Development of Cell-Active Inhibitors and Activity-Based Probe of Cysteine Cathepsins

Dibyendu Dana
Graduate Center, City University of New York

Recommended Citation
Dana, Dibyendu, "Development of Cell-Active Inhibitors and Activity-Based Probe of Cysteine Cathepsins" (2014). CUNY Academic Works.
https://academicworks.cuny.edu/gc_etds/414

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/gc_etds
Part of the Biochemistry Commons, and the Organic Chemistry Commons

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
Development of Cell-Active Inhibitors and Activity-Based Probe of Cysteine Cathepsins

by

Dibyendu Dana

A Dissertation Submitted to the Graduate Faculty in Chemistry in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy, The City University of New York

2014
This manuscript has been read and accepted for the Graduate Faculty in Chemistry in satisfaction of the dissertation requirements for the degree of Doctor of Philosophy.

Dr. Sanjai Kumar
Chair of Examining Committee

Dr. Maria C. Tamargo
Executive Officer

Dr. Yu Chen

Dr. Emmanuel Chang

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

Development of Cell-Active Inhibitors and Activity-Based Probe of Cysteine Cathepsins

by

Dibyendu Dana

Adviser: Professor Sanjai Kumar

Cysteine cathepsins are an important class of enzymes that coordinate a variety of important cellular processes, and are implicated in various types of human diseases. Still however, many of their cellular function remain poorly understood. Chemical biology approaches employing small molecules can be utilized for this purpose. Unfortunately small molecule probes that are cell-permeable and non-peptidyl in nature are scarcely available.

In this work, first a library of sulfonyloxiranes is synthesized. From this library, 2-(2-ethylphenylsulfonyl)oxirane is identified as a selective inhibitor of cysteine cathepsins. Cell-based study reveals that 2-(2-ethylphenylsulfonyl)oxirane is a cell-permeable, covalent, and irreversible inhibitor of cathepsin B with modest efficacy.

Next, a hybrid-design approach is undertaken to develop a highly potent and selective peptidyl vinylsulfonates inhibitor (KD-1) of human cathepsin L. Studies involving human breast carcinoma MDA-MB-231 cells establishes that this inhibitor can successfully block intracellular cathepsin L activity, and retards the cell-migratory potential of these highly metastatic cells. This work has been further extended to develop an activity-based probe (KDP-1) of cathepsin L by suitable modification of KD-1 inhibitory scaffold. KDP-1 has been found to be so far the most potent activity-based probe of cathepsin L. Cell-based studies are currently underway to
demonstrate the cellular efficacy of KDP-1. Further, KDP-1 is anticipated to find extensive applications in proteome-wide analysis of cathepsin L activity in both normal and diseased cells.
Acknowledgements

I would like to express my deepest gratitude to my research advisor Professor Dr. Sanjai Kumar; you have been an incredible mentor for me. I would like to thank you for giving me an opportunity to pursue my graduate study under your vigilant supervision that has helped me to grow as a research scientist. Your advice on both research as well as on my career have always been priceless. I would also like to thank my committee members Professor Dr. Yu Chen and Professor Dr. Emmanuel Chang for providing me with their invaluable inputs about my research. I would especially like convey my gratitude to Professor Dr. Gopal Subramaniam for his great support throughout my entire graduate study. I would also like to convey my gratitude to Professor Dr. Dipak K. Dey and Professor Dr. Nalini Ravishankar for their trust in my abilities.

A special thanks to all my lab mates for being so cordial and supportive. I would especially like to mention Anibal Davalos Morinigo, Nisar Afzal and Daniel Andre Novoa for their great enthusiasm and support in all the projects that we have conceived together. Also, I would like to express my thanks to all my friends for being so supportive and helpful during my entire graduate study.

I would especially like to express my deepest gratitude to my mother, father and my grandma without whom I would not have achieved anything that I have today. Finally, I would like to appreciate my beloved wife Susmita for being so patient, caring and supportive in the moments when I was left alone with queries but no one to answer.
Table of Contents

Title...........................................................................................................................................i

Approval Page..........................................................................................................................iii

Abstract.....................................................................................................................................iv

Acknowledgments....................................................................................................................vi

Table of Contents.....................................................................................................................vii

List of Figures............................................................................................................................xi

List of Tables.............................................................................................................................xiii

List of Schemes..........................................................................................................................xiv

List of Abbreviations....................................................................................................................xv

List of Appendices.......................................................................................................................xvii

Chapter 1: Introduction to Cysteine Cathepsins

1.1 Background............................................................................................................................2

1.2 Regulation and Structure.......................................................................................................3

1.3 Substrate Binding and Specificity........................................................................................5

1.4 Mechanism of Proteolysis.....................................................................................................7

1.5 Localization and Physiological Significance......................................................................8

1.6 Roles in Human Pathology................................................................................................10

1.7 Existing Classes of Reported Cysteine Cathepsin Inhibitors..........................................11
Chapter 2: Part A. Aryl Sulfonyloxiranes as New Class of Cell-Active Inhibitors of Cysteine Cathepsins

2.1 Conceptual Design..........................................................................................................................20

2.2 Chemistry........................................................................................................................................21

2.3 Results and Discussion......................................................................................................................22

2.4 Demonstration of Intracellular Activity of Aryl Sulfonyloxiranes.............................................33

2.5 Conclusions.......................................................................................................................................38

2.6 Experimental....................................................................................................................................40

2.6.1 Synthesis.......................................................................................................................................40

2.6.2 Enzymology, physiochemical, and mass spectrometric studies.............................................45

Part B. Serendipitous Discovery: Base-Catalyzed Synthesis of Allyl and Vinyl Phenyl Sulfones

2.7. Background......................................................................................................................................56

2.8. Discovery.........................................................................................................................................56

2.9 Concluding Remarks........................................................................................................................64

2.10 Experimental..................................................................................................................................66
Chapter 3: Development and Evaluation of Peptidyl Vinylsulfonate as:

I. A Cell-Active and Specific Inhibitor of Cathepsin L (KD-1)

3.1 Introduction ..........................................................................................................................74

3.2 Rationale ................................................................................................................................75

3.3 Chemistry ..............................................................................................................................77

3.4 Results and Discussion .........................................................................................................78

3.5 Conclusion .............................................................................................................................83

3.6 Experimental .........................................................................................................................85

3.6.1 Synthesis ............................................................................................................................85

3.6.2 Screening and Enzymology ...............................................................................................95

3.6.3 Biological and Cell-Based Studies ....................................................................................101

II. An Activity-Based Probe of Cathepsin L (KDP-1)

3.7 Transformation of Potent Specific Cathepsin L Inhibitor to Activity-Based Probe..106

3.8 Chemistry .............................................................................................................................107

3.9 Results and Experimental Propositions ...............................................................................108

3.10 Future Directions .................................................................................................................111

3.11 Experimental .......................................................................................................................112
3.11.1 Synthesis..................................................................................................................112

3.11.2 Enzyme Kinetics......................................................................................................115

References.........................................................................................................................188
List of Figures

Figure 1.1. Classification of proteases........................................................................................................2
Figure 1.2. Structure of cysteine cathepsin.................................................................................................3
Figure 1.3. Substrate binding to cysteine cathepsin active site.................................................................6
Figure 1.4. Pictorial representation of endopeptidase and exopeptidase....................................................7
Figure 1.5. Proposed mechanism of hydrolysis of substrate peptide by cathepsin L...............................8
Figure 1.6. Pictorial representation of postulated inhibitor binding to cathepsin pockets......................12
Figure 2.1. Hybrid design of aryl sulfonoxirane as cysteine cathepsin inhibitor.....................................20
Figure 2.2. General chemical structure of inhibitory scaffold.................................................................23
Figure 2.3. Inhibitor 5-mediated time-dependent inactivation of cathepsin B...........................................26
Figure 2.4. Determination of inactivation kinetic constants of inhibitor 5.............................................27
Figure 2.5. A mass spectrometric analysis of 5-inhibited cathepsin B complex.....................................28
Figure 2.6. Computational analysis of 5-inhibited cysteine cathepsin B complex.................................30
Figure 2.7. Demonstration of cell-active nature of inhibitor 5.................................................................35
Figure 2.8. Inhibitor 5 suppresses cathepsin B-mediated E-cadherin degradation...............................36
Figure 2.9. Inhibitor 5 reduces metastatic potential of invasive MDA-MB-231 cancer cells.............38
Figure 2.10. Half-life estimation of inhibitor 5 under physiological conditions......................................50
Figure 2.11. X-ray crystal structure of (E)-((3-phenylprop-1-en-1-yl)sulfonyl)benzene.....................58
Figure 3.1. Development of cathepsin L specific inhibitor KD-1............................................................76
Figure 3.2. Time-dependent inhibition of human cathepsin L by KD-1....................................................79
Figure 3.3. Cell active and anti-collagenase activity of KD-1.................................................................82
Figure 3.4. KD-1 reduces cell-migratory potential of MDA-MB-231 cells.............................................84
Figure 3.5. Chemical structure of synthesized aryl vinylsulfonates.......................................................97
Figure 3.6. IC_{50} determination of KD-1 against cathepsin L................................. 98

Figure 3.7. Determination of second order inactivation rate constant of KD-1.................. 99

Figure 3.8. Determination of second order inactivation rate constant of KDP-1..............109

Figure 3.9. Procedure involved in intracellular detection of cathepsin L activity using KDP-
1........................................................................................................................................110
List of Tables

Table 1.1. Different classes of existing cysteine cathepsins inhibitors……………………………13

Table 2.1. Chemical structure of inhibitory library…………………………………………………………26

Table 2.2. Selectivity of inhibition of compound 5 towards cysteine cathepsins………………..32

Table 2.3. $^1$H and $^{13}$C NMR Chemical Shifts for vinyl phenyl sulfone, 2 and allyl phenyl sulfone, 9………………………………………………………………………………………………………………………59

Table 2.4. Treatment of 1 with 2-alkylacetaldehydes under varying NaH concentrations……61

Table 3.1. Second order enzyme inactivation rate constants ($k_{\text{inact}}$) of cathepsin L inhibitor KD-1 against a panel of closely-related enzymes……………………………………………………………80
List of Schemes

Scheme 2.1. Synthetic routes utilized for the development of sulfonyloxiranes………………..22

Scheme 2.2. % yield obtained with variety of arylacetaldehydes………………………………..57

Scheme 2.3. Proposed mechanism of the formation of allylic sulfone 9 when catalytic excess of sodium hydride is used………………………………………………………………………..63

Scheme 2.4. Exclusive formation of allyl sulfone under varying base equivalents………………64

Scheme 3.1. Synthetic route utilized in the development of cathepsin L inhibitor, KD-1………78

Scheme 3.2. Synthetic route adopted for the development of KDP-1…………………………….108
List of Abbreviations

AMC  7-amino-4-methylcoumarin
Boc  tert-butoxycarbonyl
Cbz  benzylxycarbonyl
Cys  cysteine
DMF  N,N-dimethylformamide
DMSO dimethylsulfoxide
DMSO-d$_6$ dimethylsulfoxide-d$_6$ deuterated
DTT  dithiothreitol
E-64 L-trans-epoxysuccinyl-leucylamido(4-guanidino)butane
EDTA ethylenediaminetetraacetic acid
ESI-MS electrospray ionization mass spectrometry
Et$_3$N triethylamine
EtOAc ethyl acetate
HEPES N-2-hydroxyethylpiperazine-N’-2-ethanesulfonic acid
HRMS high resolution mass spectrometry
Leu leucine
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>TLC</td>
<td>thin layer chromatography</td>
</tr>
<tr>
<td>IC50</td>
<td>Inhibitor concentration when the activity of a given enzyme is reduced to 50%</td>
</tr>
</tbody>
</table>
List of Appendices

Appendix A: NMR Spectra of Synthesized Arylsulfonyloxiranes...............................116

Appendix B: NMR Spectra of Synthesized Allyl/Vinyl Sulfones and Crystal Structure of (E)-
            ((3-phenylprop-1-en-1-yl)sulfonyl)benzene (2)...............................127

Appendix C: NMR Spectra of Synthesized Vinyl Sulfonates.....................................156
Chapter 1

Introduction to Cysteine Cathepsins
1.1 BACKGROUND

Proteases are a class of enzymes that catalyze the hydrolysis of their protein substrates. They are found in almost all types of living organisms that include mammals, viruses, bacteria, protozoa, plants and fungi. Proteases have been categorized based on their active site catalytic residues that they use for catalysis. There are four existing classes of proteases: aspartate protease, metalloprotease, serine protease and cysteine protease (Figure 1.1). The emergence of several newly discovered proteases necessitate further classification to differentiate proteases within their distinct catalytic category into evolutionary families and clans. MEROPS is an evolving database where emergence of newly discovered proteases is regularly indexed. Cysteine proteases have been further categorized in three structurally distinct subclasses: the papain family, the caspase family, and the picornaviridae family. The papain family of enzymes represent majority of the cysteine proteases. Cysteine cathepsins belong to the papain family of enzyme. There are eleven members of cysteine cathepsins (cathepsin B, C/J/dipeptidyl peptidase I, F, H, L, K, O, S, W, V and Z/X/P) that are primarily housed in membrane-bound organelles called lysosomes. They have unique catalytic machinery and tissue specific expression pattern and they mediate many essential physiological processes. Aberrant expressions of cysteine cathepsins have been associated with different types of diseases, such as bone loss, autoimmune disease, metabolic syndrome and metastatic cancer. Thus, developing inhibitors of cysteine cathepsins have potential to (a) treat a wide range of
diseases, (b) assessing unknown functions of cysteine cathepsins in various cellular contexts. It is thus imperative to develop new classes of inhibitors to address the aforementioned issues.

1.2 Regulation and Structure

Cysteine cathepsins are synthesized as preproenzymes which are composed of a signal peptide, a propeptide and a catalytic domain (Figure 1.2A).\(^6\) Signal peptides are composed of 10-20 amino acid residues, and are responsible for translocation into the endoplasmic reticulum during mRNA translation.\(^7\) Cleavage of signal peptide yields propeptide\(^8\) that plays important role in cathepsin biology.\(^9\) The role of propeptide segment is manyfold: (a) it assists in proper folding of enzyme,\(^7\) (b) it uses very specific mannose-6-phosphate receptor (M6PR) pathway to direct localization to endosomal/lysosomal compartment,\(^7\) and (c) it acts as a high-affinity reversible inhibitor of cognate enzymes that prevents any aberrant zymogen activity.\(^10\) Cleavage of propeptide releases proteolytically active mature cathepsins which vary in length between 214 amino acid residues to 260 amino acid residues.\(^6\)

Cysteine cathepsins are relatively small monomeric proteins with molecular weights in the range of 24–35 kDa.\(^11\) Cathepsin C is the only exception to this and is a tetramer with a molecular weight of 200 kDa.\(^11\) Determination of papain structure was the first structural elucidation of papain-like cysteine proteases. Except cathepsin O and W\(^7\) three dimensional structures of all the cysteine cathepsins have been solved (cathepsin B\(^12\), cathepsin L\(^13\), cathepsin K\(^14\), cathepsin S\(^15\), cathepsin V\(^16\), cathepsin F\(^17\), cathepsin H\(^18\), cathepsin X\(^19\), and cathepsin C\(^20\)). All cysteine cathepsins share the common fold of papain like structure. Cathepsin L has been chosen as a prototypical of cysteine cathepsin class of enzymes for structural elucidation (Figure 1.2.B).\(^13\) Cathepsin L is a bilobal protein that consists of two structurally distinct domains; left
(L) domain and right (R) domain (Figure 1.2.B). The structure of the active cathepsin L resembles a closed book with the spine at the front\(^{13}\) where two domains form a V-shaped active site at the interface. Catalytic site of the enzyme lies within the active site that is composed of a Cys25 (from left domain) and His163 (from right domain). Together they form a thiolate-imidazolium ion pair which imparts enhanced nucleophilicity to the cysteine residue.\(^{21}\) The nucleophilic thiolate residue in turn participates in cleavage of the scissile peptide bond. The left domain contains three α-helixes. Among these, central α-helix is longest in length and is composed of 30 amino acid residues.\(^{6}\) On the other hand, right domain represents a β-barrel like
shape which includes a shorter α-helical motif.\textsuperscript{13} The propeptide binds to the active site cleft which makes it inaccessible for substrate binding\textsuperscript{9,22} (Figure 1.2).\textsuperscript{13}

The proteolytic activities of mature cathepsins are also regulated by few physiological factors. For example, cysteine cathepsins are optimally active under acidic environment of lysosome. Most of them lose activity at neutral pH, except cathepsin S which is active even at slight alkaline pH.\textsuperscript{23} It should also be noted that active site cysteine is prone to oxidation. Thus endosome accumulates cysteine to retain the reducing environment with in the cell.\textsuperscript{24} Finally, endogenous inhibitors of cysteine cathepsins maintain the balance of protease activity under normal physiological condition.\textsuperscript{25}

### 1.3 Substrate Binding and Specificity

The seminal work by Berger and Schechter in 1967 was the first elucidation of substrate binding to the active site of cysteine protease.\textsuperscript{26} Their study examined the influence of substrate chain length on substrate kinetics. Based on their discovery they have developed a nomenclature in which individual enzyme pockets have been designated as ‘S’ and substituents approaching from the peptide substrate have been depicted as ‘P’ (Figure 1.3.A). Three decades later, Turk et al.\textsuperscript{27} revisited and redefined the substrate binding to the active site of papain like cysteine proteases (Figure 1.3.B).\textsuperscript{13} The active site is comprised of amino acid residues approaching from both the lobes.\textsuperscript{11} The L-lobe contributes two shorter loops (Gln19-Cys25, Arg59-Tyr67) connected by a disulfide bridge (Cys22-Cys63) thus making it an additional loop closure.\textsuperscript{6} Further, the R-lobe contributes two larger loops\textsuperscript{6,13} (Leu134-His159, Asn175-Ser205) reminiscent of the lid of the β-barrel hydrophobic core.\textsuperscript{6} The substrate binds along the active site cleft in an extended conformation\textsuperscript{6,27} allowing its side chains to make alternate interactions with
both lobes. Loops from left lobe accommodate P3, P1 and P2’ residues while loops from R-lobe engage in favorable binding interactions with P2 and P1’. Among these subsites, S2 is a deep pocket whereas S1 and S1’ offer just a binding surface. Substrate binding beyond S3 and S2’ pocket do not include main chain interaction. Thus these regions are termed as areas but not as sites.

The substrate specificity of cysteine cathepsins are governed by physical interactions between enzyme and the substrates. Depending on the substrate cleavage site, cysteine cathepsins have been categorized as either an endopeptidase or an exopeptidase, or both (Figure 1.4). Peptidase that cleaves within a substrate polypeptide chain has been coined as endopeptidase. On the other hand, exopeptidase hydrolyzes the terminal amino acid residue of

---

Figure 1.3. [A] Pictorial representation of substrate binding sites according to Schechter and Berger (1967). Enzyme pockets have been assigned as “S” and substrate peptide binding to corresponding pockets have been assigned as “P”. [B] Revised substrate binding to papain like cysteine proteases. Papain numbering has been followed. Red-dotted line designates the bond that gets hydrolyzed by cysteine cathepsins.
the substrate peptide. Among cysteine cathepsins five of them are strict endopeptidase (cathepsin L, S, K, V and F) and three cathepsins serve exclusively as exopeptidase (cathepsin H, X and C)\(^7\). The peptidase activity of cathepsin O and W is yet to be explored. Depending on the cellular environment, cathepsin B can exert both endo and exopeptidase activity. Exopeptidases have further been categorized based on the nature of terminal amino acid residue and substrate peptide length that they cleave. Cathepsin B is a carboxypeptidase whereas cathepsin X acts as a carboxy monopeptidase.\(^{19}\) Cathepsin C is an aminodipeptidase\(^ {28}\) and cathepsin H is an amino peptidase.\(^ {18}\)

### Figure 1.4

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Diagram A" /></td>
<td><img src="image2.png" alt="Diagram B" /></td>
</tr>
</tbody>
</table>

**Figure 1.4.** Pictorial representation of [A] endopeptidase, and [B] exopeptidase. Enzyme pockets have been assigned as “S” and substrate peptide binding to corresponding pockets have been assigned as “P”. Red-dotted line designates the bond that gets hydrolyzed by active site cysteine residue of cysteine cathepsins.

### 1.4 Mechanism of Proteolysis

The cysteine (Cys25) and the histidine (His163) residues of the active site of cathepsin L form an ion pair which confers enhanced nucleophilicity to cysteine residue. The mechanism of proteolysis by papain like cysteine cathepsins occurs in a step-wise process (Figure 1.5): (1) Nucleophilic cysteine (Cys25) attacks the electrophilic carbonyl center of substrate peptide and forms a tetrahedral anionic intermediate. This intermediate gets stabilized via oxyanion hole
interactions (2) His163 acts as a general acid and assists cleavage of the peptide bond via protonation of leaving amide nitrogen. This leads to expulsion of the C-terminal peptide fragment, thereby facilitating the formation of thioacyl enzyme intermediate, and (3) Deacylation of thioacyl enzyme intermediate completes the substrate turnover process, releasing the free enzyme.

**Figure 1.5.** Proposed mechanism of hydrolysis of substrate peptide by cathepsin L.

1.5 Localization and Physiological Significance

Originally, cysteine cathepsins were considered as redundant class of enzymes, primarily housed in lysosome, and involved in non-specific terminal protein degradation. Recent studies have revealed that cysteine cathepsins are not only expressed in lysosome but also in cytoplasm,
nucleus and in extracellular spaces as well. While, cathepsin B, L, V, O, H, C, X and F are ubiquitously expressed in human tissues, cathepsin K, S and W have a more restricted expression pattern. Variable expression patterns of cysteine cathepsins have implicated them to several important physiological processes. For example, cysteine cathepsins play important roles in immunological responses, prohormones activation, bone remodeling, angiogenesis and many other important physiological processes. Cathepsin L, F, S and V participate in MHC II-mediated antigen representation that is recognized by CD4+ T cells. Cathepsin C activates various serine proteases and granzyme A and B that helps in limiting acute viral infection. Cathepsin W is selectively expressed in natural killer cells and in CD8+ lymphocites and plays crucial role in cytotoxic killing of abnormal cells. Cysteine cathepsins are also involved in prohormone activation. Cathepsin B has been associated with activation of precursors of β-galactoside, renin and trypsin. Further, cathepsin B, L and K have been associated with thyroxine production, while cathepsin H has been implicated in pulmonary surfactant-associated protein processing in human lungs.

Also, cysteine cathepsins are implicated in skeletal remodeling. Cathepsins K, L, F, S, V and B have been associated with elastin degradation with varying elastolytic activity. Cathepsin K is predominantly expressed in osteoclast and plays crucial role in bone resorption under normal and pathological conditions.

In addition to aforementioned roles, cysteine cathepsins also take part in other physiological processes. Cathepsin L has been implicated to cell-cycle progression and keratinocyte differentiation and in hair follicle cycle. Moreover, cysteine cathepsins degrade proapoptotic Bcl-2 homolog protein bid and other antiapoptotic Bcl-2 family of proteins which therefore triggers apoptosis in a synergistic manner.
1.6 Roles in Human Pathology

It is now well established that over-expression and aberrantly regulated activity of cysteine cathepsins play a major role in promoting many hallmarks of cancer.\textsuperscript{39} Often, these require cysteine cathepsins to be overexpressed in the extracellular space. Under normal physiology, extracellular space is protected from cysteine cathepsins by neutral pH, endogenous inhibitors (for example, stefins, cystatins, kininogens, and thyropins)\textsuperscript{6} and reducing environment. Nevertheless, pericellular acidification and pathophysiologically reduced anticathepsin concentrations can help to enhance the activity of cathepsins.\textsuperscript{7} Elevated level of cathepsins expression and activity are the key mediators of invasive growth and angiogenesis during multistage tumorigenesis.\textsuperscript{40} Aberrantly active of cysteine cathepsins initiate proteolytic cascade that activate urokinase-type plasminogen activator, matrix metalloproteinases and plasminogen.\textsuperscript{6} Moreover, these active proteases in conjunction with cysteine cathepsins degrade E-cadherin and extracellular matrix proteins that is key structural barrier that cells must cross to reach the vasculature.\textsuperscript{41} Subsequently, intracellular cysteine cathepsins aid the process by degrading collagen, as these cells accumulate extracellular matrix via endocytosis.\textsuperscript{41c, 42}

The expression of cysteine cathepsins (e.g. B, F, and L) has indeed been found to be significantly upregulated in solid tumors from various origins, such as breast, skin, colorectal, pancreatic, ovarian, brain, head and neck.\textsuperscript{39b, 43} More importantly, a well-defined role for cysteine cathepsins in promoting cell invasion and cancer metastasis has recently emerged.\textsuperscript{44} A high level expression of cysteine cathepsin B was observed in invasive tumor cells derived from metastatic cancer patients of distinct origins.\textsuperscript{45} Deletion of either cysteine cathepsins, B, L or S led to decreased tumor invasion with tumors progressively returning to a more benign form.\textsuperscript{43} In a separate study, cysteine cathepsin X was able to compensate in part for the absence of cathepsin
B in a mouse model for mammary cancer. In this study, inhibition of cysteine cathepsin X with a neutralizing antibody resulted in a drastic loss of invasive behavior in cathepsin B null cultured cells. E-cadherin, an important adherens junction protein that maintains cell-cell adhesion in epithelial cells, was recently identified as a direct proteolytic target of cysteine cathepsins B, L and S. Since it is well established that functional loss of E-cadherin promotes cell invasion and metastatic behavior of cancer cells, an effective inhibition of this family of enzymes is anticipated to provide significant therapeutic benefit in reducing metastatic potential of solid tumors. These studies have led to an interesting hypothesis that promotes a polypharmacological approach to target pro-tumorigenic members of cysteine cathepsins collectively for the development of effective anti-cancer agents.

### 1.7 Existing Classes of Reported Cysteine Cathepsin Inhibitors

Pathological significance has prompted cysteine cathepsins as promising drug targets. Although there are no drugs in use; few are in clinical trial. However, past three decades of effort has enabled deciphering key protein-ligand interaction that could initiate development of drug like molecules for cysteine cathepsins. Several different types of inhibitors have been developed which primarily targets the active site of cysteine cathepsins. Based on their mechanism of action, these inhibitors have been categorized in different classes. In general, the inhibitory scaffold is made of three components: (a) recognition region, which engage in physical interactions with enzyme binding pockets, (b) Address region, which generally confers additional interactions, and (c) electrophilic warhead, which traps the nucleophilic cysteine residue of active site. A cartoon diagram illustrating the design of inhibitor is shown below (Figure 1.6).
Here is the key representative classes of inhibitors enlisted in following page (Table 1.1).
Table 1.1. Different classes of existing inhibitors.

<table>
<thead>
<tr>
<th>Compound class</th>
<th>Structure</th>
<th>Activity</th>
</tr>
</thead>
</table>
| Peptidyl epoxide<sup>a</sup> | ![Peptidyl epoxide Structure](image) | $k_{2nd} = 96250 \text{ M}^{-1} \text{s}^{-1}$ (Cat L)  
Not cell active |
| Peptidyl aldehyde<sup>b</sup> | ![Peptidyl aldehyde Structure](image) | $IC_{50} = 20 \text{ nM}$ (Cat K)  
Cell active |
| Acylaminopropanone<sup>c</sup> | ![Acylaminopropanone Structure](image) | $K_i = 22 \text{ nM}$ (Cat K)  
No cell data available |
| Diacylcarbohydrazide<sup>d</sup> | ![Diacylcarbohydrazide Structure](image) | $k_{2nd} = 3.1 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ (Cat K)  
Cell active |
| Miraziridine<sup>e</sup> | ![Miraziridine Structure](image) | $IC_{50} = 2.05 \mu\text{M}$ (Cat B)  
No cell data available |
| Peptidyl diazomethyl ketones<sup>f</sup> | ![Peptidyl diazomethyl ketones Structure](image) | $k_{2nd} = 1.5 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ (Cat L)  
No in vivo data available |
| Peptidyl hydroxamate<sup>g</sup> | ![Peptidyl hydroxamate Structure](image) | $k_{2nd} = 3.54 \times 10^6 \text{ M}^{-1} \text{s}^{-1}$ (Cat L)  
Non selective towards serine proteases |
Peptidyl chloromethyl ketone\textsuperscript{h} 
\[ k_{2nd} = 2.15 \times 10^{-7} \text{ M}^{-1} \text{ s}^{-1} \text{ (Cat L)} \]
Not cell active

Peptidyl fluoromethyl ketone\textsuperscript{i} 
\[ k_{2nd} = 3.9 \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1} \text{ (Cat B)} \]
Cell active

Azapenone\textsuperscript{j} 
\[ K_i \text{ app} = 0.041 \text{ nM (Cat K)} \]
Cell active

Non-peptidyl nitrile\textsuperscript{k} 
\[ IC_{50} = 0.2 \text{ nM (Cat K)} \]
Cell active

Peptidyl vinyl sulfone\textsuperscript{l} 
\[ IC_{50} = 2.6 \text{ nM (Cat L)} \]
No in vivo data

Peptidyl acyloxymethyl Ketone\textsuperscript{m} 
\[ k_{2nd} = 2.96 \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1} \text{ (Cat L)} \]
Cell active
Thiosemicarbazone\textsuperscript{n}  
IC\textsubscript{50} = 30.5 nM (Cat L)  
Cell active

Thiocarbazate\textsuperscript{o}  
IC\textsubscript{50} = 56 nM (Cat L)  
No in vivo data available

Oxooocarbazate\textsuperscript{p}  
IC\textsubscript{50} = 6.9 nM (Cat L)  
Cell active

Fused ring thiosemicarbazone\textsuperscript{q}  
IC\textsubscript{50} = 164 nM (Cat L)  
No Cell data available

Modified β-lactum\textsuperscript{r}  
IC\textsubscript{50} = 0.5 nM (Cat K)  
No Cell data available

Organotellurium (IV)\textsuperscript{s}  
\( k\textsubscript{2nd} = 36,000 \text{ M}^{-1} \text{s}^{-1} \) (Cat B)  
Cell active

1, 2, 4-Thiadiazole\textsuperscript{t}  
\( k\textsubscript{2nd} = 5630 \text{ M}^{-1} \text{s}^{-1} \) (Cat B)  
Not Cell active
1.8 Perspective of Synthesized Inhibitors and Activity-Based Probe

While numerous studies of cysteine cathepsins have clearly established them as important therapeutic targets for the development of anti-cancer agents, progress towards development of small molecule inhibitory agents that are non-peptidyl in nature remains slow. Currently, a majority of the reported inhibitory agents of cysteine cathepsins are peptidyl- or peptidomimetic-based in nature, making them poorly cell permeable. Small molecule inhibitory agents of cysteine cathepsins that are uncharged and non-peptidyl in nature are therefore desirable so that they can be further developed as suitable biological inhibitory reagents, in addition to lead chemotypes for drug development. Furthermore, if the mechanism of inactivation could be
designed to produce a covalent and irreversible chemistry with the catalytic active site
nucleophilic Cys residue, then the inhibitors could also be developed as activity-based probes
(ABPs) for the functional analysis of cysteine proteases in complex proteomes. Activity-based
probes have indeed found extensive applications in functional annotation of enzymes in their
native biological environments.\textsuperscript{68}
Chapter 2

Part A:

Aryl Sulfonyloxiranes as New Class of Cell-Active Inhibitors of Cysteine Cathepsins

Part B:

Serendipitous Discovery: Base-Catalyzed Synthesis of Allyl and Vinyl Phenyl Sulfones
Chapter 2

Part A:

Aryl Sulfonyloxiranes as New Class of Cell-Active Inhibitors of Cysteine Cathepsins
2.1 Conceptual Design

The advent of E-64 from *Aspergillus japonicus*, an epoxysuccinyl-derived inhibitory agent, was a critically important stepping stone for deciphering the undocumented function of cysteine cathepsins family of enzymes.\(^6\) Although, E-64 is a generic non-selective inhibitor of cysteine cathepsins, structural investigation of the E-64-inhibited complex have subsequently led to the development of selective cysteine cathepsin B inhibitors (e.g. CA030 and CA074).\(^7\) The marked success of E-64 and its subsequent analogues as effective inhibitors of cysteine cathepsins relies on the exclusive reactivity of an active site cysteine residue with the electrophilic “warhead” epoxy group. The catalytic mechanism of cysteine cathepsins involves participation of this invariant and reactive cysteine residue.\(^7\) A close visual inspection of structure of CA030 and CA074 in complex with cysteine cathepsins revealed that there exists well-defined binding pockets, S2, and S1’, surrounding the S1 site (Schechter and Berger nomenclature)\(^7\) that prefer to accommodate aromatic hydrophobic groups.\(^7\) It was therefore hypothesized that an arylsulfonyl-substituted oxiranyl moiety could serve as a small molecule

---

Figure 2.1. Hybrid design of hypothesized aryl sulfonyloxirane as cysteine cathepsin inhibitor.
covalent inactivator of cysteine cathepsins, targeting the key residues of the core enzyme active site (Figure 2.1). This approach envisions that (a) an appropriate binding of inhibitory compounds within the active site of cysteine cathepsin will allow the reactive Cys residue to be entrapped by the activated sulfonyloxirane group; and (b) the sulfonyl group will likely mimic the potential h-bonding interactions that the carbonyl group α to epoxy moiety exhibits in the CA-030-cysteine cathepsin B inhibited complex70a; and (c) the sulfonyl group will further enhance the electrophilicity of oxiranyl carbon, compared to a carbonyl group, thereby facilitating the entrapment of the active site Cys residue. If successfully demonstrated, this class of molecules could serve as a new type of chemical motif on which selective cysteine cathepsin inhibitors could be developed in future, exploiting the surrounding S3 and S’ pockets. Such a strategy has indeed been utilized to develop potent and selective inhibitors of cysteine cathepsins.73

2.2 Chemistry

Sulfonyloxiranes were synthesized according to scheme 2.1. Reaction of chloromethyl phenyl sulfone with corresponding aldehyde in presence of potassium hydroxide yield to 1st generation of sulfonyl oxiranes74 (Scheme 2.1 A). 2nd class of sulfonyloxiranes were synthesized using a different protocol (Scheme 2.1 B). Reaction of 2-bromoethanol with substituted thiophenol in presence of base yield to arylthioethanol. Oxidation of arylthioethanol by hydrogen peroxide (30%) provided 2-hydroxy arylsulfonylethane. Treatment of 2-hydroxy arylsulfonylethane with mesyl chloride in presence of triethylamine at 0°C leads to arylsulfonylethene. Further conversion of alkene intermediate to aryl sulfonyloxirane was carried out using n-BuLi in presence of tert-butyl hydroperoxide.75 3rd generation of compounds were synthesized according to a literature reported protocol (Scheme 2.1 C).76 Arylthioallyl
intermediate was achieved by reaction between thiophenol and allyl bromide in presence of potassium carbonate. MCPBA mediated oxidation of the acquired intermediate lead to aryl sulfonyloxirane with an extra methylene linker. Thiirane was eventually synthesized from the intermediate sulfonyloxirane using ammonium thio cyanate.

### 2.3 Results and Discussion

A small library of sulfonyloxirane compounds was therefore synthesized. Cysteine cathepsin B, a promoter of cancer metastasis, was chosen as the initial inhibitory target to evaluate the efficacy of a time-dependent loss of enzyme activity (Table 2.1). The initial screening was performed by incubating a fixed concentration of compounds 1-13 with cathepsin B under pseudo-first order conditions. A small aliquot of the incubation mixture was withdrawn at a fixed time interval and the remaining cathepsin B activity was measured in a large volume of assay mixture (17.5 fold dilution) containing cathepsin B substrate. The IC$_{50}$ values were
deliberately not estimated for this library, since they are not considered a good yardstick for measuring a time-dependent irreversible loss of enzyme activity due to a non-equilibrium binding mechanism.  

Compounds 1 and 2 which bear a non-polar 4-fluorophenyl and a propyl substituent respectively at the second carbon of the oxirane ring were non-inhibitory. This result indicated that a second substitution at the oxirane moiety in this class of molecules was perhaps detrimental to their inhibitory efficacies. Furthermore, compound 12, which incorporates a methylene linker between the arylsulfonyl group and oxirane moiety, was also inert. We surmised that the lack of reactivity of 12 could be due to lack of assistance from a general acid appropriately positioned in the cathepsin B active site to promote nucleophilic opening of oxirane ring by Cys29-Sγ residue. So, a thiirane analogue, 13, that potentially contains a better leaving group thiolate (compared to alkoxide in 12), was synthesized and was also found to be ineffective. This again indicated that appropriate positioning of arylsulfonyl group with respect to the reactive oxirane group within the cysteine cathepsin active site was critically important for its inhibition. Efforts became focused on synthesizing analogues of 2-(arylsulfonyl)oxiranes where functional variations of the aryl ring was further investigated.

![Figure 2.2. General chemical structure of inhibitory scaffold.](image-url)
Table 2.1. Chemical structure of inhibitory library.

<table>
<thead>
<tr>
<th>Compound ID</th>
<th>Chemical Structure (Name)</th>
<th>% Inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1" alt="Structure 1" /> 2-(4-fluorophenyl)-3-(phenylsulfonyl) oxirane</td>
<td>NI*</td>
</tr>
<tr>
<td>2</td>
<td><img src="image2" alt="Structure 2" /> 2-(phenylsulfonyl)-3-propyloxirane</td>
<td>NI*</td>
</tr>
<tr>
<td>3</td>
<td><img src="image3" alt="Structure 3" /> 2-((4-chlorophenyl)sulfonyl)oxirane</td>
<td>44</td>
</tr>
<tr>
<td>4</td>
<td><img src="image4" alt="Structure 4" /> 2-(4-(trifluoromethoxy)phenylsulfonyl)oxirane</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td><img src="image5" alt="Structure 5" /> 2-(2-ethylphenylsulfonyl)oxirane</td>
<td>&gt;98</td>
</tr>
<tr>
<td></td>
<td>Chemical Structure</td>
<td>Molecular Formula</td>
</tr>
<tr>
<td>---</td>
<td>--------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>6</td>
<td><img src="" alt="2-(m-tolylsulfonyl)oxirane" /></td>
<td>C9H8O2S</td>
</tr>
<tr>
<td>7</td>
<td><img src="" alt="2-(2,5-dimethylphenylsulfonyl)oxirane" /></td>
<td>C14H12O2S</td>
</tr>
<tr>
<td>8</td>
<td><img src="" alt="2-(3-methoxyphenylsulfonyl)oxirane" /></td>
<td>C14H13O3S</td>
</tr>
<tr>
<td>9</td>
<td><img src="" alt="2-(2,5-dimethoxyphenylsulfonyl)oxirane" /></td>
<td>C15H15O3S</td>
</tr>
<tr>
<td>10</td>
<td><img src="" alt="2-(3,5-dichlorophenylsulfonyl)oxirane" /></td>
<td>C13H6Cl2O2S</td>
</tr>
<tr>
<td>11</td>
<td><img src="" alt="2-(naphthalen-3-ylsulfonyl)oxirane" /></td>
<td>C15H11O2S</td>
</tr>
</tbody>
</table>
Among the synthesized library, compound 5 was found to be most effective in inhibiting

Figure 2.3.  [A] Progress curves indicating a time-dependent loss of cathepsin B activity. Compound 5 was incubated with cathepsin B under pseudo-first order condition, and the progress curves were obtained at appropriate times of inactivation (red: 0 min, blue: 11 min, green: 17 min, black: 25 min, and pink: 41 min) using a chromogenic substrate Z-RR-para-nitroanilide. [B] Concentration-dependent loss of cathepsin B activity with time. The points are experimental, and the lines joining them are fitted to obtain the pseudo-first order inactivation rate constants ($k_{obs}$).
cysteine cathepsin B activity. The effectiveness of compound 5 was evident from a progressive loss of enzyme activity when it was incubated with active cathepsin B (Figure 2.3). No gain in cathepsin B activity was observed even after dilution of 5-inactivated cathepsin B complex, thereby indicating that the mechanism of inhibition was perhaps covalent and irreversible in nature (data not shown). To assess the inhibitory efficacy, inactivation experiments were performed at appropriate concentrations of compound 5 under pseudo-first order conditions, and the experimental data thus obtained were analyzed using a simple two-step inhibition model (Figure 2.4.)

\[
E + I \rightleftharpoons EI \rightarrow E - I
\]

**Figure 2.4.** [A] Kitz-Wilson plot of 5-mediated cathepsin B inactivation kinetics. The data points are fitted to a line that yields (a) \( K_i = 86 \pm 3 \ \mu M \) for the first equilibrium step of inhibitor binding, and (b) \( k_i = 0.21 \pm 0.01 \ \text{min}^{-1} \) for the first order inactivation rate constant. [B] The rate of 5-mediated cathepsin B inhibition is decreased in the presence of leupeptin (40 nM; blue line), a known competitive inhibitor of cathepsin B, compared to the control (5% DMSO; red line).
where $K_i$ represents the reversible equilibrium binding constant for the first step, and $k_i$ is the first order inactivation rate constant for the second irreversible step. To obtain these two parameters, the following procedure was adopted. The pseudo-first order rate constants of inactivation ($k_{\text{obs}}$) were obtained at appropriate concentration of 5 ($[I]>>[\text{cathepsin B}]$), and a Kitz-Wilson analysis was performed to fit the experimentally obtained data (Figure 2.4.A). This procedure yielded the thermodynamic binding constant ($K_i = 86 \pm 3 \ \mu\text{M}$) for the first equilibrium binding step, and the first order inactivation rate constant ($k_i = 0.21 \pm 0.01 \ \text{min}^{-1}$) for the second irreversible step. We further investigated if the time-dependent inhibition of cathepsin B by 5 could be alleviated in the presence of a known active site-directed cathepsin B inhibitor, leupeptin ($K_i = 5 \ \text{nM}$). This is mainly because if inhibitor 5 were to be an active-site directed agent, the rate of 5-mediated cathepsin B inactivation reaction would be retarded in presence of a
competitive inhibitor competing for the binding to the same active site. A protection from inhibition of cathepsin B inhibition was indeed observed in presence of leupeptin, thereby indicating that 5 occupied the core cathepsin B active site that harbors the low pKₐ nucleophilic Cys29 residue (Figure 2.4B). A MALDI-based mass spectrometric analysis of the 5-inhibited cathepsin B complex showed the expected mass increase of 212.1 amu in the cathepsin B protein. No further mass increase was observed even after prolonged incubation of 5 with cathepsin B. This result indicates that the mode of inhibition by 5 was indeed covalent and irreversible in nature and the stoichiometry of inhibition was likely 1:1 (Figure 2.5). This was important to establish since most cysteine cathepsins do contain multiple cysteines, in addition to the nucleophilic active site cysteine residue.

**Computational Analysis of 5-inhibited cysteine cathepsin B complex**

To understand the potential molecular interactions of 5 with cysteine cathepsin B in the covalently inhibited complex, computational modeling studies were performed. The CA073-inhibited complex of cathepsin B (PDB id: 2DC6) was appropriately modified to incorporate the inhibitor 5 covalently linked to Cys29-Sγ of cathepsin B active site. A molecular dynamic simulation (100 ps) was performed on the inhibitory complex and the trajectories of the molecular dynamics simulation were analyzed. Several interactions between the inhibitor 5 and cathepsin B active site were noteworthy (Figure 2). While the O3 of sulfonyl oxygen formed a h-bond (3.4 Å) with the side chain of Gln23-Nε – a key oxyanion hole residue, its O2 atom was engaged in a robust hydrogen bond interactions (2.8 Å) with a tightly-bound active site water molecule (Wtr12). The OH group of inhibitor was also stabilized with two h-bond interactions, one with the backbone carbonyl of a S2 site residue Gly198 (3.1 Å) and the other with an active site water (Wtr56) molecule (2.5 Å). The phenyl moiety of the inhibitor was found to be
Figure 2.6. Stereoscopic view of the energy-minimized structure of the covalently modified 5-cathepsin B complex. The low pKa active site Cys29 residue is the site for 5-mediated covalent modification. For clarity, only the inhibitor 5 (yellow) and the key cathepsin B active site residues are shown. The key h-bonding distances of 5 with cathepsin B are shown in green. The intramolecular h-bonding interactions among cathepsin B active site residues near the inhibitor binding site are depicted in red (heavy atom h-bond donor and acceptor atoms cut off is 3.6 Å). The position of the primed and non-primed binding pockets on cathepsin B surface is shown in pink text with underline.

involved in a hydrophobic interaction with a non-polar patch containing Met196, Gly197, and Gly198 residues of S2 pocket. In addition, the ortho-ethyl group interacted with CH-π interactions with the aromatic side chains of Trp221 and His110 from S2’ pocket, albeit weakly. A conclusive design of future inhibitory ligand that are selective towards individual members of the cysteine cathepsin family will require structural elucidation of 5-inhibited
cysteine cathepsin complex using x-ray crystallography. Our modeling study suggests that ortho-
substituted aryl groups containing a negatively charged group could prove to be selective
inhibitory agents towards cathepsin B. Notably, cathepsin B is a unique member of cysteine
cathepsins in that it exhibits endopeptidase, peptidylidipeptidase and carboxypeptidase
activities.\textsuperscript{82} This broad substrate specificity of cathepsin B has been ascribed to its unique
structural feature that contains a flexible occluding loop (residues 104-126). Since Try221,
His110, and His111 are part of this occluding loop, securing potential interactions from this loop
could result in enhanced selectivity and potency towards cathepsin B. Such features could be
exploited in designing selective inhibitors of cathepsin B.

**Compound 5 is inert to external nucleophile and inhibits cysteine cathepsins
selectively.**

To evaluate if the presence of an external nucleophile could compromise the inhibitory
efficacy of 5, we carried out cathepsin B inactivation kinetics in the presence of a strong
nucleophile, azide ion (1 mM). No measurable change in inactivation kinetics was evident
thereby indicating that compound 5 is relatively inert to an external nucleophile and unleashes its
covalent reactivity only upon binding to the cathepsin B active site (data not shown). The
selectivity of inhibition of compound 5 towards cysteine cathepsins was investigated next. To do
so, a panel of distinct family of enzymes were chosen and their inhibition by 5 was assessed
(Table 2.2). While inhibition of cathepsin B by 5 was about 1.9-fold more selective compared to
Table 2.2. Selectivity of inhibition of compound 5 towards cysteine cathepsins as assessed by the 2\textsuperscript{nd} order enzyme inactivation rate constants.

<table>
<thead>
<tr>
<th>Enzyme (Enzyme Family)</th>
<th>2\textsuperscript{nd} Order Inactivation Rate Constant (M\textsuperscript{-1} Min\textsuperscript{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cathepsin B (Cysteine Cathepsin)</td>
<td>2441 ± 202\textsuperscript{a}</td>
</tr>
<tr>
<td>Cathepsin F (Cysteine Cathepsin)</td>
<td>1300 ± 208\textsuperscript{b}</td>
</tr>
<tr>
<td>Cathepsin S (Cysteine Cathepsin)</td>
<td>1196 ± 96\textsuperscript{b}</td>
</tr>
<tr>
<td>Cathepsin D (Aspartyl Cathepsin)</td>
<td>65 ± 5\textsuperscript{b}</td>
</tr>
<tr>
<td>Cathepsin G (Serine Cathepsin)</td>
<td>NI\textsuperscript{c}</td>
</tr>
<tr>
<td>Trypsin (Serine Protease)</td>
<td>NI\textsuperscript{c}</td>
</tr>
<tr>
<td>hPTP1B (Protein Tyrosine Phosphatase)</td>
<td>NI\textsuperscript{c}</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Calculated from the \(k_i\) and \(K_i\) parameters obtained using Kitz-Wilson Analysis. \textsuperscript{b}Determined by measuring pseudo-first order rate constant of inactivation \(k_{obs}\) first and dividing the \(k_{obs}\) by the inhibitor concentration. \textsuperscript{c}NI: No apparent time-dependent inhibition even at a concentration as high as 1 mM during a 1-hour inactivation reaction.

other two prominent members of cysteine cathepsins, namely cathepsin F and cathepsin S, an aspartyl-utilizing cathepsin D was only weakly (1/38\textsuperscript{th} fold compared to cysteine cathepsin B) inhibited. Cathepsin G and trypsin protease that use a Ser nucleophile for catalysis remained inert to inhibition. These experiments suggested that compound 5 is selective towards cysteine cathepsins that utilize the uniformly conserved and reactive Cys residue for catalysis. The ability of compound 5 to inhibit protein tyrosine phosphatases (PTPs) was also evaluated. This enzyme
utilizes a low pKa Cys residue in nucleophilic catalysis. Although 2-(arylsulfonyl)-oxiranes were previously reported to be modest inhibitors of PTPs\textsuperscript{75}, this enzyme also remained inert to inhibition by 5. This outcome was observed with the prototypical PTP, hPTP1B, even at a concentration as high as 1 mM and after a prolonged incubation of one hour. While a conclusive analysis of this selectivity pattern towards cysteine cathepsins will require structural elucidation in the 5-inhibited cathepsin B complex, it appears that the presence of an ethyl group at the ortho position of 2-(phenylsulfonyl)oxirane is significantly favored for binding at the cysteine cathepsin active site.

2.4 Demonstration of Intracellular Activity of Aryl Sulfonyloxiranes

Before proceeding to evaluate the cellular efficacy of 5 to inhibit cysteine cathepsins in human cells, it was important to establish its hydrolytic stability at physiological pH. This consideration was prompted by the oxirane functionality in 5 which could become hydrolytically unstable as a result of the adjacent electron-withdrawing sulfonyl moiety. This could render the inhibitor unsuitable for biological applications. Using NMR-based experiments, the half life of 5 at physiological pH (pH = 7.4) was found to be 3.6 hours (Experimental section. Figure 2.10). This study suggested that compound 5 was relatively stable at physiological pH, and therefore could be used as an important tool in biological applications.

Cysteine cathepsins play important functional roles in both intracellular and extracellular signaling events.\textsuperscript{83} Contrary to popular belief, recent studies strongly suggest that within the cell, active cysteine cathepsins play regulatory roles in distinct cellular compartments other than lysosomes, such as nucleus, mitochondrial matrix, cytoplasm, and plasma membrane.\textsuperscript{84} For example, a mis-regulated and an enhanced proteolytic activity of cytoplasmic cysteine cathepsins
have shown to be linked to tumor invasion.\textsuperscript{83} Indeed, the invasion potential of both human melanoma and prostate carcinoma cells was found to be significantly reduced only when these cells were subjected to cell membrane permeable cysteine cathepsin inhibitors, thereby suggesting the importance of intracellular cathepsins and their effective inhibition.\textsuperscript{85} Similarly, a study by Sloane et al. in human breast carcinoma BT549 cells suggested that only cell-permeable cysteine cathepsin inhibitors were able to reduce the intracellular proteolysis in these cells.\textsuperscript{42c} These investigations together underscore the importance of developing non-peptidyl and cell permeable small molecule inhibitors of cysteine cathepsins; so that their precise function can be annotated in both normal physiology and human diseases. Unfortunately, cell-permeable and non-peptidyl inhibitory agents of cysteine cathepsins remain scarcely available today.

Thus in an initial approach, we assessed whether 5 had the potential to inhibit intracellular cysteine cathepsin B activity in living cells. The metastatic MDA-MB-231 human breast cell line was chosen for this study since it has been shown that cysteine cathepsin B activity is upregulated in these cells.\textsuperscript{86} A highly efficient and specific cathepsin B substrate, Z-RR-AMC ($k_{\text{cat}}/K_m=10^5$ M$^{-1}$s$^{-1}$), previously employed in cell-based experiments, cell-based experiments, was chosen for this study.\textsuperscript{87} While efficient proteolysis of the substrate was evident in control MDA-MB-231 cells, a significant loss of intracellular cathepsin B activity was observed in 5-treated cells (Figure 2.7.A). Furthermore, intracellular cathepsin B activity was shown to decrease in a dose-dependent manner (Figure 2.7.B).

In a second approach, the usefulness of 5 to inhibit cellular cysteine cathepsin activity in cells was also investigated using a physiologically relevant proteolytic protein target of cysteine cathepsin B and S, namely E-cadherin, an adherens junction protein.\textsuperscript{43} The integrity of this epithelial adhesion protein is critically important in maintaining both intercellular contacts and
Figure 2.7. Effect of inhibitor 5 on cathepsin B activity in breast cancer MDA-MB-231 cells. [A] Cells were incubated with vehicle (DMSO 0.1%) and 20 μM of 5 for 20 minutes, followed by an additional 10 minutes of incubation with the substrate Z-RR-AMC (25 μM net) of 5. Images show a significant loss of signal intensity (green) from the enzymatic turnover of the fluorophore cathepsin B substrate Z-RR-AMC (shown by red arrow in the bottom panel) at 20 μM of 5. The control experiment without 5 indicates that a high level of substrate turnover occurs in these cells, thereby indicating a robust cathepsin B activity (white arrow – top panel). Scale bar = 10 μm.[B] The dose-dependent inhibition of intracellular cysteine cathepsin B activity with inhibitor 5 as measured by average fluorescence intensity over the entire cell. Note: The fluorescent intensity values are mean ± SD calculated over a population of at least 5-8 cells for each condition.

*Statistical significance was evaluated by the Student’s t-test.
Figure 2.8. A physiological substrate of cysteine cathepsins, E-cadherin, is enriched in 5-treated MDA-MB-231 breast cancer cells. [A] E-cadherin expression levels as detected by immunofluorescence in MDA-MB-231 cells. Cells were treated with control (0.1% DMSO) and cysteine cathepsin-B inhibitor 5 (20 μM) for 20 mins, fixed, and then stained for E-cadherin. DIC and merged fluorescence images of E-cadherin (red) and nuclear stain Hoechst 33342 (blue signal) are shown. Green and yellow arrows show enhanced E-cadherin levels (red) near the cell periphery (green arrow) and perinuclear region (yellow arrow) respectively in 5-treated cell, compared to control (0.1% DMSO). Scale bar = 10 μm. [B] A dose-dependent increase of E-cadherin expression levels from immunocytochemistry as measured by average fluorescence intensity over the cell-periphery. The fluorescence intensity values are mean ± SD calculated over a population of 28 cells for each condition. *Statistical significance was evaluated by the Student’s t-test.
cellular morphology. While it has been shown that a mere loss of cell-cell contacts alone is not sufficient to promote metastasis, a loss of E-cadherin alone has been shown to be sufficient to induce this behavior through epithelial-to-mesenchymal transition (EMT), enhanced motility, resistance to apoptosis, and invasiveness. The majority of human tumors are carcinomas that originate from cells of epithelial tissues where a complete or a partial loss of E-cadherin occurs as the tumor progressively becomes malignant. Thus, the loss of E-cadherin has been directly linked to invasive migratory behavior of cancer cells. In a study involving metastatic MDA-MB-231 breast cancer cells which expresses very low levels of E-cadherin, the re-expression of E-cadherin led to the restoration of normal E-cadherins-mediated cell-cell junctions, and a reversion to a normal epithelial morphology. We therefore evaluated the levels of E-cadherin in MDA-MB-231 cells using an E-cadherin antibody in the presence and absence of inhibitor 5. As anticipated, a dose-dependent increase in the levels of E-cadherin was observed in a population of 5-treated MDA-MB-231 cells, as compared to control cells treated only with DMSO (Figure 2.8.A and B). It should also be noted that the overall cellular morphology of metastatic MDA-MB-231 cells also changed upon inhibitory treatment (Figure 2.7.A and 2.8.A). This perhaps is not too surprising since cellular enrichment of E-cadherin is known to have a direct impact on cellular morphology. Further experiments will be required to investigate the molecular basis of altered cellular morphology in MDA-MB-231 cells upon inhibitory treatment. A western blot analysis of 5-treated MDA-MB-231 cells was also performed that exhibited an enrichment of E-cadherin levels in 5-treated cells, as compared to the control cells (Appendix D). Collectively, these studies demonstrated that 5 was effective in inhibiting the intracellular activity of cysteine cathepsins in the invasive MDA-MB-231 breast cancer cells.
An enrichment of E-cadherin levels as a result of inhibition of cysteine cathepsins by 5 is anticipated to exhibit a reduced cell migratory behavior. To investigate this, the cell motility assay was carried out in presence and absence of inhibitor 5 (control, 0.05% DMSO) on highly metastatic breast cancer MDA-MB-231 cells. As anticipated, a dose-dependent decrease in cell migratory behavior of metastatic MDA-MB-231 cells was observed (Figure 2.9). This experiment supported the notion that compound 5 inhibits cell migration potential of these highly metastatic cells. Future experiments will be planned to further optimize the inhibitory and anti-cell migratory potential of this class of compounds.

2.5 Conclusions.

We have developed a new class of small molecule inhibitors of cysteine cathepsins. A detailed kinetic and mass spectrometric analysis of cysteine cathepsin B inhibition by 5 was
performed to determine the kinetic parameters and mode of inhibition. The covalent inhibition of compound 5 was found to be selective towards cysteine cathepsins. The mechanistically distinct enzymes, such as aspartyl cathepsin D, serine cathepsin G, serine protease trypsin and a low \( \text{pK}_a \) cysteine-containing enzyme protein tyrosine phosphatase 1B, were not inhibited by 5. Molecular modeling studies of 5-inhibited cysteine cathepsin B complex indicated that 5 interacted favorably with the key active site residues surrounding S1 and S2 pockets. The inhibitor 5 was shown to be relatively stable to hydrolysis, cell-permeable, and effective in inhibiting intracellular cysteine cathepsin B activity in metastatic human breast carcinoma MDA-MB-231 cells. We also demonstrated that compound 5 was effective in reducing the migratory potential of the metastatic MDA-MB-231 breast cancer cell lines. This class of inhibitory agents could find applications in construction of selective activity-based probe of cysteine cathepsins, and in the development of selective inhibitory agents for individual members of cysteine cathepsins.\(^{94}\)
2.6 Experimental

2.6.1 Synthesis

General. $^{13}$C NMR and $^1$H spectra were recorded on a Brucker DPX 400 MHz FT NMR with automatic sample changer. Chemical shifts (δ) are reported in parts per million (ppm) and referenced to CDCl$_3$ (7.26 ppm for $^1$H and 77.0 ppm for $^{13}$C), DMSO (2.50 ppm for $^1$H and 39.5 ppm for $^{13}$C), (CD$_3$)$_2$CO (2.05 ppm for $^1$H and 29.9, 206.7 for $^{13}$C) and CD$_3$OD (4.87, 3.31 for $^1$H and 49.1 for $^{13}$C). Coupling constants (J) are reported in Hertz (Hz) and multiplicities are abbreviated as singlet (s), doublet (d), doublet of doublets (dd), triplet (t), triplet of doublets (td), and multiplet (m). The following materials were purchased from Matrix Scientific (Columbia, SC, USA) m-thiocresol, 3-methoxybenzenethiol, 3,5-dichlorothiophenol, 2,5-dimethylthiophenol, 4-(trifluoromethoxy)thiophenol, 2-ethylthiophenol and 2,5-dimethoxythiophenol. Thiophenol, 2-naphthalenethiol, chloromethyl phenyl sulfone, 4-fluorobenzaldehyde, butyraldehyde, disodium salt of silica gel (Merck, grade9385, 230-400 mesh, 60 Å) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Sodium hydroxide, hydrochloric acid, sodium chloride, sodium sulfate, ammonium chloride, and solvents were purchased from Fisher Scientific Inc (Pittsburgh, PA, USA).

2-(4-fluorophenyl)-3-(phenylsulfonyl)oxirane (I): The synthesis of this compound was achieved in a single step by appropriate modification of a published protocol$^{74}$. The procedure involved a 1-hour long reaction of 4-fluorobenzaldehyde with (chloromethylsulfonyl)benzene at room temperature in the presence of a phase transfer catalyst, tetra-n-butylammonium bromide (10 mol%), in toluene and potassium hydroxide solution. The final purification was performed using silica gel-based flash chromatography procedure that yielded compound I as a white solid.
R_f = 0.28 Hexane: EtOAc (9:1). Yield: 90%. \(^1\)H NMR (400 MHz, CDCl\(_3\)) \(\delta\) ppm: 4.16 (d, \(J = 1.4\) Hz, 1H), 4.58 (d, \(J = 1.4\) Hz, 1H), 7.06 (m, 2H), 7.24 (m, 2H), 7.64 (m, 2H), 7.75 (m, 1H), 7.99 (m, 2H). \(^{13}\)C NMR (500 MHz, CDCl\(_3\)); \(\delta\) ppm 56.94, 70.95, 116.11 (d, \(J_{C-F} = 22.3\) Hz, 2C intensity), 128.00 (d, \(J_{C-F} = 8.5\) Hz, 2C intensity), 128.53 (d, \(J_{C-F} = 3.3\) Hz, 1C), 128.88 (2C intensity), 129.56 (2C intensity), 134.70, 136.77, 163.49 (d, \(J_{C-F} = 250\) Hz, 1C intensity). LRMS [ESI-MS +ve] calculated for (C\(_{14}\)H\(_7\)FO\(_3\)S + Na): 301.03, observed: 301.25.

2-(phenylsulfonyl)-3-propyloxirane (2): This compound was synthesized using the same synthetic protocol as outlined in the synthesis of 1. The final purification was performed using silica gel-based flash chromatography procedure that yielded compound 2 as a viscous oil. R_f = 0.22 Hexane: EtOAc (9:1). Yield: 85%. \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) ppm: 0.94 (t, \(J = 7.4\) Hz, 3H) 1.47 (m, 2H), 1.57 (m, 1H), 1.66 (m, 1H), 3.63 (m, 1H), 3.90 (d, \(J = 1.6\) Hz, 1H), 7.58 (m, 2H), 7.68 (m, 1H), 7.91 (m, 2H). \(^{13}\)C NMR (500 MHz, CDCl\(_3\)) \(\delta\) ppm: 13.63, 18.93, 32.22, 57.79, 68.30, 128.69 (2C intensity), 129.41 (2C intensity), 134.43, 137.08. LRMS [ESI-MS +ve] calculated for (C\(_{11}\)H\(_{14}\)O\(_3\)S + Na): 249.06, observed: 249.25

2-(4-chlorophenylsulfonyl)oxirane (3): This compound was synthesized by following an earlier published protocol.\(^{75}\) The final purification was performed using silica gel-based flash chromatography procedure that yielded compound 3 as a white powder. R_f = 0.50 (hexane/EtOAc = 9/1). Yield: 63%. \(^1\)H NMR (400 MHz, CDCl\(_3\)) \(\delta\) ppm: 3.07 (dd, \(J = 5.4\) and 3.7 Hz, 1H), 3.37 (dd, \(J = 5.4\) and 2.1 Hz, 1H), 4.05 (dd, \(J = 3.7\) and 2.1 Hz, 1H), 7.52 (d, \(J = 8.60\) Hz, 2H), 7.81 (d, \(J = 8.60\) Hz, 2H). \(^{13}\)C NMR (400 MHz, CDCl\(_3\)) \(\delta\) ppm: 45.54, 63.25, 129.84 (2C intensity), 130.31 (2C intensity), 135.25, 141.55. LRMS [ESI-MS +ve] calculated for (C\(_8\)H\(_7\)ClO\(_3\)S + H): 218.99, observed: 219.00
2-((4-trifluoromethoxy)phenylsulfonyl)oxirane (4). This compound was synthesized by following a published protocol. The final purification was performed using silica gel-based flash chromatography procedure that yielded compound 4 as white semi-solid. \( R_f = 0.29 \) (hexane/EtOAc = 19/1). Yield: 40%. \(^1\)H NMR (400 MHz, CDCl\(_3\) \( \delta \) ppm: 3.16 (dd, \( J = 5.3 \) and 3.6 Hz, 1H), 3.46 (dd, \( J = 5.3 \) and 2.0 Hz, 1H), 4.14 (dd, \( J = 3.6 \) and 2.0 Hz, 1H), 7.43 (d, \( J = 8.90 \) Hz, 2H), 8.01 (d, \( J = 8.8 \) Hz, 2H). \(^{13}\)C NMR (400 MHz, CDCl\(_3\) \( \delta \) ppm: 45.58, 63.24, 120.20(q, \( J_{C-F} = 283 \) Hz), 121.10 (2C intensity), 131.22 (2C intensity), 134.99, 153.71. LRMS [ESI-MS +ve] calculated for (C\(_9\)H\(_7\)F\(_3\)O\(_4\)S + H): 269.01, observed: 269.00.

2-((2-ethylphenylsulfonyl)oxirane (5). This compound was synthesized by following a published protocol. The final purification was performed using silica gel-based flash chromatography procedure that yielded compound 5 as a pale yellow viscous oil. \( R_f = 0.28 \) (hexane/EtOAc = 19/1). Yield: 72%. \(^1\)H NMR (400 MHz, CDCl\(_3\) \( \delta \) ppm: 1.32 (t, \( J = 7.4 \) Hz, 3H), 3.09 (q, \( J = 7.4 \) Hz, 2H), 3.11 (DD, \( J = 5.4 \) and 3.7 Hz, 1H), 3.40 (dd, \( J = 5.4 \) and 2.1 Hz, 1H), 4.19 (dd, \( J = 3.8 \) and 2.1 Hz, 1H), 7.42 (m, 2H), 7.61 (td, \( J = 7.7 \) and 1.2 Hz, 1H), 7.99 (dd, \( J = 8.0 \) and 1.2 Hz, 1H). \(^{13}\)C NMR (400 MHz, CDCl\(_3\) \( \delta \) ppm: 16.15, 26.59, 44.87, 63.65, 126.64, 130.81, 131.33, 134.57, 134.92, 145.86. LRMS [ESI-MS +ve] calculated for (C\(_{10}\)H\(_{12}\)O\(_3\)S + H): 213.06, observed: 213.08.

2-(m-tolylsulfonyl)oxirane (6). This compound was synthesized by following a published protocol. The final purification was performed using silica gel-based flash chromatography procedure that yielded compound 6 as a colorless viscous oil. \( R_f = 0.39 \) (hexane/EtOAc = 9/1). Yield: 70%. \(^1\)H NMR (400 MHz, CDCl\(_3\) \( \delta \) ppm: 2.46 (s, 3H), 3.12 (dd, \( J = 5.4 \) and 3.7 Hz, 1H), 3.41 (dd, \( J = 5.4 \) and 2.0 Hz, 1H), 4.13 (dd, \( J = 3.7 \) and 2.0 Hz, 1H), 7.49 (m, 2H), 7.75 (s, 2H).
$^{13}$C NMR (400 MHz, CDCl$_3$) δ ppm: 139.82, 136.66, 135.39, 129.33, 129.06, 125.92, 63.25, 45.35, 21.34. LRMS [ESI-MS +ve] calculated for (C$_9$H$_{10}$O$_3$S + H): 199.05, observed: 199.25

2-(2,5-dimethylphenylsulfonyl)oxirane (7): This compound was synthesized by following an earlier published protocol. The purification using silica gel column chromatography yielded 7 as colorless viscous oil. R$_f$ = 0.32 (hexane/ EtOAc = 9/1). Yield: 68%. $^1$H NMR (400 MHz, CDCl$_3$) δ ppm: 2.39 (s, 3H), 2.67 (s, 3H), 3.10 (dd, $J$ = 5.3 and 3.7 Hz, 1H), 3.88 (dd, $J$ = 5.3 and 2.1 Hz, 1H), 4.18 (dd, $J$ = 3.7 and 2.1 Hz, 1H), 7.25 (d, $J$ = 7.5 Hz, 1H), 7.35 (d, $J$ = 7.5 Hz, 1H), 7.79 (s, 1H). $^{13}$C NMR (400 MHz, CDCl$_3$) δ ppm: 20.33, 20.80, 44.68, 63.36, 131.03, 132.83, 134.87, 135.22, 136.31, 136.78. LRMS [ESI-MS +ve] calculated for (C$_{10}$H$_{12}$O$_3$S + H): 213.06, observed: 213.08

2-(3-methoxyphenylsulfonyl)oxirane (8): This compound was synthesized by following a published protocol. The purification using silica gel column chromatography yielded 8 as a brownish viscous oil. R$_f$ = 0.27 (hexane/ EtOAc = 9/1). Yield: 70%. $^1$H NMR (400 MHz, CDCl$_3$) δ ppm: 3.13 (dd, $J$ = 5.4 and 3.7 Hz, 1H), 3.42 (dd, $J$ = 5.4 and 2.0 Hz, 1H), 3.89 (s, 3H), 4.13 (dd, $J$ = 3.7 and 2.0 Hz, 1H), 7.23 (dd, $J$ = 7.6, 2.5, and 1.9 Hz, 1H), 7.42 (dd, $J$ = 1.9 and 1.6 Hz, 1H), 7.52 (m, 2H). $^{13}$C NMR (400 MHz, CDCl$_3$) δ ppm: 160.15, 137.94, 130.54, 121.27, 120.94, 112.93, 63.25, 55.78, 45.31. LRMS [ESI-MS +ve] calculated for (C$_{10}$H$_{12}$O$_4$S + H): 215.04, observed: 214.83

2-(2,5-dimethoxyphenylsulfonyl)oxirane (9): This compound was synthesized by following a published protocol. The purification using silica gel column chromatography yielded 9 as a white semi-solid. R$_f$ = 0.24 Hexane: EtOAc (9:1). Yield: 78 %. $^1$H NMR (500 MHz, (CD$_3$)$_2$CO) δ ppm: 3.19 (dd, $J$ = 5.6 and 3.7 Hz, 1H), 3.31 (dd, $J$ = 5.6 and 2.0 Hz, 1H), 3.83 (s, 3H), 3.97
(s, 3H), 4.75 (q, $J = 3.7$ and 2.0 Hz, 1H), 7.28 (m, 3H). $^{13}$C NMR (500 MHz, (CD$_3$)$_2$CO) δ ppm: 44.56, 55.46, 56.45, 56.47, 62.31, 114.07, 114.71, 121.83, 151.82, 153.37. LRMS [ESI-MS +ve] calculated for (C$_{10}$H$_{12}$O$_3$S + H): 245.05, observed: 245.08

2-(3,5-dichlorophenylsulfonyl)oxirane (10): This compound was synthesized by following a published protocol. The purification using silica gel column chromatography yielded 10 as a white powder. $R_f = 0.67$ (hexane/ EtOAc = 9/1). Yield: 45%. $^1$H NMR (400 MHz, CDCl$_3$) δ ppm: 3.18 (dd, $J = 5.4$ and 3.7 Hz, 1H), 3.48 (dd, $J = 5.4$ and 2.0 Hz, 1H), 4.16 (dd, $J = 3.7$ and 2.0 Hz, 1H), 7.69 (t, $J=1.8$ Hz, 1H), 7.82 (d, $J=1.8$ Hz, 2H). $^{13}$C NMR (400 MHz, CDCl$_3$) δ ppm: 45.54, 63.19, 127.18 (2C intensity), 134.61, 136.54 (2C intensity), 139.72. LRMS [ESI-MS +ve] calculated for (C$_8$H$_6$Cl$_2$O$_3$S + Na): 274.93, observed: 275.00

2-(naphthalen-3-ylsulfonyl)oxirane (11): This compound was synthesized by following a published protocol. The purification using silica gel column chromatography yielded 11 as a white powder. $R_f = 0.27$ (hexane/ EtOAc = 19/1). Yield: 78%. $^1$H NMR (400 MHz, CDCl$_3$) δ ppm: 3.15 (dd, $J = 5.4$ and 3.7 Hz, 1H), 3.47 (dd, $J = 5.4$ and 2.1 Hz, 1H), 4.19 (dd, $J = 3.7$ and 2.1 Hz, 1H), 7.66 (m, 1H), 7.72 (m, 1H), 7.91 (dd, $J = 8.6$ and 1.8 Hz, 1H), 7.96 (d, $J = 8.6$ Hz, 1H), 8.03 (dd, $J = 8.4$ and 1.0 Hz, 1H), 8.06 (d, $J = 8.4$ Hz, 1H), 8.53 (d, $J = 1.0$ Hz, 1H). $^{13}$C NMR (400 MHz, CDCl$_3$) δ ppm: 45.43, 63.39, 123.09, 127.85, 128.06, 129.60, 129.71, 129.76, 131.02, 132.20, 133.69, 135.75. LRMS [ESI-MS +ve] calculated for (C$_{12}$H$_{10}$O$_3$S + H): 235.05, observed: 235.08

2-((phenylsulfonyl)methyl)oxirane (12): This compound was synthesized by following the literature-reported protocol and its characterization was found to be consistent with the previously reported NMR spectra.
2-((phenylsulfonyl)methyl)thiirane (13): This compound was synthesized by following the literature-reported protocol and its characterization was found to be consistent with the previously reported NMR spectra.

2.6.2 Enzymology, physiochemical, and mass spectrometric studies

**General.** Following enzymes and their substrates were purchased from Enzo Life Sciences (Farmingdale, NY, USA): cathepsin B (BML-SE198), cathepsin D (BML-SE199), cathepsin S (BML-SE453), cathepsin F (BML-SE568), cathepsin G (BML-SE283), human PTP1B (BML-SE332), Z-Arg-Arg-pNA (BML-P138), Z-Arg-Arg-AMC (BML-P137), Z-Phe-Arg-AMC (BML-P139), Suc-Ala-Ala-Pro-Phe-pNA (BML-P141), Mca-Gly-Lys-Pro-Ile-Leu-Phe-Phe-Leu-Lys(Dnp)-D-Arg-NH₂ (BML-P145), Z-Val-Val-Arg-AMC (BML-P199). N-alpha-benzoyl-DL-arginine-p-nitroanilide (BAPNA) (B4875), trypsin (T1426), leupeptin, para-nitrophenylphosphate and bovine serum albumin (A2153) were purchased from Sigma-Aldrich (St. Louis, MO, USA). All the biological grade buffers were used and were purchased from Sigma-Aldrich (St. Louis, MO, USA). All enzyme kinetics experiments (unless otherwise stated) were carried out at 30 °C in appropriate buffer condition with 5% DMSO concentration. The mass spectra were acquired using a MALDI-microMX mass spectrometer (Waters Corp, Milford, MA) instrument equipped with an N₂ laser operating at 10 Hz, using delayed extraction and positive ion linear mode analysis. For measuring the initial rates of enzyme catalyzed reaction, a temperature-controlled steady-state arc lamp fluorometer equipped with Felix 32 software (Photon Technology Instrument, Birmingham, NJ, USA) and a UV-Vis spectrophotometer (Model λ25; Perkin Elmer Inc., Waltham, MA, USA) was used.
Methodologies for assessing time-dependent inactivation of cysteine cathepsin B and inhibitor screening procedure:

The inactivation reaction of human liver cathepsin B (net 175 nM) with inhibitor was performed under pseudo first order condition ([I]>>[E]) in a 0.5 ml eppendorf tube maintained at 30 °C in a temperature controlled bath. Freshly made Inhibitor stock solution (in DMSO) was used for each experiment. A previously reported cathepsin B assay procedure was employed. Cathepsin B enzyme was first activated for 10 minutes in sodium phosphate buffer (100 mM, pH 6.0) containing 5 mM DTT and 1 mM Na2EDTA. The inactivation reaction was initiated by the addition of inhibitor where DMSO concentration was maintained at 5%. After suitable time intervals, an aliquot of 10 μl of incubation mixture was withdrawn and added to an assay mixture (net assay volume 175 μl) that contained 3mM of Z-RR-pNA (Km = 0.9 mM) substrate at 30 °C. A progress curve was recorded at 405 nM, and the enzyme activity was determined by measuring the initial rates of substrate turnover. For initial screening experiments, the assay protocol remained the same except that the concentration of all inhibitors was fixed at 500 μM and the residual enzyme activity was measured after 10 minutes of reaction.

Method for estimation of kinetic parameters for 5-mediated cathepsin B inhibition:

To obtain Ki and ki parameters for the two step inhibition model, the following procedure was adopted. A sequence of independent cathepsin B inactivation experiments were performed at 25, 50, 75, 100, and 150 μM of 5 under pseudo-first order condition ([I]>>[Cathepsin B]) and enzyme activities were monitored at appropriate time intervals. The data thus obtained at each concentration were plotted and fitted to the following equation:

\[ A_t = A_f - (A_f-A_0) e^{(-k_{obs} t)} \]

where

- \( A_t \) = enzyme activity at time t of cathepsin B inactivation,
- \( A_f \) = final cathepsin B activity at
infinite time of inactivation, \( A_0 \) = initial enzyme activity at zero time of inactivation, and \( k_{\text{obs}} \) = pseudo-first order rate constant of inactivation. Kitz-Wilson analysis is then performed to estimate the reversible equilibrium binding constant \( K_i \) for the first step, and the first order inactivation rate constant \( (k_i) \) for the second irreversible step, as described elsewhere. For demonstration that compound 5 is active site-directed, the inactivation reaction of 5 with cathepsin B was performed in presence of an established competitive inhibitor of cathepsin B, leupeptin (40 nM) and the residual cathepsin B activity was measured, as described above. Such a methodology has been previously utilized to demonstrate the active site-directed nature of inhibition for previously uncharacterized covalent and irreversible inhibitory agents.

**Cathepsin F inhibition assay:**

A previously reported enzyme assay procedure was adopted with appropriate modifications. The enzyme was activated for 10 minutes in 50 mM potassium phosphate buffer at pH 6.5 containing 2.5 mM DTT, 2.5 mM Na\(_2\)EDTA, and 1.0 mg/ml BSA at 25 °C. The inactivation reaction was initiated by addition of inhibitor 5 to activated human cathepsin F (net 1.1 nM). After suitable time intervals, an aliquot of 10 μl of incubation mixture was withdrawn and the enzyme activity was measured in an assay buffer (net 175 μl) containing 3 μM of Z-FR-AMC fluorogenic substrate \( (K_m = 0.44 \mu M) \) at 25° C.

**Cathepsin S inhibition assay:**

For assaying cathepsin S activity, a previously published protocol was followed with appropriate modification. Thus, cathepsin S enzyme was activated first in the 100 mM potassium phosphate buffer (pH 6.5) containing 5 mM DTT, 5 mM Na\(_2\)EDTA for 10 minutes. The inactivation reaction of human Cathepsin S (net 134.6 nM) with inhibitor 5 (75 μM) was
initiated. After suitable time intervals, an aliquot of 3 μl of incubation mixture was withdrawn and the enzyme activity was measured in assay buffer (175 μl; 8% DMSO) containing 250 μM of Z-VVR-AMC (K_m >100 μM^{102}; Ex/Em: 365/440 nm) fluorogenic substrate.

**Cathepsin G inhibition assay:**

A previously reported assay conditions were followed^{103}. The enzyme was activated in 20 mM Tris-HCl buffer (pH 7.4) containing 500 mM NaCl for 15 minutes. The inactivation reaction of human Cathepsin G (net 7.66 nM) with inhibitor 5 was initiated. After suitable time intervals, an aliquot of 10 μl of incubation mixture was withdrawn and the enzyme activity was measured in assay buffer containing 3.4 mM Suc-AAPF-pNA cathepsin G substrate (K_m = 1.7 mM).

**Cathepsin D inhibition assay:**

The enzyme was activated in 50 mM sodium acetate buffer (pH 4) containing 5 mM DTT and 1 mg/ml bovine serum albumin. The inactivation reaction of human Cathepsin D (net 20 nM) with inhibitor was initiated. After suitable time intervals, an aliquot of 1 µl of incubation mixture was withdrawn and added to an assay mixture (net assay volume of 175 µl) that contained 40 µM of Mca-Gly-Lys-Pro-Ile-Leu-Phe-Phe-Arg-Leu-Lys(Dnp)-D-Arg-NH₂ [Mca=(7-methoxycoumarin-4-yl)acetyl] cathepsin D substrate (K_m = 3.7µM). Enzyme activity was monitored by measuring fluorescent enhancement over the time.

**Serine protease (trypsin) inhibition assay:**

The assay procedure was adopted from an earlier published protocol^{104}. Trypsin enzyme solution was freshly made using 1 mM HCl just before the inactivation reaction. Inactivation reaction was initiated by the addition of inhibitor to a pre-activated trypsin (net concentration of
190-200 U/μl). After suitable time intervals, an aliquot of 2 μl incubation mixture were withdrawn and added to a 67 mM sodium phosphate assay buffer (net assay volume of 175 μl) which contained 3 mM N-alpha-benzoyl-DL-arginine-p-nitroanilide (BAPNA, \( K_m = 0.94 \) mM) as a substrate. Enzyme activity was monitored by measuring the initial rate of substrate consumption at 410 nm.

**Protein tyrosine phosphatase 1B (PTP1B) inhibition assay:**

The time-dependent inactivation kinetics of 5 with PTP1B was performed as described before.

**Mass-spectral studies:**

Cathepsin B was activated and incubated with inhibitor under conditions similar to those used for kinetics studies for 10 minutes. Both 5-inhibited cathepsin B inhibitory complex and control (5% DMSO) were prepared for matrix-assisted laser desorption/ionization mass spectrometry using the ultra-thin layer essentially as published.\(^{105}\) Samples were saturated in a 1:9 ratio with a saturated solution of recrystallized \( \alpha \)-cyano-4-hydroxycinnamic acid in 50% formic acid/ 33% isopropyl alcohol/ 17% water (v/v/v; Fisher Scientific- all HPLC grade). Acceleration potential of 12,000 V and an extraction pulse of 1,950 V with a delay time for 410 ns were used. Spectra were smoothed and baseline corrected using MassLynx (Waters Corp) and calibrated pseudo-internally using equine apomyoglobin.

**Hydrolytic stability of inhibitor 5:**

To study the hydrolytic stability of compound 5, NMR experiment was carried out in following manner. First, sodium phosphate buffer (10 mM) of pH 7.4 was prepared. A portion (5
ml) of this buffer was lyophilized and re-dissolved in deuterium oxide. This step was repeated thrice so as to ensure a complete H/D exchange. The buffer components were re-dissolved in deuterium oxide (5 ml). The hydrolysis experiment was carried out under pseudo first order reaction condition. A stock solution of 5, freshly made in deuterated DMSO was added to deuterated phosphate buffer (total volume 500 μl; net concentration of 5 at 2.5 mM) and NMR spectra were recorded after appropriate time intervals. The rate of hydrolysis was measured by comparing the integration values of initial and final compounds. The rate constant of hydrolysis was obtained by fitting the data to a single exponential decay curve \( I_t = I_0 e^{-kt} \) and the half was calculated using the following equation: \( t_{1/2} = 0.693/k \).

**Figure 2.10.** Estimation of the half life of compound 5 at physiological pH 7.4. The relative peak intensity of 5 was plotted against time and the data was fitted To a single exponential. The \( t_{1/2} \) was estimated to be 3.6 hours.
Molecular Modeling

The molecular modeling studies were performed on a Dell PowerEdge 6000 server operating on Red Hat Enterprise Linux ES release 4. The CH$_3$CONH$_2$CysCONHCH$_3$ peptide covalently attached to inhibitor 5 at the terminal aliphatic carbon was built with open oxirane ring in Builder module of Maestro 7.5 software and partial charges on each atoms were determined in Jaguar (Schrodinger Inc., New York, NY, USA) using density-functional theory (DFT) calculation with B3LYP/6-31G* basis set. The coordinate file of cathepsin B (PDB id: 2DC6) was downloaded and appropriately modified at pH 7.0 to include inhibitor 5 covalently linked to Cys29Sγ residue using Insight II software (Accelrys Inc., SanDiego CA, USA). The partial charges on the atoms of inhibitor were assigned from DFT calculations and the complex was soaked with a 15 Å spherical shell of water with Cys29Sγ as the center. The resulting complex was energy minimized (10,000 steps of Steepest Descent). A molecular dynamics run (100 ps) was then carried out on the energy-minimized complex using Consistent Valence Forcefield (CVFF) force field in Discover standalone program (Accelrys Inc., San Diego CA, USA). The trajectories from molecular dynamic run were saved every 2 ps and analyzed. During the energy minimization procedure and the molecular dynamics simulations, no constratins were imposed on 5-cathepsin B inhibitor complex.

Cell culture:

All the cell culture reagents were purchased from Invitrogen (Carlsbad, CA). MDA-MB-231 cells (American Type Culture Collection, Manassas, VA) were cultured on 10-cm plates (BD Falcon™) at 37°C and 5% CO$_2$ in Iscove’s Modified Dulbecco’s medium with L-glutamine (IMDM), 10% fetal bovine serum, and antibiotics (1% penicillin/streptomycin and 0.5 μg/ml
fungizone) as described earlier. Cells were passaged at 1:3 ratios every 3–4 days and subcultured by 0.25% trypsin-EDTA treatment. Cells were re-plated one day prior to experiment so that the cell density was 70-80% on the day of the experiment.

**Confocal microscopy:**

For live-cell imaging, cells were seeded to Poly-D-lysine coated 35mm-glass bottom dishes (MatTek; Ashland, MA) and grown overnight in complete medium. Prior to imaging, cells were treated in 15 mM HEPES containing phenol-red (PR)-free complete medium. To reduce photo-bleaching, 0.6u of Oxyrase enzyme per ml (Oxyrase Inc.) was added to the PR-free medium containing 10 mM sodium-lactate, as previously described. Cells were incubated with either dimethylsulfoxide (DMSO, 0.1% v/v) as vehicle control, or varying concentrations of 5 for 20 minutes at 37°C, followed by an additional 10 minutes of incubation with the substrate Z-RR-AMC (25 μM net) of 5. Images of cells were acquired using an inverted Leica TCS-SP5 confocal microscope (Leica; IL) by Plan apochromat 63x1.4 oil objective lens (numerical aperture 0.6-1.4). An argon ion laser (25%) was used to generate an excitation at 488 nm, and pinholes were typically set to 1–1.5 Airy units. Results for every condition are the average over three independent experiments in each of which 8-10 cells were analyzed by Image J.

**Immunofluorescence:**

Cells, grown on 35-mm glass coverslips in 6-well plate for 24 hrs, were treated with the vehicle (0.1% DMSO), or appropriate concentrations of cathepsin B inhibitor for 20 minutes at 37°C. After treatment they were fixed in 3.7% PFA at room temperature for 10 minutes followed by 30 sec incubation in microtubule-stabilizing and cell-permeabilizing buffer (80 mM PIPES, 1 mM MgCl₂, 4 mM EGTA, 0.5% Triton-X100) at room temperature. The samples were then
washed thrice with phosphate-buffered saline (PBS) and incubated in immunoblocker (2% BSA, 0.1% sodium azide) for 15 minutes at 37°C.

For immunostaining, cells were incubated overnight at 4°C with monoclonal mouse antibody E-cadherin (1:100) (Santa Cruz Biotechnology, Santa Cruz, CA) followed by incubation with secondary Alexa 594-conjugated donkey anti-mouse antibody (1:300) (Invitrogen, Carlsbad, CA) for 2 hrs at 37°C. DNA was stained with Hoechst 33342 (1 mg/ml in PBS). Cells were mounted on clean glass slides using mounting media DABCO (Invitrogen). Samples were examined with a Plan-Neofluar 40x oil immersion objective lens (numerical aperture 0.75) using a Zeiss Axio Imager-M2 upright fluorescence microscope (Zeiss, Heidelberg, Germany). Results for every condition are the average of three independent experiments in each of which 28 cells were analyzed by ImageJ software.

**Western-blot analysis:**

Cell lysate proteins were (50 µg) resolved by an 8% SDS-PAGE on Duracryl followed by electrophoretic transfer to 15x15 cm Immobilon-P (Millipore, Bedford, MA) nitrocellulose membranes. After blocking with 5% nonfat dry milk, membranes were incubated overnight at 4°C with E-cadherin (1:500) (Santa Cruz Biotechnology, Santa Cruz, CA, USA) and DM1A-α-tubulin mouse monoclonal antibody (1:5000) (Sigma-Aldrich, St. Louis, MO, USA), followed by incubation with peroxidase-conjugated secondary antibodies and chemiluminescence detection (Supersignal West Pico, Pierce).

**Cell Motility Assay:**

The effect of inhibitor 5 on cell migratory behavior of MDA-MB-231 cells was evaluated by following a slight modification of an earlier published protocol. Thus MDA-MB-231 cells
were plated as concentric circles on a 10-well glass slide pre-coated with 1% BSA through a 10-hole cell sedimentation manifold (CSM, Inc., Phoenix, AZ) followed by incubation at 37 °C and 5% CO₂. After removal of the manifold, cells were treated with control vehicle (0.05% DMSO) or appropriate concentrations of inhibitor 5, and incubated for 6 hours at 37 °C. Images were acquired by a digital camera (Moticam 2000) attached to an inverted Nikon Diaphot microscope both before (t0) and after inhibitory treatment (t6). Motility of cells was calculated by the change in total area (in square micrometers) by Image J software. Each reported value is the average of triplicate measurements with corresponding standard deviation value (S.D.).
Chapter 2

Part B.

Serendipitous Discovery: Base-Catalyzed Synthesis of Allyl and Vinyl Phenyl Sulfones
2.7. Background

A diverse group of important synthetic intermediates can be widely accessed through the chemistry of the aryl sulfonyl functional group. Vinyl and allyl sulfonyl esters are important members of this class of compounds and are used as intermediates in a variety of organic transformations. For example, vinyl aryl sulfonyl esters have been successfully utilized to develop potent and selective inhibitory agents and activity-based probes of cysteine proteases. Similarly, allyl aryl sulfonyl esters are convenient intermediates for synthesizing natural products, such as \((-\)-erythrodiene], squalene, and coenzyme Q10.

Most synthetic methods for preparation of vinyl sulfonyl esters start with an alkene or 1-haloalkene derivative and involve multiple steps. One of the challenges in the synthesis of vinyl sulfonyl esters lies in the increased preference for \(\beta, \gamma\) over \(\alpha, \beta\) unsaturation. This often leads to isomeric mixtures of vinyllic and allylic sulfonyl esters. Isomerically pure vinyl sulfonyl esters are easily synthesized in systems where \(\beta, \gamma\) unsaturation is not possible. Most allylic sulfonyl esters are either prepared by treating vinyllic sulfonyl esters with a base, or by a direct conversion of allylic alcohols or alkenes in the presence of a metal catalyst. Recently, the reaction of benzylic alcohols with sulfinyl chlorides and arenensulfonyl cyanides with allylic alcohols were utilized to produce allylic sulfonyl esters.

2.8 Discovery

In our effort to make vinyl sulfonyl esters as inhibitors of certain classes of therapeutically important enzymes, we followed established literature protocols for analogous reactions that involved the use of NaH with sulfonyl methylphosphonate. Thus, I was treated with 2-phenylacetaldehyde in the presence of 2.0 equiv NaH at 0 °C resulting in a quantitative
conversion to allylic sulfone 9 without any trace of the desired vinylic sulfone 2 (Scheme 2.2). Even when we reduced the number of equivalents of NaH to 1.1 equiv, only allylic sulfone 9 formed, again with no vinylic sulfone 2. The use of one equiv of NaH however yielded predominantly the vinylic sulfone 2 (4:1 vinylic:allylic mixture). Surprisingly, a further reduction of NaH (0.9 equiv) led to exclusive formation of vinylic sulfone 2, with no trace of the allylic sulfone 9. Importantly, all reactions went to completion within an hour, and any additional time of reaction did not affect either the regioselectivity of the products or their yield. This is an interesting finding since many of the literature methods involve significantly higher reaction times, and often give products with compromised regioselectivity.114

Our efforts to ascertain the structure of 2 was hampered by previously misreported spectral characteristics of this compound114b, 120. Therefore, a complete structural analysis by NMR was undertaken to unambiguously establish the structures of 2 and 9 (Table 2.2 for complete 1H and 13C assignment for 2 and 9). The 1H-1H scalar coupling between the olefinic protons was more than 15 Hz indicating a trans geometry for both compounds. The relative position of the double bond with respect to the phenyl ring was established from 1H-13C long
range correlation (HMBC) experiments (Appendix B). In addition, we also detected NOE between methylene protons and the ortho protons of the adjacent phenyl ring in 2 while a similar NOE was observed between one olefinic proton and the ortho protons of the adjacent phenyl ring in 9. A complete assignment of $^1$H and $^{13}$C NMR peaks for 2 and 9 also established the characteristic spectral differences between the α, β unsaturation versus the β, γ unsaturation. For example, the methylene carbon in 9 experiences the –I effect of the sulfone and is 23 ppm downfield shifted compared to the methylene carbon in 2. The -R effect of the sulfone can be observed on the β olefinic carbon in 2 which is substantially downfield shifted and observed at 145.42 ppm. The structure of 2 was also confirmed by X-ray crystallography (Figure 2.11).\textsuperscript{121}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2_11.png}
\caption{X-ray crystal structure of 2. ORTEP drawing with complete numbering of atoms, viewed approximately normal to the alkene plane.}
\end{figure}
**Table 2.2.** $^1$H and $^{13}$C NMR Chemical Shifts for vinyl phenyl sulfone, 2 and allyl phenyl sulfone, 9.$^a$

![Diagram of vinyl phenyl sulfone and allyl phenyl sulfone](image)

<table>
<thead>
<tr>
<th>Compound 2 $^{13}$C shifts (ppm)</th>
<th>Compound 9 $^{13}$C shifts (ppm)</th>
<th>Compound 2 $^1$H shifts (ppm), J values in Hz</th>
<th>Compound 9 $^1$H shifts (ppm), J values in Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 131.53</td>
<td>60.51</td>
<td>6.26 (dt, 15.1, 1.7)</td>
<td>3.96 (dd, J=7.6, 1.2)</td>
</tr>
<tr>
<td>2 145.42</td>
<td>115.12</td>
<td>7.16 (dt, 15.1, 6.5)</td>
<td>6.10 (dt, 15.8, 7.6)</td>
</tr>
<tr>
<td>3 37.64</td>
<td>139.25</td>
<td>3.56 (dd, 6.5, 1.7)</td>
<td>6.36 (dt, 15.9, 1.2)</td>
</tr>
<tr>
<td>a 140.5</td>
<td>138.30</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>b 127.64</td>
<td>128.55</td>
<td>7.86 (dd, 8.4, 1.4)$^b$</td>
<td>7.88 (dd, 8.2, 1.3)$^b$</td>
</tr>
<tr>
<td>c 129.27</td>
<td>129.14</td>
<td>7.54 (tt, 7.5, 1.2)$^b$</td>
<td>7.54 (tt, 7.5, 1.3)$^b$</td>
</tr>
<tr>
<td>d 133.34</td>
<td>133.84</td>
<td>7.62 (tt, 7.5, 1.3)$^b$</td>
<td>7.64 (tt, 7.5, 1.3)$^b$</td>
</tr>
<tr>
<td>a’ 136.21</td>
<td>135.73</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>b’ 128.84</td>
<td>126.64</td>
<td>7.13 (dd, 8.2, 1.4)$^b$</td>
<td>7.28 (m)</td>
</tr>
<tr>
<td>c’ 128.89</td>
<td>128.70</td>
<td>7.32 (tt, 7.5, 1.4)$^b$</td>
<td>7.30 (m)</td>
</tr>
</tbody>
</table>
Chemical shifts were measured relative to TMS at 20 °C on a 500 MHz NMR spectrometer. Solvent is CDCl₃. To make comparison of shifts easier, the numbering of carbons is kept the same in both compounds and is shown in the structures above. Complete assignments were made using 1D ¹H and ¹³C as well as 2D homo- and hetero-nuclear correlation and NOESY experiments.

These protons exhibit second-order splitting, the appearance of which may deviate by a small extent from the pattern mentioned here depending on the spectrometer frequency.

After clearly establishing the structures of 2 and 9, we used six additional arylacetaldehyde substrates to probe the generality of the regio- and stereochemistry (Scheme 2.2). Thus, when a panel of 2-arylacetaldehydes, shown in scheme 1, were synthesized and subsequently treated with 1 using 0.9 and 1.1 equiv of NaH, the corresponding vinyl and allyl sulfones in high yields were obtained with exclusive E stereochemistry. To probe the importance of aryl substitution at the 2-position of acetaldehyde, the aliphatic aldehyde butanal was first examined. Thus, when butanal was allowed to react with 1 in the presence of 0.9 equiv of NaH, only the vinylic derivative 16 was observed with diminished stereoselectivity (Table 2.3). At 1.1 equivalents of NaH, a reaction of butanal with 1 gave allyl sulfone, but a substantial amount of vinyl sulfone was also present (Table 2.3). When 2.2 equiv. of NaH was used, the reaction gave allylic sulfone 19 as the major product, again with diminished stereoselectivity. Table 2.3 lists the distribution of regioisomers with their E/Z ratios obtained for a panel of 2-alkylacetaldehydes. With 0.9 equiv of NaH, exclusive formation of the vinylic sulfones was observed but again the corresponding allylic sulfones formed with diminished regio- and stereoselectivity when 1.1

| d’   | 127.11 | 128.57 | 7.26 (tt, 7.5, 1.4)ᵇ | 7.28 (m) |

ᵃ Chemical shifts were measured relative to TMS at 20 °C on a 500 MHz NMR spectrometer. Solvent is CDCl₃. To make comparison of shifts easier, the numbering of carbons is kept the same in both compounds and is shown in the structures above. Complete assignments were made using 1D ¹H and ¹³C as well as 2D homo- and hetero-nuclear correlation and NOESY experiments.

ᵇ These protons exhibit second-order splitting, the appearance of which may deviate by a small extent from the pattern mentioned here depending on the spectrometer frequency.
equiv of NaH was used. These results, taken together, indicate that the base-controlled regio- and stereochemistry is sensitive to the nature of the α-substituent of the acetaldehyde.

**Table 2.3.** Treatment of 1 with 2-alkylacetaldehydes under varying NaH concentrations.

<table>
<thead>
<tr>
<th>R</th>
<th>With 0.9 equiv NaH</th>
<th>With 1.1 equiv NaH</th>
</tr>
</thead>
<tbody>
<tr>
<td>-CH₂CH₃</td>
<td>100% 16 (5/1), 0% 19</td>
<td>58% 16 (13/1), 42% 19 (2/1)</td>
</tr>
<tr>
<td>-CH(CH₃)₂</td>
<td>100% 17 (1/0), 0% 20</td>
<td>67% 17 (1/0), 33% 20 (1/1)</td>
</tr>
<tr>
<td>-(CH₂)₄CH₃</td>
<td>100% 18 (7:3), 0% 21</td>
<td>72% 18 (7/3), 28% 21 (1/1)</td>
</tr>
</tbody>
</table>

*Peaks corresponding to each isomer were integrated and normalized from the NMR spectrum of the crude mixture. ¹H NMR, ¹³C NMR, and mass spectrometric characterization of 16 and 19 are included in the experimental section. Complete characterization of NMR peaks was not carried out for compounds 17, 18, 20, and 21 as these compounds were previously reported and characterized.*

To probe if a strong base such as NaH was essential for the observed regio- and stereoselectivity, we carried out the reactions of 1 with 2-phenylacetaldehyde and 1 with butanal in the presence of a weaker base, potassium tert-butoxide (1.1 equiv). In both cases, the exclusive regioselectivity was lost as the reaction with 2-phenylacetaldehyde yielded a 11:14...
mixture of 9 and 2 while butanal gave a 17:3 mixture of 16 and 19. To assess if a base stronger than NaH could drive an exclusive conversion of butanal to allylic sulfones 19, a reaction of 1 with butanal was carried out in the presence of 1.1 equiv of n-butyllithium. Surprisingly, the reaction yielded exclusively the vinyl sulfones 16, without any trace of corresponding allylic sulfones. This perhaps could be due to the compromised stability of n-butyllithium reagent in the reaction mixture under our experimental conditions (typically 0-10 °C, 1.5 hrs). Thus, it appears that an optimal basicity of NaH is required for this regio- and stereoselective transformation.

Considering the results from Table 2.3, it is clear that aryl substitution at the γ position of the vinylic sulfones plays a crucial role for regio- and stereoselectivity, and only a catalytic amount of NaH is necessary for the formation of allyl sulfone from vinyl sulfone. To gain further insight into the mechanism of this transformation, a computational analysis was undertaken [All calculations were carried out in vacuum using Spartan program. (Spartan '10, version 1.1.0, Wavefunction Inc., Irvine, CA, 2011). The structures of the E and Z isomers of 2, 9, 16, and 19 were minimized using MM2 calculations. The minimized structures were then further subjected to DFT calculations using B3LYP/6-311G* basis set to obtain the equilibrium geometries and total energies. From the energy-minimized structures of 2 and 16, one of the methylene protons was removed to create the anion.]. The standard heat of formation (ΔH°f) by T1 method\textsuperscript{124} as well as total energy (ΔE) by DFT methods showed that allylic sulfones 9 and 19E were thermodynamically more stable than their corresponding vinylic sulfones 2 and 16E by a comparable amount [ΔΔH°f favors 9 and 19E by -13.1 kJ/mol and -14.9 kJ/mol respectively, and ΔE favors 9 and 19E by -20.4 kJ/mol and -24.1 kJ/mol respectively]. The energy-minimized geometry for the sulfones as well as the ΔE value for the conversion of 2 to 9 agrees closely with the estimated values reported by Lee et al.\textsuperscript{125}. To understand why only
vinylic sulfone 2 would convert easily to allylic sulfone 9 with catalytic NaH but not the vinylic sulfone 16, we compared the differences between the neutral sulfones and the anions obtained after proton abstraction to approximate the energy barrier for this conversion. DFT calculations at the B3LYP/6-311G* level showed that the difference is lower by 41.3 kJ/mol for 2 compared to 16E. It can be inferred that the resonance stabilization of the resulting negative charge by the aromatic ring (23 – Scheme 2.3) plays a dominant role in the ease with which the allylic sulfone forms. Based on our experimental observation and computational analysis, a plausible mechanism for the quantitative conversion of 1 to 2 and 9 is proposed. The reaction proceeds through a cyclic Wittig intermediate 22 which readily breaks down to yield the vinylic sulfone 2. When a catalytic excess of base is present, an acidic proton is abstracted from 2 to form a resonance-stabilized carbanion 23. This carbanion acts as a base to abstract another proton from 1 to form the thermodynamically stable product 9, thus regenerating carbanion 23.

**Scheme 2.3.** Proposed mechanism of the formation of allylic sulfone 9 when catalytic excess of sodium hydride is used.
which can subsequently turn over another substrate molecule 2. Catalytic participation of NaH has been observed previously.\textsuperscript{127}

Having established the effect of conjugation from one phenyl moiety and to gain further support for our postulated mechanism, we hypothesized that 2,2-diphenylacetaldehyde will form the allyl sulfone even more favorably. We therefore synthesized 2,2-diphenylacetaldehyde\textsuperscript{128} and carried out a reaction with 1 using 0.9 equiv NaH. Not surprisingly, the reaction exclusively gave the allylic sulfones 24 (Scheme 2.4), a compound that has been previously prepared by other methods.\textsuperscript{129} When this reaction was repeated with 0.5 equiv of NaH, it still resulted in a quantitative conversion of the 2,2-diphenyl-acetaldehyde to 24 (50\% overall yield) with no trace of the vinyl sulfone. Indeed the $\Delta E$ calculations by DFT methods show that formation of 24 is favored by 35.9 kJ/mol over its corresponding vinylic isomer (i.e. about twice the amount calculated for 9). Both experimental results and theoretical calculations confirm that an extended conjugation favors the formation of the allylic sulfones over the corresponding vinylic analogues.

![Scheme 2.4](image)

**Scheme 2.4.** Exclusive formation of allyl sulfone under varying base equivalent.

### 2.9 Concluding Remarks

In conclusion, we have described a convenient general route to synthesize vinylic and allylic sulfones that contain an aromatic group on the $\gamma$-carbon of the sulfones. The reactions
described here proceed in just 1 hr time, do not require any metal catalyst, and yield exclusively a single product with high regio- and stereoselectivity. Importantly, this can be achieved simply by changing the amount of NaH used in the reaction mixture. Thus, while a catalytic excess of NaH (1.1 equiv) produces the allylic sulfones in quantitative yield upon reaction of \(1\) with 2-arylacetaldehydes, a sub-stoichiometric amount of NaH (0.9 equiv) produces only the corresponding vinylic sulfones.\(^{130}\)
2.10 Experimental

General \(^{13}\)C NMR and \(^1\)H spectra were recorded on a Brucker DPX 400 MHz FT NMR with automatic sample changer. Chemical shifts (\(\delta\)) are reported in parts per million (ppm) and referenced to CDCl\(_3\) (7.26 ppm for \(^1\)H and 77.0 ppm for \(^{13}\)C), (CD\(_3\))\(_2\)CO (2.05 ppm for \(^1\)H and 29.9, 206.7 for \(^{13}\)C). Coupling constants (\(J\)) are reported in Hertz (Hz) and multiplicities are abbreviated as singlet (s), doublet (d), doublet of doublets (dd), triplet (t), triplet of doublets (td), and multiplet (m).

Synthesis

Typical procedure for synthesis of 2-8: A solution of diethyl ((phenylsulfonyl)methyl)phosphonate (500 mg, 1.7 mmol, 1 equiv) in dry tetrahydrofuran (2 mL) was added dropwise to a stirring suspension of sodium hydride (39 mg, 1.53 mmol, 0.9 equiv) in dry tetrahydrofuran (5 mL) at 0 °C under inert condition. After 30 minutes a solution of \(\alpha\)-substituted acetaldehyde (1.7 mmol, 1.0 equiv) in dry tetrahydrofuran (3 mL) was added to the reaction mixture. Stirring was continued for an hour. The reaction was quenched by addition of 5 mL of sodium bisulfate (5 wt% solution). The solution was concentrated under vacuum, diluted with water (10 mL) and extracted with ethylacetate (3 x 5 mL). The combined organic extracts were washed with saturated aqueous sodium bicarbonate and brine, dried over sodium sulfate, filtered and concentrated under vacuum. Purification by flash column chromatography (silica gel; 97:3 hexane:ethylacetate) provided the desired product.

Compounds 9-15 were synthesized by following the same protocol used in the synthesis of 2-8 except that 1.1 equiv (44.9 mg, 1.87 mmol) of NaH was used. For the synthesis of 24, E and Z
isomers of 16-21, appropriate amount of NaH was utilized (see Table 2). The characterizations of compounds are as follows:

(E)-((3-phenylprop-1-en-1-yl)sulfonyl)benzene (2). Yield: 81%. [see Table 1 for $^1$H and $^{13}$C NMR assignments]. LRMS [ESI-MS +ve] calculated for (C$_{15}$H$_{14}$O$_2$S+Na): 281.1, observed: 281.3.

(E)-4-(3-(phenylsulfonyl)allyl)-1,1'-biphenyl (3). Yield: 74%. $^1$H NMR (400 MHz, (CDCl$_3$) δ ppm: 3.60 (d, J = 6.6 Hz, 2H), 6.30 (d, J = 15.2 Hz, 1H), 7.14-7.66 (m, 13H), 7.88(m, 2H). $^{13}$C NMR (100 MHz, (CDCl$_3$) δ ppm: 37.27, 127.04, 127.39, 127.61, 127.67, 128.82, 129.28, 129.29, 131.64, 133.37, 135.21, 140.12, 140.47, 140.58, 145.30. LRMS [ESI-MS +ve] calculated for (C$_{19}$H$_{16}$O$_2$S+Na): 357.1, observed: 357.2

(E)-1-(3-(phenylsulfonyl)allyl)naphthalene (4). Yield: 71%. $^1$H NMR (500 MHz, (CD$_3$)$_2$CO) δ ppm: 4.14 (d, J = 6.4 Hz, 2H), 6.51(d, J =15.1 Hz, 1H), 7.20-8.09 (m, 13H). $^{13}$C NMR (125 MHz, (CDCl$_3$) δ ppm: 34.36, 123.77, 125.71, 125.85, 126.27, 127.10, 127.35, 127.76, 128.76, 129.35, 131.68, 131.97, 133.28, 133.53, 134.07, 141.22, 145.47. LRMS [ESI-MS +ve] calculated for (C$_{19}$H$_{16}$O$_2$S+Na): 331.1, observed: 331.2

(E)-1-methyl-2-(3-(phenylsulfonyl)allyl)benzene (5). Yield: 66%. $^1$H NMR (500 MHz, CDCl$_3$) δ ppm: 2.20 (s, 3H), 3.56 (dd, J = 6.2 and 1.4 Hz, 2H), 6.13 (dt, J = 15.2 and 1.4 Hz, 1H), 7.05 (d, J = 6.9 Hz, 1H), 7.16 (m, 4H), 7.53 (t, J =7.6 Hz, 2H), 7.61 (t, J = 7.6 Hz, 1H), 7.85 (d, J = 7.6 Hz, 2H). $^{13}$C NMR (100 MHz, (CDCl$_3$) δ ppm: 19.3, 35.4, 126.5, 127.4, 127.6, 129.2, 129.6, 130.6, 131.5, 133.3, 134.6, 136.3, 140.6, 145.1. LRMS [ESI-MS +ve] calculated for (C$_{16}$H$_{16}$O$_2$S+Na): 295.1 Observed: 295.0 Calculated Mass for (C$_{16}$H$_{16}$O$_2$S+Na): 295.1 Observed: 295.0
(E)-1-chloro-2-(3-(phenylsulfonyl)allyl)benzene (6). Yield: 69%. $^1$H NMR (500 MHz, CDCl$_3$) δ ppm: 3.63 (dd, J = 6.3 and 1.4 Hz, 2H), 6.23 (tt, J = 15.0 and 1.4 Hz, 1H), 7.13 (dt, J = 15.0 and 6.3 Hz, 1H), 7.17 (m, 1H), 7.22 (m, 2H) 7.37 (m, 1H), 7.53 (t, J =7.7 Hz, 2H), 7.62 (t, J = 7.7 Hz, 1H), 7.86 (d, J = 7.7 Hz, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ ppm: 35.4, 127.3, 127.6, 128.8, 129.3, 129.9, 130.8, 131.9, 133.3, 134.1, 134.3, 140.4, 143.7. LRMS [ESI-MS +ve] calculated for (C$_{15}$H$_{13}$ClO$_2$S): 315.0 Observed: 315.0

(E)-1-methyl-4-(3-(phenylsulfonyl)allyl)benzene (7). Yield: 75%. $^1$H NMR (400 MHz, CDCl$_3$) δ ppm: 2.32 (s, 3H), 3.51 (dd, J = 6.6 and 1.8 Hz, 2H), 6.25 (dt, J = 15.0 and 1.8 Hz, 1H), 7.00 (d, J = 8.0 Hz, 2H) 7.11 (d, J = 8.0 Hz, 2H), 7.13 (dt, J = 15.0 and 6.6 Hz, 1H), 7.52 (tt, J = 7.6 and 1.4 Hz, 2H), 7.61, (tt, 7.6 and 1.4 Hz, 1H), 7.86 (dd, J = 7.6 and 1.4 Hz, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ ppm: 21.0, 37.2, 127.6, 128.7, 129.3, 129.6, 131.3, 133.1,133.3, 136.7, 140.6, 145.8. LRMS [ESI-MS +ve] calculated for (C$_{16}$H$_{16}$O$_3$S + Na): 295.1 Observed: 295.0

(E)-1-methoxy-4-(3-(phenylsulfonyl)allyl)benzene (8). Yield: 73%. $^1$H NMR (400 MHz, CDCl$_3$) δ ppm: 3.50 (dd, J = 6.6 and 1.3 Hz, 2H), 3.79 (s, 3H), 6.23 (dt, J = 14.9 and 1.3 Hz, 1H), 6.84 (d, J = 8.6 Hz, 2H) 7.03 (d, J = 8.6 Hz, 2H), 7.13 (dt, J = 14.9 and 6.6 Hz, 1H), 7.53 (tt, J = 7.6 and 1.3 Hz, 2H), 7.61, (tt, 7.6 and 1.3 Hz, 1H), 7.86 (dd, J = 7.6 and 1.3 Hz, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ ppm: 36.8, 55.3, 114.3, 127.6, 128.1, 128.3, 129.3, 131.2, 133.3, 140.6, 145.9, 158.7. LRMS [ESI-MS +ve] calculated for (C$_{16}$H$_{16}$O$_3$S+Na): 311.1 Observed: 311.0

(cinnamylsulfonyl)benzene (9). Yield: 94%. [see Table 1 for $^1$H and $^{13}$C NMR assignments]. LRMS [ESI-MS +ve] calculated for (C$_{15}$H$_{14}$O$_2$S+Na): 281.1, observed: 281.2

(E)-4-(3-(phenylsulfonyl)prop-1-en-1-yl)-1,1'-biphenyl (10). Yield: 83%. $^1$H NMR (500 MHz, (CD$_3$)$_2$CO) δ ppm: 4.12 (d, J = 7.6 Hz, 2H), 6.22 (m, 1H), 6.53 (d, J =15.9 Hz, 1H), 7.32-7.51
(m, 5H), 7.59-7.77 (m, 7H), 7.92 (d, J = 7.4 Hz, 2H). \^13C NMR (125 MHz, (CD3)2CO) δ ppm: 59.66, 116.24, 126.69, 127.09, 127.10, 127.51, 128.37, 128.90, 129.14, 133.67, 135.32, 137.96, 139.30, 140.24, 140.77. LRMS [ESI-MS +ve] calculated for (C\textsubscript{19}H\textsubscript{16}O\textsubscript{2}S+Na): 357.1, observed: 357.2

(E)-1-(3-(phenylsulfonyl)prop-1-en-1-yl)naphthalene (11). Yield: 81%. \(^1H\) NMR (400 MHz, (CDCl\textsubscript{3}) δ ppm: 4.09 (dd, J = 7.65, 1.0 Hz, 2H), 6.13 (m, 1H), 7.08 (d, J = 15.6 Hz, 1H), 7.38-7.96 (m, 12H). \(^13C\) NMR (100 MHz, (CDCl\textsubscript{3}) δ ppm: 60.69, 118.51, 123.50, 124.30, 125.57, 125.99, 126.26, 128.56, 128.68, 128.86, 129.21, 130.84, 133.46, 133.53, 133.80, 136.98, 138.28. LRMS [ESI-MS +ve] calculated for (C\textsubscript{19}H\textsubscript{16}O\textsubscript{2}S+Na): 331.1, observed: 331.2

(E)-1-methyl-2-(3-(phenylsulfonyl)prop-1-en-1-yl)benzene (12). Yield: 69%. \(^1H\) NMR (500 MHz, (CDCl\textsubscript{3}) δ ppm: 2.13 (s, 3H), 3.98 (d, J = 7.6 Hz, 2H), 5.98(dt, J =15.5, 7.6 Hz, 1H), 6.56 (d, J = 15.5 Hz, 1H), 7.10 (dd, J = 7.4, 2.0 Hz, 1H), 7.16 (td, J = 7.4, 2.0 Hz, 1H), 7.19 (td, J = 7.4, 2.0 Hz, 1H), 7.34 (dd, J = 7.4, 2.0 Hz, 1H), 7.54 (tt, J = 7.5, 1.3 Hz, 2H), 7.64 (tt, J = 7.5, 1.3 Hz, 1H), 7.90 (dd, J = 7.5, 1.3 Hz, 2H). \(^13C\) NMR (100 MHz, (CDCl\textsubscript{3}) δ ppm: 19.5, 60.7, 116.6, 125.9, 126.2, 128.4, 128.6, 129.1, 130.3, 133.8, 134.9, 135.6, 137.3, 138.4. LRMS [ESI-MS +ve] calculated for (C\textsubscript{16}H\textsubscript{16}O\textsubscript{2}S+Na): 295.1, observed: 295.0

(E)-1-chloro-2-(3-(phenylsulfonyl)prop-1-en-1-yl)benzene (13). Yield: 84%. \(^1H\) NMR (400 MHz, (CDCl\textsubscript{3}) δ ppm: 4.00 (d, J = 7.6 Hz, 2H), 6.12 (tt, J =15.7, 7.6 Hz, 1H), 6.71 (d, J = 15.7 Hz, 1H), 7.21 (m, 2H), 7.31 (m, 1H), 7.46 (m, 1H), 7.55 (tt, J = 7.4, 1.4 Hz, 2H), 7.65 (tt, J = 7.4, 1.4 Hz, 1H), 7.89 (dd, J = 7.4, 1.4 Hz, 2H). \(^13C\) NMR (100 MHz, CDCl3) δ ppm: 60.5, 118.2, 127.0, 127.1, 128.5, 129.2, 129.5, 129.7, 133.2, 133.8, 133.9, 135.4, 138.2. LRMS [ESI-MS +ve] calculated for (C\textsubscript{15}H\textsubscript{13}ClO\textsubscript{2}S): 315.0 Observed: 315.0
(E)-1-methyl-4-(3-(phenylsulfonyl)prop-1-en-1-yl)benzene (14). Yield: 82%. \(^1\)H NMR (400 MHz, CDCl\(_3\)) \(\delta\) ppm: 2.34 (s, 3H), 3.94 (d, \(J = 7.6\) Hz, 2H), 6.05 (dt, \(J = 15.6\) and 7.6 Hz, 1H), 6.32 (d, \(J = 15.6\) Hz and 1H) 7.11 (d, \(J = 8.0\) Hz, 2H), 7.18 (d, \(J = 8.0\) Hz, 2H), 7.53 (tt, \(J = 7.6\) and 1.4 Hz, 2H), 7.63 (tt, 7.6 and 1.4 Hz, 1H), 7.88 (dd, \(J = 7.6\) and 1.4 Hz, 2H). \(^{13}\)C NMR (100 MHz, CDCl\(_3\)) \(\delta\) ppm: 21.3, 60.6, 114.0, 126.6, 128.6, 129.1, 129.4, 133.0, 133.7, 138.4, 138.6, 139.2. Calcuarded Mass for (C\(_{16}\)H\(_{16}\)O\(_2\)S+Na): 295.1 Observed: 295.0

(E)-1-methoxy-4-(3-(phenylsulfonyl)prop-1-en-1-yl)benzene (15). Yield: 85%. \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) ppm: 3.81 (s, 3H), 3.93 (d, \(J = 7.7\) Hz, 2H), 5.96 (tt, \(J = 15.6, 7.7\) Hz, 1H), 6.31 (d, \(J = 15.6\) Hz, 1H), 6.84 (dd, \(J = 8.7\) Hz, 2H), 7.23 (d, \(J = 8.7\) Hz, 2H), 7.54 (t, \(J = 7.7\) Hz, 2H), 7.64 (t, \(J = 7.7\) Hz, 1H), 7.88 (d, \(J = 7.7\) Hz, 2H). \(^{13}\)C NMR (100 MHz, CDCl\(_3\)) \(\delta\) ppm: 55.3, 60.6, 112.6, 114.1, 127.9, 128.5, 128.6, 129.1, 133.7, 138.5, 138.7, 159.9. LRMS [ESI-MS +ve] calculated for (C\(_{16}\)H\(_{16}\)O\(_3\)S+Na): 311.1 Observed: 311.0

(E)-(pent-1-en-1-ylsulfonyl)benzene (16E). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) ppm: 0.92 (m, 3H), 1.50 (m, 2H), 2.23 (m, 2H), 6.33 (d, \(J = 15.0\) Hz, 1H), 7.00 (m, 1H) 7.54 (m, 3H), 7.87 (m, 2H). \(^{13}\)C NMR (125 MHz, CDCl\(_3\)) \(\delta\) ppm: 13.6, 20.9, 33.4, 127.5, 129.3, 130.4, 133.3, 140.7, 147.1. LRMS [ESI-MS +ve] calculated for (C\(_{11}\)H\(_{14}\)O\(_2\)S+Na): 233.1, observed: 233.3.

(Z)-(pent-1-en-1-ylsulfonyl)benzene (16Z). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) ppm: 0.92 (m, 3H), 1.50 (m, 2H), 2.64 (m, 2H), 6.33 (m, 2H), 7.54 (m, 3H), 7.92 (m, 2H). \(^{13}\)C NMR (125 MHz, CDCl\(_3\)) \(\delta\) ppm: 13.7, 21.9, 29.6, 127.2, 129.3, 130.5, 133.4, 141.8, 147.5. LRMS [ESI-MS +ve] calculated for (C\(_{11}\)H\(_{14}\)O\(_2\)S+Na): 233.1, observed: 233.2.

(E)-(pent-2-en-1-ylsulfonyl)benzene (19E). \(^1\)H NMR (500 MHz, CD\(_3\)OD) \(\delta\) ppm: 0.88 (t, \(J = 7.52\)Hz, 3H), 1.98 (m, 2H), 3.87 (d, \(J = 7.31\) Hz, 2H), 5.37(m, 1H), 5.56 (m, 1H), 7.71 (m, 2H),
7.71 (m, 1H), 7.86 (m, 2H). $^{13}$C NMR (125 MHz, MeOD-D4) $\delta$ ppm: 12.0, 25.2, 59.1, 115.2, 128.3, 128.8, 133.5, 138.3, 142.8. LRMS [ESI-MS +ve] calculated for (C$_{11}$H$_{14}$O$_2$S+Na): 233.1, observed: 233.2.

(Z)-(pent-2-en-1-ylsulfonyl)benzene ($^{19}$Z). $^1$H NMR (500 MHz, CD$_3$OD) $\delta$ ppm: 0.71 (t, J = 7.55 Hz, 3H), 1.77 (m, 2H), 3.98 (d, J = 8.01, 2H), 5.37 (m, 1H), 5.70 (m, 1H), 7.62 (m, 2H), 7.71 (m, 1H), 7.89 (m, 2H). $^{13}$C NMR (125 MHz, MeOD-D$_4$) $\delta$ ppm: 12.3, 20.1, 54.2, 114.5, 128.2, 128.9, 133.6, 138.6, 140.7. LRMS [ESI-MS +ve] calculated for (C$_{11}$H$_{14}$O$_2$S+Na): 233.1, observed: 233.2.

(3-(phenylsulfonyl)prop-1-ene-1,1-diyl)dibenzene ($^{24}$) $^1$H NMR (500 MHz, CDCl$_3$) $\delta$ ppm: 3.93 (d, J = 7.9 Hz, 2H), 6.15 (t, J = 7.9 Hz, 1H), 6.67 (d, J = 7.6 Hz, 2H), 7.16 (m, 2H), 7.23 (t, J = 7.6 Hz, 2H), 7.27 (m, 4H), 7.51 (t, J = 7.6 Hz, 2H), 7.64 (t, J = 7.6 Hz, 1H), 7.80 (d, J = 7.6 Hz, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) $\delta$ ppm: 57.5, 114.0, 127.4, 127.8, 128.2, 128.3, 128.4, 128.5, 129.1, 129.2, 130.1, 133.7, 137.7, 138.5, 140.7, 149.7. LRMS [ESI-MS +ve] calculated for (C$_{21}$H$_{18}$O$_2$S+Na): 357.1, observed: 357.1.
Chapter 3

Development and Evaluation of Peptidyl Vinylsulfonate as:

I. A Cell-Active and Specific Inhibitor of Cathepsin L (KD-1)

II. An Activity-Based Probe of Cathepsin L
Chapter 3

Development and Evaluation of Peptidyl Vinylsulfonate as:

I. A Cell-Active and Specific Inhibitor of Cathepsin L (KD-1)
3.1 Introduction

Cathepsin L is a prominent member of cysteine cathepsin family of enzymes whose aberrant function has been linked to metastatic cancer. Cathepsin L promotes metastatic behavior by efficiently degrading and hydrolyzing various components of the extracellular matrix (ECM), such as laminin, fibronectin, and collagen. Numerous studies strongly indicate that cathepsin L activates signaling pathways that drive invasion in metastatic cancer. Its inhibition in invasive cancer is considered an attractive strategy for anti-cancer drug development.

Despite the linkage of cathepsin L to metastatic cancer, its functions in appropriate cellular contexts also remain largely unknown. Although it is generally believed that in normal cells cathepsin L is endosomal or lysosomal, nuclear and cytosolic roles have also recently emerged in other cells. In addition, numerous contradictory roles of cathepsin L have been reported. For example, in glioblastoma nuclear cathepsin L activity is involved in regulating the activity of transcription factors in a manner that promotes aberrant proliferation. In contrast, a pro-apoptotic role was attributed to cytosolic cathepsin L in human glioma cells. In the context of bone tissues, cathepsin L seems to be involved in both bone growth and bone loss processes. Such examples suggest that the underlying role of cathepsin L depends on its activation level in distinct cell-types and its specific microenvironment. Determination of the cellular level of activated cathepsin L would require a cell-active, potent and selective small molecule probe. In recent years, activity-based probes have emerged as powerful tools for such tasks. Acquisition of a suitable activity-based cathepsin L probe however requires that first a small molecule inhibitor be developed that (a) is cell-permeable, (b) inhibits cathepsin L activity potently and selectively, (c) utilizes a covalent irreversible chemistry for inhibition, and (d) is
non-basic in nature. The property of being non-basic is especially important for homogenous
distribution of probe in cell or else an undesired accumulation (and thus non-selective inhibition
of other cysteine cathepsins) will occur in the acidic environment of the lysosome\textsuperscript{138}. While
several potent and reversible modulators of cathepsin L activity exist, only a few irreversible
inactivating agents have been reported so far; none of which are suitable for development of a
cell-active activity-based cathepsin L probe. In this study, using a hybrid-design approach, the
development of a highly potent, selective, non-basic, and active-site directed irreversible
inhibitor (KD-1, Figure 3.1) of cathepsin L activity is reported.

3.2 Rationale

The active site architecture of cathepsin L includes substrate binding cavities made of
both non-prime and prime site pockets, arranged on two sides of the scissile amide bond\textsuperscript{27}. The
molecular framework of a cathepsin L-specific inhibitor was conceptualized so that the
functional groups occupied both the regular non-prime site (i.e. S1, S2, S3) and the prime site
(S') binding pockets. Earlier studies involving x-ray crystallography indicated that the S’ pocket
of cathepsin L could provide a strategic site for designing a selective and potent inhibitor\textsuperscript{139}. A
visual analysis of a covalent and reversible inhibitor, Z-Phe-Tyr-(O-\textit{tert} Bu)-C(O)CHO, bound to
cathepsin L indicated that a non-polar cluster from the S’ pocket, involving Trp189, Leu144,
Ala138, and Gly139 residues, formed critically important binding interactions with the \textit{tert}-butyl
group.\textsuperscript{139b} In another related study, it was noted that an aromatic moiety suitably placed could
perhaps form more effective inhibitory interactions with the S’ residues of cathepsin L.\textsuperscript{140} Thus,
we sought to first screen for a ligand that would contain an electrophilic warhead and a
functionalized aromatic group suitable for interacting with the prime site residues. Although
vinylsulfones have been utilized as an electrophilic warhead, the vinylsulfonate ester moiety has not yet been explored for the development of a covalent and irreversible inhibitor of cathepsin L. This was perhaps surprising since it was carefully noted that vinylsulfonate ester moiety displays a far superior reactivity as a Michael acceptor than vinyl sulfones (and sulfonamides). In addition, we surmised that an additional S-O bond present in vinylsulfonate (but not in vinylsulfone) could place the functionalized aryl group into the S’ pocket for optimal interaction. To test this hypothesis for cathepsin L inhibition, the inhibitory potency of phenyl vinylsulfonate and phenyl vinylsulfone towards cathepsin L was determined. As anticipated, phenyl vinylsulfonate exhibited 3-fold better potency than phenyl vinylsulfone under identical conditions.

**Figure 3.1.** A hybrid-design approach in the development of a highly potent time-dependent inhibitor (KD-1) of human cathepsin L. To generate KD-1, a competitive inhibitor CBZ-Phe-Leu-CONH₂ of modest potency (Kᵢ=14 µM), known to target S1, S2, and S3 pockets of cathepsin L, is conjugated to a prime site (S’)-targeting inhibitory ligand I at its C-terminal.
experimental conditions (data not shown). This finding strengthened the possibility that an effective S'-targeting ligand could be constructed. Thus a focused library of aryl vinylsulfonate ester compounds was synthesized and screened against cathepsin L (see experimental section). This screen identified 4-bromophenyl vinylsulfonate ester (Figure 3.1, Compound 1) as the most effective lead aryl vinylsulfonate ligand.

Torkar et al.\textsuperscript{58} recently reported Cbz-Phe-Leu-CONH\textsubscript{2} (Figure 3.1A) as a modest inhibitor (K\textsubscript{i} = 14 µM) of cathepsin L. This compound is non-basic and inhibits cathepsin L activity by presumably occupying the non-prime pockets (i.e. S1, S2, and S3). We hypothesized that a hybrid design that incorporated the key features of 1 and Cbz-Phe-Leu-CONH\textsubscript{2} could render a highly potent and selective inhibitor of cathepsin L. Thus KD-1 was conceived (Figure 3.1) and synthesized.

3.3 Chemistry

KD-1 was synthesized according to scheme 3.1. Oxidation of N-Boc-L-Leucinol lead to corresponding aldehyde which was utilized for the next step without any further purification\textsuperscript{143}. Wittig intermediate (II) was synthesized from ethyl methanesulfonate by treating with butyllithium followed by addition of diethyl chlorophosphate in anhydrous THF under inert condition\textsuperscript{144}. Treatment of intermediate II with sodium hydride followed by addition of N-Boc-L-Leucinal (II) in anhydrous THF under inert condition yielded Boc-protected leucine vinylsulfonates (III). Deprotection of Boc group was carried out by treating intermediate III with 25% TFA in DCM. The free amine thus obtained was coupled to commercially available Cbz-protected phenylalanine to generate intermediate IV. Treatment of IV with tetrabutyl ammonium iodide in acetone led to sulfonate salt which was utilized for the next step without any further
purification. The sulfonate salt, thus obtained, was converted to sulfonyl chloride by treating with sulfonyl chloride in presence of triphenylphosphine. The intermediate sulfonyl chloride was treated with 4-bromo phenol in presence of triethyl amine to obtain final compound KD-1.

3.4 Results and Discussion

To get an estimate of potency, an experiment was performed to obtain the IC$_{50}$ value. As anticipated, a highly potent inhibition was observed (IC$_{50} = 3.6 \pm 0.1$ nM). Since the IC$_{50}$ value is not an accurate indicator of potency of a time-dependent inhibitor, a dose-dependent experiment was performed at various inhibitor concentrations and the progress curves were recorded under pseudo-first order conditions (i.e. [KD-1] >> [Cat. L]) (Figure 3.2). Analysis of the progress curves revealed near diffusion-controlled inactivation kinetics (second order inactivation rate constant ($k_{inact}$) = 4.3 x 10$^6$ M$^{-1}$s$^{-1}$). The mode of inhibition was irreversible in nature because
enzyme activity could not be recovered, even after a very large dilution of the inactivated reaction mixture (data not shown).

**Figure 3.2.** Time-dependent inhibition of human cathepsin L by KD-1. Progress curves for cathepsin L-catalyzed hydrolysis of fluorescent substrate, Z-Phe-Arg-(7-amino-4-methylcoumarin, in the absence and presence of appropriate concentration of KD-1.

The selectivity of KD-1 towards cathepsin L enzyme was assessed next. To demonstrate this, a panel of recombinant proteases and phosphatases were evaluated for their inhibition using steady-state inhibition assays. This panel of enzymes included cysteine cathepsins, an aspartyl cathepsin, a serine cathepsin, a cysteine containing enzyme PTP, and a serine protease. The second order inactivation rate constants were thus determined by performing inactivation kinetics under a pseudo-first order condition, as described before (Table 3.1). In the panel of enzymes tested, KD-1 displayed a remarkable selectivity towards cathepsin L.
Table 3.1. Second order enzyme inactivation rate constants ($k_{\text{inact}}$) of cathepsin L inhibitor KD-1 against a panel of closely-related enzymes.

<table>
<thead>
<tr>
<th>Enzyme Family</th>
<th>Enzyme</th>
<th>$k_{\text{inact}}$ ($M^{-1}s^{-1}$)</th>
<th>RSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cysteine Cathepsins</td>
<td>Cathepsin L</td>
<td>$4.3 \times 10^6$</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Cathepsin K</td>
<td>$4.4 \times 10^4$</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Cathepsin B</td>
<td>96</td>
<td>44,000</td>
</tr>
<tr>
<td></td>
<td>Cathepsin S</td>
<td>$3.5 \times 10^5$</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Cathepsin H</td>
<td>NI</td>
<td>NA**</td>
</tr>
<tr>
<td>Aspartyl Cathepsin</td>
<td>Cathepsin D</td>
<td>NI</td>
<td>NA**</td>
</tr>
<tr>
<td>Serine Cathepsin</td>
<td>Cathepsin G</td>
<td>NI*</td>
<td>NA**</td>
</tr>
<tr>
<td>Protein Tyrosine Phosphatase</td>
<td>hptp1b</td>
<td>NI*</td>
<td>NA**</td>
</tr>
<tr>
<td>Serine Protease</td>
<td>Trypsin</td>
<td>NI*</td>
<td>NA**</td>
</tr>
</tbody>
</table>

Relative Selectivity Factor (RSF) = $k_{\text{inact}}$ (Cathepsin L) / $k_{\text{inact}}$ (Other enzyme); NI* = No Inhibition at 0.1 mM of KD-1 in 1 h; NA** = Not Applicable.

While a 13- and 100-fold selectivity was observed against most closely homologous cysteine cathepsins S and K respectively, a very high selectivity (44,000 fold) was observed against
cysteine cathepsin B. Furthermore, no reactivity was observed against other panel of enzymes (e.g. aspartyl cathepsin D, serine cathepsin G, and a low pKa cysteine containing enzyme hPTP1B, and serine protease trypsin). A remarkable selectivity profile observed for KD-1 in these experiments is noteworthy and should bode well for biological applications.

The ability of KD-1 to decrease the activity of cathepsin L towards collagen type I, a physiological substrate of cathepsin L, was investigated next\textsuperscript{145}. As expected, KD-1 protected the degradation of human collagen type I by cathepsin L \textit{in vitro} (Figure 3.2A). To further demonstrate that KD-1 was cell permeable and therefore suitable for use in biological contexts, an experiment involving metastatic breast cancer MDA-MB-231 cells was performed. This cell line is known to overexpress cathepsin L\textsuperscript{146}, thereby promoting invasion. A cell-permeable fluorescent substrate of cathepsin L, Z-Phe-Arg-(7-amino-4-methylcoumarin, was utilized to assess the intracellular inhibitory efficacy of KD-1. A near complete loss of intracellular substrate turnover (as indicated by loss of green fluorescence) was evident in KD-1-treated cells, compared to control (Figure 3.2B). This experiment demonstrated that KD-1 was indeed cell permeable and was able to successfully block cathepsin L activity in intact cells. To the best of our knowledge this is the first example where an inhibitor developed against cathepsin L is (i) highly potent and selective (ii) non-basic (iii) hydrolytically stable, and most importantly (iv) cell active. Such desirable traits are anticipated to lend themselves well in delineating precise roles of cathepsin L biology in specific tissue-types, and the development of a selective activity-based cathepsin L probe.

Finally, the functional effect of KD-1 on migratory behavior of highly metastatic MDA-MB-231 cells was investigated using a wound-healing assay. The metastatic cells are highly motile because E–cadherin, an important adhesive protein of cell-cell junctions, is a proteolytic
Figure 3.3: [A] Inhibitor KD-1 protects a physiological substrate of cathepsin L, human collagen type I, from proteolytic degradation in vitro. The α, β, and γ bands of collagen I are shown by blue arrows (left) [B] Effect of KD-1 on cathepsin L activity in a metastatic breast cancer cell line. MDA-MB-231 cells were incubated with (a) vehicle (DMSO 0.05%), and (b) KD-1 (750 nM) overnight (22 hrs).

Intracellular inhibition of cathepsin L activity was assessed by exposing the cell to a cell-permeable fluorogenic substrate, Z-Phe-Arg-(7-amino-4-methylcoumarin) (2 minutes). The signal intensity (green) that emerges from the enzymatic turnover of the fluorophore substrate (white arrow) is lost when inhibitor KD-1 is present. This experiment demonstrates that KD-1 is cell permeable, and is inhibiting intracellular cathepsin L activity in these cells. The results are representative of 3 independent experiments. Scale bar = 25 μm.

target of overabundant cathepsin L. Most cancer cells lose E-cadherin function as they become migratory, and hence metastatic. When MDA-MB-231 cells were treated with KD-1 (compared with DMSO-treated control cells) and the wound-healing assay was initiated, it was observed that cell motility decreased in KD-1-treated cells, compared to control (Figure 3.4). This result is
consistent with cellular inhibition of cathepsin L by KD-1 since an enhanced integrity of cell-cell junctions is expected in these cells due to maintenance of appropriate levels of E-cadherin. Further experiments are currently being planned to systematically investigate the effect of KD-1 on other invasive cancer cell types.

3.5 Conclusions

Using a hybrid-design approach and employing a vinylsulfonate ester functionality, a highly potent and selective cathepsin L inhibitory agent KD-1 has been developed. KD-1 successfully blocks intracellular cathepsin L activity in intact cells, and possesses anti-migratory potential. It is anticipated that KD-1 will find extensive applications in probing the undocumented roles of cathepsin L in diverse cellular microenvironments and human diseases.147
Figure 3.4. [A] KD-1 reduces migratory potential of highly metastatic MDA-MB-231 Breast cancer cells, as assessed by the wound healing assays. MDA-MB-231 cells were grown overnight in a 6-well plastic dish before treating with DMSO (control, 0.05%) or KD-1 (750 nM). After creating the wounds by drawing a pipet across a confluent field of cells, images were acquired at 0 h and 22 h. Representative images of three independent experiments are shown here. The dotted yellow lines indicate the edge of the wound at time zero, and the height of the yellow arrows are indicative of wound width. [B] The wound closure, a direct measurement of cell migratory potential, was calculated using the following equation:

\[
\text{% Wound Closure} = \frac{(A_0 - A_{22})}{A_0} \times 100\%
\]

where \(A_0\) is the area of the wound measured immediately after the scratch, and \(A_{22}\) is the area of the wound after 22 h.

The figure represents a statistical analysis from 3 independent experiments using the unpaired Student’s t test (*, \(p<0.0001\)). The data were evaluated as the mean ±1 standard deviation. The image analysis was performed by ImageJ software.
3.6 Experimental

3.6.1 Synthesis

**General:**   $^1$H NMR spectra were recorded at either 400 or 500 MHz using CDCl$_3$ or DMSO as the solvent. $^{13}$C NMR spectra were recorded at either 125 or 100 MHz using CDCl$_3$ or DMSO as the solvent. Chemical shifts (δ) are reported in parts per million (ppm) and referenced to CDCl$_3$ (7.26 ppm for $^1$H and 77.0 ppm for $^{13}$C), DMSO (2.50 ppm for $^1$H and 39.5 ppm for $^{13}$C). Coupling constants (J) are reported in Hertz (Hz) and multiplicities are abbreviated as singlet (s), doublet (d), doublet of doublets (dd), triplet (t), triplet of doublets (td), and multiplet (m). The mass spectra were acquired using a 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC/MS (Agilent Technologies, Santa Clara, CA, USA). 4-ethynylbenzyl alcohol was purchased from Santa Cruz Biotechnology, Inc. (Dallas, Texas, USA). O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) was purchased from peptide international (Louisville, Kentucky, USA). All anhydrous solvents were purchased from Sigma-Aldrich (St. Louis, MO, USA). All other materials were purchased from Fisher Scientific Inc. (Pittsburgh, PA, USA).

**Synthesis of library of aryl vinylsulfonate ester compounds (1-27):**

Synthesis of all aryl vinylsulfonates were carried out using a literature reported protocol.

4-Bromophenyl ethenesulfonate (1): Pale yellow semi solid. $R_f = 0.46$. Hexane: EtOAc (8:2). Yield: 61%. $^1$H NMR (400 MHz CDCl$_3$) δ ppm: 6.19 (d, $J = 10.0$Hz, 1H), 6.36 (d, $J = 15.1$Hz, 1H), 6.67 (dd, $J = 15.1$Hz, 10.0Hz, 1H), 7.12 (d, $J = 10.2$Hz, 2H), 7.50 (d, $J = 10.2$Hz, 2H). $^{13}$C
NMR (100 MHz CDCl₃) δ ppm: 120.9, 123.7, 131.7, 132.3, 133.0, 148.3. HRMS [ESI-MS +ve] calculated for (C₈H₇BrO₃S+H)⁺: 262.9378, observed: 262.9375.

Quinoline-8-yl ethenesulfonate (2): Colorless solid Rᵣ = 0.21. Hexane: EtOAc (8:2). Yield: 65%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.05 (d, J= 10.2Hz, 1H), 6.35 (d, J= 15.1Hz, 1H), 7.11 (dd, J= 15.1Hz and 10.2Hz, 1H), 7.50 (dd, J= 5.2Hz and 5.3Hz, 1H), 7.56 (m, 1H), 7.72 (d, J= 5.3Hz, 1H), 7.81 (d, J= 5.3Hz, 1H), 8.25 (d, J= 5.3Hz, 1H), 8.98 (s, 1H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 122.1, 123.3, 126.2, 127.2, 129.7, 130.2, 133.4, 136.1, 141.3, 145.3, 150.9. HRMS [ESI-MS +ve] calculated for (C₁₁H₉NO₃S+H)⁺: 236.0376, observed: 236.0375.

5-Nitroquinoline-8-yl ethenesulfonate (3): Colorless solid. Rᵣ = 0.24. Hexane: EtOAc (8:2). Yield: 68%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.15 (d, J= 8.2Hz, 1H), 6.45 (d, J= 16.1Hz, 1H), 7.11, (dd, J= 16.1Hz and 8.2Hz, 1H), 7.74 (dd, J= 4.1Hz and 8.2Hz, 1H), 7.83 (d, J= 8.5Hz, 1H), 8.44 (d, J= 8.5Hz, 1H), 9.07 (d, J= 8.3Hz, 1H), 9.11 (d, J= 4.1Hz, 1H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 121.5, 122.9, 124.8, 124.9, 131.3, 132.4, 133.2, 141.2, 143.7, 149.7, 151.9. HRMS [ESI-MS +ve] calculated for (C₁₁H₈N₂O₅S+Na)⁺: 281.0227, observed: 281.0227.

[1, 1'-Biphenyl]-4-yl ethenesulfonate (4): white powder. Rᵣ = 0.28. Hexane: EtOAc (9:1). Yield: 76%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.19 (d, J= 10.1Hz, 1H), 6.41 (d, J= 16.6Hz, 1H), 6.70 (dd, J= 10.1Hz and 16.6Hz, 1H), 7.30 (m, 2H), 7.33-7.47 (m, 3H), 7.57 (m, 4H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 122.6, 127.1, 127.8, 128.5, 128.9, 131.8, 132.1, 139.8, 140.6, 148.7. HRMS [ESI-MS +ve] calculated for (C₁₄H₁₂O₃S+Na)⁺: 283.0399, observed: 283.0399.

4'-Nitro-[1, 1'-biphenyl]-4-yl ethenesulfonate (5): yellow powder. Rᵣ = 0.25. Hexane: EtOAc (8:2). Yield: 72%. ¹H NMR (400MHz, DMSO-D₆) δ ppm: 6.35 (d, J= 16.8Hz, 1H), 6.43 (d, J= 10.1 Hz, 1H), 7.30 (dd, J= 10.1Hz and 16.8Hz, 1H), 7.46 (d, J= 8.7Hz, 2H), 7.90 (d, J= 8.5Hz,
2H), 7.98 (d, J= 8.7Hz, 2H), 8.32 (d, J= 8.7Hz, 2H). $^{13}$C NMR (100MHz, DMSO-D$_6$) δ ppm: 123.1, 124.1, 128.0, 129.1, 132.2, 133.6, 136.9, 145.1, 146.9, 149.5. HRMS [ESI-MS +ve] calculated for (C$_{14}$H$_{11}$NO$_5$S+H): 328.0250, observed: 328.0247.

Methyl 2-((vinylsulfonyl) oxy) benzoate (6): Viscous liquid $R_f = 0.51$. Hexane: EtOAc (8:2). Yield: 67%. $^1$H NMR (500MHz, CDCl$_3$) δ ppm: 3.93 (s, 3H), 6.16 (d, J= 10.1Hz, 1H), 6.37 (d, J= 14.9Hz, 1H), 6.82 (dd, J= 14.91Hz and 10.1Hz, 1H), 7.39 (m, 2H), 7.57 (t, J= 10.4Hz, 1H), 7.97 (d, J= 10.4Hz, 1H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 52.5, 124.1, 124.8, 127.2, 131.3, 132.1, 132.7, 133.7, 147.8, 164.9. HRMS [ESI-MS +ve] calculated for (C$_{10}$H$_{10}$O$_5$S+Na)$^+$: 265.0141, observed: 265.0141.

Methyl 4-((vinylsulfonyl) oxy) benzoate (7): white powder. $R_f = 0.59$. Hexane: EtOAc (8:2). Yield: 62%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 3.92 (s, 3H), 6.23 (d, J= 10.2Hz, 1H), 6.39 (d, J= 14.9Hz, 1H), 6.73 (dd, J= 14.9Hz and 10.2Hz, 1H), 7.31 (d, J= 10.1 Hz, 2H), 8.08 (d, J= 10.1Hz, 2H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 52.4, 122.2, 129.2, 131.5, 131.9, 132.3, 152.6, 165.9. HRMS [ESI-MS +ve] calculated for (C$_{10}$H$_{10}$O$_5$S+H)$^+$: 243.0322, observed: 243.0322.

4-Cyanophenyl ethenesulfonate (8): Yellow powder. $R_f = 0.20$. Hexane: EtOAc (9:1). Yield: 73%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 6.25 (d, J= 9.9Hz, 1H), 6.43 (d, J= 16.6Hz, 1H), 6.70 (dd, J= 16.6Hz and 9.9Hz, 1H), 7.37 (d, J= 8.7Hz, 2H), 7.72 (d, J= 8.6Hz, 2H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 111.4, 117.6, 123.3, 131.8, 132.7, 134.1, 152.2. HRMS [ESI-MS +ve] calculated for (C$_9$H$_7$NO$_3$S+H): 210.0219, observed: 210.0221.

4-Nitrophenyl ethenesulfonate (9): white powder. $R_f = 0.49$. Hexane: EtOAc (8:2). Yield: 69%. $^1$H NMR (500MHz, CDCl$_3$) δ ppm: 6.30 (d, J= 9.7Hz, 1H), 6.46 (d, J= 14.8Hz, 1H), 6.74 (dd, J= 14.8Hz and 9.7Hz, 1H), 7.44 (d, J= 9.8Hz, 2H), 8.30 (d, J= 9.8Hz, 2H). $^{13}$C NMR (125MHz,
CDCl₃) δ ppm: 123.1, 125.6, 131.5, 133.2, 146.3, 153.5. HRMS [ESI-MS +ve] calculated for (C₈H₇NO₅S+H)⁺: 230.0118, observed: 230.0118.

5-Fluoro-2-nitrophenyl ethenesulfonate (10): yellow viscous liquid. Rf = 0.29. Hexane: EtOAc (9:1). Yield: 56%. ¹H NMR (400Hz, CDCl₃) δ ppm: 6.27 (d, J= 9.6 Hz, 1H), 6.52 (d, J= 15.8 Hz, 1H), 6.76 (dd, J= 15.8 Hz and 9.6 Hz, 1H), 7.23 (m, 1H), 7.31 (m, 1H), 8.17 (m, 1H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 113.5 (d, J_C-F = 26.1 Hz), 115.0 (d, J_C-F = 23.2 Hz), 128.1, 132.1, 133.2, 139.0, 143.8, 165.3 (d, J_C-F = 258.2 Hz). HRMS [ESI-MS +ve] calculated for (C₈H₆FNO₅S+Na)⁺: 269.9843, observed: 269.9841.

4-Chlorophenyl ethenesulfonate (11): pale yellow viscous liquid. Rf = 0.27. Hexane: EtOAc (9:1). Yield: 68%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.21 (d, J= 9.9Hz, 1H), 6.39 (d, J= 16.7Hz, 1H), 6.67 (dd, J= 16.7Hz and 9.9Hz, 1H), 7.19 (d, J= 9.0, 2H), 7.36 (d, J= 9.0, 2H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 123.7, 129.9, 131.7, 132.3, 133.1, 147.7. HRMS [ESI-MS +ve] calculated for (C₈H₇ClO₃S+H)⁺: 218.9877, observed: 218.9877.

2-Chlorophenyl ethenesulfonate (12): pale yellow viscous liquid. Rf = 0.25. Hexane: EtOAc (9.5:0.5). Yield: 68%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.18 (d, J= 10.4Hz, 1H), 6.37 (d, J= 16.6Hz, 1H), 6.78 (dd, J= 16.6Hz and 10.4Hz, 1H), 7.25 (t, J= 7.7Hz, 1H), 7.30 (t, J= 7.7Hz, 1H), 7.42 (d, J= 8.1Hz, 1H), 7.44 (d, J= 8.1Hz, 1H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 124.5, 127.1, 128.1, 128.2, 130.9, 131.8, 132.4, 145.3. HRMS [ESI-MS +ve] calculated for (C₈H₇ClO₃S+Na)⁺: 240.9697, observed: 240.9699.

3-Chlorophenyl ethenesulfonate (13): colorless viscous liquid. Rf = 0.25. Hexane: EtOAc (9.5:0.5). Yield: 70%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.21 (d, J= 9.9Hz, 1H), 6.40 (d, J= 16.7Hz, 1H), 6.67 (dd, J= 16.7Hz and 9.9Hz, 1H), 7.16 (m, 1H), 7.26 (m, 1H), 7.31 (m, 2H). ¹³C
NMR (100MHz, CDCl₃) δ ppm: 120.6, 122.8, 127.7, 130.6, 131.8, 132.3, 135.1, 149.6. HRMS [ESI-MS +ve] calculated for (C₈H₇ClO₃S+H)⁺: 218.9877, observed: 218.9877.

4-Chloro-2-fluorophenyl ethenesulfonate (14): colorless viscous liquid. Rᵥ = 0.48 Hexane: EtOAc (8:2). Yield: 63%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.22 (d, J= 10.0Hz, 1H), 6.39 (d, J= 16.6Hz and 10.0Hz, 1H), 7.21 (m, 2H), 7.34 (m, 1H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 125.2 (d, J_C-F= 3.6 Hz, 1C), 125.8, 131.7, 132.4, 133.5 (d, J_C-F= 8.8 Hz, 1C), 135.4 (d, J_C-F= 12.6 Hz, 1C), 154.1 (d, J_C-F= 250.1 Hz, 1C).HRMS [ESI-MS +ve] calculated for (C₈H₆ClFO₃S+Na)⁺: 258.9602, observed: 258.9604.

4-Chloro-3-fluorophenyl ethenesulfonate (15): colorless viscous liquid. Rᵥ = Hexane: EtOAc (9:1). Yield: 64%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.25 (d, J= 9.9 Hz, 1H), 6.41 (d, J=16.6 Hz, 1H), 6.68 (dd, J= 16.6 Hz and 9.9 Hz, 1H), 7.01 (m, 1H), 7.12 (m, 1H), 7.43 (m, 1H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 117.9 (d, J_C-F= 21.1 Hz, 1C), 125.2 (d, J_C-F= 3.6 Hz, 1C), 125.8, 131.1, 131.6, 132.8, 148.2 (d, J_C-F= 9.3 Hz), 151.0 (d, J_C-F= 251.8 Hz) HRMS [ESI-MS +ve] calculated for (C₈H₆ClFO₃S+H)⁺: 236.9783, observed: 236.9783.

4-Chloro-3-(trifluoromethyl) phenyl ethenesulfonate (16): colorless viscous liquid. Rᵥ = 0.32. Hexane: EtOAc (9:1). Yield: 57%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.27 (d, J= 9.8Hz, 1H), 6.42 (d, J= 16.6Hz, 1H), 6.71 (d, J= 16.6Hz and 9.8Hz, 1H), 7.40 (m, 1H), 7.55 (m, 2H). ¹³C NMR (100MHz, CDCl₃) δ ppm: 121.7 (q, J_C-F= 121.9 Hz), 122.2 (q, J_C-F= 270.1 Hz), 126.9, 131.1, 132.5, 133.0, 133.1, 147.3 HRMS [ESI-MS +ve] calculated for (C₈H₆ClF₃O₃S+Na)⁺: 308.9570, observed: 308.9572.

4-Chloro-2, 3, 5, 6-tetrafluorophenyl ethenesulfonate (17): colorless liquid. Rᵥ = 0.24 Hexane: EtOAc (9.8:0.2). Yield: 66%. ¹H NMR (400MHz, CDCl₃) δ ppm: 6.35 (d, J= 9.9Hz, 1H), 6.56
(\text{d}, J= 16.5\text{Hz}, 1\text{H}), 6.81 (dd, J= 16.5\text{Hz} \text{ and} 9.9\text{Hz}, 1\text{H}), \text{^{13}C NMR (100MHz, CDCl}_3\text{) \text{ \delta ppm: 131.8, 133.1, 140.3, 143.1, 143.2, 145.6}\text{ HRMS [ESI-MS +ve] calculated for (C}_8\text{H}_3\text{ClF}_4\text{O}_3\text{S+Na}^{+}): 312.9320, \text{observed: 312.9323.}\text{ }

4-Chloro-3-methylphenyl ethenesulfonate (18): colorless viscous liquid. \text{R}_f = 0.32. \text{Hexane: EtOAc (9:1). Yield: 80\%. }\text{^1H NMR (400MHz, CDCl}_3\text{) \text{ \delta ppm: 2.35 (s, 3H), 6.19 (d, J= 10.3Hz, 1H), 6.35 (d, J= 16.7Hz, 1H), 6.68 (dd, J= 10.3Hz \text{ and} 16.7 \text{Hz}, 1\text{H}), 7.00 (dd, J= 2.9Hz \text{ and} 8.8 \text{Hz}, 1\text{H}), 7.10 (d, J= 2.8Hz, 1H), 7.30 (d, J= 8.6Hz, 1H). }\text{^13C NMR (100MHz, CDCl}_3\text{) \text{ \delta ppm: 20.2, 120.9, 124.6, 130.1, 131.8, 132.3, 133.1, 138.1, 147.6. HRMS [ESI-MS +ve] calculated for (C}_9\text{H}_9\text{ClO}_3\text{S+Na}^{+}: 254.9853, \text{observed: 254.9855.}\text{ }

4-Chloro-3, 5-dimethylphenyl ethenesulfonate (19): white powder. \text{R}_f = 0.43. \text{Hexane: EtOAc (8:2). Yield: 59\%. }\text{^1H NMR (500MHz, CDCl}_3\text{) \text{ \delta ppm: 2.27 (s, 6H), 6.10 (d, J= 5.5Hz, 1H), 6.28 (d, J= 20.0Hz, 1H), 6.59 (dd, J= 20.0Hz \text{ and} 5.5Hz, 1\text{H}), 6.87 (s, 2H). }\text{^13C NMR (100MHz, CDCl}_3\text{) \text{ \delta ppm: 20.8, 121.8, 131.9, 132.1, 133.3, 138.6, 146.9 HRMS [ESI-MS +ve] calculated for (C}_10\text{H}_11\text{ClO}_3\text{S+Na}^{+}: 269.0010, \text{observed: 269.0010.}\text{ }

4-Chloro-2-nitrophenyl ethenesulfonate (20): pale yellow viscous liquid. \text{R}_f = 0.24 \text{Hexane: EtOAc (9:1). Yield: 72\%. }\text{^1H NMR (400MHz, CDCl}_3\text{) \text{ \delta ppm: 6.29 (d, J= 9.9Hz, 1H), 6.49 (d, J= 16.7Hz, 1H), 6.81 (dd, J= 16.7Hz \text{ and} 9.9Hz, 1\text{H}), 7.52 (d, J= 8.9, 1\text{H}), 7.65 (dd, J= 8.9Hz \text{ and} 2.5Hz, 1\text{H}), 8.01 (d, J= 2.5Hz, 1\text{H}). }\text{^13C NMR (100MHz, CDCl}_3\text{) \text{ \delta ppm: 126.1, 126.7, 131.9, 133.1, 133.4, 134.5, 139.8, 142.5. HRMS [ESI-MS +ve] calculated for (C}_8\text{H}_6\text{ClNO}_5\text{S+Na}^{+}: 285.9547, \text{observed: 285.9547.}\text{ }

3, 4-Dichlorophenyl ethenesulfonate (21): colorless viscous liquid. \text{R}_f = 0.33. \text{Hexane: EtOAc (9:1). Yield: 67\%. }\text{^1H NMR (400MHz, CDCl}_3\text{) \text{ \delta ppm: 6.25 (d, J= 9.9Hz, 1H), 6.41 (d, J=}
16.6Hz, 1H), 6.68 (dd, J= 16.6Hz and 9.9Hz, 1H), 7.12 (dd, J= 8.8Hz and 2.7Hz, 1H), 7.39 (d, J= 2.7Hz, 1H), 7.47 (d, J= 8.8Hz, 1H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 121.9, 124.5, 131.2, 131.5, 131.7, 132.8, 133.4, 147.6. HRMS [ESI-MS +ve] calculated for (C$_8$H$_6$Cl$_2$O$_3$S+H)$^+$: 252.9487, observed: 252.9489.

2, 4, 5-Trichlorophenyl ethenesulfonate (22): white powder. R$_f$ = 0.48 Hexane: EtOAc (8:2). Yield: 59%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 6.26 (d, J= 8.5Hz, 1H), 6.45 (d, J= 16.1Hz, 1H), 6.77 (dd, J= 16.1Hz and 8.5Hz, 1H), 7.56 (s, 1H), 7.57 (s, 1H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 125.9, 126.3, 131.4, 131.9, 132.1, 132.7, 143.8. HRMS [ESI-MS +ve] calculated for (C$_8$H$_5$Cl$_3$O$_3$S+H)$^+$: 286.9098, observed: 286.9101.

2, 4, 6-Trichlorophenyl ethenesulfonate (23): white powder. R$_f$ = 0.44 Hexane: EtOAc (9:1). Yield: 63%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 6.26 (d, J= 10.3Hz, 1H), 6.56 (d, J= 16.5Hz, 1H), 6.94 (d, J= 16.5 Hz and 10.3Hz, 1H), 7.4 (s, 2H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 126.7, 128.3, 128.8, 130.6, 131.4, 139.5. HRMS [ESI-MS +ve] calculated for (C$_8$H$_5$Cl$_3$O$_3$S+Na)$^+$: 308.8917, observed: 308.8919.

Methyl 5-chloro-2-((vinylsulfonyl) oxy) benzoate (24): pale yellow viscous liquid. R$_f$ = 0.23. Hexane: EtOAc (9.5:0.5). Yield: 60%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 3.95 (s, 3H), 6.18 (d, J= 10.0Hz, 1H), 6.37 (d, J= 16.7Hz, 1H), 6.81 (dd, J= 16.7Hz and 10.0Hz, 1H), 7.34 (d, J= 8.7Hz, 1H), 7.50 (d, J= 8.7Hz, 1H), 7.93 (s, 1H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 52.7, 125.5, 126.1, 131.8, 132.4, 132.9, 133.5, 133.88, 146.2, 163.6. HRMS [ESI-MS +ve] calculated for (C$_{10}$H$_9$ClO$_5$S+Na)$^+$: 298.9751, observed: 298.9751.

4-Fluorophenyl ethenesulfonate (25): colorless viscous liquid. R$_f$ = 0.25. Hexane: EtOAc (9:1). Yield: 69%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 6.10 (d, J= 10.0Hz, 1H), 6.37 (d, J= 16.7Hz,
1H), 6.66 (dd, J= 10.0Hz and 16.7Hz, 1H), 7.10 (m, 2H), 7.22 (m, 2H). $^{13}$C NMR (100 MHz, CDCl$_3$) δ ppm: 116.6 (d, $J_{C:F}$ = 29.0Hz.), 124.0 (d, $J_{C:F}$ = 9.2 Hz.), 131.8, 132.1, 145.1 (d, $J_{C:F}$ = 3.5Hz.), 161.7 (d, $J_{C:F}$ = 247.3Hz, 1C). HRMS [ESI-MS +ve] calculated for (C$_8$H$_7$FO$_3$S+H)$^+$: 203.0173, observed: 203.0174.

4-Iodophenyl ethenesulfonate (26): colorless viscous liquid. $R_f$ = 0.29. Hexane: EtOAc (9:1). Yield: 64%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 6.19 (d, $J$ = 10.0Hz, 1H), 6.37 (d, $J$ = 16.8Hz, 1H), 6.65 (dd, $J$ = 16.8Hz and 10.0Hz, 1H), 6.99 (d, $J$ = 8.7Hz, 2H), 7.70 (d, $J$ = 8.7Hz, 2H). $^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 91.9, 124.4, 131.7, 132.3, 139.1, 149.2. HRMS [ESI-MS +ve] calculated for (C$_8$H$_7$IO$_3$S+Na)$^+$: 332.9053, observed: 332.9056.

2-Ethylphenyl ethenesulfonate (27): pale yellow viscous liquid. $R_f$ = 0.37. Hexane: EtOAc (9:1). Yield: 64%. $^1$H NMR (400MHz, CDCl$_3$) δ ppm: 1.22 (t, $J$ = 7.6Hz, 3H), 2.71 (q, $J$ = 7.6Hz, 2H), 6.14 (d, $J$ = 9.9Hz, 1H), 6.38 (d, $J$ = 16.6Hz, 1H), 6.72 (dd, $J$ = 16.6Hz and 9.9Hz, 1H), 7.23 (m, 4H).$^{13}$C NMR (100MHz, CDCl$_3$) δ ppm: 14.2, 23.1, 121.9, 127.1, 127.4, 130.1, 131.2, 132.7, 137.0, 147.6. HRMS [ESI-MS +ve] calculated for (C$_{10}$H$_{12}$O$_3$S+Na)$^+$: 235.0399, observed: 235.0401.

**The synthetic scheme utilized in the development of cathepsin L inhibitor KD-1**

Inhibitor KD-1 was synthesized by following scheme 3.1

(i). Synthesis of ethyl (diethoxyphosphoryl)methanesulfonate (I). This compound was synthesized by following the literature reported protocol$^{144}$ and its characterization was found to be consistent with the previously reported NMR spectra$^{149}$. 

---

92
(ii) Synthesis of tert-butyl (4-methyl-1-oxopentan-2-yl)carbamate (II). This compound was synthesized by following the literature reported protocol and its characterization was found to be consistent with the previously reported NMR spectra\(^{143}\).

(iii). Synthesis of ethyl (E)-3-((tert-butoxycarbonyl)amino)-5-methylhex-1-ene-1-sulfonate (III). This compound was prepared using a literature reported protocol with appropriate modification\(^{130}\). A solution of ethyl (diethoxyphosphoryl)methanesulfonate (483 mg, 1.86 mmol, 1 equiv) in anhydrous tetrahydrofuran (4 mL) was added dropwise to a stirring suspension of sodium hydride (58 mg, 2.41 mmol) in anhydrous tetrahydrofuran (3 mL) at 0°C under inert condition. After 30 min, a solution of N-BOC-L-leucinal in anhydrous tetrahydrofuran (5 mL) was added to the reaction mixture. Stirring was continued for an hour. The reaction was quenched by an addition of 5 mL of saturated ammonium chloride solution. The solution was concentrated under vacuum, diluted with water (10 mL), and extracted with ethylacetate (3×25 mL). The combined organic extracts were washed brine and dried over sodium sulfate, filtered, and concentrated under vacuum. Purification by flash column chromatography (silica gel; 90:10 hexane/ethylacetate) provided 419 mg (yield: 70%) of product. \(R_f = 0.6\) (80:20 hexane/ethylacetate). \(^1\)H NMR (500 MHz, CDCl\(_3\)) \(\delta\) ppm: 0.94 (m, 6H), 1.37 (t, \(J = 6.9\) Hz, 3H), 1.44 (s,11H), 1.70 (m1H), 4.16 (q, \(J = 6.9\) Hz, 2H), 4.37 (m,1H), 4.73 (m,1H), 6.31 (d, 15.2 Hz, 1H), 6.79 (dd, \(J = 15.2\) and 5.4 Hz, 1H). \(^{13}\)C NMR (125 MHz, CDCl\(_3\)) \(\delta\) ppm: 15.1, 22.0, 23.1, 24.8, 28.3, 43.4, 49.7, 67.3, 80.1, 124.4, 149.5, 155.3. HRMS [ESI-MS +ve] calculated for (C\(_{14}\)H\(_{27}\)NO\(_5\)S + Na): 344.1508, observed 344.1528.

(iv). Synthesis of Ethyl (E)-3-(2-(((benzylxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (IV)
To a stirred solution of the BOC-protected amine III (180 mg, 0.56 mmol) was added 25% solution of trifluoroacetic acid in dichloromethane (5 ml) and stirred for 1 hr. The solution was removed in vacuo. The residue was dissolved in DMF(2 ml). To this was added a cocktail of N-Cbz-L-phenylalanine (168 mg, 0.56 mmol), HCTU (232mg, 0.56 mmol) and N,N-diisopropylethylamine (375μl, 2.15 mmol) as a solution in anhydrous DMF (5 ml) at 0°C. Once the addition was done, ice bath was removed and the reaction was stirred overnight. The reaction was then diluted with 50 ml of ethyl acetate and washed with saturated ammonium chloride solution (25 ml), saturated sodium bicarbonate solution, and with brine (25 ml). The organic layer was dried over sodium sulfate and concentrated in vacuo. Purification by flash column chromatography (silica gel; 80:20 hexane/ethylacetate) provided 169 mg (yield: 60%) of product. Rf=0.21(80:20 hexane/ethylacetate). 1H NMR (400 MHz, CDCl3) δ ppm: 0.86 (m, 6H), 1.31 (m, 2H), 1.37 (t, J= 6.9 Hz, 3H), 1.51 (m,1H), 3.05 (dd, J= 14.4 and 7.7 Hz, 1H), 3.11 (dd, J= 14.4 and 6.4 Hz, 1H), 4.13 (m, 2H), 4.38 (m, 1H), 4.62 (m, 1H), 5.08 (m, 2H), 5.35 (bs, 1H), 5.94 (d, 15.4 Hz, 2H), 6.61 (dd, J= 15.4 and 5.2 Hz, 1H), 7.17 (m, 2H), 7.32 (m,8H). 13C NMR (100 MHz, CDCl3) δ ppm: 14.9, 21.9, 22.7, 24.6, 38.2, 42.7, 48.0, 56.5, 67.0, 67.3, 124.5, 127.6, 128.2, 128.4, 128.6, 129.2, 129.2, 135.8, 135.9, 147.7, 156.1, 170.6. HRMS [ESI-MS +ve] calculated for (C26H34N2O6S + H)\(^+\): 503.2217, observed 503.2218.

(v). Synthesis of 4-bromophenyl (E)-3-((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (KD-1)

This compound was prepared using a literature reported protocol with appropriate modification. To a stirred solution of IV (150 mg, 0.3 mmol) in acetone was added tetrabutyl ammonium iodide (45 mg, 0.3 mmol) and refluxed for overnight. The mixture was concentrated in vacuo and was dissolved in anhydrous dichloromethane (2 ml). This was added to a stirred
solution of triphenyl phosphine (157 mg, 0.6 mmol) and sulfuryl chloride (53 μl, 0.664 mmol) in anhydrous dichloromethane (4 ml) at 0°C. Once the addition was done, ice bath was removed and the reaction was stirred overnight. The mixture was then concentrated in vacuo and passed through silica (oven dried) using hexane and ethyl acetate (90:10) to recover sulfonyl chloride as white solid. Subsequently, the sulfonyl chloride (100mg, 0.23 mmol) was dissolved in anhydrous dichloromethane (3 ml). 4-bromophenol (60 mg, 0.345 mmol) was added to the solution followed by addition of triethylamine (100 μl, 0.69 mmol) at 0°C. The mixture was stirred for one hr and diluted with ethyl acetate (30 ml) followed by washing with saturated ammonium chloride solution (15 ml), saturated sodium bicarbonate solution (15ml ×3), and with brine (15 ml). The organic layer was dried over sodium sulfate and concentrated in vacuo. Purification by flash column chromatography (silica gel; 90:10 hexane/ethylacetate) provided 40 mg (overall yield: 22%) of the product as white solid. R f =0.37 (80:20 hexane/ethylacetate). 1H NMR (400 MHz, CDCl 3 ) δ ppm: 0.82 (m, 6H), 1.22 (m, 2H), 1.41 (m, 1H), 3.0 (dd, J= 13.7 and 7.9 Hz, 1H), 3.11 (dd, J= 13.7 and 6.4 Hz, 1H), 4.38 (m,1H), 4.52 (m, 1H), 5.07(m, 2H), 5.4 (bs,1H), 5.93(bs, 1H), 6.03(d, J=15.1 Hz, 1H), 6.46(dd, J= 15.1 and 5.2 Hz, 1H), 7.08 (m, 2H), 7.13 (m, 2H), 7.33 (m, 8H), 7.51 (m, 2H). 13C NMR (100 MHz, CDCl 3 ) δ ppm: 21.8, 22.5, 24.5, 38.3, 42.2, 48.2, 56.6, 67.4 120.8, 123.7, 124.5, 127.6, 128.2, 128.4, 128.6, 129.1, 129.1, 132.9, 135.8, 135.8, 148.3, 150.1, 156.1, 170.6. HRMS [ESI-MS +ve] calculated for (C 30 H 33 BrN 2 O 6 S +H) + : 629.1322, observed 629.1315.

3.6.2. Screening and Enzymology

General. Following enzymes and their substrates were purchased from Enzo Life Sciences (Farmingdale, NY, USA): cathepsin L (BML-SE201), cathepsin K (BML-SE553), cathepsin B
(BML-SE198), cathepsin D (BML-SE199), cathepsin S (BML-SE453), cathepsin H (BML-SE200), cathepsin G (BML-SE283), human PTP1B (BML-SE332), Z-Arg-Arg-pNA (BML-P138), Z-Gly-Pro-Arg-AMC (BML-P142), Z-Phe-Arg-AMC (BML-P139), H-Arg-AMC.2HCl (BML-P135), Suc-Ala-Ala-Pro-Phe-pNA (BML-P141), Mca-Gly-Lys-Pro-Ile-Leu-Phe-Phe-Arg-Leu-Lys(Dnp)-D-Arg-NH₂ (BML-P145), Z-Val-Val-Arg-AMC (BML-P199). N-alpha-Benzoyl-DL-arginine-p-nitroanilide (BAPNA) (B4875), trypsin (T1426), para-nitrophenylphosphate were purchased from Sigma–Aldrich (St. Louis, MO, USA). All the biological grade buffers were used and were purchased from Sigma–Aldrich (St. Louis, MO, USA). All enzyme kinetics experiments (unless otherwise stated) were carried out at 30°C in appropriate buffer condition with 5% DMSO concentration. For measuring the initial rates of enzyme catalyzed reaction, a temperature-controlled steady-state arc lamp fluorometer equipped with Felix 32 software (Photon Technology Instrument, Birmingham, NJ, USA) and a UV–Vis spectro-photometer (Modelk25; Perkin Elmer Inc., Waltham, MA, USA) was used.

**Screening of aryl vinylsulfonate library compounds:**

The inactivation reaction of human liver cathepsin L (net 50 nM) by 10 μM of inhibitor was performed under pseudo first order condition ([I]>>[E]) in a 0.5 ml eppendorf tube maintained at 30°C in a temperature controlled bath. A previously reported cathepsin L assay procedure was employed. Cathepsin L enzyme was first activated for 10 min in sodium acetate buffer (400 mM, pH 5.5) containing 8 mM DTT and 4 mM Na₂EDTA. The inactivation reaction was initiated by the addition of inhibitor where DMSO concentration was maintained at 5%. After suitable time intervals, an aliquot of 8μl of incubation mixture was withdrawn and added to an assay mixture (net assay volume 200μl) that contained 5 μM of Z-FR-AMC (Kₘ= 2.2 μM) substrate at 30°C. A progress curve was recorded (excitation: 365 nm; emission: 440
nm) and the enzyme activity was determined by measuring the initial rates of substrate turnover (Figure 3.5).

**Figure 3.5.** Chemical structures of the synthesized aryl vinylsulfonate library (1-27) targeted for cathepsin L inhibition. Time-dependent inhibition experiments were performed under pseudo-first order condition at 10 µM inhibitor concentration, as described in the main text. In parenthesis, % inhibition is calculated from two independent experiments.

**Determination of IC$_{50}$ of KD-1 against cathepsin L:**
Briefly IC$_{50}$, experiment was performed by subsequent addition of active cathepsin L (160 pM) and 5 μM of Z-FR-AMC in acetate buffer (pH 5.5) containing KD-1 (concentration ranging from 0 pM to 15 nM) at 30°C. Net DMSO concentration was maintained at 8%. The reaction was monitored fluorometrically (excitation: 365 nm; emission: 440 nm) for 100 seconds at 30°C. The experimental data thus obtained were plotted using SigmaPlot software and fitted using Enzyme Kinetics module to obtain IC$_{50}$ value. The IC$_{50}$ value thus obtained was 3.6±0.1 nM (Figure 3.6).

**Determination of kinetic inactivation parameters of KD-1 against cathepsin L:**

The inactivation reaction of human liver cathepsin L (net 160 pM) by appropriate concentrations of KD-1 was performed under pseudo first order condition ([I]>>[E]) in a fluorescence quartz cuvette maintained at 30°C. Briefly, the inactivation reaction was performed by subsequent addition of active cathepsin L (160 pM) and 5 μM of Z-FR-AMC in acetate buffer (pH 5.5) containing KD-1 (concentration ranging from 1 nM to 20 nM). Net DMSO concentration was maintained at 8%. The reaction was monitored fluorometrically (excitation:
365 nm; emission: 440 nm) for 3 minutes at 30°C. Subsequently, the progress curves were analyzed using non-linear regression according the following equation.\(^{58}\)

\[ P = v_z [1 - \exp(-k_{obs} \cdot t)] / k_{obs}; \]

where \(P\) is the product, \(v_z\) is the initial velocity (at time zero), and \(k_{obs}\) is the pseudo-first order rate constant.

Pseudo-first order rate constants \(k_{obs}\) were plotted against appropriate inhibitor concentrations \([I]\), and 2nd order inactivation rate constant \(k_{inact}\) was calculated from the slope of the plot according to following equation.\(^{58}\) (Figure 3.7);

\[ k_{obs} = k_{inact} [I] / [1 + ([S]/K_M)] \]

where \(k_{obs}\) is the pseudo-first order rate constant, \([I]\) is the inhibitor concentration, \([S]\) is the substrate concentration, with a given \(K_M\). The data were analyzed using KaleidaGraph (version 3.52).

**Figure 3.7.** Determination of second order inactivation rate constant \((k_{inact})\) for KD-1-mediated inhibition of human liver cathepsin L. The \(k_{obs}\) obtained from the progress curve fit were plotted against appropriate concentration of [KD-1]. The slope of the line yielded the enzyme inactivation rate constant \(k_{inact}\).
Selectivity profile of KD-1 against other enzymes

Determination of kinetic inactivation parameters of KD-1 against cathepsin S

A previously reported cathepsin S assay procedure was followed with suitable modification \(^{152}\). Thus, recombinant human cathepsin S enzyme was activated in the 100 mM potassium phosphate buffer (pH 6.5) containing 2.5 mM DTT, 2.5 mM Na\(_2\)EDTA for 10 min. The inactivation reaction of cathepsin S (net 1 nM) by appropriate concentrations of KD-1 was performed under pseudo first order condition ([I]>>[E]) in a fluorescence quartz cuvette maintained at 30°C. Briefly, the inactivation reaction was initiated by subsequent addition of active cathepsin L and 50 μM of Z-VVR-AMC (K\(_M\) = 17.7 μM) in phosphate buffer (pH 6.5) containing KD-1 (concentration ranging from 20 nM to 200 nM). Net DMSO concentration was maintained at 8%. The reaction was monitored fluorometrically (excitation: 365 nm; emission: 440 nm) for 3 minutes at 30°C. The progress curves were analyzed using the same equations used for cathepsin L assay.

Cathepsin K inhibition assay

For assaying cathepsin K activity, a previously published protocol was followed with appropriate modification \(^{153}\). Thus, recombinant human cathepsin K enzyme was activated in the 50 mM sodium acetate buffer (pH 5.5) containing 2.5 mM DTT, 2.5 mM Na\(_2\)EDTA for 10 min. The inactivation reaction of cathepsin K (net 50 nM) with inhibitor KD-1(250 nM) was initiated under pseudo first order condition ([I]>>[E]). After suitable time intervals, an aliquot of 10 μl of incubation mixture was withdrawn and the enzyme activity was monitored fluorometrically in assay buffer (200 μl; 7.5% DMSO) containing 200 μM of Z-GPR-AMC (K\(_M\) =68 μM; Ex/Em: 365/440 nm). The data thus obtained was plotted and fitted to the following equation:
$A_t = A_f - (A_f - A_0) \exp (-k_{obs} t t)$

where $A_t$ = enzyme activity at time t of cathepsin K inactivation, $A_f$ = final cathepsin K activity at infinite time of inactivation, $A_0$ = initial enzyme activity at zero time of inactivation, and $k_{obs}$ = pseudo-first order rate constant of inactivation. The data were analyzed using KaleidaGraph (version 3.52).

**Cathepsin H inhibition assay**

For assaying cathepsin H activity, a previously published protocol was followed with appropriate modification $^{151a}$. Thus, human liver cathepsin H enzyme was activated in the 200 mM KH$_2$PO$_4$ / 200 mM Na$_2$HPO$_4$ buffer (pH 6.8) containing 5 mM DTT, 2 mM Na$_2$EDTA for 10 min. The inactivation reaction of cathepsin H (net 330 nM) with inhibitor KD-1 (100 μM) was initiated under pseudo first order condition ($[I]>>[E]$). After suitable time intervals, an aliquot of 10 μl of incubation mixture was withdrawn and the enzyme activity was monitored fluorometrically in assay buffer (200 μl; 8% DMSO) containing 300 μM of H-R-AMC.2HCl ($K_M = 150$ μM; Ex/Em: 365/440 nm).

*Assays involving cathepsin B, cathepsin D, cathepsin G, trypsin and PTP1B were carried out according to the literature reported protocol $^{94}$.

**3.6.3. Biological and Cell-Based Studies**

**Inhibition of Cathepsin L’s Collagenase Activity by KD-1**

This experiment was carried out by following slight modification of a literature reported protocol $^{154}$. Human skin type I collagen (Calbiochem) was dissolved in 100 mM acetic acid for
24 h at room temperature to give a final concentration of 1.0 mg/ml. In 500 µl eppendorf tubes, different concentrations of KD-1 (0 nM, 300 nM, 500 nM and 1000 nM) were incubated with 200 nM of active cathepsin L (Assay Condition: 100 mM sodium acetate pH 5.5, 1 mM EDTA, 3 mM DTT, 2% DMSO, 30°C) for 30 minutes. Reaction was initiated by addition of 6 µg of collagen (Total reaction volume 26 µl). Reactions were quenched after 6 days by addition of 10 µl of Laemmli’s SDS- Sample Buffer (4X, reducing) and heated to 95°C for 10 minutes. 20 µl of reaction mixtures were (3.3 µg of type I collagen) loaded to 4-15% gradient Bio-Rad Mini-Protean® TGX™ Gels. Gels were stained in Coomassie Brilliant Blue, G250 and imaging was performed using scanner.

**Cell culture:** All the cell culture reagents were purchased from Invitrogen (Carlsbad, CA, USA). MDA-MB-231cells (American Type Culture Collection, Manassas, VA) were cultured on 10-cm plates (BD biosciences) at 37°C and 5% CO₂ in Iscove’s Modified Dulbecco’s medium with L-glutamine (IMDM), 10% fetal bovine serum, and antibiotics (1% penicillin/streptomycin and 0.5µg/ml fungizone) as described earlier. Cells were passaged at 1:3 ratio every 3–4 days and subcultured by 0.25% trypsin-EDTA treatment. Cells were re-plated one day prior to experiment so that the cell density was 80-90% on the day of the experiment.

**Confocal microscopy**

For live-cell imaging, cells were trypsinized at ~ 80% confluence, seeded to Poly-D-lysine coated 35mm-glass bottom dishes (MatTek; Ashland, MA) and grown overnight in complete medium, as described previously (Dana et al. 2013). Cells were treated with (0.05% Dimethyl-sulfoxide, DMSO) or Cathepsin-L inhibitor (750nM) inhibitors overnight, as
mentioned in the text. Prior to imaging, cells were treated with the substrate (Z-FR-AMC) for 2-3 mins.

Images of live cells were acquired using an inverted Leica TCS-SP5 confocal microscope (Leica; IL) by Plan apochromat 63x1.4 oil objective lens. An argon ion laser (25%) was used to generate excitation at 488nm, and pinholes were typically set to 1–1.5 Airy units.

Wound-healing cell migration assay: For wound healing experiments, MDA-MB-231 cells were grown to confluence in a 6-well plate. Prior creating wound with a sterile pipette tip, the cells were treated with the vehicle control (0.05% Dimethyl-sulfoxide, DMSO) or Cathepsin-L inhibitor (750nM). Floating cells were removed by washing once with PBS. Images were captured by 4x objective at different time intervals.

Images were captured right after the scratch (‘wound’), t=0h, to record the initial area of the wounds, followed by image acquisition at t=22hrs to measure the recovery of the wounded area by cell migration under different conditions. The area of the ‘wound’ was quantified by ImageJ software1.4 by polygon selection mode. The migration of cells towards the denuded area (‘wound’) was expressed as a percentage of wound closure as:

\[
\text{% of wound closure} = \left[ \frac{A_{t=0\, h} - A_{t=\Delta\, h}}{A_{t=0\, h}} \right] \times 100\%,
\]

where \(A_{t=0\, h}\) is the area of the wound measured immediately after the scratch, and \(A_{t=\Delta\, h}\) is the area of the wound after 22 hrs.

Statistical analysis
Each experiment was performed a minimum of three times, and differences between groups were determined using the unpaired Student’s $t$ test. The data were evaluated as the mean $\pm 1$ s.d.
Chapter 3

Development and Evaluation of Peptidyl Vinylsulfonate as:

II. An Activity-Based Probe of Cathepsin L (KDP-1)
3.7 Transformation of Potent Specific Cathepsin L Inhibitor to Activity-Based Probe

Functional roles of cysteine cathepsins in cancer cells often remain poorly understood. Although, many biochemical techniques including traditional genomic and proteomic have been utilized, often they provide the information only about the change in expression levels of the proteins. These parameters do not correlate with the cysteine cathepsins activity as they are expressed as inactive zymogens or exist in complex with their endogenous inhibitors. For example, a protease could be present in low abundance but may still be actively turning over its substrate. Because the activity of a protein is directly correlated to its function in vivo and given the fact that many proteins are post-translationally modified, an activity-based assessment of an enzyme is more directly relevant to its functional roles. Thus, the growing interest lies in correlating enzyme activity with its biological manifestation, such as oncogenic transformation or metastasis.

In this study, we seek to develop the first small molecule probe of cathepsin L that is highly sensitive, cell permeable, non-basic, and exquisitely selective. A photoaffinity-based cathepsin L probe has recently been reported. However, it suffers from severe limitations, which makes it totally incompatible for in vivo monitoring of cathepsin L activity in a natural cellular environment. These caveats mainly includes its (i) non-cell permeability that requires labelling experiments to be performed only after cell lysis, (ii) low affinity (Kᵢ = 3.6 µM) and poor selectivity against analogous cathepsins (~only 3-6 fold against cathepsin S, K, and B). Thus, it is critically important to develop a small molecule cathepsin L inhibitor first which will covalently and irreversibly inactivate cathepsin L by targeting active site catalytic cysteine.
residue. The inhibitory agent must also be selective, highly potent, stable, cell permeable, and non-basic in nature. We have developed a novel non-basic small molecule inhibitor of cathepsin L (KD-1) that is highly potent, selective and cell permeable. Careful investigation revealed that inclusion of an alkynyl moiety at 4-position of Cbz group will presumably retain the inhibitor activity towards cathepsin L while rendering it with a handle for a plausible copper-based click chemistry. The developed probe KDP-1 would find extensive usage in a variety of research applications, relevant to cathepsin-based anti-cancer drug development endeavors. These, for instances, may include (a) in situ tumor models for probing cathepsin L function, (b) validation of cathepsin L as a viable therapeutic target in individual cancer types, (c) assessment of in vivo pharmacodynamics property and efficacy of drug candidates, and (d) estimation of drug’s selectivity and off-target reactivity\textsuperscript{137}, (e) investigating the importance of targeting cathepsin L alone in cancer cells that overexpress both cathepsin L and B; this is especially important given the fact these two share a common physiological protein substrate of extracellular matrix, a protein known to promote metastasis\textsuperscript{43} and (f) profiling the selectivity of cathepsin L-directed anti-cancer drug candidates in vivo.\textsuperscript{137,159}

3.8 Chemistry

Synthesis of KDP-1 was carried out according to scheme 3.2. Synthesis of P2 was achieved by treating 4-ethynyl benzylalcohol with triphosgene with sodium carbonate under anhydrous condition following a literature reported protocol.\textsuperscript{160} This was used for the next step without further purification. Chloroformate (P2) was dropwise added to a stirred slurry of methyl ester of phenyl alanine and sodium carbonate in water/toluene mixture to obtain intermediate P3.\textsuperscript{161} Ester hydrolysis was carried out by treating P3 with lithium hydroxide in
equimolar mixture of water and THF\textsuperscript{162}. The free acid P4 thus obtained, was utilized in further steps to obtain KDP-1.

\begin{center}
\textbf{Scheme 3.2.} Synthetic route adopted for the development of KDP-1.
\end{center}

\textbf{3.9 Results and Experimental Propositions}

The inhibitory efficacy of KDP-1 was evaluated against cathepsin L by following enzyme kinetics method. A, dose-dependent experiment was performed at various inhibitor concentrations and the progress curves were recorded under pseudo-first order conditions (i.e. [KDP-1]>>[Cat. L]). Analysis of the progress curves\textsuperscript{58} established KDP-1 as one of the most potent activity based probe known for cathepsin L (figure 3.5) (second order inactivation rate constant ($k_{\text{inact}}$) = 4.3 x 10\textsuperscript{5} M\textsuperscript{-1}s\textsuperscript{-1}).
Figure 3.8. [A] Time-dependent inhibition of human cathepsin L by KDP-1. Progress curves for cathepsin L-catalyzed hydrolysis of fluorescent substrate, Z-Phe-Arg-(7-amino-4-methylcoumarin), in the absence and presence of appropriate concentration of KDP-1. [B] Determination of second order inactivation rate constant ($k_{\text{inact}}$) for KDP-1-mediated inhibition of human liver cathepsin L. The $k_{\text{obs}}$ obtained from the progress curve fit were plotted against appropriate concentrations of KDP-1. The slope of the line yielded the enzyme inactivation rate constant $k_{\text{inact}}$. 

The effectiveness of KDP-1 to report the intracellular cathepsin L activity will be demonstrated in variety of carcinoma and normal cell lines. Detection of covalently modified cathepsin L will be accomplished using click chemistry approach. This technique offers an unique tag-less strategy to assess enzyme activity at proteome level in live cells. Click chemistry protocol allows formation of a triazole ring between two functional units (i.e. between acetylene and azide). Initially recombinant human liver cathepsin L will be incubated with KDP-
Incubate cell permeable cathepsin L probe, KDP-1, with metastatic and non-metastatic cancer cells

1. KDP-1 will covalently and irreversibly modify cathepsin L via formation of a C-S bond with active site catalytic cysteine residue. The covalently modified enzyme will be subjected to aqueous friendly click chemistry protocol\textsuperscript{164} using rhodamine-azide. The rhodamine labelled protein will be serially diluted and the samples will be resolved using 10% SDS-PAGE. Labeled protein resolved on a denaturing gel will be visualized directly using Typhoon scanner (GE Healthcare). This will allow detecting minute quantity of active cathepsin L (as low as pmoles-

\textbf{Figure 3.9.} The procedure involved in profiling intracellular cathepsin L activity in metastatic and non-metastatic cancer cells.
fmoles). Once it is demonstrated in vitro; cell-based experiment will be carried out using MDA-MB-231 cell lines. The process involves incubation of KDP-1 with MDA-MB-231 cells. Once the in vivo labelling is completed, the cells will be lysed and subjected to click chemistry followed by quantification of active enzyme by resolving it on SDS-PAGE (figure 3.6).

### 3.10 Future Directions

The developed activity based probe KDP-1 could be further improved by changing the position of alkyne and amino acid substituents. Another latest strategy involving palladium-mediated Suzuki-Miyuara coupling could be utilized for tag-less labelling experiments. In this strategy, no further modification on inhibitory scaffold will be needed. KD-1-mediated covalently modified enzyme could be subjected to a palladium catalyzed Suzuki-Miyuara cross coupling reaction with a fluorophore containing boronic acid under physiological conditions. This will enable the inhibitory scaffold to label active enzyme without affecting its inhibitory potency.
3.11 Experimental

3.11.1 Synthesis

**General.** $^1$H NMR spectra were recorded at either 400 or 500 MHz using CDCl$_3$ or DMSO as the solvent. $^{13}$C NMR spectra were recorded at either 125 or 100 MHz using CDCl$_3$ or DMSO as the solvent. Chemical shifts (δ) are reported in parts per million (ppm) and referenced to CDCl$_3$ (7.26 ppm for $^1$H and 77.0 ppm for $^{13}$C), DMSO (2.50 ppm for $^1$H and 39.5 ppm for $^{13}$C). Coupling constants ($J$) are reported in Hertz (Hz) and multiplicities are abbreviated as singlet (s), doublet (d), doublet of doublets (dd), triplet (t), triplet of doublets (td), and multiplet (m). The mass spectra were acquired using a 6520 Accurate-Mass Quadrupole Time-of-Flight (Q-TOF) LC/MS (Agilent Technologies, Santa Clara, CA, USA). 4-ethynylbenzyl alcohol was purchased from Santa Cruz Biotechnology, Inc. (Dallas, Texas, USA). O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) was purchased from peptide international (Louisville, Kentucky, USA). All anhydrous solvents were purchased from Sigma-Aldrich (St. Louis, MO, USA). All other materials were purchased from Fisher Scientific Inc. (Pittsburgh, PA, USA).

**Synthesis of Activity based probe KDP-1**

**Synthesis of 4-ethynylbenzyl carbonochloridate (P2)**

Synthesis of P2 was carried out using a literature reported protocol$^{160}$. This was used for the next step without further purification.

Synthesis of Methyl (((4-ethynylbenzyl)oxy)carbonyl)phenylalaninate (P3):
Synthesis of this compound was carried out using a literature reported protocol\textsuperscript{161} and its characterization was found to be consistent with the previously reported NMR spectra.\textsuperscript{162}

Synthesis of (((4-ethylbenzyl)oxy)carbonyl)phenylalanine (P4)

Synthesis of this compound was carried out using a literature reported protocol and was used for the next step without further purification.\textsuperscript{162}

Synthesis of Ethyl (E)-3-2-(((4-ethynylbenzyl)oxy)carbonylamino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (P5)

To a stirred solution of the BOC-protected amine \textit{III} (248 mg, 0.773 mmol) was added 25% solution of trifluoroacetic acid in dichloromethane (5 ml) and stirred for 1 hr. The solution was removed in vacuo. The residue was dissolved in DMF(2 ml). To this was added a cocktail of P4 (250 mg, 0.773 mmol), HCTU (351mg, 0.85 mmol) and \textit{N,N}-diisopropylethylamine (337μl, 1.93 mmol) as a solution in anhydrous DMF (5 ml) at 0°C. Once the addition was done, ice bath was removed and the reaction was stirred overnight. The reaction was then diluted with 50 ml of ethyl acetate and washed with saturated ammonium chloride solution (25 ml), saturated sodium bicarbonate solution, and with brine (25 ml). The organic layer was dried over sodium sulfate and concentrated in vacuo. Purification by flash column chromatography (silica gel; 80:20 hexane/ethylacetate) provided 256 mg (yield: 63%) of product. \textit{R}_{f} =0.25(70:30 hexane/ethylacetate) \textsuperscript{1}H NMR (400 MHz CDCl\textsubscript{3}) δ ppm: 0.87 (m, 6H), 1.31 (m, 2H), 1.38 (t, \textit{J}=7.0 Hz, 3H), 1.50 (M, 1H), 3.02 (dd, \textit{J}= 13.5 Hz and 8.1 Hz, 1H), 3.09 (s, 1H), 4.14 (m, 2H), 4.36 (m, 1H), 4.62 (m, 1H), 5.08 (m, 2H), 5.33 (bs, 1H), 5.75 (bs, 1H), 5.94 (d, \textit{J}= 15.2 Hz, 1H), 6.60 (dd, \textit{J}=15.2 Hz and 5.1 Hz, 1H), 7.18 (d, \textit{J}= 7.0Hz, 2H), 7.3 (m, 5H), 7.48 (d, \textit{J}=8.2 Hz, 2H). \textsuperscript{13}C NMR (100 MHz CDCl\textsubscript{3}) δ ppm: 14.90, 21.82, 22.65, 24.60, 38.27, 42.74, 48.03, 56.64,
Synthesis of 4-bromophenyl (E)-3-(2-(((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-sulfonate (KDP-1)

This compound was prepared using a literature reported protocol with appropriate modification.\textsuperscript{150} To a stirred solution of P5 (300 mg, 0.57 mmol) in acetone was added tetrabutyl ammonium iodide (210.5 mg, 0.57 mmol) and refluxed for overnight. The mixture was concentrated in vacuo and was dissolved in anhydrous dichloromethane (4 ml). This was added to a stirred solution of triphenyl phosphine (299 mg, 1.14 mmol) and sulfuryl chloride (101.3 μl, 1.254 mmol) in anhydrous dichloromethane (6 ml) at 0°C. Once the addition was done, ice bath was removed and the reaction was stirred overnight. The mixture was then concentrated in vacuo and passed through silica (oven dried) using hexane and ethyl acetate (90:10) to recover sulfonyl chloride as white solid. Subsequently, the sulfonyl chloride (260 mg, 0.503 mmol) was dissolved in anhydrous dichloromethane (3 ml). 4-bromophenol (130 mg, 0.76 mmol) was added to the solution followed by addition of triethylamine (140 μl, 1.01 mmol) at 0°C. The mixture was stirred for one hr and diluted with ethyl acetate (30 ml) followed by washing with saturated ammonium chloride solution (15 ml), saturated sodium bicarbonate solution (15 ml ×3), and with brine (15 ml). The organic layer was dried over sodium sulfate and concentrated in vacuo. Purification by flash column chromatography (silica gel; 90:10 hexane/ethy lacetate) provided 125 mg (after three steps yield: 40%) of the product as white solid. $R_f = 0.52$ (70:30 hexane/ethylacetate). \textsuperscript{1}H NMR (400 MHz CDCl\textsubscript{3}) δ ppm: 0.82 (m, 6H), 1.22 (m, 2H), 1.41 (m, 1H), 2.99 (dd, $J$= 13.5 and 8.4 Hz, 1H), 3.10 (dd, $J$= 13.5 and 8.4 Hz, 1H), 3.10 (s, 1H), 4.37 (m,
1H), 4.53 (m, 1H), 5.06 (m, 2H), 5.44 (bd, \( J=7.1 \) Hz, 1H), 5.89 (bd, \( J=5.2 \) Hz, 1H), 6.04 (d, \( J=15.1 \) Hz, 1H), 6.47 (dd, \( J= 15.1 \) and 5.3 Hz, 1H), 7.08 (m, 2H), 7.13 (m, 2H), 7.28 (m, 5H), 7.49 (m, 4H) \(^{13}\)C NMR (100 MHz CDCl\(_3\)) \( \delta \) ppm: 21.85, 22.53, 24.57, 38.38, 42.34, 48.26, 56.65, 66.75, 77.83, 83.16, 120.84, 122.20, 123.86, 124.45, 127.64, 127.88, 129.09, 129.15, 132.36, 132.94, 135.86, 136.57, 148.36, 149.96, 155.93, 170.54). LRMS [ESI-MS +ve] calculated for (C\(_{32}\)H\(_{33}\)BrN\(_2\)O\(_6\)S +Na\(^+\)) : 676.57, observed 676.66.

### 3.11.2 Enzyme Kinetics

Inhibitory efficacy of KDP-1 was assessed against cathepsin L by following the same protocol that has been adopted for inhibitor KD-1.
Appendix A: NMR Spectra of Synthesized Arylsulfonyloxiranes.

2-(4-Fluorophenyl)-3-(phenylsulfonyl)oxirane (I):

$^1$H NMR:

$^{13}$C NMR:

$^{13}$C DEPT45:
2-(Phenylsulfonyl)-3-propyloxirane (2):

$^1$H NMR:

$^{13}$C NMR:
2-(4-Chlorophenylsulfonyl)oxirane (3).

$^1$H NMR:

$^{13}$C NMR:
2-((4-Trifluoromethoxy)phenylsulfonyl)oxirane (4).

$^1$H NMR:

$^{13}$C NMR:
2-(2-Ethylphenylsulfonyl)oxirane (5).

$^1$H NMR:

$^{13}$C NMR:
2-(m-Tolylsulfonyl)oxirane (6).

$^1$H NMR:

$^{13}$C NMR:
2-(2,5-Dimethylphenylsulfonyl)oxirane (7):

$^1$H NMR:

$^{13}$C NMR:
2-(3-Methoxyphenylsulfonyl)oxirane (8):

$^1$H NMR:

$^{13}$C NMR:
2-(2,5-Dimethoxyphenylsulfonyl)oxirane (9):

\[ \text{H}_2\text{CO} \begin{array}{c} \text{S} \\ \text{O} \end{array} \text{OCH}_3 \]

\(^1\text{H NMR:}\)

\[^{13}\text{C NMR:}\]
2-(3,5-dichlorophenylsulfonyl)oxirane (10):

$^1$H NMR:

$^{13}$C NMR:
(j) 2-(Naphthalen-3-ylsulfonyl)oxirane (11):

$^1$H NMR:

$^{13}$C NMR:
Appendix B: NMR Spectra of Synthesized Allyl/Vinyl Sulfones and Crystal Structure of (E)-((3-phenylprop-1-en-1-yl)sulfonyl)benzene (2)

Figure: $^1$H and $^{13}$C NMR spectrum of 2.
Figure: $^1$H-$^1$H COSY spectrum of 2. Observed crosspeaks establish the connectivity between protons 1, 2, and 3 as well as between the ortho, meta, and para protons of the phenyl rings.
Figure: $^{13}$C-$^1$H heteronuclear COSY spectrum of 2. Observed crosspeaks establish the assignment of protons and the carbons directly attached to it. C3 and H3 are not shown in the figure and their crosspeak is observed at (37.64, 3.56).
Figure: $^1$H-$^{13}$C HMBC spectrum of 2. Peaks marked A to G are assigned as follows. A: H$_c$-C$_a$ B: H$_c$'-C$_{a'}$ C: H$_2$-C$_{a'}$ D: H$_3$-C$_2$ E: H$_3$-C$_{a'}$ F: H$_3$-C$_1$ G: H$_3$-C$_{b'}$. The long range $^1$H and $^{13}$C connectivity illustrated through peaks A to G help establish the assignment of quaternary carbons a and a' as well as distinguish the two rings. Olefinic proton 2 and aliphatic proton 3 and show three bond coupling to C$_{a'}$ and C$_{b'}$ respectively which clearly establishes the position of double bond with respect to the phenyl ring.
Figure: $^1$H and $^{13}$C NMR spectrum of 3.
Figure: $^1$H and $^{13}$C NMR spectrum of 4.
Figure: $^1$H and $^{13}$C NMR spectrum of 5.
Figure: $^1$H and $^{13}$C NMR spectrum of 6.
Figure: $^1$H and $^{13}$C NMR spectrum of 7.
Figure: $^1$H and $^{13}$C NMR spectrum of 8.
Figure. $^1$H and $^{13}$C NMR spectrum of 9.
Figure: $^1$H-$^1$H COSY spectrum of 9. Observed crosspeaks establish the connectivity between protons 1, 2, and 3 as well as between the phenyl ring protons b, c, and d.
Figure: $^1$H-$^{13}$C heteronuclear HSQC spectrum of 9. Observed crosspeaks establish the assignment of protons and the carbons directly attached to it. H1 and C1 are not shown in the figure and their crosspeak is observed at (3.96, 60.51).
Figure: $^1$H-$^{13}$C HMBC spectrum of 9. Peaks marked A to E are assigned as follows. A: $\text{H}_c$-$\text{C}_a$ B: $\text{H}_3$-$\text{C}_b$ C: $\text{H}_{c''}$-$\text{C}_a'$. D: $\text{H}_3$-$\text{C}_b'$. E: $\text{H}_2$-$\text{C}_a'$. The long range $^1$H and $^{13}$C connectivity illustrated through peaks A to E help establish the assignment of quaternary carbons a and a’ as well as distinguish the two rings. Olefinic protons 2 and 3 show three bond coupling to $\text{C}_a'$ and $\text{C}_b'$ respectively which clearly establishes the position of double bond with respect to the phenyl ring.
Figure: $^1$H and $^{13}$C NMR spectrum of 10.
Figure: $^1$H and $^{13}$C NMR spectrum of 11.
Figure: $^1$H and $^{13}$C NMR spectrum of 12.
Figure: $^1$H and $^{13}$C NMR spectrum of 13.
Figure: $^1$H and $^{13}$C NMR spectrum of 14.
Figure: $^1$H and $^{13}$C NMR spectrum of 15.
Figure: $^1$H and $^{13}$C NMR spectrum of 16E and 16Z.
Figure: $^1$H-$^1$H COSY spectrum of the mixture of 16E and 16Z
Figure: $^{13}$C-$^1$H heteronuclear correlation spectrum of the mixture of 16E and 16Z. Aliphatic $^1$H-$^{13}$C crosspeaks for both compounds are illustrated in the figure.
Figure: $^1$H and $^{13}$C NMR spectrum of 19E and 19Z.
Figure: $^1$H-$^1$H COSY spectrum of the mixture of 19E and 19Z
Figure: $^{13}$C-$^1$H heteronuclear correlation spectrum of the mixture of 19E and 19Z. Aliphatic $^1$H-$^{13}$C crosspeaks for both compounds are illustrated in the figure.
Figure: $^1$H and $^{13}$C NMR spectrum of 24.
X-ray Structure Determination of Compound 2

A clear fragment of 2 with approximate dimensions 0.18 mm x 0.40 mm x 0.51 mm, cut from a cluster of crystals obtained by crystallization at room temperature from ethyl acetate/hexanes, was used for the X-ray crystallographic analysis. The X-ray intensity data were measured on a Bruker Smart Breeze CCD system equipped with a graphite monochromator at 100(2) K, cooled by an Oxford Cryosystems 700 Series Cryostream. A total of 1464 frames were collected. The total exposure time was 4.07 hours. The frames were integrated with the Bruker SAINT software package using a narrow-frame algorithm. The integration of the data using a monoclinic unit cell yielded a total of 7242 reflections to a maximum θ angle of 27.26° (0.78 Å resolution), of which 2900 were independent (average redundancy 2.497, completeness = 99.8%, Rint = 2.07%, Rsig = 2.62%) and 2850 (98.28%) were greater than 2σ(F2). The final cell constants of \(a = 5.6179(5) \text{ Å}, b = 7.6160(6) \text{ Å}, c = 15.3045(13) \text{ Å}, \beta = 95.8260(10)°\), volume = 651.43(10) Å\(^3\), are based upon the refinement of the XYZ-centroids of 4601 reflections above 20 σ(I) with 5.349° < 2θ < 54.435°. Data were corrected for absorption effects using the numerical method (SADABS). The ratio of minimum to maximum apparent transmission was 0.915. The calculated minimum and maximum transmission coefficients (based on crystal size) are 0.8883 and 0.9583.

The structure was solved and refined using the Bruker SHELXTL Software Package, using the space group Pn, with \(Z = 2\) for the formula unit, \(C_{15}H_{14}O_{2}S\). The final anisotropic full-matrix least-squares refinement on F2 with 163 variables converged at \(R1 = 2.80\%\), for the observed data and \(wR2 = 7.00\%\) for all data. The goodness-of-fit was 1.065. The largest peak in the final difference electron density synthesis was 0.231 e\(^{-}/\text{Å}^3\) and the largest hole was -0.179 e\(^{-}/\text{Å}^3\) with an RMS deviation of 0.040 e\(^{-}/\text{Å}^3\). On the basis of the final model, the calculated density was 1.317 g/cm\(^3\) and F(000), 272 e\(^{-}\).
Cambridge Crystallographic Data Centre deposition number for 2: CCDC 916842. The data can be obtained free from Cambridge Crystallographic Data Centre via http://www.ccdc.cam.ac.uk/data_request/cif.

Figure S28. ORTEP\textsuperscript{1} drawing of 2 with complete numbering of atoms, viewed approximately normal to the alkene plane. Selected bond lengths (Å) and angles (°): S1-O1, 1.443(1); S1-O2, 1.443(1); S1-C1, 1.753(2); S1-C4, 1.760(2); C1-C2, 1.320(2); C2-C3, 1.504(2); C3-C10, 1.525(2); O1-S1-O2, 118.37(8); O1-S1-C1, 109.14(7); O2-S1-C1, 107.07(8); O1-S1-C4, 108.38(7); O2-S1-C4, 108.07(7); C1-S1-C4, 105.01(7); C2-C1-S1, 120.9(1); C1-C2-C3, 124.6(2); C2-C3-C10, 110.3(1).

Appendix C: NMR Spectra of Synthesized Vinyl Sulfonates

4-Bromophenyl ethenesulfonate (1):

Figure: $^1$H and $^{13}$C NMR spectrum of 1
Quinolin-8-yl ethenesulfonate (2):

Figure. $^1$H and $^{13}$C NMR spectrum of 2.
5-Nitroquinolin-8-yl ethenesulfonate (3):

![Chemical Structure](attachment:image)

Figure. $^1$H and $^{13}$C NMR spectrum of 3.
[1,1’-Biphenyl]-4-yl ethenesulfonate (4):

Figure. $^1$H and $^{13}$C NMR spectrum of 4.
4’-Nitro-[1, 1’-biphenyl]-4-yl ethenesulfonate (5):

Figure. $^1$H and $^{13}$C NMR spectrum of 5.
Methyl 2-((vinylsulfonyl) oxy) benzoate (6):

Figure. $^1$H and $^{13}$C NMR spectrum of 6.
Methyl 4-((vinylsulfonyl) oxy) benzoate (7):

![Chemical Structure](image)

Figure. $^1$H and $^{13}$C NMR spectrum of 7.
4-Cyanophenyl ethenesulfonate (8):

Figure. $^1$H and $^{13}$C NMR spectrum of 8.
4-Nitrophenyl ethenesulfonate (9):

![Chemical Structure](image)

**Figure.** $^1$H and $^{13}$C NMR spectrum of 9.
5-Fluoro-2-nitrophenyl ethenesulfonate (10):

Figure. $^1$H and $^{13}$C NMR spectrum of 10
4-Chlorophenyl ethenesulfonate (11):
Figure. \(^1\)H and \(^{13}\)C NMR spectrum of 11

2-Chlorophenyl ethenesulfonate (12):

![Chemical Structure](image)

Figure. \(^1\)H and \(^{13}\)C NMR spectrum of 12.
3-Chlorophenyl ethenesulfonate (13):

Figure. $^1$H and $^{13}$C NMR spectrum of 13
4-Chloro-2-fluorophenyl ethenesulfonate (14):

Figure. $^1$H and $^{13}$C NMR spectrum of 14
4-Chloro-3-fluorophenyl ethenesulfonate (15):

Figure. $^1$H and $^{13}$C NMR spectrum of 15.
4-Chloro-3-(trifluoromethyl) phenyl ethanesulfonate (16):

![Chemical Structure](image)

Figure. $^1$H and $^{13}$C NMR spectrum for FN-16
4-Chloro-2, 3, 5, 6-tetrafluorophenyl ethenesulfonate (17):

![Chemical Structure](image1)

Figure. $^1$H and $^{13}$C NMR spectrum of 17
4-Chloro-3-methylphenyl ethenesulfonate (18):

Figure. $^1$H and $^{13}$C NMR spectrum of 18
4-Chloro-3, 5-dimethylphenyl ethenesulfonate (19):

Figure. $^1$H and $^{13}$C NMR spectrum of 19
4-Chloro-2-nitrophenyl sulfonate (20):

![Chemical Structure]

Figure. $^1$H and $^{13}$C NMR spectrum of 20.
3, 4-Dichlorophenyl ethenesulfonate (21):

![Chemical Structure]

Figure. $^1$H and $^{13}$C NMR spectrum of 21.
2, 4, 5- Trichlorophenyl ethenesulfonate (22):

Figure. $^1$H and $^{13}$C NMR spectrum of 22.
2, 4, 6-Trichlorophenyl ethenesulfonate (23):

\[
\begin{align*}
\text{Cl} & \quad \text{Cl} \\
\text{Cl} & \quad \text{Cl} \\
\end{align*}
\]

\[
\begin{align*}
\text{S} & \quad \text{O} \\
\text{O} & \quad \text{S} \\
\end{align*}
\]

Figure. \(^1\)H and \(^13\)C NMR spectrum of 23.
Methyl 5-chloro-2-((vinylsulfonyl) oxy) benzoate (24):

![Chemical Structure Image]

Figure. $^1$H and $^{13}$C NMR spectrum of 24
4-Fluorophenyl ethenesulfonate (25):

Figure. $^1$H and $^{13}$C NMR spectrum of 25.
4-Iodophenyl ethenesulfonate (26):

![Chemical Structure](image)

Figure. $^1$H and $^{13}$C NMR spectrum of 26.
2-Ethylphenyl ethenesulfonate (27):

Figure. $^1$H and $^{13}$C NMR spectrum of 27
Ethyl (E)-3-((tert-butoxycarbonyl)amino)-5-methylhex-1-ene-1-sulfonate (III)

Figure. $^1$H and $^{13}$C NMR spectrum of III.
Ethyl(E)-3-(2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (IV)

Figure. $^1$H and $^{13}$C NMR spectrum of IV.
4-bromophenyl(E)-3-((2-((benzyloxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (KD-1)

Figure. $^1$H and $^{13}$C NMR spectrum of KD-1.
Ethyl (E)-3-(2-(((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-1-sulfonate (P5)

Figure: $^1$H and $^{13}$C NMR spectrum of P5.
4-bromophenyl (E)-3-(2-(((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenylpropanamido)-5-methylhex-1-ene-sulfonate (KDP-1)

Figure: $^1$H and $^{13}$C NMR spectrum of KDP-1.
References:


198


Human cathepsin B is a metastable enzyme stabilized by specific ionic interactions associated with the active site. *Biochemistry* **1994**, *33* (49), 14800-6.


121. Cambridge Crystallographic Data Centre deposition No. 916842. The data can be obtained free from Cambridge Crystallographic Data Centre via [http://www.ccdc.cam.ac.uk/data_request/cif](http://www.ccdc.cam.ac.uk/data_request/cif).

alkenes to aldehydes catalyzed by iron(III) porphyrins with triflate as a counter anion. *Chem Commun (Camb)* 2011, 47 (39), 10963-5.


exhibit abnormal skin and bone development and show increased resistance to osteoporosis following ovariectomy. *International journal of experimental pathology* 2004, 85 (2), 85-96.


