor ARS-853 has been discovered that has weak binding affinity to K-Ras^{G12C} (K_i 200 ± 90 μM) but fast k_{inact} 0.05 s^{-1} (k_{inact}/K_i 250 ± 20 $M^{-1}s^{-1}$) due to the electrophile activation by K-Ras.[14] When profiled against 2,740 cellular cysteines this inhibitor displayed remarkable selectivity and covalently modified Cys^{12} of K-Ras and only two other proteins: FAM213A and Reticulon-4.[15] Further improvement of ARS-853 has led to ARS-1620 (k_{inact}/K_i 1100 ± 300 $M^{-1}s^{-1}$) active *in vivo*.[16] Thus, it is rational to propose that improving k_{inact}/K_i values of compound **13** to ~250 $M^{-1}s^{-1}$ may yield useful covalent probes with cellular activity. This can be achieved by improving the K_i of the fragment and/or by improving the $k_{inact.}$ Further use of virtual docking of covalent fragments to discover covalent enzyme inhibitors will be reported in the future.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

This work was supported in part by National Institutes of Health Awards R01GM115632 (to A.V.S.) and T32GM105538 (to S.G.K.), William & Ella Owens Research Foundation, and the University of Houston. A.V.S. is a Pew Scholar in the Biomedical Sciences, supported by the Pew Charitable Trusts. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors declare that they have no conflicts of interest with the contents of this article.

REFERENCES

(1). Kathman SG, Xu Z, Statsyuk AV J. Med. Chem 2014; 57: 4969–4974. [PubMed: 24870364]
(2). Backus KM, Correia BE, Lum KM, et al. Nature; 2016: 534: 570–574. [PubMed: 27309814]
(3). Kathman SG, Span I, Smith AT et al. *J. Am. Chem. Soc*. 2015; 137: 12442–12445. [PubMed: 26371805]
(4). Steinbrecher TB, Dahlgren M, Cappel D, et al. J. Chem. Inf. Model 2015; 55: 2411–2420. [PubMed: 26457994]
(5). Newberry RW, Raines RT Acc. Chem. Res 2017; 50: 1838–1846. [PubMed: 28735540]
(6). London N, Miller RM, Krishnan S et al. *Nat. Chem. Biol*. 2014; 10:1066–1072. [PubMed: 25344815]
(7). Janowski R, Kozak M, Jankowska E, et al. *J. Pept. Res*. 2004; 64: 141–150. [PubMed: 15357669]
(8). Li YY; Fang J; Ao GZ *Expert. Opin. Ther. Pat*. 2017; 27: 643–656. [PubMed: 27998201]

(9). Kathman SG; Statsyuk AV Medchemcomm 2016; 7: 576–585. [PubMed: 27398190]
(10). Palmer JT, Rasnick D, Klaus JL et al. *J. Med. Chem.* 1995; 38: 3193–3196. [PubMed: 7650671]
(11). Ehmke V, Winkler E, Banner DW, et al. Chemmedchem 2013, 8: 967–975. [PubMed: 23658062]
(12). Hardegger LA, Hardegger LA, Berndt K et al. *Angew. Chem. Int. Edit.* 2011; 50: 314–318.
(13). Tian WX, Tsou CL Biochemistry 1982; 21: 1028–1032. [PubMed: 7074045]
(14). Hansen R, Peters U, Babbar A et al. *Nat. Struct. & Mol. Biol.* 2018; 25: 454–462. [PubMed: 29760531]
(15). Patricelli MP, Janes MR, Li LS et al. Cancer Discov 2016; 6 :316–29. [PubMed: 26739882]
(16). Janes MR, Zhang J, Li LS, Cell 2018; 172 :578–589. [PubMed: 29373830]

Highlights

- Virtual docking of 1648 covalent fragments is performed.
- Covalent inhibitor of Cathepsin L is identified.
- Cathepsin L inhibitor did not inhibit USP8, Calpain I, Cathepsin B.

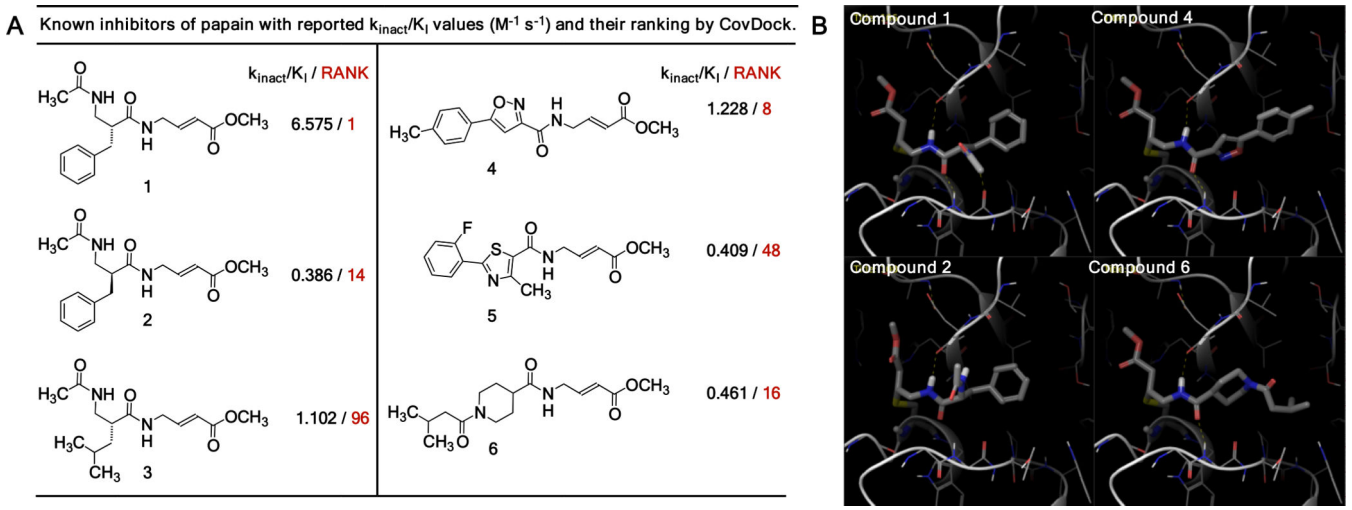


Figure 1.

Covalent docking of known inhibitors of papain using Shrodinger CovDock at the active site cysteine (Cys^{25}) of papain (1KHQ:1.6 Å). A) Known inhibitors of papain with reported k_{inact}/K_i values and their rankings in CovDoc. B) Docking poses of compounds **1, 2, 4,** and **6** at the papain active site.

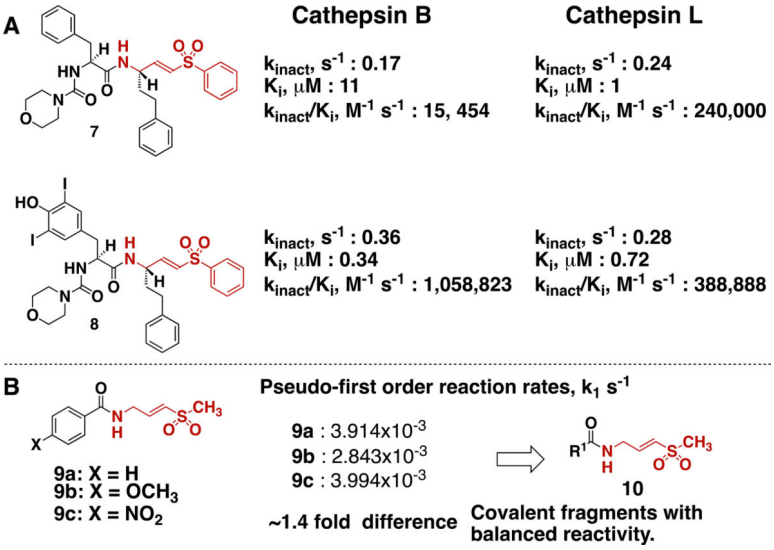


Figure 2.

Selection of electrophile for covalent fragment library. A) k_{inact} and K_i values of previously reported inhibitors of cathepsin B and L, suggesting that vinyl sulfones are safe electrophiles. B) Pseudo-first-order reaction rates of vinylsulfones **9a-c** with N-acetylcysteine methylester at pD 8.0 as measured by NMR spectroscopy.

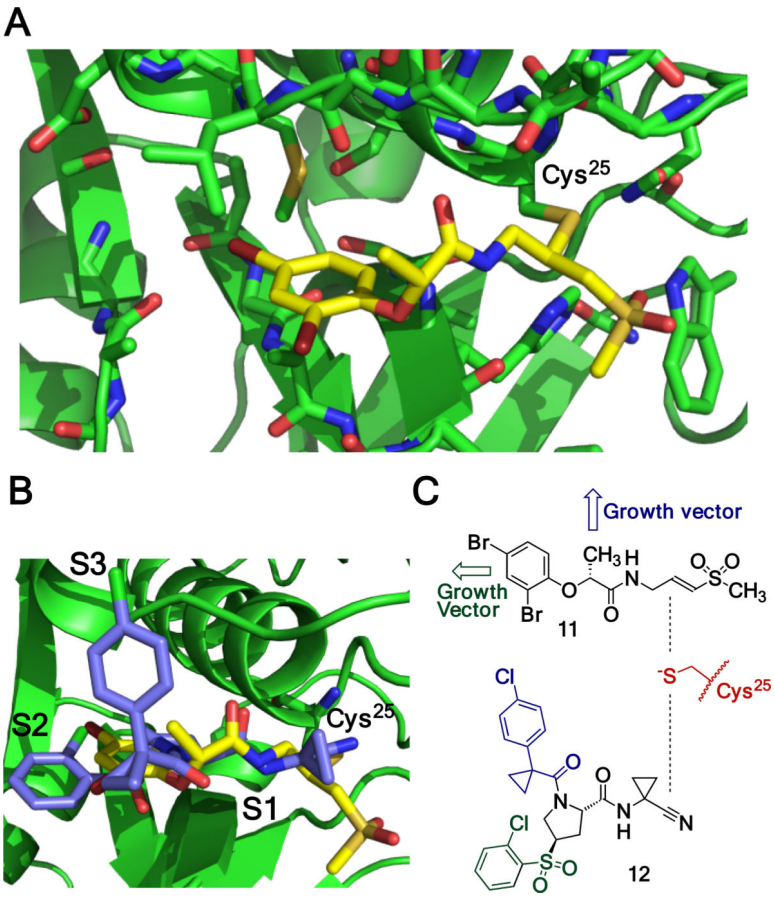


Figure 3.

Compound **11** is a covalent inhibitor of cathepsin L. A) Docked pose of compound **11**. B) Structural overlay of compound **11** (yellow) and nitrile-based inhibitor **12** (green) bound to cathepsin L active site cysteine (Cys^{25}, PDB-2xu1). C) Analysis of compound **11** and nitrile **12**, showing potential growth vectors for covalent fragment **11**.

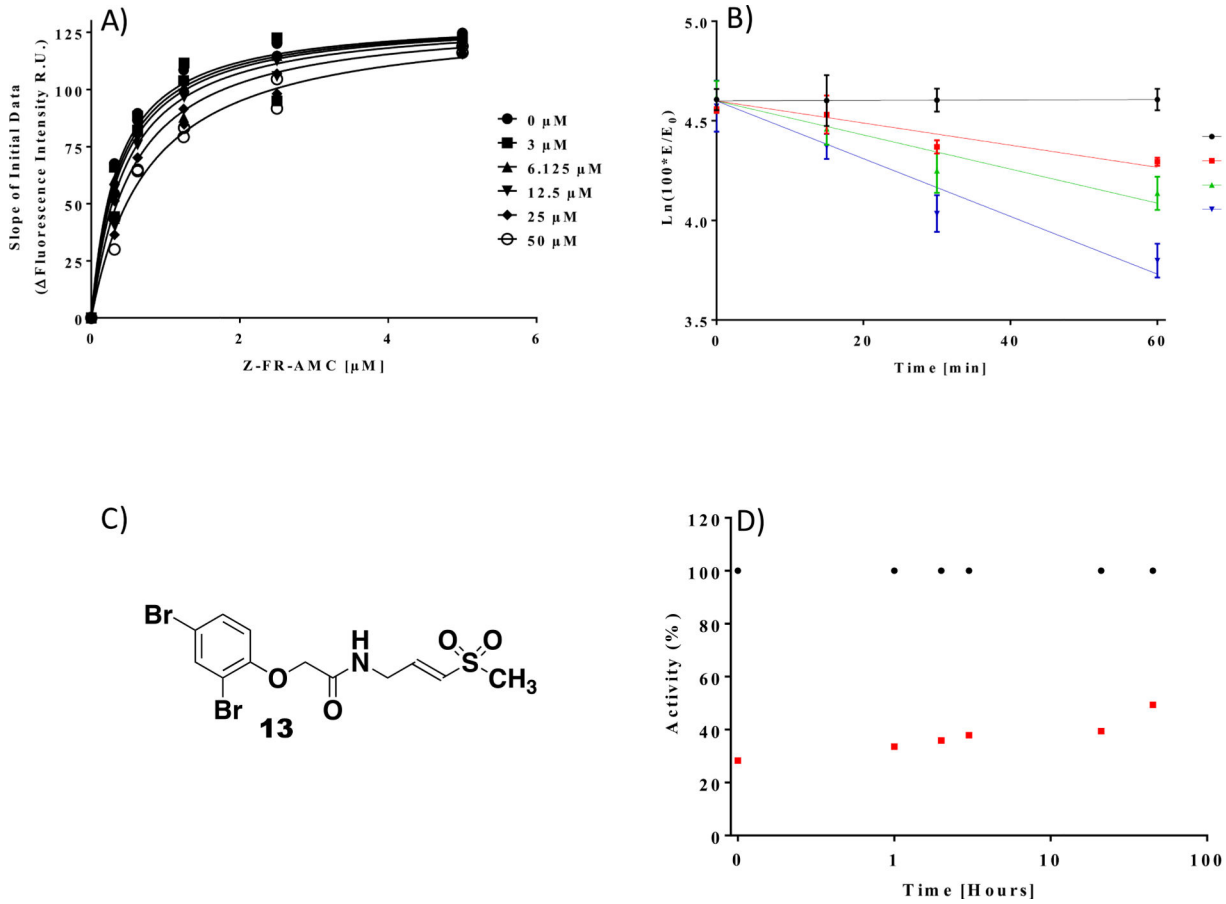


Figure 4.

Determination of K_i, k_{inact} and reversibility studies. A) K_m determination for Cathepsin L with its substrate in the presence of varying concentrations of compound **13** shows dose dependent inhibition of the enzyme and is used to determine a K_i value of 59 ± 10 μM. (B) Pseudo-first-order inhibition plots of compound **13** at varying concentrations for 15, 30, or 60 minute preincubation time. (C) The Structure of Compound 13 (D) Irreversibility studies. Compound **13** (50 μM) was preincubated with Cathepsin L (10 nM), followed by zeba spin column. Flow through aliquots were tested for enzyme activity recovery at different time points. ● DMSO controls, □ compound **13**.