ABSTRACT

Kinetic Characterization of Thiosemicarbazones as Cysteine Protease Inhibitors and Their Potential Use as Therapeutic Agents against Metastatic Cancer and Chagas' Disease

Gustavo E. Chavarria Nolasco, Ph.D.

Mentor: Mary Lynn Trawick, Ph.D.

Human cathepsins L and K are cysteine proteases that play important roles in physiological and pathological processes, such as cancer metastasis, bone resorption and neurodegenerative diseases. This research has focused on the evaluation of synthetic thiosemicarbazones that could inhibit their proteolytic activities with the objective of preventing cancer metastasis. In addition, cruzain, a cathepsin L-like cysteine protease found in *Trypanosoma cruzi,* is a validated target for pivotal roles in the parasitic invasion in Chagas' disease, a condition that could be fatal if not treated. Currently, there is no effective treatment against the disease, which is rapidly extending to non-endemic areas in the United States and Europe. More than 150 synthetic thiosemicarbazones (obtained through a collaborative study) were tested against cathepsins L, K and cruzain. This work presents preliminary *in vitro* analysis of these compounds in order to characterize their potency and mode of inhibition. A number of potent inhibitors was found for each enzyme. A smaller subset of thiosemicarbazones were found to be selective. Results showed that compound **1**, one of the most potent inhibitors in this library is a slow binding, slowly reversible, competitive inhibitor of these targets.

Furthermore, **1** was able to delay and partially inhibit the activation of procathepsin K under acidic conditions. Similar results were found with 3-bromo-3'-hydroxybenzophenone thiosemicarbazone (**8**). Compounds **8** and **1** inhibited the *in vitro* type I collagenase activity of cathepsin L in a time-dependent manner and type IV collagenase activity of cathepsin K.

Analogs, **1**, **8**, **156**, **157**, and **168** were also used in cell studies. These compounds were able to delay cell migration and cell invasion in MDA-MB-231 cells, a type of carcinoma breast cancer cell line. It was determined that cell invasion and cell migration were inhibited in a concentration dependent manner.

Lastly, analog **17** was also found to be a slow reversible, slow binding, competitive inhibitor of cruzain. This compound was also able to inhibit the collagenase activity of recombinant cruzain when human type I collagen was used as a substrate.

Kinetic studies and molecular modeling indicate the best thiosemicarbazone inhibitors form a reversible covalent bond with each enzyme.

Kinetic Characterization of Thiosemicarbazones as Cysteine Protease Inhibitors and Their Potential Use as Therapeutic Agents against Metastatic Cancer and Chagas' Disease

by

Gustavo E. Chavarria, B.S.

A Dissertation

Approved by the Department of Chemistry and Biochemistry

_____

Patrick J. Farmer, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Doctor of Philosophy

Approved by the Dissertation Committee

_____

Mary Lynn Trawick, Ph.D., Chairperson

_____

Sung-Kun Kim, Ph.D.

_____

Bryan F. Shaw, Ph.D.

_____

Kevin G. Pinney, Ph.D.

_____

Jaime L. Diaz-Granados, Ph.D.

Accepted by the Graduate School
December 2012

_____

J. Larry Lyon, Ph.D., Dean

*Page bearing signatures is kept on file in the Graduate School.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

| | |
|---|---|
| [E] | Enzyme concentration |
| $[E_T]$ | Total enzyme concentration |
| [I] | Inhibitor concentration |
| [S] | Substrate concentration |
| $\Delta G°$ | Gibbs Free Energy |
| °C | Degrees Celsius |
| 3-D | 3-Dimensional |
| A | Alanine |
| AA | Amino acid |
| ABL | V-ABL Abelson Murine Leukemia Viral Oncogene Homolog 1 |
| ABR | Assay Buffer for Reversibility Studies (Cathepsin L) |
| Abz | 3-Aminobenzoic Acid |
| ACI | Type I Collagen in Acetic Acid |
| ACIV | Type IV Collagen in Acetic Acid |
| AD | Anno Domini |
| ADT | Androgen Deprivation Therapy |
| AKR | Assay Buffer for Reversibility Studies (Cathepsin K) |
| Ala | Alanine |
| AMC | 7-Amino-4-Methylcoumarin |
| Arg | Arginine |

| | |
|---|---|
| Asn | Asparagine |
| Asp | Aspartate |
| ATCC | American Type Culture Collection |
| ATG | Start codon |
| AZ | Arizona |
| BC | Before Christ |
| BCL | Apoptosis regulator proteins |
| BCR-ACL | Tyrosine-Kinase Inhibitors |
| bFGF | Basic Fibroblast Growth Factor |
| Bid | Interacting-Domain Death Agonist |
| BRCA | Breast Cancer Type 1 Susceptibility Protein, |
| BT474 | Human Breast Carcinoma Cell Line |
| BZ | Benznidazole |
| C | Cysteine |
| C- | Carboxy Terminal |
| C=O | Carbonyl, Carbonyl Bond |
| C4-S | Chondroitin 4-Sulfate |
| C4-S | Chondroitin 4-Sulfate |
| cABL | Oncogene |
| CATH | Protein Structure Classification |
| Cbz | Carboxybenzyl |
| CD20 | B-Lymphocyte Antigen |
| CD33 | Siglec-3 |

| | |
|---|---|
| Cdc42 | Cell Division Control Protein 42 |
| CK | Cathepsin K |
| CKR | Cathepsin L Stock Solution for Reversibility Studies (Cathepsin K) |
| CL | Cathepsin L |
| CLC | Cathepsin L Stock Solution for Collagenase Studies |
| CLIP | Class II-Associated Invariant Chain Peptide |
| CLR | Cathepsin L Stock Solution for Reversibility Studies (Cathepsin L) |
| C-N | Carbon-Nitrogen Angle |
| COPD | Chronic Obstructive Pulmonary Disease |
| CRES | Cystatin-Related Epididymal Spermatogenic |
| CRT | Conformal Radiation Therapy |
| CTC | Stop Codon |
| CTX-I | C-Telopeptide Of Type I Collagen |
| Cys | Cysteine |
| CZ | Cruzain |
| D | Aspartate |
| DKK-1 | Dickkopf-Related Protein 1 |
| DMEM | Dulbeco's Modfied Eagle's Medium |
| DMSO | Dimethylsulfoxide |
| DNA | Deoxyribonucleic Acid |
| DTT | Dithiothreitol |
| E | Glutamate |
| E | Enzyme |

| | |
|---|---|
| E-64 | Irreversible Inhibitor Of Cysteine Proteases |
| EC | Enzyme Classification |
| ECM | Extracellular Matrix |
| EDTA | Ethylenediaminetetraacetic acid |
| EGF | Epidermal Growth Factor |
| EGFR | Epidermal Growth Factor Receptor |
| ELISA | Enzyme-Linked Immunosorbent Assay |
| Erk | Extracellular-Signal-Regulated Kinases |
| ES | Michaelis-Menten Complex |
| ET-1 | Endothelin |
| F | Phenylananine |
| FAK | Focal Adhesion Kinase-1 |
| FBS | Fetal Bovine Serum |
| FDA | United States Food And Drug Administration |
| G | Glycine |
| g | Gram |
| g | Gram |
| G1-S | Post-Mitotic Phase To Synthesis Phase |
| GEF | Guanine Nucleotide Exchange Factors |
| Gln | Glutamine |
| Glu | Glutamate |
| Gly | Glycine |
| H | Histidine |

| | |
|---|---|
| h | Hour |
| HER2/neu | Human Epidermal Growth Factor Receptor 2 |
| His | Histidine |
| HIV | Human Immunodeficiency Virus |
| HPV | Human Papillomavirus |
| HUVEC | Human Umbilical Vein Endothelial Cells |
| I | Isoleucine |
| I | Inhibitor |
| $IC_{50}$ | Half Maximal Inhibitory Concentration |
| ID | Identification |
| IFA | Indirect Immunofluorescen Assay |
| IGF-1 | Insulin-Like Growth Factor 1 |
| IL-1α | Interleukin 1 Alpha |
| Ile | Isoleucine |
| IMRT | Intensity-Modulated Radiation Therapy |
| ING | Inhibitor Of Growth 1 Gene |
| IORT | Intraoperative Radiation Therapy |
| JNK | C-Jun N-Terminal Kinases |
| K | Lysine |
| kDa | Kilodalton |
| $K_I$ | Inhibition Constant |
| $K_I^{app}$ | Apparent Inhibition Constant |
| KIT | Kit |

| | |
|---|---|
| $K_M$ | Michaelis-Menten Constant |
| $k_{obs}$ | Rate Constant For Conversion From The Initial Velocity Phase To The Steady State Velocity Phase. |
| $k_{off}$ | Rate Of Dissociation |
| $k_{on}$ | Rate Of Association |
| l | Liter |
| L- | Left Domain |
| LDL | Low-Density Lipoprotein |
| Leu | Leucine |
| LHRH | Luteinizing Hormone-Releasing Hormone |
| Lys | Lysine |
| M | Methionine |
| M | Molar |
| MAPK | Mitogen-Activated Protein Kinases |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| Met | Hepatocyte Growth Factor Receptor |
| Met | Methionine |
| MHC II | Major Histocompatibility Complex Ii |
| min | Minute |
| mM | Millimolar |
| MMP | Metalloprotease |
| MOPS | 3-(N-morpholino)propanesulfonic acid |
| mRNA | Messenger Ribonucleic Acid |

| | |
|---|---|
| MW | Molecular Weight |
| MW | Molecular Weight |
| N | Asparagine |
| N- | Amino |
| NaCl | Sodium Chloride |
| NADPH | Nicotinamide Adenine Dinucleotide Phosphate (Reduced) |
| NaOAc | Sodium Acetate Buffer |
| NaOAc | Sodium Acetate Buffer |
| NaOH | Sodium Hydroxide |
| NaOH | Sodium Hydroxide |
| NC1D | Noncollagen Domain |
| NF-κB | Nuclear Factor Kappa-Light-Chain-Enhancer Of Activated B Cells |
| NGF | Nerve Growth Factor |
| NH | Amino Group |
| NIH3T3 | Mouse Embryonic Fibroblast Cell Line |
| nM | Nanomolar |
| $NO_2$ | Nitrogen Dioxide |
| Nx | Nifurtimox |
| o-But | O-Butyl |
| OPG | Osteoprotegerin |
| P | Proline |
| p38 | P38 MAPK |
| p53 | Tumor Protein 53 |

| | |
|---|---|
| PAGE | Polyacrylamide gel electrophoresis |
| PCR | Polymerase Chain Reaction |
| PDB | Protein Data Bank |
| PDGF | Platelet-Derived Growth Factor |
| PDGFRA | Alpha-Type Platelet-Derived Growth Factor Receptor |
| pH | Measure Of Hydrogen Ion Concentration |
| Phe | Phenylananine |
| PI3K | Phosphatidylinositol 3-Kinases |
| PIM | Proto-Oncogene Serine/Threonine-Protein Kinase Pim-1 |
| pM | Picomolar |
| Pro | Proline |
| Q | Glutamine |
| R | Arginine |
| R | Radical |
| R- | Right Domain |
| $r^2$ | Coefficient Of Determination |
| Rac | A G-Protein |
| RANK | Receptor Activator Of Nuclear Factor K B |
| RANKL | Receptor Activator Of Nuclear Factor Kappa-B Ligand |
| Ras | A G-Protein |
| Rho | A G-Protein |
| RTK | Receptor Tyrosine Kinases |
| S | Serine |

| | |
|---|---|
| S | Substrate |
| s | Second |
| SA-C | Solution A with DMSO (SA-C) |
| SA-I | Solution A with Inhibitor in DMS |
| SAR | Structure-Activity Relationship |
| SCCA | Squamous Cell Carcinoma Antigen |
| SCOP | Structural Classification Of Proteins |
| SDS | Sodium Dodecyl Sulfate |
| SE | Standard Error |
| sec | Second |
| Ser | Serine |
| Src | Proto-Oncogene Tyrosine-Protein Kinase |
| SRP | Signal Recognition Particle |
| *T. cruzi* | Trypanosoma Cruzi |
| t-But | Tertiary-Butyl |
| TGF-β | Transforming Growth Factor Beta |
| Thr | Threonine |
| TIC | Type I Collagen |
| TIC | Type I Collagen |
| TIVC | Type IV Collagen |
| TIVC | Type IV Collagen |
| TM | Trademark |
| TNF | Tumor Necrosis Factors |

| | |
|---|---|
| TNF-α | Tumor Necrosis Factor-Alpha |
| Trp | Tryptophan |
| TUB | Tubulin |
| Tyr | Tyrosine |
| U87MG | Human Glioblastoma-Astrocytoma, Epithelial-Like Cell Line |
| uPA | Urokinase-Type Plasminogen Activator |
| USA | United States of America |
| V | Valine |
| $v$ | Velocity |
| Val | Valine |
| VEGF | Vascular Endothelial Growth Factor |
| VEGFR | Vascular Endothelial Growth Factor Receptor |
| $V_{MAX}$ | Maximum Velocity |
| $v_{max}$ | Catalytic Rate Of An Enzyme Without The Presence Of An Inhibitor (Lowest Concentration, $IC_{50}$ Determination) |
| $v_{min}$ | Catalytic Rate Of An Enzyme With The Presence Of An Inhibitor (Highest Concentration, $IC_{50}$ Determination) |
| $v_o$ | Initial Velocity |
| $v_s$ | Steady-State Velocity |
| W | Tryptophan |
| WNT | Wingless Gene |
| X | Nonspecific Amino Acid |
| X | Inhibitor Concentration |

| | |
|---|---|
| Y | Tyrosine |
| Y | Fractional Activity |
| Z-FR-AMC | N-Carbobenzoxy-L-Phenylalanyl-Larginine Amide, Hydrochloride |
| α | Alpha |
| β | Beta |
| γ | Gamma |
| μ | Micro |
| μM | Micromolar |
| μm | Micrometer |

ACKNOWLEDGMENTS

It is hard to say in a couple of sentences how I feel about this moment. My dissertation is the final product of several hours of patience and unconditional attention from my friends, coworkers, family and myself.

First, and foremost, I would like to thank Dr. Mary Lynn Trawick. I honestly can say, without any doubts, she is the most patient person I have ever met in my life. Thank you Dr. Trawick, for your guidance, knowledge, and patience. I have learned so much from you and I am really thankful for the opportunity you gave me in 2007.

I would like to acknowledge the financial support I received from OXiGENE, Inc in order to carry out the experimental design of my research. Additionally, these project could have not been accomplished without collaborative work between Dr. Kevin G. Pinney and the members of his groups who synthesized the tested compounds.

The present work could have not been here without the extraordinary help Ms. Alison Garzone. Ali, thank you so much for reading my dissertation. I will be eternally thankful for the amazing work you did by editing, and suggestions to improve my writing techniques.

Dr. Carlos Manzanares: Thank you Dr. Manzanares. You gave me the opportunity to reach of one my biggest dreams in life, research. Your class was one of the courses I enjoyed the most in graduate school. Also, your leading role is beyond inspirational and I am so proud that I had the opportunity to meet you.

Amanda Charlton-Sevcik and Dr. Tracy Strecker: Your friendship is probably the best help I could have received from you. Amanda, we were in the biochem group for so

many years and we shared successes and failures, and at the end, we made it together. Thank you for your sweet treats, your friendship and your moral support when I needed it. Tracy Strecker: you are an extremely intelligent and humble person. I learned so much from you, and I do not think I would have ever made it without your help.

Trawick group (TG): This long journey was also possible with the help of former Trawick group members, Dr. Milenka Arispe and Dr. Sam Chen. Thank you. I would also like to thank Justin Tidmore, and Samuel Odutola.

Undergraduate students: Jake Wiley, Melinda Soeung, Morgan Cooper, Jay Johnson, and Joe Guillory. I was blessed because I had the opportunity to work with amazing students through my graduate student experience. Mentoring all of you was inspirational and taught me so much about my personality. Each one of you gave me important life lessons and I believe it was very beneficial for me to work with you all. Good luck and hope you achieve your goals in life.

My acknowledgments would not be completed if I don't mention the wonderful staff of the department of Chemistry and Biochemistry and Baylor University. My deep appreciation is to Ms. Nancy Kallus, Ms. Adonna Cook and Ms. Virginia Hynek. Also, a special thank you to Dr. Michelle Nemec and Dr. Craig Moehnke for their technical support that helped me with my research.

My dissertation committee: Dr. Sung Kun Kim, Dr. Bryan Shaw, Dr. Jaime Diaz-Granados and Dr. Kevin Pinney. Your suggestions and guidance will always be appreciated. I am thankful I had the opportunity to work with you. Your research and ethics are incomparable and helped me respect you as scientists.

My friends: Denka, Helena, Austin, Ivanna, Nelson, Alfredo, Otsmar, Tara, Pilar, Addy, Ricardo, Ashley, Rema. You guys became like my family in Waco and I will remember you for the rest of my life. Our conversations, adventures, tears, and joys, made us stronger and your friendships are, without a doubt, the best gift you could have ever given to me. Jose Boquin deserves a special mention. You introduced me to this wonderful university, and I will always thank your for your help and friendship.

Evelyn Chavarria: my mom. A paragraph is not enough to tell you how much I love you. You have always been the best example I have had in life about personal success. Your life inspired me and I would not be the man I am without your love, guidance and support. La quiero mucho madre.

Lidia Nolasco: "mamá Lidia". You are the reason I am here. " My other mom". I have always admired your integrity and education. My professional success is the product of your hard word and love. Gracias mamá Lidia por sus consejos. Su consentido.

German Chavarria Umanzor: my father. My role model. You will always be my only father and I miss you a lot. You left a huge gap in my heart, but I am pretty sure you are in heaven taking care of me. I used to dislike your music, old people's music'; but now, your records are my most valuable treasure. Papa, lo extraño mucho.

Bertha (Tita) y Olga Chavarria. You both are not just my aunts. I always see you as my substitute mothers. There are no words to thank you for what you have done for me. Your love, support, advice, and guidance exceed what any person could have ever wished for. Javier and Layla: the new generation of Chavarrias'. You represent the future of our family and God's gift to our family.

Eliana Argueta. I guess you are my favorite relative in the family. We get along so much and I wish nothing but happiness for you and your family.

Mauricio Chavarria: the MD. You became an inspiring role when I was a little child and I hope to be as good as you are in your profession.

Nolasco family: Simply. Thank you!

This list of acknowledgments would not be complete if I did not mention the support I have had from this incredible country, the United States of America. I am infatuated by your culture, your people, and everything that makes this country what it is. Thank you for opening your doors to me and I hope the prosperity you have gained will never end. God bless America!

God: you're the last one in the list. But it is because you're the most important in the thank you list. My faith in you brought me here and I will be your son. My love for you is unconditional and I am proud to say that I am your follower. Thank you for your blessings and actions. You had a reason for me in life and I am always trying to understand your decision. God, you're good!

# DEDICATION

To

My parents

German Chavarria Umanzor

and

Lidia Nolasco

This is the product of your love and dedication

# CHAPTER ONE

## Introduction

### *Statement of Purpose and Significance*

Science has always been fascinated by the extraordinary ability of enzymes to catalyze numerous reactions in living organisms. Their catalytic activity has been a major challenge in drug discovery due to their vital role in pathological diseases and conditions. Therefore, enzymes are often the main target in drug discovery. Examples of enzymes implicated in pathological conditions include the kinases, proteins that have a pivotal role in cancer.[1] To date, there are cases where treatments have been accepted for the treatment of specific conditions. For example, the Food and Drug Administrations have approved several medications for the treatment of the Human Immunodeficiency Acquired (HIV) that directed toward enzymatic targets. The list of enzymes include: nucleoside reverse transcriptase, HIV integrase and proteases.[2]

Proteases comprise one of the most studied groups due to their implications in diseases. The cysteine cathepsins are proteases and have become a 'hot' target in research due to their participation in numerous processes. Two members of this family, cathepsins L and K have been extensively characterized and studied. Pioneer cathepsin K research in drug discovery has lead to the discovery of selective inhibitors of this protease that are currently in clinical trials as potential anti-osteoporosis agents. To date, cathepsin L inhibitors are not currently in clinical trials.

Cathepsins L and K play important roles in bone metastasis and cathepsin L overexpression is associated with a number of metastatic cancers. The main cause of

death in patients with cancer is its metastatic spread. Overexpression of cathepsins L and K promote the degradation of the components of the extracellular matrix in cancer microenvironments.

The main objective of this research project was to examine a library of small molecules as inhibitors of individual enzymes, cathepsin L and cathepsin K, for potential use in targeted therapies as anti-metastatic agents. The ideal chemotherapeutic agent should meet several criteria, including: potency, selectivity, solubility, and be non-toxic in patients. It should produce only minor side effects, and exhibit stability when administered with long shelf life (chemically stable), etc. One of the major disadvantages when using non-specific chemotherapy as a cancer treatment is the considerable list of negative side effects (i.e. vomiting, hair loss, skin rash). In targeted therapy, the inhibition of one specific enzyme or protein has the potential to limit these negative side effects.

The discovery of cathepsins L and K inhibitors involve several considerations. One the first steps in the search of these compounds was the creation of a library of synthetic compounds.[3–7] These compounds were synthesized by members of Dr. Kevin G. Pinney laboratory. Thiosemicarbazone analogs were first applied to the cathepsin L-like enzyme cruzain[8,9], and then to cathepsin L as enzyme inhibitors.[10–13] Assay optimization is a key step in drug discovery. Discrepancies in results might lead to misinterpretations and false positives. Thus, the necessity of solid and reliable assays was mandatory. The analysis of their structure-activity relationships SARs was extremely helpful in order to achieve these goals. SARs helped to optimize pre-clinical studies. They also helped to identify possible selectivities and lead to the design of more

analogs that could be potent inhibitors. The characterization of these compounds, including the kinetic mechanism of action is important for the design and study of possible candidates that can function as effective drugs in the treatment of cancer metastasis. This involved numerous evaluations. Fluorometric based assays were utilized to study various assay parameters in inhibitory activities, reversibility studies, determination of $K_{I,}$ and the mechanism and mode of inhibition.

We also explored the application of thiosemicarbazones as potential inhibitors of the activation of procathepsin K which occurs under acidic conditions. The inhibition of the activation of procathepsin K was explored by using one of the lead thiosemicarbazone inhibitors of cathepsin K. Experiments using human type IV collagen, a natural substrate of cathepsin K, were carried out in order to determine if thiosemicarbazones were capable of inhibit the outstanding collagenase activity of cathepsin K using molecular biology techniques and fluorescent protein stains.

Results indicated that a number of thiosemicarbazones are lead compounds as cathepsin L inhibitors. We also explored the inhibition of the type I collagenase activity of cathepsin L by using a good inhibitor of this protease that was found in the library of inhibitors.

Cancer metastasis is characterized by enhanced cell invasion and cell migration. Two dimensional cell based assays were used to determine the inhibitory ability of these compounds within the extracellular matrices of a metastatic breast cancer cell line.

Another cysteine protease, cruzain, is the most potent hydrolase found in *Trypanosoma cruzi*, the responsible of Chagas' disease. Chagas' disease, once upon considered as a disease affecting countries in development, is rapidly spreading

throughout developed countries. The parasitic infection is now considered a threat in some European countries and the United States. Cruzain is mainly involved in cell host invasion and organ invasion due to its proteolytic capability of degrading surrounding tissues.

On the other hand, the treatment of Chagas' disease, a condition that has become chronic in millions of people in South America, is a major challenge. There is no effective treatment for the eradication of the disease and efforts in the search for better alternatives are non-existing. The disease was considered a neglected condition, but this is rapidly changing due to the increasing number of patients in the United States and Europe. We evaluated many analogs in the library of thiosemicarbazones against cruzain in the search for therapeutic agents against Chagas' disease.

These studies were part of an ongoing collaboration between the Trawick and Pinney group (Baylor University). Synthesis of synthetic thiosemicarbazones and biological evaluations were carried out by members of Dr. Kevin G. Pinney group and Dr. Mary Lynn Trawick group respectively.

### *General Overview of Cancer*

Defining cancer can be quite difficult. There is not a unique definition due to the complexity of the disease. Many factors can affect the specific definition of cancer, such as the origin, organ affected as well as others. Thus, a combination of these concepts will help to understand the basics of the disease. Cancer is a generic name that numerous diseases share in their pathology: growth of malignant cells that invade healthy tissues and organs. Thus, tumors are the collection of malignant cells. Additionally, a malignant tumor has the ability to invade healthy organs and tissues.[14,15] Perhaps, the complexity of

cancer is the biggest obstacle when finding appropriate treatments. Several factors promote the development and spread of this disease in normal tissues; thus a specific route for cancer treatment seems a long term objective. The combination of chemotherapy, radiotherapy, surgery and targeted therapies increases survival rates in cancer patients. Nowadays, chemotherapy is one of the most used therapies in the treatment of cancer.

### *Types of Cancer and Organ Distribution*

As previously stated, cancer is a generic term that comprises more than 100 diseases with similar characteristics. However, a general form to identify them is by defining the organ(s) that are mainly affected. In general, the most common cancers distribute within oral cavities, endocrine, digestive, respiratory genital, urinary, bones, heart, skin, breast, eye, and brain systems. Also, lymphoma, myeloma, and leukemia are also other types of cancer.[16] However, organs can be affected by more than one type of cancer. For example, there are several different types of brain cancers affecting children (gliomas, astrocytomas, brain stem, ependymoma, medulloblastomas and other).[17]

### *History of Cancer*

Cancer has been prevalent for centuries. Nevertheless, early civilizations did not understand the causes and origins of this disease. The oldest written description of cancer was originated in 3000 B.C. in the Egyptian civilization. However, the term cancer was used for the first time by Hippocrates in 460 B.C. The Greek physician used the terms “carcinos” and “carcinoma” when referring to ulcer-based tumors. Modern terminutesology, “cancer” and “oncos”, were first used by Celsus (28-50 B.C.) and Galen

(130-200 A.D.), respectively. "Cancer" and "oncos" are the Greek words for crab and swelling.[18,19]

*Statistics of Cancer*

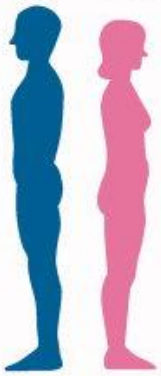
Cancer and its variations combined represent the second leading cause of death in the United States.[20] More than 550,000 patients died in 2010 of cancer.[21] It is also important to observe the economical impact of cancer in society. Cancer accounted for more than 115 billion dollars in lost productivity when other indirect costs, such as care giving and job wages, are included, the costs are much higher. The total cost of cancer was calculated to be $232 billion in 2000. Projections estimate this number can increase up to $308 billion in 2020.[22]

Lung and bronchus cancers account for 14% of new cases in the United States in 2012 (468610 cases).[21] However, breast and prostate cancer are the leading types of cancer for the same period in women and men respectively. Prostate cancer represents an alarming thirty percent of new cases within the male population. In the female population, thirty percent of new cases of breast cancer account for in the United States.

Other leading types of cancer found in the male population are: colon, melanoma, leukemia, and pancreas. Similarly, new cases of cancer in females are distributed among uterus, thyroid, melanoma and colon. Figure 1 shows the ten leading cancer types only in the United States.[21]

Estimated New Cases*

| Males | | | Females | | |
|---|---|---|---|---|---|
| Prostate | 241,740 | 29% | Breast | 226,870 | 29% |
| Lung & bronchus | 116,470 | 14% | Lung & bronchus | 109,690 | 14% |
| Colon & rectum | 73,420 | 9% | Colon & rectum | 70,040 | 9% |
| Urinary bladder | 55,600 | 7% | Uterine corpus | 47,130 | 6% |
| Melanoma of the skin | 44,250 | 5% | Thyroid | 43,210 | 5% |
| Kidney & renal pelvis | 40,250 | 5% | Melanoma of the skin | 32,000 | 4% |
| Non-Hodgkin lymphoma | 38,160 | 4% | Non-Hodgkin lymphoma | 31,970 | 4% |
| Oral cavity & pharynx | 28,540 | 3% | Kidney & renal pelvis | 24,520 | 3% |
| Leukemia | 26,830 | 3% | Ovary | 22,280 | 3% |
| Pancreas | 22,090 | 3% | Pancreas | 21,830 | 3% |
| **All Sites** | **848,170** | **100%** | **All Sites** | **790,740** | **100%** |



Estimated Deaths

| Males | | | Females | | |
|---|---|---|---|---|---|
| Lung & bronchus | 87,750 | 29% | Lung & bronchus | 72,590 | 26% |
| Prostate | 28,170 | 9% | Breast | 39,510 | 14% |
| Colon & rectum | 26,470 | 9% | Colon & rectum | 25,220 | 9% |
| Pancreas | 18,850 | 6% | Pancreas | 18,540 | 7% |
| Liver & intrahepatic bile duct | 13,980 | 5% | Ovary | 15,500 | 6% |
| Leukemia | 13,500 | 4% | Leukemia | 10,040 | 4% |
| Esophagus | 12,040 | 4% | Non-Hodgkin lymphoma | 8,620 | 3% |
| Urinary bladder | 10,510 | 3% | Uterine Corpus | 8,010 | 3% |
| Non-Hodgkin lymphoma | 10,320 | 3% | Liver & intrahepatic bile duct | 6,570 | 2% |
| Kidney & renal pelvis | 8,650 | 3% | Brain & other nervous system | 5,980 | 2% |
| **All Sites** | **301,820** | **100%** | **All Sites** | **275,370** | **100%** |

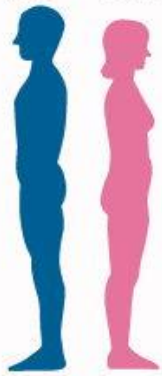


Figure 1. Ten Leading Cancer Types for the Estimated New Cancer Cases and Deaths by Sex, United States, 2012 (directly reproduced from Siegel , page 13)

*Causes of Cancer*

Advances in science and technology have opened the opportunity to study in depth some of the cancer that attacks our modern society. Research also has demonstrated that cancer can be triggered by internal and external factors that influence the development of malignant tumors.

*Oncogenes*

In general, an oncogene is a gene that promotes the development of cancer (tumorigenesis). Numerous genes have the potential of becoming oncogenes and extensive research has been done to study the role of oncogenes in the development of cancer.[23] The most common oncogenes found in human cancer belong to the Rat Sarcoma family (RAS). The family is a well known series of related proteins that participate in cellular signaling, replication stress and activation of DNA damage responses.[24] The Proto-Oncogene Proteins c-pim-1 family (PIM) constitutes a family of related proteins, serine/threonine kinases, which are associated with the prognosis of solid tumors, especially prostate cancer. However, they are considered as proto-oncogenes.[25] The Hepatocyte Growth Factor Receptor Met, a tyrosine kinase oncogene has been strongly related to the progress of basal-like breast carcinoma.[26] Other oncogenes have been identified and several reviews show the importance of these targets. The Human Epidermal Growth Factor Receptor 2 HER2/neu is a proto-oncogene; however, its overexpression in normal cells can catalyze the activity of this protein as a oncogene.[27] Other proteins that have been found to have oncogene-properties are BCL-ABL (Proto-Oncogene Proteins fusioned with Abelson proto-oncogene), Tyrosine-Protein Kinase (KIT), and Platelet-Derived Growth Factor Receptor, α-polypeptide (PDGFRA).[28–30]

*Tumor-Suppressor Genes*

In contrast tumor-suppressor genes have the ability to arrest the development of tumors in human bodies.[31] Thus, tumor-suppressor genes are also known as anti-oncogenes. Tumor-suppressor gene research is more challenging due to the difficulty to

restore the anti-oncogene activity of these targets. Some of these genes include p53, a much studied tumor-suppressor gene with a high activity against genome mutation, and ING genes, tumor-suppressor genes in the formation of head and neck carcinomas.[32,33] Another example of a tumor-suppressor gene is breast cancer type 1 susceptibility protein BCRA.[34]

*Viruses as Causes of Cancer*

Viruses have become one the biggest centers of attention for scientists for their ability to induce cancer in humans. The discovery of viruses as cancer-promoter agents one hundred years ago opened an undiscovered world that viruses that could attack healthy cells. Hepatitis B and C viruses can promote liver cancer.[35] Another herpes virus, the Epstein-Barr virus has been linked with non-Hodgkin lymphomas and respiratory system cancer.[36] Patients with human immunodeficiency virus (HIV) have higher risks to develop Kaposi sarcoma and non-Hodgkin lymphomas among others.[37] Perhaps, the best known case is related to the human papilloma virus (HPV).[38] This virus is the reason for a high number of female patients with genital cancers including cervix, vulva, vagina, and anus. Furthermore, the carcinogen action of HPV has been extended in the male population (penis cancer) as well.[39] New studies have related oral cancer (tongue, tonsil, and throat) with this virus.[40]

*Non-Biological Carcinogens*

Oncogenes, tumor-suppressor genes, and Human papillomavirus (HPV) are considered biological carcinogens due to their expression (or sometimes overexpression) or transmission among mammalian organisms. Nevertheless, the existence of hundreds

of carcinogens is well known and studied, and they account for a significant number of new cases of cancer every year. Physical, chemical, and external biological agents are classified as carcinogens and human exposure to these agents is now restricted. The most common carcinogens found on daily basis are tobacco, sunlight, and coal-tars. Cancer is also promoted by exposure to other agents or material such as arsenic, coal production, secondhand smoke, sunbeds, X-radiation, and asbestos.[41–44]

## *Cancer Treatments*

Medicinal research has achieved important goals in the fight against cancer. Death rates have decreased and cancer patients have higher chances of overcoming this terrible disease. Generally, cancer treatments can be divided into five major categories: Surgery, hormone therapy, radiation, immunotherapy, targeted therapy, and chemotherapy.

### *Surgery*

Overall, surgery is one the most common techniques in cancer treatment. The main objective of surgical operations is the removal of the tumor along with the blood vessel network that has been formed around the tumor and the lymph nodes. This technique has been greatly evolved from early days when radical surgeries were carried out (like complete mastectomies with women affected with breast cancer) to modern techniques where less invasive surgeries are performed. Endoscopic surgeries are efficient tools when it comes to surgeries. This technique utilizes modern technology (exploratory surgery) to remove tumors in the colon, esophagus, and breast. Other

modern methods that are classified as surgeries are cryosurgery (by using liquid nitrogen), and laser surgery.[45–48]

*Hormone Therapy*

Cancer is often strongly hormone dependent, especially in breast (estrogen) and prostate (testosterone) cancers. The discovery of drugs such as Tamoxifen, has shown that patients with breast cancer have higher probabilities of overcoming their conditions.[49] Similarly, male patients with prostate cancer have been treated with Androgen Deprivation Therapies (ADT).[50] Briefly, testosterone is one of the main responsible in abnormal cell growth in the prostate. Aromatase inhibitors (LHRH or Luteinizing Hormone-Releasing Hormone) are also considered as great promises for the treatment of cancer in male and female patients.[51]

*Radiation*

Another technique that is greatly utilized for the treatment of cancer is radiation, also known as radiotherapy. Today's technology offers radiotherapy devices with excellent precision when treating a tumor without affecting neighboring organs or tissues. Conformal radiation therapy (CRT), intensity-modulated radiation therapy (IMRT), stereotactic radiosurgery, stereotactic radiation therapy, and intraoperative radiation therapy (IORT) are also other used techniques that belong to this category.[52–55]

*Immunotherapy*

Immunotherapy is the biological analogy of chemotherapy (also discussed in this chapter). It is also known as biological response modifier therapy, biologic therapy, or biotherapy. The idea is simple: researchers are looking for natural components found in

the human body that control the growth of abnormal cells.[56] Examples of these agents are interferons, interleukins, or cytokines.[57–60] The production of monoclonal antibodies that could target cancer cells has been extensively promoted. Currently, Bavituximab, is in clinical trials. The antibody was designed against the membrane phospholipid phosphatidylserine and has been utilized in the treatment of lung cancer.[61] Rituximab (a monoclonal antibody against CD20, a protein highly expressed in B cells) and trastuzumab (a monoclonal antibody against HER2/neu, a proto-oncogene) are FDA approved immunotherapy agents that are currently used in the treatment of lymphoma and breast cancer, respectively.[62–64] Furthermore, prostate cancer patients now have the option to get treated with Sipuleucel-T, a vaccine for hormone-refractory prostate cancer.[65]

*Targeted Therapies*

The latest discoveries in cancer treatment can be compiled under this category. The objective of these agents is to control the growth, division, and life-cycle of cancer cells. Currently, there are three sub-categories for targeted therapies. Growth signal inhibitors, angiogenesis inhibitors, and apoptosis-inducing drugs. Growth signal inhibitors are compounds that specifically target growth factors and other proteins, such as receptors, that are closely involved in cell signaling. Gefitinib, imatinib, cetuximab, dasatinib, and nilotinib are some of these compounds that showed good efficacy when treating patients with lung cancer, kidney cancer, and glioblastomas.[66,67] Angiogenesis inhibitors are aiming to target the generation of blood vessels that are formed around tumors. Angiogenesis is the biological process that generates new networks of blood vessels. Bevacizumab, an angiogenesis inhibitor, has been extensively used for the

treatment of numerous cancers, including kidney, lung, and advanced colorectal.[68] Apoptosis-inducing drugs are specifically designed to promote programmed cell death in cancer cells. Apoptosis, which is the process of programmed cell death, is another biological process for target-based therapies.[69] Table 1 summarizes some of these anti-cancer agents and their targets.[70]

Table 1. Selected Anti-Cancer Agents that are Used as Targeted Therapeutic Agents[70]

| Hematologic Malignancies | | Solid Tumors | |
|---|---|---|---|
| Molecular Target | Anti-Cancer Agent | Molecular Target | Anti-Cancer Agent |
| CD20 | Rituximab<br>$^{90}$Y-Ibritumomab<br>$^{131}$I-Tositumomab | HER2/neu | Lapatinib<br>Trastuzumab |
| CD33 | Gemtuzumab ozogamicin | EGFR | Erlotinib<br>Gefitinib<br>Cetuximab<br>Panitumumab |
| CD52 | Alemtuzumab | VEGFR | Sorafenib<br>Sunitinib |
| BCR-ACL | Imatinib<br>Dasatinib | VEGF | Bevacizumab |

*Chemotherapy*

One of the most important therapies in the fight against cancer is chemotherapy. This therapy relies on the power of chemical agents in the eradication of cancer cells by altering their DNA or affecting important steps in cell growth and replication. Aminopterin, for example, is a potent inhibitor in DNA replication and is used for the treatment of leukemia.[71] Methotrexate, is also used for the treatment of choriocarcinoma (a form of uterus cancer).[72] The list of chemotherapeutic agents is extensive and several reviews offer more complete information about each one of them. However, undesired

effects accompany these treatments. Some side-effects include nausea, vomiting, and hair loss. Thus, researchers are always looking for new chemotherapeutic agents that overcome these problems. Chemotherapeutic agents are usually divided into seven categories. Table 2 shows the classification of chemotherapeutic agents, their targets, and some examples for each type of compounds.[73]

### *Cancer Metastasis*

It is well known that the main reason for death within patients with any kind of cancer is the metastasis of the tumor. Thus, it is important to understand the biochemistry and cell biology of metastasis in order to identify potential targets with new and more powerful agents. Metastasis is the process when a cancer tumor (primary) spreads and invades other organs or tissues (metastasis). However, the metastatic tumor conserves identical features to the primary tumor, that is, overexpression or underexpression of proteins, morphology, and similar chromosomal composition. Once the tumor undergoes the metastatic process, survival rates decrease dramatically and treatment becomes inefficient.[74] Metastasis can occur in almost every case of cancer. Lungs are the most common target organ for metastasis. Several cancers including breast, colon, kidney, melanoma, ovary, pancreas, prostate, rectum, stomach, thyroid, and uterus metastasize to this organ. Lung cancer accounts for the highest number of new cases and deaths among cancer patients. Other metastatic organs are liver, bones, peritoneum, and adrenal glands.[69,75]

Table 2. Classification of Chemotherapeutic Agents

| Type | Mechanism of action | Examples | References |
|---|---|---|---|
| Alkylating agents | Alters DNA structure | Cyclophosphamide<br>Busulfan<br>Temozolomide<br>Cisplatin | 76–78 |
| Antimetabolites | Inhibits DNA replication | Cladribine<br>Floxuridine<br>Fludarabine<br>Methotrexate | 77,79 |
| Anti-tumor antibiotics | Inhibits enzyme involved with DNA replication | Daunorubicin<br>Doxorubicin<br>Epirubicin<br>Mitomycin-C | 80,81 |
| Topoisomerase inhibitors | Inhibits topoisomerases | Topoisomerase I inhibitors<br>Topotecan<br>Irinotecan<br>Topoisomerase II inhibitors<br>Etoposide<br>Teniposide | 82,83 |
| Mitotic inhibitors (Tubulin-binding agents) | Inhibit cell mitosis or enzymes related to cell reproduction | Paclitaxel<br>Docetaxel<br>Vinblastine<br>Vinorelbine | 84 |
| Corticosteroids | Hormone-type therapy | Methylprednisolone<br>Dexamethasone | 85–87 |
| Others | | Bortezomib (L-asparaginase inhibitor) | 88 |

### *The Metastatic Mechanism*

Several reviews have been written about the metastatic cycle. One of them is presented by Geiger and Peeper.[89] The complexity of this mechanism can be summarized into eight steps. Invasion and migration (steps one through three) will be further discussed in chapter two.

1. Tumor cells experience an epithelial-mesenchymal transition.
2. Tumor cells degrade the extracellular matrix (ECM) by using proteolytic enzymes.
3. Tumor invades surrounding tissues or organs.
4. Tumor intravasation into a new or pre-existing network of blood vessels
5. Tumor cells are transported through the blood vessel system.
6. Tumor cells undergo extravasation at the final destination (metastasis site)
7. Tumor cells could undergo a log phase (asymptomatic phase)
8. Tumor cells start the formation of a new tumor in the metastatic site. This process requires ECM remodeling and angiogenesis.
9. Cells surrounding the acidic microenviroment undergo anoikis or programmed cell death.

### *Proteolytic Enzymes (Proteases)*

One of the key steps in cancer metastasis is the role of proteolytic enzymes in the migration and invasion of cancer cells. Thus, the importance of proteases as anti-cancer targets has been observed by several researchers.[90]

*Classification of Proteases*

A general classification divides proteases by their active site. Usually, proteases are divided into four categories: serine proteases, aspartate proteases, cysteine proteases, and metalloproteases.[91] Table 3 summarizes the main features and differences for each sub classification of these proteases.

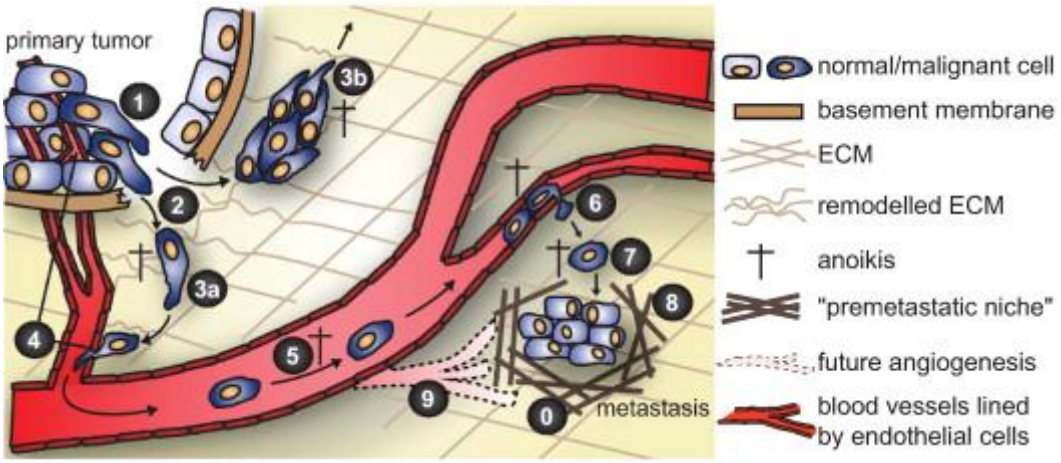



Figure 2. The Metastatic Process. (directly reproduced from Geiger, page 294)[89]

*The Cathepsins (Cysteine Proteases)*

Currently, there are fifteen well characterized cathepsins found in nature.[92] Interestingly, they do not belong to the same classification as serine, aspartate, or cysteine proteases. Two cathepsins are classified as serine proteases (cathepsins A and G);[93,94] two cathepsins are aspartyl proteases (cathepsins D and E),[95,96] and the rest of them are classified as cysteine proteases (cathepsins B, C, F, H, K, L, O, S, V, W, and X).[92] The cathepsins are considered as peptidases; that is, they have the ability of cleave peptidic bonds. Expression of the cathepsins is mainly found in the lysosomal, cellular organelles responsible of processing cellular debris.[97] Cathepsin activity is usually maximal under

acidic conditions. Cysteine proteases are also classified by their specific target. Cathepsins L, S, K, V, and F are endopeptidases, while cathepsin C has exopeptidase and aminopeptidase activities; cathepsin X is an exopeptidase; and cathepsin H is an aminopeptidase and exopeptidase. Cathepsin B has both exopeptidase and endopeptidase activities. The present work will focus on two cathepsin cysteine proteases, cathepsin K, and cathepsin L, and cruzain, a cathepsin L-like enzyme. Cruzain, a parasitic protease found in *Trypanosoma cruzi* will be discussed in detail in Chapter 4.

Table 3. Classification of Hydrolases

| Protease | Main feature/mechanism of action | Examples | References |
|---|---|---|---|
| Serine/Threonine proteases | Cleave peptidic bonds by using a serine as a nucleophilic residue in the active site | Trypsin<br>Elastase<br>Chymotrypsin | 98,99 |
| Aspartate proteases | Two highly conserved aspartate residues are found in the active site (acid-base mechanism) | Pepsins<br>Renins<br>Cathepsins | 100 |
| Cysteine proteases | A cysteine residue is the key amino acid in the catalytic activity of the enzyme | Cathepsins<br>Papain<br>Caspases | 92,101 |
| Metalloproteases | Their active site contains a metal, usually zinc or cobalt | MMP1<br>MMP8<br>MM14 | 102 |

*Classification of the Cathepsins*

Proteins can be classified according to their structural characteristics. The Structural Classification of Proteins (SCOP) and the CATH protein structure classification are usually well accepted classification systems for macromolecules. Table 4 shows how cysteine proteases are classified based on their structural features. The

SCOP system classifies cysteine proteases as alpha + beta proteins. The fold is named cysteine proteases, and the superfamily is called cysteine proteases. Finally, they all belong to the papain family. A similar classification can be found under the CATH system.[103,104]

Table 4. SCOP Classification of Cysteine Proteases

| SCOP | Name |
|---|---|
| Classification | Alpha + beta |
| Fold | Cysteine proteinases |
| Superfamily | Cysteine proteinases |
| Family | Papain |

*Distribution of Cysteine Cathepsins*

Seven cysteine proteases (cathepsins B, L, H, O, F, X, and C) are ubiquitously distributed in mammalians. Liver, for example, expresses five of these proteins (cathepsins F and C are not expressed). Other organs that express these cysteine proteases are thyroid gland, kidneys, spleen, and placenta. In contrast, cathepsins S, K, V, and W are expressed in specific organs or cells. For example, cathepsin K is usually expressed in osteoclasts and macrophages; cathepsin S is expressed in the heart and spleen. Similar cases are can be seen for cathepsin V and W that are expressed in cornea and lymph nodes, respectively.[105] Table 5 shows cysteine proteases that are found ubiquitously in mammalians organs.

Table 5. Expression of Cysteine Proteases that are Distributed Ubiquitously (Adapted from Berdowska)

| Cathepsin | Liver | Thyroid Gland | Kidney | Spleen | Placenta | Ovary | Lung |
|---|---|---|---|---|---|---|---|
| B | + | + | + | + | - | - | - |
| L | + | + | + | - | - | - | - |
| H | + | - | + | + | - | - | - |
| O | + | - | - | + | + | + | - |
| X | + | - | + | - | + | - | + |
| C | - | - | + | + | + | - | + |

*Biological Roles of Cysteine Proteases*

Proteases play an important role in biological processes. Specifically, cysteine proteases are quite involved in the degradation and processing of important macromolecules involved in several cycles.[105,106] Table 6 summarizes some of the most important biological processes where cathepsins are primarily involved. The most important biological functions are related to protein processing, immunity, bone remodeling, and differentiation of keratinocytes, reproduction processes, and apoptosis.

Additionally, cathepsin H has been linked with the atherogenesis processes, and the transformation of LDL and protein C macromolecules.[107] Similarly, cathepsin X is linked to immune system processes and inflammation.[108] Cathepsin V is able to degrade plasminogen, releasing subproducts that are angiostatin-like proteins.[109] It has been demonstrated that cathepsins L, S, and K might participate in obesity and weight loss in animal studies. The importance of cysteine protease in pathological roles will be discussed separately.[110]

Table 6. Selected Biological Roles of Cysteine Proteases.

| Function | Specific activity | Cathepsin | References |
|---|---|---|---|
| Protein activation | Proteolysis of thyroglobulin<br>Activation of β-galactosidase, renin and trypsin<br>Activation of granzymes A and B<br>Activation of SP-C in lung | B, L, K<br>B<br>C<br>H | 111–113 |
| Mediation of antigen presentation | Generation of antigenic peptides<br>Mediation of MHC II antigen<br>Release of CLIP by Ii processing | B, L, S<br>S, L<br>S, L, F, V | 114,115 |
| Bone remodeling | Bone remodeling (resorption and formation) | K | 116 |
| Hair follicle cycle | Anomalies in hair follicle cycle and skin morphology<br>Terminal differentiation of keratinocytes | L | 117 |
| Reproduction cycles | Spermatogenesis<br>Oogenesis and embryogenesis<br>Embryo partition | L | 118 |
| Apoptosis | Granzymes A and B (implicated in apoptosis)<br>TNF-α induced apoptosis<br>Activation of procaspase 3 | C<br>B<br>L | 119 |

*Mechanism of Action of Cysteine Proteases*

There are three key active site residues: Cys25 (L), His159 (R), and Asn175 (R). The former two residues form a thiolate-imidazolium ion that is stabilized by Asn175 via hydrogen bonds. The difference between serine and cysteine proteases is cysteine residues are present in an ionized form. The thiolate moiety attacks the carbonyl amide carbon of the substrate to form a tetrahedral intermediate. Then, the complex converted into an acyl enzyme, releases the C-terminal portion of the substrate. A molecule of

water hydrolyzes the acyl-enzyme intermediate by forming a second tetrahedral complex. Lastly, the complex, enzyme-intermediate, releases the intact enzyme and the N-terminal portion of the substrate. (Figure 3).[106]

*General Structure of Cysteine Proteases*

Cysteine proteases are expressed with three well defined portions. They all contain a signal peptide, a propeptide, and the mature domain. The signal peptides are short polypeptides having between 15 (cathepsin K) and 24 residues (cathepsin C). The propeptides are medium size polypeptide regions that range between 38 (cathepsin X) and 251 residues (cathepsin F).[120]

Finally, the mature domain varies in length depending on each protein. Cathepsin F contains the shortest amino acid sequence –ironically, this protease's pro-region is bigger than the mature form- and cathepsin B is the biggest protein with 260 amino acid residues.[106] The catalytic active site is composed of three amino residues: Cys25 (L), His159 (R), and Asn175 (R) (papain numbering). Cys25, which acts as nucleophile during peptide hydrolysis, is surrounded by a conserved polypeptide made of 9 residues: Cys-Gly-Ser-**Cys**-Trp-Ala-Phe-Ser. The second residue, His159, is embedded in conserved hexapeptide that has the sequence: Gly-**His**-X-X-X-Gly, where X could be any aliphatic hydrophobic residue (Val, Ile, or Leu). Lastly, Asn175, the third amino acid, composes a tripeptide with the form **Asn**-Ser-Trp. Both Cys25 and His159 interact with two polar and one hydrophobic residue (Gln19, Gly68, Trp183), the N-terminus Pro2, and some cysteine residues as well.

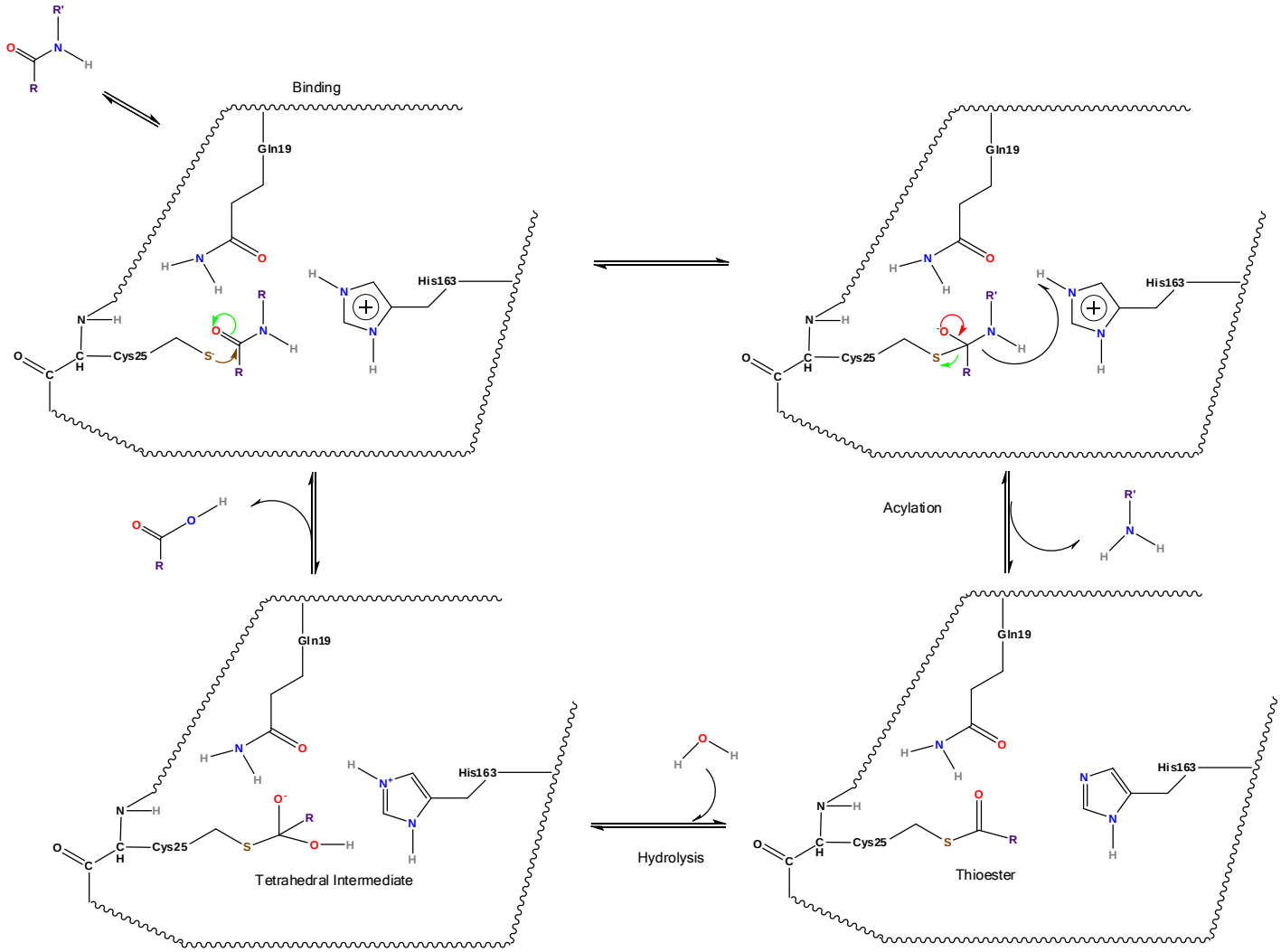



Figure 3. Mechanism of Action of Cysteine Proteases

Deeper analysis of sequence alignment reveals that cysteine proteases can also be divided into three cathepsin-like: cathepsin B-like, cathepsin L-like, and cathepsin F-like.

For example, cathepsins S, K, V, and L are cathepsin L-like. Furthermore, Karrer's work found specific characteristics for each subclass. Cathepsin L-like proteases share the ERF/WNIN motif. The research group found cathepsins, with the exception of cathepsin B, have proregions made with 107 amino acids containing a highly conserved hexapeptide (ERFNIN) using cloning techniques.[121] Similarly, cathepsin F-like subfamily contains the ERFNAQ/A motif. Finally, cathepsins C, O, and X cannot be classified as a cathepsin-like subclass.

*Crystal Structures of Cysteine Proteases*

Crystal structures of nine members of human cathepsins have been elucidated. The list includes cathepsins, B, L, K, H, X, V, C, S, and F. Crystal structures for cathepsins O, W, and X have not yet been resolved.[92,122–130]

Cysteine proteases are formed by two domains: left (L) and right (R). The L-domain's main characteristic is formed by three α-helices. The longest helix is called the central helix, and is a polypeptide of more than 30 residues long. The R-domain could be described as a β-barrel forming a coil-type motif. The barrel is enclosed by an α-helix at the bottom of the macromolecule. One of the components of the catalytic triad, His159, is located at the top of the barrel. Overall, both domains open up, forming what is called the active-site cleft. This is the part of the protease where Cys25 and His159 are located. Cys25 comes from the N-terminus of the L-domain while His159 comes from the C-terminus of the R-domain.

The active site surface is made of residues coming from four loops, two from each domain. The L-domain loops are shorter and disulfide-bond connected. The R-domain loops, on the other hand, are larger in length and are placed at the top of the β-barrel motif.

However, endopeptidases and exopeptidases have subtle differences in their crystal structures that help their substrate affinity. First, the active-site in endopeptidases is located between both domains. Nevertheless, exopeptidases have fewer binding sites due to the fact that more loops are occupying space of the active-site. Figure 4 shows General representation of cysteine proteases. Three active residues (Cys25, His159 and Asn175) are shown in ball and stick form. The structure is colored according to the following: α-Helices are red, β-sheets are cyan, turns are green, and coils are white. (PDB ID: 1ME3).[131]

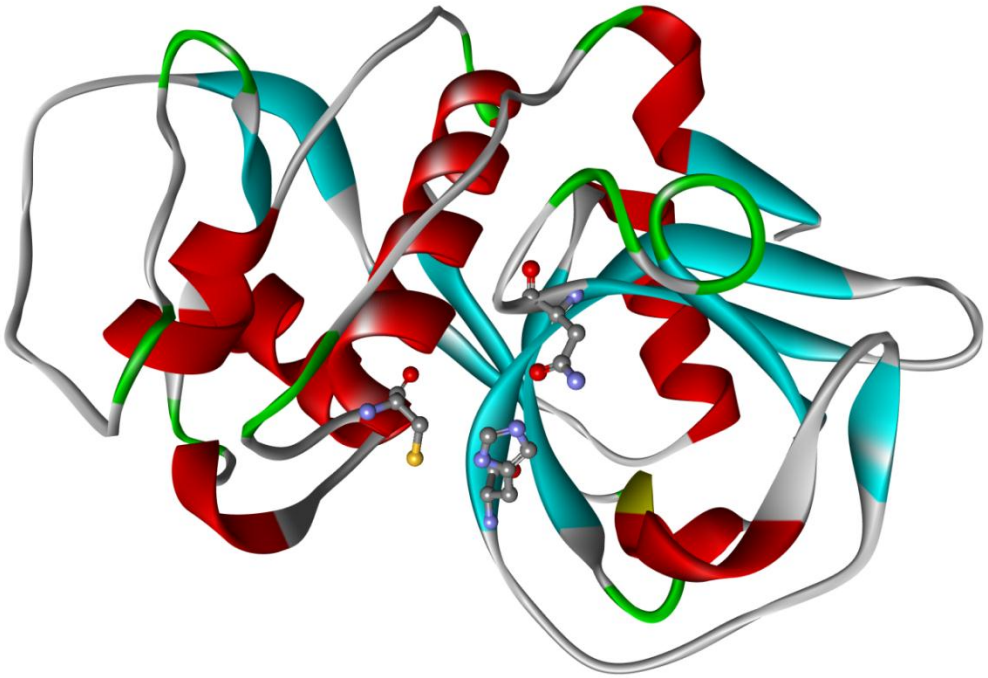


Figure 4. General Representation of Cysteine Proteases

*Relationship between Cysteine Proteases and their Substrates: The Active Site*

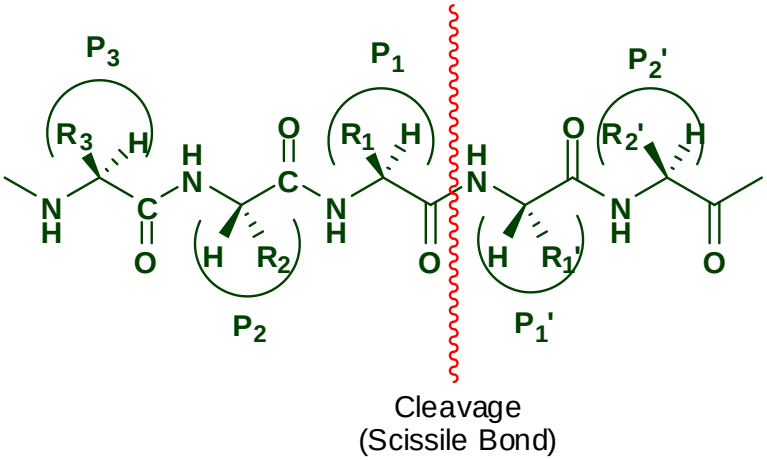
Standard nomenclature for cysteine proteases was established by Schechter and Berger forty-five years ago. They used papain to determine binding sites between the hydrolase and its substrates. This nomenclature is still used and accepted among cysteine proteases. The structures of the cathepsin family form several different subsites: $S_n$, $S_{n+1}$, …, $S_n'$ and $S_{n+1}'$. Nonprimed subsites refer to the N-terminal while primed nomenclature refers to the C-terminal of the substrate. Figure 5 is a two-dimensional representation of papain binding sites. In general, five substrate binding sites are distributed around the catalytic active site. S1, S3, and S2' are located at the left of the active site, while S2 and S1' can be found at the right side of the active site. S2, S1, and S1' are substrate binding sites but S2 is the only one that forms a pocket. The other substrate-binding sites are just considered areas. Four loops, two from each domain, are involved in the substrate-binding process: Gln19-Cys25, Arg59-Tyr67 (L-domain), and Leu134-His159, Asn175-Ser205 (R-domain). A disulfide bond made between Cys22 and Cys63 is also found.[92] Figure 5 shows the Schechter-Berger notation for binding sites in proteases[132,133] and the schematic presentation of substrate binding sites for cysteine proteases is shown in a view along the active site cleft.[97]

*Characteristics of Procathepsins*

Cysteine proteases are expressed as preproenzymes. The peptide signal, with a length average of approximately eighteen amino acids, binds to the signal recognition particle (SRP).

The propeptide region is a polypeptide that varies in length and differs from every cathepsin (38-251 amino acids). The importance of the propeptide can be summarized in three different functions.[120]

**A**




**B**

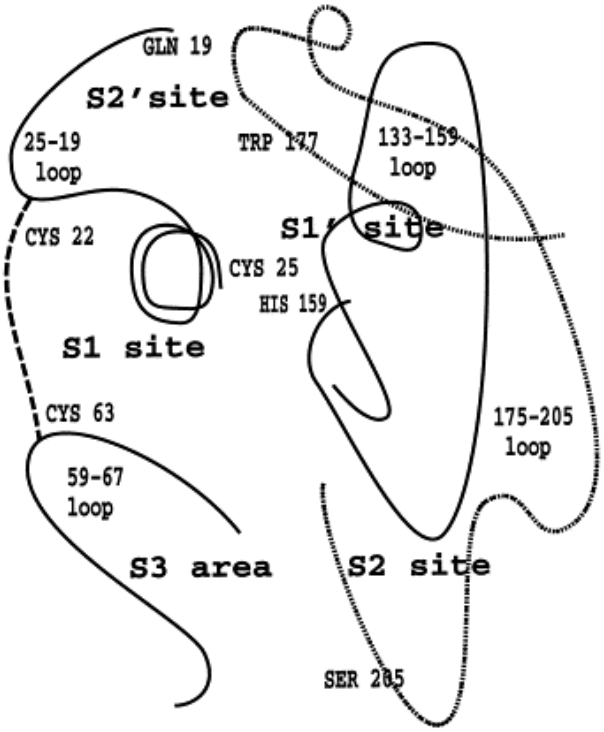



Figure 5. **A**. Schechter-Berger Notation for Binding Sites in Proteases.(adapted from Fersht).[132,133] **B**. Schematic Presentation of Substrate Binding Sites for Cysteine Proteases. (directly reproduced from reference Turk, page 100)

1. Protein fold. Some of these proregions function as a moiety for protein folding of the mature form.
2. Transporter. The proregion can help to transport the mature form of the enzyme to the lysosomes by using specific mannose-6-phosphate receptor pathway.[134]
3. Inhibitors. The proregion acts a potent reversible inhibitor of the enzyme and undesired activation of the cathepsins is prevented.[135]

### *Similarities of Proenzymes*

Four proenzymes of the cathepsin family have been elucidated (procathepsins B, L, K, and X).[136–139] All of them showed that their proregions bind to the mature forms of the proteases blocking the active-site binding site preventing substrate-enzyme interaction. Hydrophobic interactions, salt bridges, and hydrogen bonds are the most common types of interaction between proregions and their respective mature forms. Interestingly, procathepsin X shows a disulfide bridge between Cys25 and one cysteine residue coming from the proregion (Cys10).[139]

### *Mechanisms of Proenzyme Activation*

Activation of cysteine proteases is simply cleaving the proregion that is connected to the mature form. In general, there are two forms by which proenzymes can be activated:

1. Enzyme-catalyzed reaction. In this mechanism, other proteins are involved in the activation of cysteine proteases. That is the case of cathepsin D, an

aspartyl protease. Cathepsins C and X do require other cathepsins (L and S) for their respective activation.[140–142]

2. pH dependent. Some cysteine proteases can undergo an autocatalytic process under acidic conditions. The mechanisms involve several other mechanisms that are uni- and bimolecular. First, the procathepsin cleaves its propeptide under acid conditions. Then, a newly mature proenzyme intervenes and helps the activation process of inactive macromolecules, generating a chain reaction that helps the activation process.[143]

### *Similarities of Propeptides Amino Acid Sequences*

Cathepsin propeptides show little identicality. This can be seen in the discrepancies found in lengths. As previously stated, cathepsin X propeptide is a 38-residue polypeptide, while cathepsin F's proregion is 6 times as big as its cathepsin X analog (251 amino acid polypeptide). However, procathepsin L-like proregions average 100 residues and contain two conserved motifs: ERF/WNIN and GNFD.[121]

### *Natural Substrates of Cysteine Proteases*

Cysteine proteases are non-specific proteases with high promiscuity. This behavior explains the wide distribution and expression of these proteases among several tissues and organs. The extracellular matrix (ECM) is composed of numerous proteins that are degraded by cysteine proteases. Important components of the ECM are proteoglycan, collagen, elastin, fibronectin, laminin, osteocalcin, and osteonectin. Table 7 shows ECM components that have been shown to be susceptible to degradation by selected cysteine proteases.[144]

Table 7. Degradation of ECM Components by Cysteine Proteases (adapted from Parks)

| Cat | Proteoglycan | Collagen | Elastin | Fibronectin | Laminin | Osteocalcin | Osteonectin |
|---|---|---|---|---|---|---|---|
| B | + | + | - | + | + | + | + |
| S | - | + | + | + | - | + | - |
| V | - | - | + | - | - | - | - |
| H | - | - | - | - | - | + | |

*Inhibitors of Cysteine Proteases*

Cysteine proteases are tightly regulated by other macromolecules found in mammalian organisms. This section will present a summary of the most import natural inhibitors found in nature. The three most important classes of cysteine protease inhibitors are cystatins, thyropins, and serpins.[145–148] However, it is important to add that cathepsins' proregions are also potent inhibitors of the mature forms. Other macromolecules that show inhibition against cathepsins are: CRES (cystatin-related epididymal spermatogenic protein), testatin, cystatin-T, sialostatin L2, and sialostatin.[92] Specific inhibitors for cathepsins L and K will be discussed in chapters two and three.

*Synthetic Inhibitors of Cysteine Proteases*

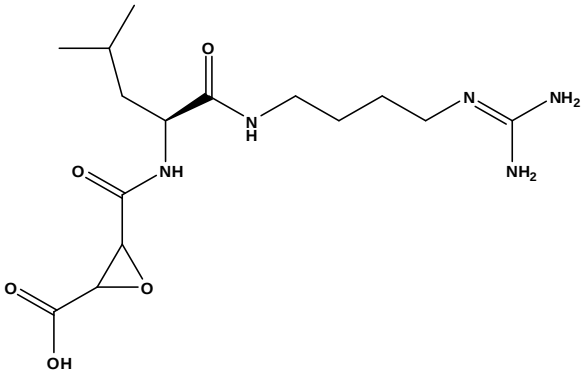
The importance of cathepsins in pathological processes, such as cancer, has led to numerous investigations with the aim to synthesize potent, yet selective, cathepsin inhibitors. The volume of generic and specific inhibitors for the cathepsins is quite large, and previous generations of inhibitors have been investigated.[149–152]

One of the most potent inhibitors of cysteine proteases is E-64, a natural compound isolated from *Aspergillus japonicas.* Chemically, E-64 is an epoxide that inhibits irreversibly numerous cysteine proteases including cathepsins B, L, K, and S.

Table 8 depicts the chemical structure of E-64 and selected $IC_{50}$ values against selected cathepsins. E-64 became the building block for analogs containing epoxides as their main moiety.[153] Other groups synthesized CA030 and CA074 as cathepsin B inhibitors.[154] CA030 is a peptidyl ethyl ester derivative of E-64, while CA074 is a methylated analog of E-64. Similar compounds have been developed for cathepsin S as well.[155]

Leupeptin, an aldehyde, is also a potent cysteine protease inhibitor. The disadvantage of this natural product is it is not specific. Leupeptin inhibits serine and cysteine proteases and therefore its use in cysteine proteases research is limited.[156] Nevertheless, other functional groups have also demonstrated generic and specific activity against cathepsins. The list includes: diacyl bis-hydrazides, diamino pyrrolidinone, aldehydes, cyclic ketones, nitriles, epoxysuccinyl derivatives, vinyl sulfones, diazomethanes, azepanone-based compounds, ally sulfones, and β-lactams.[106] Synthetic inhibitors for cathepsins L and K will be reviewed in chapters two and three.

Table 8. Chemical Structure of E-64 and Selected $IC_{50}$ Values against Cathepsins.[156]




| Cathepsin | $IC_{50}$ (nM) |
|---|---|
| L | 2.5 |
| S | 4.1 |
| K | 1.4 |

*Thiosemicarbazones as Potential Cancer Chemotherapeutic Agents*

Thiosemicarbazones, derivatives of semicarbazones, are a well known type of synthetic compounds. They have received special attention due to their inhibitory activity of biological processes. Research has been conducted to test their antibacterial, antiviral, antifungal and antineoplastic effects.[157] However recent studies have confirmed that thiosemicarbazones have the ability to inhibit human and parasitic cysteine proteases.[8–13]

*Importance of Cysteine Proteases in Medicine*

A general overview has been described throughout about this chapter about biological and chemical properties of human cysteine proteases. However, cysteine proteases are also found in other organisms, such as bacteria, protozoa, and plants (papain, the main cysteine protease, is a potent hydrolase found in papaya, a tropical fruit). Table 9 summarizes some of the roles of cysteine proteases in human pathological conditions.[105]

Table 9. Cysteine Proteases and Their Participation in Human-Related Diseases

| Disease | Description | Cathepsin | References |
|---|---|---|---|
| *Osteoporosis* | Bone-related disease. Characterized by excessive fragility in human bones. Osteoclasts, bone cells that are responsible for bone resorption, express high concentrations of cathepsin K. Cathepsin K is a powerful hydrolase that can catalyze several types of collagen and other proteins. Deficiency of cathepsin K produces a pycnodysostosis. | K | [158] |

Table 9. Cysteine Proteases and Their Participation in Human-Related Diseases (Continued)

| Disease | Description | Cathepsin | References |
|---|---|---|---|
| *Rheumatoid arthritis* | This is an inflammatory-related disease that is characterized by excessive degradation of type II collagen, one of the cartilage's main components. Overexpression of cathepsins L, B, S, and K contributes to this condition. | K, L, B, S | [159,160] |
| *Inflammatory and immune diseases* | Myasthenia gravis, inflammatory bowel disease, and asthma are classified under this category. *In vivo* research has concluded that selective inhibition of cathepsin S might be a potential therapeutical target in the combat against inflammatory diseases. Furthermore, overexpression of cathepsin V, a cathepsin-L like protease, has been found in patients with myasthenia gravis, an autoimmune disease. | S, V | [161,162] |
| *Atherosclerosis* | This condition is characterized by a reduction of blood flow in arteries. Cathepsins S and K are strongly linked to the degradation of elastin. Low levels of elastin reduce the elasticity of arteries and veins | S, K | [163,164] |
| *Periodontitis* | Strong evidence relates cathepsins L and B to this condition which is characterized by the inflammation of gums and destruction of connective tissues supporting the teeth. | L, B | [165] |

Table 9. Cysteine Proteases and Their Participation in Human-Related Diseases (Continued)

| Disease | Description | Cathepsin | References |
|---|---|---|---|
| *Respiratory system-related diseases* | Overexpression of numerous cathepsins (B, S, L, and H) catalyzes a chronic obstructive pulmonary disease (COPD). COPD is characterized by a chronic bronchitis and emphysema. | L, B, S | 166,167 |
| *Pancreatitis* | *In vivo* research with mice demonstrated that cathepsin B prematurely activates trypsinogen, one of the most potent serine hydrolases found in the pancreas. A disruption in the trypsinogen activation cycles produces acinar cell necrosis. | B | 168 |
| *Psoriasis* | Expression of procathepsins L and B have been found in patients with this condition. Cathepsin S is also overexpressed in human keratinocytes. | L, B | 169–171 |
| *Papillon-Lefevre and Haim-Munk syndromes* | Both syndromes are gene-related diseases. Characteristics of these conditions are teeth loss and hyperkeratosis of specific epithelial tissues. Research has demonstrated a cathepsin C mutation gene is responsible for these syndromes. | C | 172 |
| *Other diseases* | Patients with Gaucher disease overexpress cathepsins B, K, and S. Cathepsin H is strongly linked with patients with asthma. Finally, Sjogren's syndrome, a rare condition characterized by xerosotomia, is cathepsin B-dependent. | B, K, S, H | 173–175 |

*General Introduction of Enzyme Kinetics and Enzyme Inhibition*

Enzymes are proteins that catalyze biochemical reactions in organisms. They utilize compounds found in nature and convert them into different products that can be used for other purposes.

Enzymatic reactions are extremely complex; however, models have been proposed to explain these processes. Nevertheless, the Michaelis-Menten model is a good approximation for many of these biochemical systems.[176]

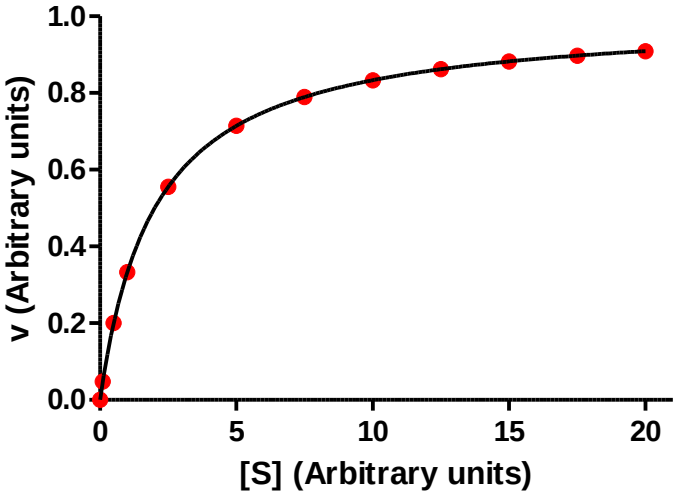
The model establishes a hyperbolic relationship between substrate concentration and enzymatic activity. Scheme 1 describes chemical reactions between an enzyme and its substrate.

First, a reversible reaction occurs between both species (E= enzyme; S= substrate), creating an intermediate, the Michael-Menten complex (ES). The seconds step of the process is a very slow irreversible process. The Michaelis-Menten complex is broken down into the final product and the release of the enzyme.

$$E + S \rightleftharpoons E{\cdot}S \rightleftharpoons E + P$$

Scheme 1. Chemical Reactions Between Enzyme and Substrate

This simple model is the basis for the derivation of the Michaelis-Menten equation (Equation 1.1), where $K_M$: Michaelis-Menten constant, S: substrate concentration, $V_{max}$: maximum enzymatic activity, and *v:* velocity. The Michaelis-Menten constant, $K_{M,}$ represents the amount of substrate that is necessary to achieve half maximal enzyme's activity. The value also includes the dissociation constant. The lower the value, the tighter the interaction is between both species (substrate and enzyme). A graphical representation of simulated data is depicted in Figure 6.




$$v = \frac{v_{MAX}\,[S]}{K_M + [S]} \quad (1.1)$$

Figure 6. Michaelis-Menten Equation. Simulated Data Give Calculated $K_M$ and $V_{max}$ Values of 2 and 1 Arbitrary Units, Respectively

*Enzyme Inhibition. Basic Concepts*

An enzyme inhibitor is a natural or synthetic compound that has the ability to reduce the catalytic activity of an enzyme. In general, inhibitors can mimic the structure of the intermediates that can be formed during the biochemical reactions (Enzyme-Substrate). Several examples of transition-state inhibitors can be found in literature and some of them have clinical applications. Hence, the importance of understanding the mechanism of action and the reaction optimal conditions are required to study the mode of inhibition. It is important for the design and study of possible candidates to function as inhibitors. Assay optimization is perhaps the most critical step in drug discovery and development. Parameters such as temperature, pH, chelating metals in solution, buffer components, and viscosity can determine the success or failure of any enzymatic assay.[177]

Several inhibitors work in a reversible fashion. Reversible inhibition can occur when the enzymatic activity can be recovered through the time of reaction. On the contrary, irreversible inhibitors interact with the target enzyme by inactivating it, and the

protein is no longer capable of catalyzing a reaction. Extensive research aims at designing reversible and irreversible inhibitors. Irreversible inhibitors are the best candidates for infectious diseases (bacterial, viral, fungal, etc.). However, this is not applicable when the target enzyme is found in the human body and participates in biological processes. Inactivation of a human protein can be an undesired consequence of a medical treatment, and patients could be in danger of suffering subsequent conditions. For example, irreversible inactivation of human cathepsin K, a potential target in osteoporosis, could lead to osteopetrosis. Symptoms of osteopetrosis include harder and denser bones, which lead to bone fracture.[178]

Reversible inhibitors can be classified based on their mechanism of action. In general, there are four cases: competitive, non-competitive, uncompetitive, and mixed reversible inhibition. Viagra, nevirapine, and camptothecin are classical examples of these modes of inhibitions.[179–181] Thus, inhibition mode studies are crucial and play a major role in drug discovery and development. Quantitative comparisons can be utilized to select the best candidates for future pre-clinical and clinical trials. Kinetics parameters such as the inhibition constant ($K_I$), $IC_{50}$ value and binding Gibbs free energy ($\Delta G^0$) of that occur when the inhibitor binds to the target protein, can quantify the structural changes of the system (enzyme, inhibitor, or both) that undergo during the reaction. An important feature in inhibitor design is its selectivity. It is not uncommon to find a compound that inhibits a family of proteins with similar characteristics (usually structure or mechanism-based). Therefore, inhibition assays are performed with different enzymes in order to investigate a possible selectivity or a generic application. *Trans*-

Epoxysuccinyl-L-leucylamido(4-guanidino)butane, or better known as E-64, is a potent irreversible cysteine protease inhibitor widely used in research.[182]

The discovery, optimization, and utilization of various compounds that can be potential inhibitor agents are the main focus of our research. The creation of a library and the analysis of their structure-activity relationship (SAR) can be extremely helpful in order to achieve these goals. SARs can help to optimize future and advanced pre-clinical and *in vivo* studies. They also help to identify possible selectivity and lead to the design of more analogs that could be potent inhibitors.[183–185]

### *Modes of Inhibition*

An inhibitor is considered as competitive when it binds to the catalytic active site of the enzyme and competes with the substrate for that site. Noncompetitive inhibition occurs when the compound binds to an established enzyme-substrate inhibitor or the free enzyme. Uncompetitive inhibition is usually found when the compound binds to the enzyme-substrate complex. Finally, mixed inhibition occurs when the inhibitor binds in more than one fashion.

Table 10. Classical Examples of Reversible Inhibitors

| Mode of Inhibition | Compound | Medical Condition | Enzyme Target |
|---|---|---|---|
| Competitive | Viagra (Sildenafil Citrate) | Erectile Dysfunction | Phosphodiesterase Type-5 Monophosphate |
| Non-competitive | Nevirapine (nucleoside) | HIV | Reverse Transcriptase |
| Uncompetitive | Camptothecin (Quinoline Alkaloid) | Cancer | Topoisomerase |

*The Definition of $IC_{50}$ and $K_I$ and Their Importance in Inhibition Studies*

There are two important concepts in enzyme inhibition that are widely used when testing potential inhibitors with several targets: $IC_{50}$ and $K_I$.

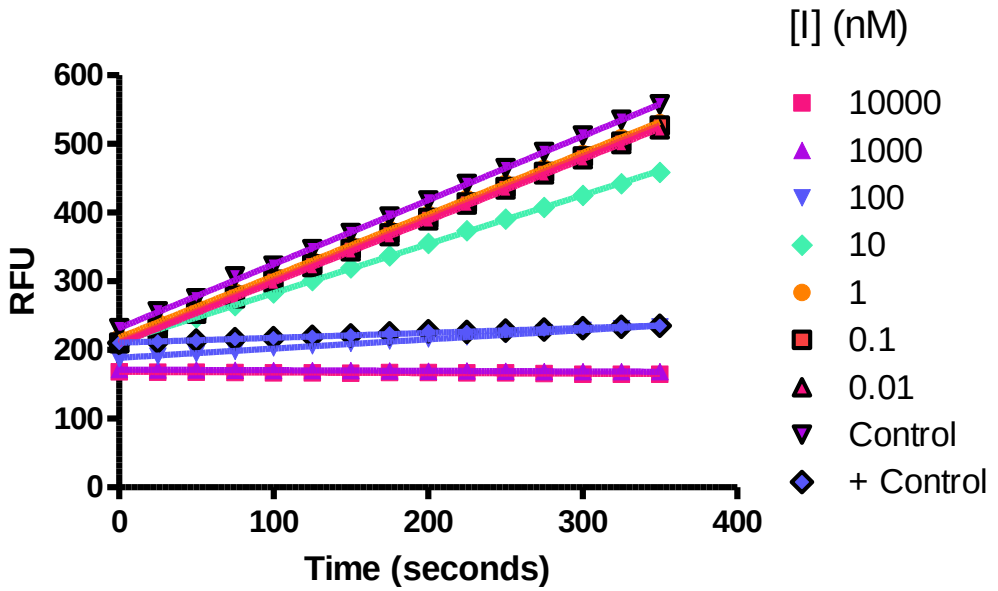
The $IC_{50}$ value or half maximum inhibitory concentration is the concentration of a specific compound (inhibitor) that is required to inhibit the enzymatic activity of a target by fifty percent. Thus, $IC_{50}$ values are concentration dependent and equation 1.2 describes the mathematical relationship between both variables. Y represents the inhibited activity (compared to control), X is log([inhibitor]) in *M*, $v_{min}$ and $v_{max}$ represent the velocities of the enzyme when it is pre-incubated with the highest and lowest inhibitor concentrations, respectively. Finally, the Hillslope value is the slope of the sigmoidal curve. The data fit into a sigmoidal curve and can be seen in Figure 7.

The physiological significance of this parameter resides in the search of compounds that at relatively lower concentrations can reduce the activity of a specific target (enzyme).

The Hillslope usually ranges between -1 and 1. Unity is considered an ideal inhibitor behavior; however, there are cases when this value is not close to unity, meaning there are other factors affecting the behavior of the compound.

The determination of $IC_{50}$ for any enzyme is method dependent. This parameter is susceptible to laboratory conditions such as temperature, buffer (type, concentration), pH, enzyme concentration, preincubation time, solvent effect, substrate inhibition, etc. Thus, the importance of establishing a validated, consistent protocol is vital for comparison and reproducibility purposes. This is also a determinant factor when comparing $IC_{50}$ values for other systems.[186]

**A**




**B**

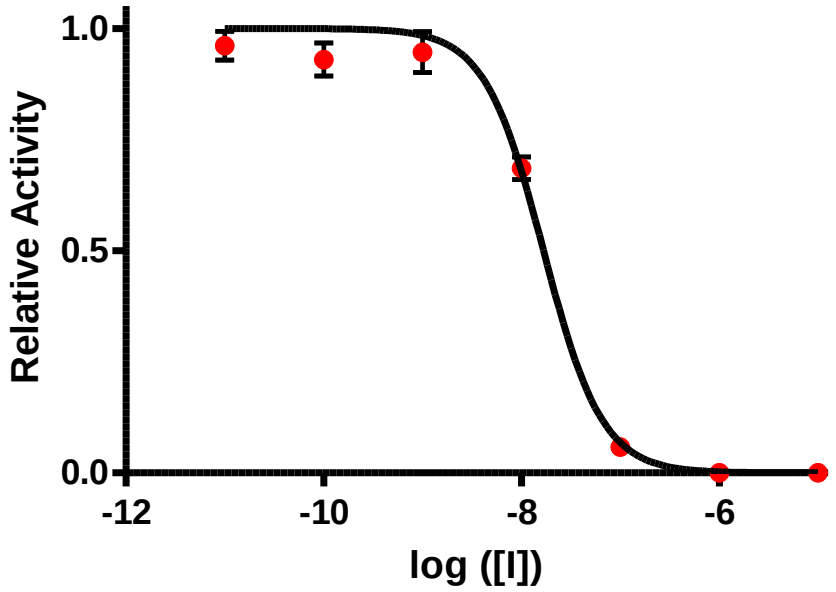



Figure 7. **A**. Inhibition of Cathepsin K Activity by an Inhibitor ($0 \leq [I] \leq 10$ μM). **B**. Calculation of $IC_{50}$ values. Real Data were Fitted into Equation 1.2, $v_{min}$ and $v_{max}$ were Given Constraints Set to 0 and 1, Respectively. Results were: $IC_{50}$: 16.7 nM, Hillslope: -1.14 and $r^2$: 0.9863

The inhibition constant ($K_I$, Equation 1.3) is another concept that correlates the effect of any inhibitor with a given enzyme (the lower the value, the better the inhibitory activity of any compound). In general, $K_I$ is the ratio of $k_{off}$ and $k_{on}$ and, the rate constants for the inhibition offset and onset of both species (enzyme and inhibitor), respectively.

$$Y = \frac{v_{MIN} + (v_{MAX} - v_{MIN})}{1 + 10^{(\log(IC_{50} - X) * Hillslope)}} \tag{1.2}$$

$$\mathrm{E} + \mathrm{I} \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} \mathrm{EI} \qquad K_I = \frac{k_{off}}{k_{on}} \tag{1.3}$$

There are several approaches to calculate this parameter. Equation 2.3 offers the simplest form to calculate $K_I$, assuming both $k_{on}$ and $k_{off}$ are already known. Unfortunately, the determination of these rate constants is not always a simple process, and several authors have come up with alternative models aiming the determination of the inhibition constant. It is also important to understand the limitation of each approach in order to guarantee the most exact value for $K_I$. The exact calculation of $K_I$ can be done by using the transient approach. Data sets are fitted into a system of differential equations (usually five or six equations). However, this approach requires the mode of inhibition being established or previously known; and a complex mathematical program.

A simple, yet well accepted approach was determined by Cheng and Prusoff in 1973 when they derived a system of equations that correlate $IC_{50}$ values and $K_I$ values.[187]

Three equations describe these relationships when a reversible inhibitor works as a competitive (Equation 1.4), noncompetitive (Equation 1.5), and uncompetitive inhibitor (Equation 1.6). The simplicity of these equations attracts the majority of researchers.

However, the mode of inhibition needs to be established when choosing the appropriate equation.

$$IC_{50} = K_I\left(1 + \frac{[S]}{K_M}\right) \tag{1.4}$$

$$IC_{50} = \frac{[S] + K_M}{\frac{K_M}{K_I} + \frac{[S]}{\propto K_I}} \tag{1.5}$$

$$IC_{50} = \propto K_I\left(1 + \frac{K_M}{[S]}\right) \tag{1.6}$$

*Slow Binding and Tight Inhibition. Determination of Mode of Inhibition*

A compound is considered to be a slow binding inhibitor when the inhibition process occurs as a response of time. This means the potential inhibitor binds to an enzyme at longer times and slowly releases throughout the course of the reaction. A visual form to determine if the compound acts as a slow binding inhibitor occurs when comparing treated and nontreated reaction progress curves. A non-linear behavior is the key characteristic of this form of inhibition. Several examples of slow binding inhibitors can be found in literature. Figure 8 shows an example of reaction progress curves of an enzyme when it is inhibited by a slow binding inhibitor.

The treated progress curve can be fitted into equation 1.7 by fixing the concentrations of substrate, inhibitor, and enzyme. The release of product is dependent of the time reaction. Fitting the data into this equation solves for three different parameters ($v_o$ , $v_s$, and $k_{obs}$); the rates $v_o$ and $v_s$ are the initial and steady-state velocities, t is the time, and $k_{obs}$ the apparent first-order rate constant for the steady state that is

formed between all the species present in the reaction (enzyme, inhibitor, enzyme-inhibitor complex, and substrate).[188]

$$P = v_s t \frac{(v_o - v_s)}{k_{obs}} (1 - e^{-k_{obs} t}) \tag{1.7}$$

The vast majority of inhibitors reported in literature have been found to work as competitive inhibitors. Nevertheless, a considerable number of compounds are noncompetitive or uncompetitive inhibitors. The mode of inhibition can be determined by using mathematical equations such as Michaelis-Menten and Lineweaver-Burk and using experimental data.[176,189]

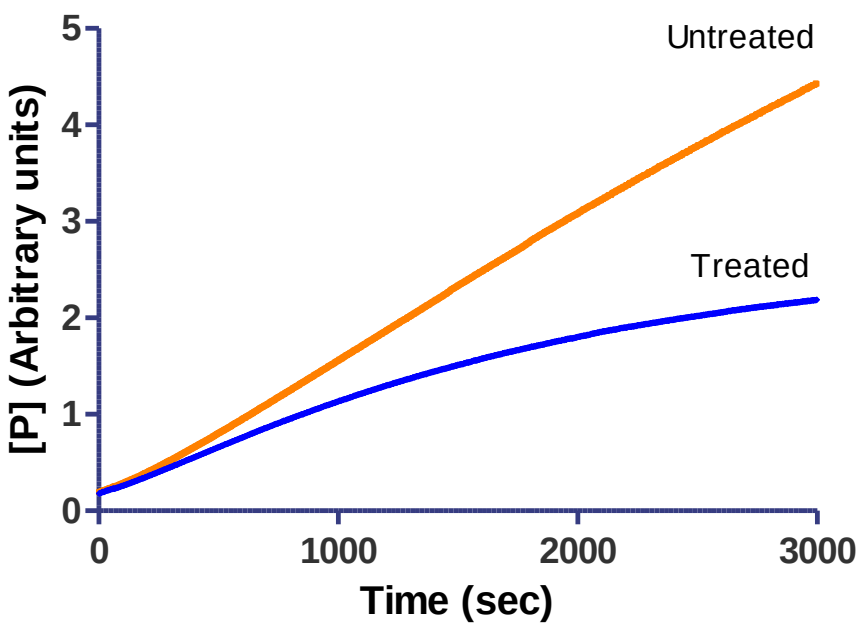



Figure 8 Typical Reaction Progress Curves of Untreated (orange) and Treated Sample with a Slow Binding Inhibitor (blue)

Alternatively, the mode of inhibition slow-binding inhibitors can be determined by using equation 1.7. The $k_{obs}$ values are substrate dependent as seen in equations 1.8,

1.9 and 1.10. Calculated rate constants can be plotted versus substrate concentration if the inhibitor concentration is constant (Figure 9).[177]

For competitive inhibitors, rate constants ($k_{obs}$) decrease hyperbolically by increasing substrate concentration (See equation 1.8). In the case of uncompetitive inhibition (equation 1.9), rate constants are substrate concentration independent. Finally, the rate constants for non-competitive inhibitors will increase hyperbolically with higher substrate concentrations (equation 1.10).

$$k_{obs} = \frac{k}{1 + \frac{[S]}{K_M}} \tag{1.8}$$

$$k_{obs} = k \tag{1.9}$$

$$k_{obs} = \frac{k}{1 + \frac{K_M}{[S]}} \tag{1.10}$$

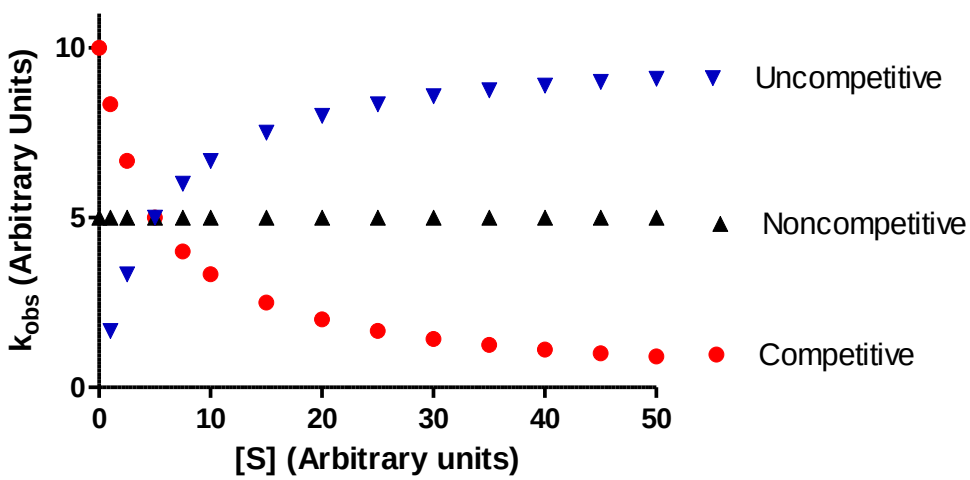



Figure 9. Determination of Mode of Inhibition for Slow-Binding Inhibitors. See Equations 1.7-1.10. Values are in Arbitrary Units. $K_M$ : 5; $k$: 10 (for competitive and uncompetitive inhibition; $k$: 5 (for noncompetitive inhibition)

*Relationship between Enzyme Concentration and $K_I$: The Morrison's Equation*

Perhaps, the best mathematical method to evaluate inhibition constants for slow, tight-binding inhibitors is the Morrison's quadratic equation (equation 1.11). The equation expresses the $K_I$ value in terms of the enzyme species concentrations. Thus, the inhibition constant is dependent of free enzyme and enzyme bound to the inhibitor. Mathematically, the Morrison's mathematical model is a quadratic equation; that is, it has two plausible solutions. Obviously, only one solution is physically possible because $K_I > 0$. Nevertheless, Morrison's equations solves the apparent inhibition constant $K_I^{app}$. Equation 1.11 solves the true inhibition constant $K_I$ by solving Equation 1.12 with a relationship between $K_I^{app}$, $K_I$, [S], and $K_M$.

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \tag{1.11}$$

$$K_I^{app} = K_I(1 + \left(\frac{[S]}{K_M}\right) \tag{1.12}$$

Equation 1.11 presents the relation between $K_I^{app}$, $K_I$, [S], and $K_M$ and inhibitor. $K_I^{app}$, $K_I$, [S], and $K_M$ Inhibited ($v_i$) and uninhibited ($v_o$) can be experimentally determined. Michaelis-Menten, and substrate concentration should be constants. Thus, the apparent inhibition constant can be then determined by solving the dependence of the fractional activity (i.e. $v_i$/ $v_o$) versus inhibitor concentration. However, other considerations, such as optimization and limitation of the model have also been studied.[190]

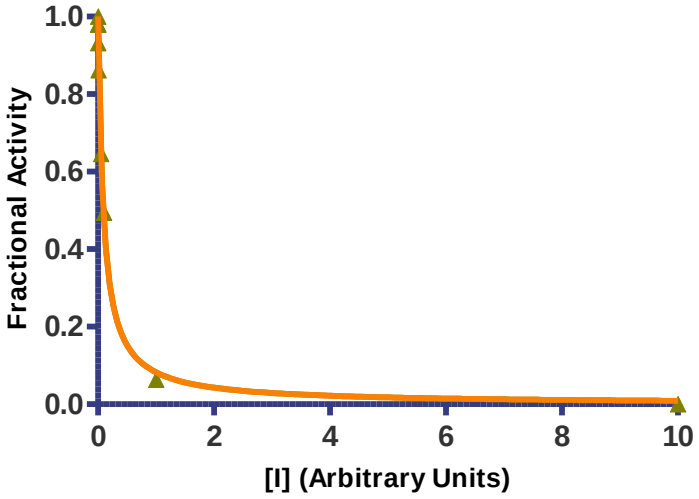



Figure 10. Graphical Representation of the Morrison's Equation. Real data was Used to Calculate $K_I^{app}$ Conditions (in μM): [S]:50; $K_M$: 1; [$E_T$]: 0.001; $v_o$: 1; $r^2$: 0.9949

# CHAPTER TWO

## Evaluation of Thiosemicarbazones as Cathepsin L Inhibitors

### *Nomenclature, Classification and Historical Background*

Human cathepsin L (EC 3.4.22.15) is a hydrolase and belongs to the sub-class of peptidases, enzymes that act on peptidic bonds.[191] Cathepsin L is also classified as an endopeptidase, that is, a hydrolase that only acts on peptidic bonds located within amino acid sequences.[192]

The official name of the protein is cathepsin L (CL). CL is also known as cathepsin L1.[193] One of the first literature references about cathepsin L is from 1977 when its purification from rat liver, characterization and catalytic activity were reported.[194]

### *Localization and Biological Roles of Cathepsin L*

Cathepsin L is a ubiquitous hydrolase found in mammals. Cloning experiments revealed that cathepsin L mRNA is expressed among several cell lines in similar quantities. Human cathepsin L was found to be encoded in the chromosome 9q21-22.[195]

The extraordinary catalytic activity of cathepsin L is quite fascinating because of its participation in numerous yet diverse biological processes. Cathepsin L is mostly considered a lysosomal protease and its distribution in human tissues and organs is quite extensive. Skin, liver, thyroid glands, and kidneys are organs where considerable concentrations of this protease have been detected.[196–199] Animal studies showed considerable amounts of cathepsin L in intervertebral disc tissue, Type A (macrophages)

and Type B (fibroblast) cells that form the synovial lining layer.[200,201] However, recent studies have shown that cathepsin L is also expressed in other cellular organelles. Active cathepsin L was detected in the nucleus of NIH3T3 cells during G1-S transition cell phase, controlling cell cycle progression.[202] Similarly, Sullivan and coworkers detected cathepsin L activity in colorectal cancer cell nuclei.[203] Cathepsin L is also found as a secreted protease in the secretory vesicles where it participates in the activation of proenkephalin confirming that cathepsin L is also selectively secreted under specific conditions.[204]

Serum, cytoplasmic and nuclear proteins, are also susceptible to be degraded by this protease.[203,205,206] Furthermore, the versatility of cathepsin L has been found in multiple processes: bone resorption, sperm maturation, intestinal neoplasia, and processing of neuropeptides into neurotransmitters and some hormones.[118,207,208]

Cathepsin L hydrolyzes thyroglobulin in the thyroid glands, releasing two important subproducts (thyroxine and thyronine). Animal studies also revealed this protease is highly involved in the processing of peptide epitopes in thymic epithelial cells. It has also been implicated in other immunogenic processes. For example, interferon-γ enhances the activity of cathepsin L in human monocytes.[209,210]

Genetically manipulated mice (cathepsin L expression was suppressed) showed these animals had problems with hair loss, epidermal hyperplasia, acanthosis, and hyperkeratinosis. Other studies support the role of cathepsin L in the differentiation of keratinocytes and its importance in hair pigmentation and the biology of hair cells.

Reproduction cycles (spermatogenesis) also require the activity of this protease. For example, Sertoli cells in the male reproductive system express considerable amounts

of cathepsin L. Cathepsin L-defficient mice showed dysfunctional seminiferous tubules and reduced spermatogenesis levels.[209–213]

Female reproductive systems require the proteolytic activity of cathepsin L. Cathepsin L has been linked to oocyte generation (oogenesis). Other studies confirmed higher concentrations of cathepsin L mRNA in placenta, specifically during implantation and before labor. Mice studies also confirmed higher amounts of cathepsin L in the visceral yolk sack.[214,215]

Finally, cathepsin L plays a crucial role in angiogenesis. The proteolytic degradation of the components of the extracellular matrix (ECM) is increased by the activity of cathepsin L.[216]

### *Natural Substrates of Cathepsin L*

Cathepsin L is one of the most potent cysteine cathepsins. Its ubiquitous distribution explains its importance. However, there is little information regarding the potency of this hydrolase in physiological conditions. Research has shown its *in vitro* efficacy in recent reports.

*In vitro* studies have proved extracellular matrix components, are optimal cathepsin L natural substrates under physiological conditions. Cathepsin L is capable of degrading proteoglycans, aggrecan, elastin, fibronectin, laminin, and osteocalcin.[217–221]

Collagen deserves special attention regarding cathepsin L activity. The collagenase activity has been well demonstrated by using several types of this macromolecule. Cathepsin L collagenase activity has been demonstrated by using types II (cartilage, vitreous humor), IX (cartilage), XI (cartilage), I (skin, tendon, bone), and IV (basal lamina).[222–224]

*Catalytic Activity of Cathepsin L and Substrate Specificity*

Cathepsin L is considered one of the most potent cysteine proteases found in mammalian organisms. As previously stated, the protease is able to process a wide range of proteins, from small peptides to high molecular weight proteins, such as type I collagen (Molecular Weight: ~ 300 kDa). Recent reports have characterized the substrate specificity of cathepsin L. The protease cleaves Arg-Ser peptidic bonds when synthetic polypeptides were used to test the hydrolase activity specificity. Substitution of the serine residue by a hydrophobic residue (Phe) reduces the hydrolytic activity of cathepsin L.

It was also demonstrated that cathepsin L is able to cleave more than one peptidic bond. The protease was able to hydrolyze both Phe-Ser and Ser-Ser peptidic bonds in the synthetic substrate Abz-KL**FSS**KQEDDNp, while Abz-KL**LSS**KQ-ED-DNp was hydrolyzed at two different sites (Leu-Ser and Ser-Ser). Basic amino acids residues at the P1 position (Arg and Lys) are the preferred targets for cathepsin L for peptidic cleavage. In constrast, substrates with acidic residues (Glu, Asp) and polar residues (Asn, Thr, Asn) at the P2 position are poor cathepsin L substrates. Proline-containing peptides are the worst substrates for the protease.[109]

Other studies confirmed the preference of cathepsin L for substrates containing hydrophobic amino acids at the P2 position, positive-charge (Lys and Arg) at the P1 position and broader specificities at P3 and P4 positions.[225]

*Structure of Cathepsin L*

Cathepsin L contains 217 residues per monomer with a molecular weight between 25 and 29 kDa.[193] Cathepsin L contains one heavy chain (MW: 20-24 kDa; 175 residues,

Residues 1-175) and one light chain (MW: 5 kDa; 42 residues, Residues 176-215).[226,227] The amino acid sequence of the heavy chain is shown in Figure 11. Peptide sequences also confirmed the high sequence identity of this enzyme when it is compared to papain up to 41%.[123,228]

| | | | | | |
|---|---|---|---|---|---|
| 1 | APRSVDWREK | GYVTPVKNQG | QCGS**C**WAFSA | TGALEGQMFR | KTGRLISLSE |
| 51 | QNLVDCSGPQ | GNEGCNGGL | DYAFQYVQDN | GGLDSEESYP | YEATEESCKY |
| 101 | NPKYSVANDT | GFVDIPKQEK | ALMKAVATVG | PISVAIDAGH | ESFLFYKEGI |
| 151 | YFEPDCSSED | MD**H**GVLVVGY | GFEST**E**SDNN | KYWLVK**N**SWG | EEWGMGGYVK |
| 201 | MAKDRRNHCG | IASAASYPTV | | | |

Figure 11. Amino acid Sequence of Cathepsin L. Legend: Blue, Heavy Chain; Red, Light Chain; Catalytic Residues, Green.[229,230]

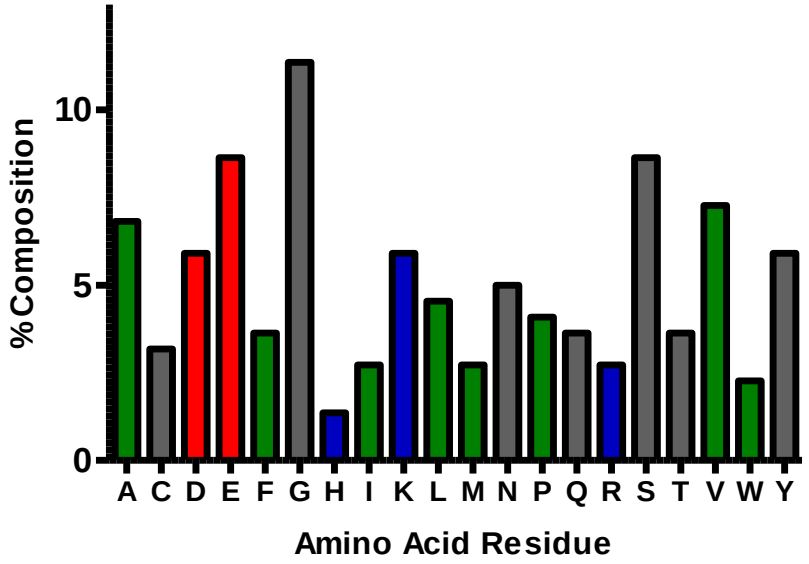



Figure 12. Composition of Human Cathepsin L. Legend= Blue: Asic Residues; Red: Acidic Residues; Green: Hydrophobic (Non Polar) Residues; Gray: Neutral, Polar Residues

The three most abundant residues are glycine, glutamate and serine (25, 19, and 19 residues, respectively). On the other hand, the least abundant residues are histidine (three residues) and tryptophan (five residues). Hydrophobic residues account for almost

one third of the total composition of cathepsin L. Approximately, 41 percent of the composition is entirely made of polar, uncharged residues.

Table 11. Amino Acid Residue Distribution of Cathepsin L

| AA | # | % | AA | # | % | AA | # | % |
|---|---|---|---|---|---|---|---|---|
| Ala | 15 | 6.82 | Ile | 6 | 2.73 | Arg | 6 | 2.73 |
| Cys | 7 | 3.18 | Lys | 13 | 5.91 | Ser | 19 | 8.64 |
| Asp | 13 | 5.91 | Leu | 10 | 4.55 | Thr | 8 | 3.64 |
| Glu | 19 | 8.64 | Met | 6 | 2.73 | Val | 16 | 7.27 |
| Phe | 8 | 3.64 | Asn | 11 | 5.00 | Trp | 5 | 2.27 |
| Gly | 25 | 11.36 | Pro | 9 | 4.09 | Tyr | 13 | 5.91 |
| His | 3 | 1.36 | Gln | 8 | 3.64 | Total | 220 | 100.0 |

The secondary structure reveals that ten α-helices constitute thirty percent of the total amino acid sequence. However, only one fifth of the sequence forms beta sheets distributed in fifteen strands.[229,230]

*Crystal Structure of Cathepsin L*

One of the first reports of the analysis of cathepsin L by X-ray crystal structures revealed important characteristics of the structure.[123] The S´ regions of cathepsin L are described as shallow depressions and show numerous similarities when compared to the S´ regions of papain. The depression is limited by a flat bottom that involves interaction of the Trp189 side chain, Gln21, Gly20, Asn18, Glu192, Trp193 and Leu144. The oxyanion hole of cathepsin L is made by three residue side chains (Gln16, Trp189, and His163) and the Cys25 backbone. The $S_2$ pocket of cathepsin L is remarkably hydrophobic, a characteristic that is found among several cysteine proteases. The bottom $S_2$ pocket of cathepsin L consists of Met70, Ala135, Leu69, Met161 and Gly164, which are hydrophobic. Detailed inspection of that $S_2$ site reveals that Asp162, Met161, Asp160,

and Ala214 build one of the walls that encapsulate the $S_2$ pocket. Furthermore, this wall is the reason the $S_2$ subsite is a narrow pocket. Met70 is the responsible for the $S_2$ pocket shallow characteristic. The $S_3$ subsite of the pocket can be found on one of the sides of the left domain of the cleft. The $S_3$ residues Gly67 and Gly68 were found to be located in the middle of the former subsite. They both form an ample substructure that is surrounded by Asn66, Glu63, Leu69 and the carbonyl oxygen of Gly61. Figure 13 shows the Crystal Structure of Human Cathepsin L. Three active residues (Cys25, His159, and Asn175) are shown in ball and stick form. The structure is colored according to the following: α-Helices are red, β-sheets are cyan, turns are green, and coils are white (PDB ID: 1ICF).[228]

*General Considerations of Procathepsin L*

Cysteine proteases are often produced in mammalian organism as precursors, also known as procathepsins. Cathepsin L is expressed as a proenzyme in eukaryotic organisms. The preproform of cathepsin L, named preprocathepsin L (MW ~ 39 kDa) is inactive and consists of three major polypeptide regions: the signal peptide, consisting of 17 residues; the proregion made of 96 amino acid residues; and the mature form (217 amino acid residues).[137,231] The proregion is located between the signal sequence and the N-terminus of the mature form. The proregion has regulatory activities. It controls the activation and activity of cathepsin L, so it acts a natural peptidic inhibitor of cathepsin L. However, it has also been suggested that both activation and processing of cathepsin L are two related, yet different processes.[232,233] The proenzyme is found to be more stable than the mature form at higher pH ranges, because the proregion helps to stabilize the macromolecule in basic environments.[142] The crystal structure of procathepsin L has also

been resolved using cloning techniques. The proregion is located over the mature portion of cathepsin L.[137]

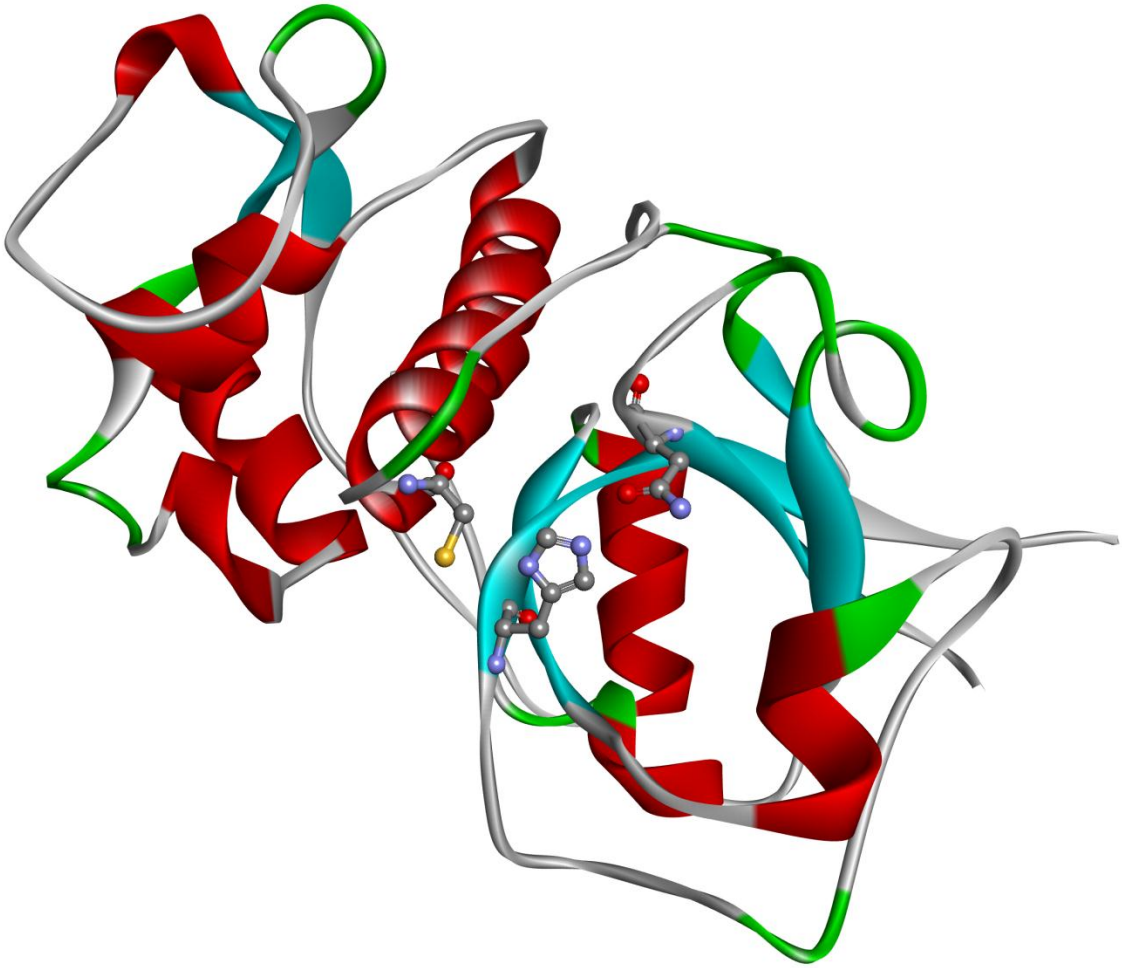


Figure 13. Crystal Structure of Human Cathepsin L (PDB ID: 1ICF)

In general, the proregion is a 75 residue globular-shaped segment and an elongated 21 residue segment. There are two major interactions sites between the proregion and cathepsin L. The first one is found by the side of the catalytic active site, and the second interaction site is with one of the connecting loops of the mature form. These findings explain the inhibitory function of the proregion against the mature form.

A hydrophobic residue (Phe71) and a small peptide region (Met75-Glu79) are the specific connecting points between the proregion and cathepsin L active site by forming numerous hydrogen bonds. The hydrophobic globular segment contains three α-helices and corresponding connecting loops. Furthermore, activation of procathepsin L can be a pH dependent mechanism. Acidic conditions promote the activation of the protease by an auto activation process. Activation of procathepsin L generates two major macromolecules: the proregion and mature cathepsin L with molecular weights of 10 and 29 kDa respectively. Further investigation has revealed the proregion of the procathepsin L can be cleaved at two different sites.[234] Figure 14 shows the crystal structure of human procathepsin L. Three active residues (Cys25, His159 and Asn175) are shown in ball and stick form. The structure is colored according to the following: α-Helices are red, β-sheets are cyan, turns are green, and coils are white, and the propetide is dark green (96 amino acid residues). PDB ID: 1CS8[137]

## *Inhibitors of Cathepsin L*

As proteases found in mammalian organisms, cathepsin L has natural inhibitors that regulate its catalytic activity and degradation. Also, numerous synthetic inhibitors have been developed in recent years to inactivate its ability due to its importance in medicine. To date, none of the possible inhibitors developed have been used in clinical trials.

### *Natural Inhibitors of Cathepsin L*

One of the first cathepsin L natural inhibitors that was discovered is α-macroglobulin.[235] This protein is extremely large (MW: 725 kDa) and due to its size, it is able to encapsulate cathepsin L within its structure, inactivating the latter one. The

cystatins are a well known class of protein inhibitors of cathepsin L.[236] These natural inhibitors might act as cysteine protease regulators due to the high catalytic activity of the cathepsins. Generally, the superfamily of cystatins is divided into three major groups: stefins, cystatins, and kininogens.[237] The stefins are found in epithelial cells and numerous cells and tissues.[238] An interesting characteristic of this subfamily is their lack of disulfide bonds and carbohydrates.[239] The inhibitory potency of stefins can be seen in reported values for stefins, which are in the picomolar range. The second family of cysteine protease inhibitors is the cystatins.

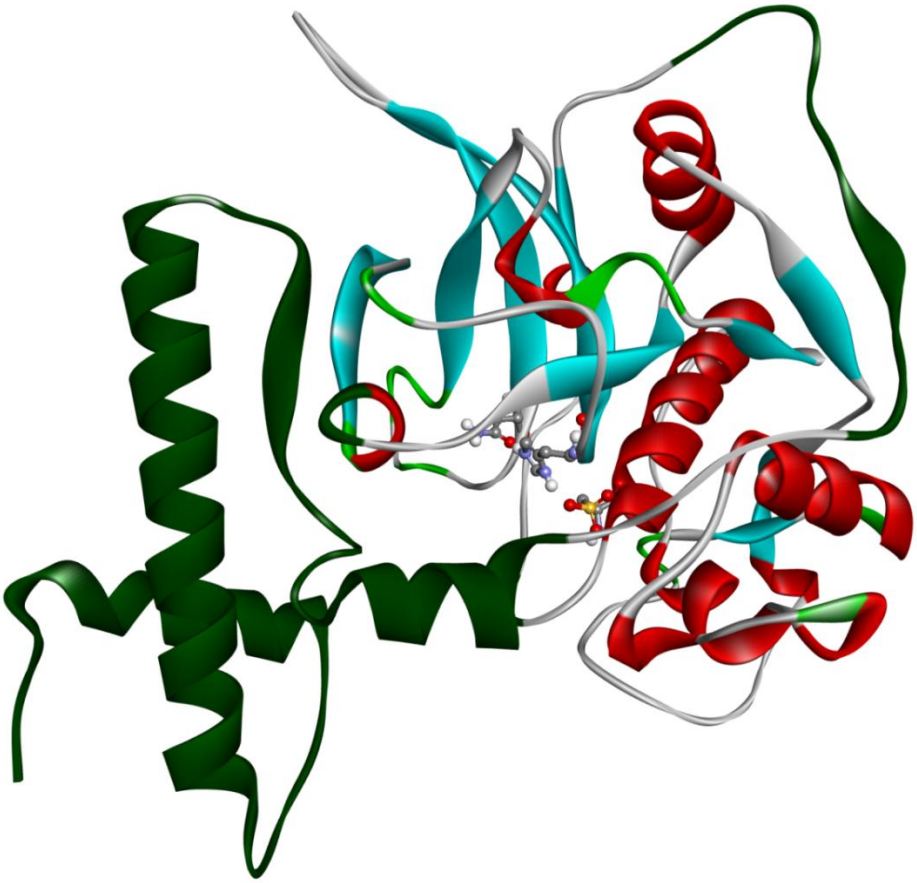


Figure 14. Crystal Structure of Human Procathepsin L (PDB ID: 1CS8)

The cystatins have similar amino acid sequences to the egg-white chicken protein found more than 30 years ago.[146] The characteristic feature of these proteins is the presence of two disulfide bond in close proximity to the carboxyl-terminus. Cystatins can be found in several fluids of mammalian organisms in a wide range of concentrations.[240] Cystatin C is one of the most abundant proteins found in seminal plasma and cerebrospinal fluid, and traces have been found in plasma, saliva, and urine. Lastly, the third family of inhibitors of proteases that belong to the cystatin superfamily is the kininogen family. These proteins have regulatory activities and function as precursors of kinins, an important class of proteins that act on vasodilation and contraction of smooth muscles tissues processes.[148]

However, other studies have confirmed there are other two different types of macromolecules that act as natural cathepsin L inhibitors. First, the thyropins, specifically the MHC class II-associated p41 invariant chain (Ii) is a potent cathepsin L inhibitor;[228] and serpins, that were originally known to inhibit serine proteases. Hurpin, a member of the serpin family, is also a known inhibitor of cathepsin L.[241] Additionally, the squamous cell carcinoma antigen (SCCA) shows inhibitory activity by inhibiting cathepsin L.

*Synthetic Inhibitors of Cathepsin L*

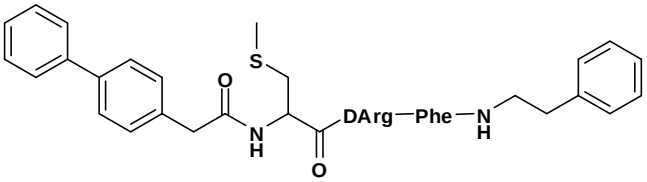
The importance of cathepsin L in physiological processes and pathological conditions has lead to a significant number of research projects that focused on potent, yet selective cathepsin L inhibitors. Many of these inhibitors have small peptides as a key moiety of their structures. Furthermore, many non-peptidic inhibitors, which are usually smaller, have better bioavailability. The large list of peptidic inhibitors includes

derivatives of: ene diones, epoxysuccinyl, nitrile, aldehydes, cyanopyrrolidines, cyclic heptapeptides, glyoxal and the cathepsin V propeptide. The list of nonpeptidic inhibitors of cathepsin L includes: Hydroperoxides, succinyl epoxides, diazomethylketones, thiocarbazates, chalcones, acyl hydrazones, thiosemicarbazones and unsaturated amides.[106,242] A sample of peptidic and non-peptidic inhibitors reported within the last ten years can be seen in Figure 15. The list of inhibitors includes oxocarbazate, nitriles, thiocarbazates, aldehydes and azepanones.

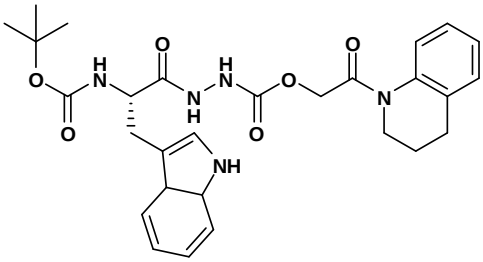
*Crystal Structures of Reversible and Irreversible Inhibitors of Cathepsin L*

Shenoy and Sivaraman recently reported the crystal structures of two inhibitors of cathepsin L. They obtained the crystal structures of cathepsin L complexed with a keto-aldehyde based and a diazomethylketone derivative inhibitors that are commercially available.[230]

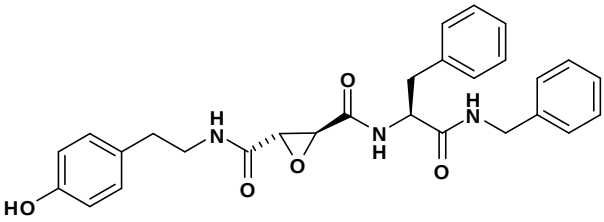
The aldehyde derivative (Z-Phe-Tyr-(O-But)-COCHO), α-keto-β-aldehyde) has shown to be a slow, tight-binding inhibitor of cathepsin L with a calculated $K_I$ value of 0.6 nM.[250] They found the reversible inhibitor binds to the active residue Cys25 via a transient covalent bond. Overall, the aldehyde forms a tetrahedral thiosemiacetal complex with cathepsin L. One of the key features of its inhibitory potency is due to the α-ketone oxygen which is in front of cathepsin L's oxyanion hole. Cys25 is covalently modified by the aldehyde group of the analog forming a thiosemiacetal moiety. Lastly, the S1 pocket is occupied by the Tyr-(O-But) moiety. The rest of the inhibitor (phenyl, and carboxybenzoxy groups) occupies the S2 and S3 pockets of the cysteine protease.



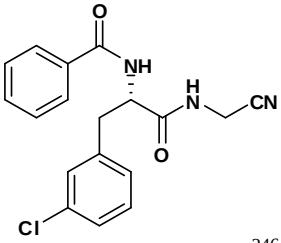

Biphenylacetyl Peptide ($K_I$: 19 nM)[243]



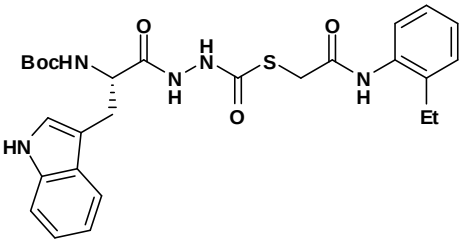

Oxocarbazate ($IC_{50}$: 6.9 nM)[244]



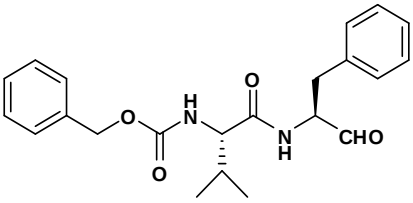

Epoxysuccinyl Peptide (CAA0225) ($IC_{50}$: 6.9 nM)[245]



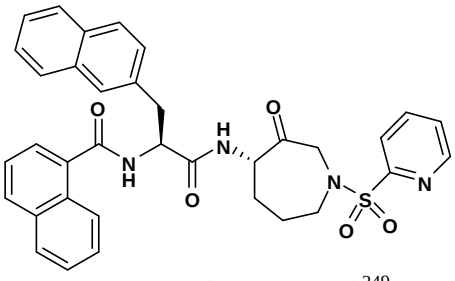

Dipeptidyl Nitrile ($pIC_{50}$: 6.6)[246]




Thiocarbazate ($IC_{50}$: 56 nM)[247]




Dipeptide Aldehyde (MDL21870) ($IC_{50}$: 2.5 nM)[248]




Azepanone ($K_I$: 0.43 nM)[249]

Figure 15. Selected Cathepsin L Non-Thiosemicarbazone Inhibitors Reported Between 2003 and 2012

The crystal structure also showed there were six hydrogen bonds between the inhibitor and the active site of the enzyme. Two of these hydrogen bonds involve His163 and Cys25, two of the active site residues.

The diazomethylketone Z-FY(*t*-Bu)-DMK is an irreversible inhibitor of cysteine proteases in *in vitro* and *in vivo* studies. The investigators found the irreversible inhibitor binds to the active residue Cys25 via covalent bond. The diazomethane methylene moiety makes a covalent bond with the thiol group of Cys25. The S1, S2 and S3 pockets are occupied by the Tyr(*t*-Bu), phenylalanine phenyl group and the carboxybenzyl groups of the inhibitor. The crystal structure also showed there were five hydrogen bonds between the inhibitor and the active site of the enzyme. Two of these hydrogen bonds are made between the inhibitor and the Gln19 side chain and Cys25 backbone. Additionally, the carbonyl oxygen of the synthetic compound was found to be facing the oxyanion hole of the enzyme.

*Role of Cathepsin L in Cancer*

The overexpression and abnormal activity of cathepsin L is also been evidenced in pathological conditions, as previously described in chapter one. Furthermore, proteolytic enzymes have become one of the field targets in the search for new and potent anticancer agents due to their importance in metastatic processes like cell invasion and migration.[251] Extensive research has demonstrated the participation of cathepsin L in malignant tumors due to its role in apoptosis, angiogenesis, cell invasion and cell migration.[252] Invasion and migration will be discussed separately.

*Apoptosis*

One of the particular characteristics of cancer cells is their ability to extend their life time by altered programmed cell death processes (apoptosis). In normal cells, apoptosis is an active, but highly regulated process. One of the proposed mechanisms of apoptosis reduction relates to cathepsin L and its ability to cleave Bid, a protein with apoptotic characteristics. Cleavage of Bid will liberate cytochrome C from the mitochondria which also induces apoptosis.[253] Interestingly, animal and *in vitro* studies using cancer cell lines give somewhat conflicting results with respect to the role of cathepsin L in apoptosis.

Cell culture with human transfected cells shows that cathepsin L has apoptotic functions.[254] Animal studies with cathepsin L knockout mice, indicate the lack of cathepsin L reduces cell proliferation and tumor growth, yet contributes to the apoptotic processes.[255]

*Angiogenesis*

Angiogenesis, or the process of formation of new blood vessels is vital for tumor growth.[256] Endostatin, a degradation product of type XVIII collagen, has angiogenetic properties. Cathepsin L has excellent collagenase activity as discussed previously. Research has shown that endostatin can be produced by degrading type XVIII collagen.[257] However, research has not provided specific evidence that supports the regulation of endostatin by cathepsin L.

Research has also investigated the role of cathepsin L in angiogenesis in animal models; however, their results are quite contradictory. Angiogenesis was reduced in mice treated with non specific cysteine protease inhibitors.[258] In contrast, animal studies

revealed that the angiogenesis process was not stopped in cathepsin L knockout mice.[255] Additionally, Urbich and coworkers reported overexpression of cathepsin L promoted the invasive characteristics of mature endothelial stem cells and neovascularization in mice was also found.[259]

### *Cathepsin L and Cancer in the Male Population*

Research has found cathepsin L is overexpressed in six of the most common types of cancers found in the United States. The list includes: prostate, bladder, colorectal, lung, melanoma and kidney cancers which constitute 70% of the new cases in 2012.[203,260–264] Specifically, cathepsin L is highly expressed in male urine, prostate cancer cell culture and male cancer tissues (colorectal, lung, melanoma, and kidney). Interestingly, the protease is downregulated in prostate cancer tissue.[105]

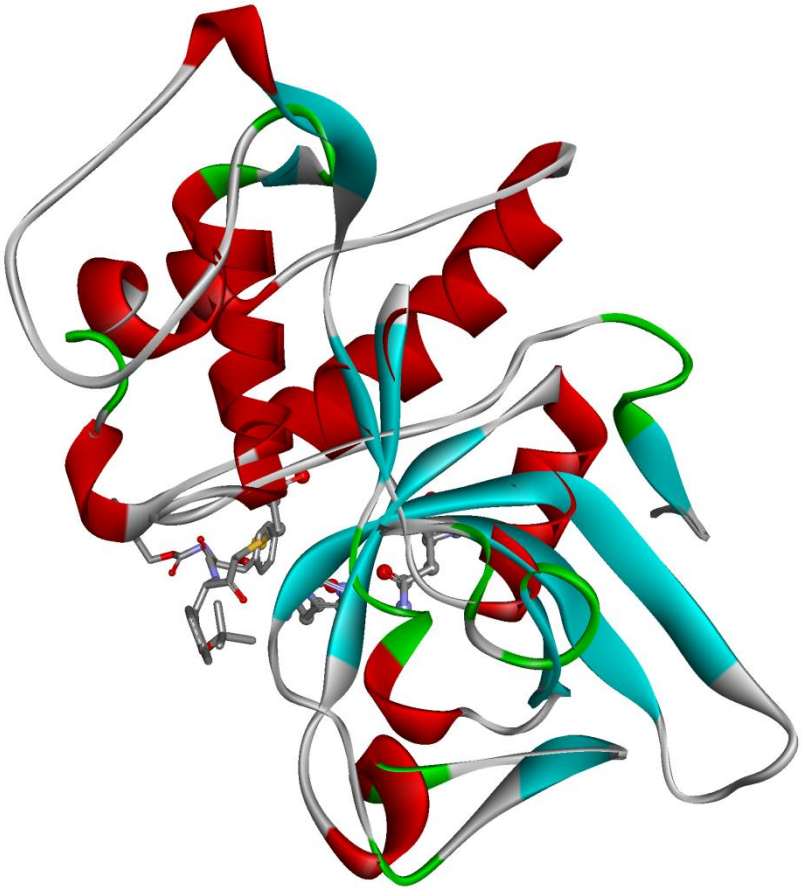
### *Cathepsin L and Cancer in the Female Population*

Similarly, upregulation of cathepsin L has been found in seven of the most common types of cancer. The list includes: breast, ovarian, colorectal, lung, melanoma, kidney and uterus cancers which constitute an alarminsg 88% percent of the new cases in 2012. Overexpression of cathepsin L is found in female cancer tissues (breast, colorectal, lung, and melanoma); female serum (ovarian and uterus) and cell culture (breast). Downregulation was also found in female cancer tissue (kidney and uterus).[265–267]

### *Cathepsin L and Other Types of Cancer*

The importance of cathepsin L in other cancers has also been reviewed and investigated. Nasopharyngeal, gastric, brain, head, neck cancer tissues also expressed elevated concentrations of this protease.[105,268–270]

**A**



**B**

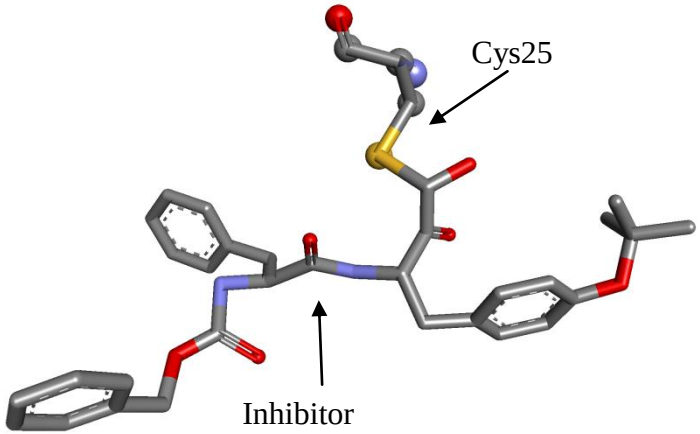
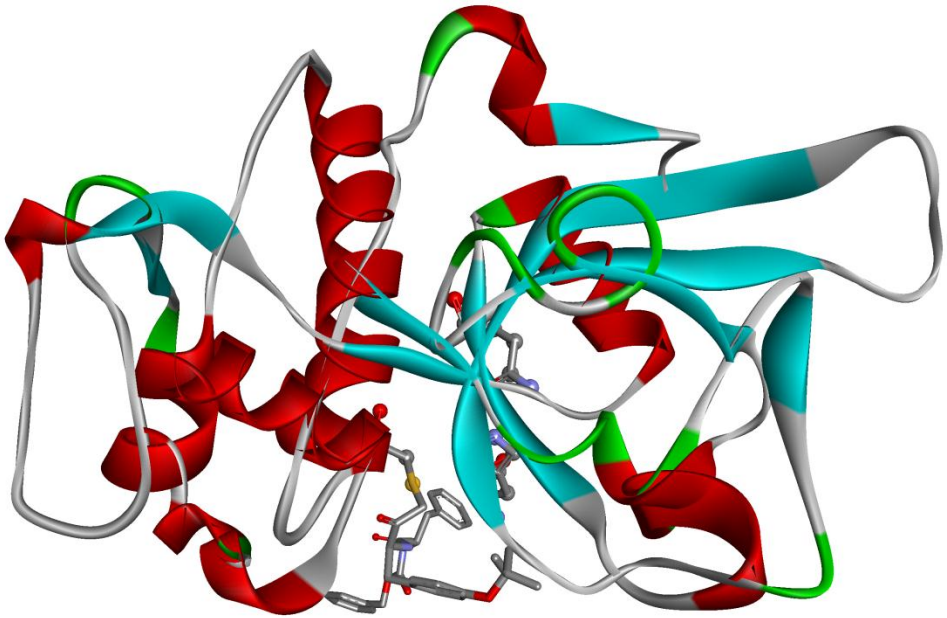



Figure 16. Crystal Structure of Human Cathepsin L with Z-Phe-Tyr(OBut)-COCHO, a Reversible, Slow, Tight Binding Inhibitor. **A**. Overall Structure. **B**. Covalent Modification of Cys25 by the Inhibitor (PDB ID: 3OF8)

**A**



**B**

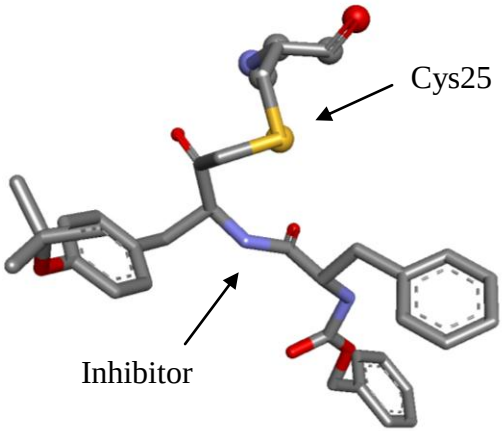



Figure 17. Crystal Structure of Human Cathepsin L with Z-F-Y(t-Bu)-DMK, an Irreversible Inhibitor. A. Overall Structure. B. Covalent Modification of Cys25 by the Inhibitor (PDB ID: 3OF9)

### *Cathepsin L and Breast Cancer*

More than 225,000 new cases of breast cancer in women are predicted to be detected in 2012. This deadly disease is the second cause of death among female cancer patients. Almost 40,000 women are expected to die in the same year. Thus, it is important to find new chemotherapeutic targets that can arrest or delay this cancer, yet with minimal side effects. As previously stated, cathepsin L is upregulated in breast cancer tissue and cell lines.

Investigators have linked the invasiveness of breast cancer with cathepsin L activity.[271–273] Therefore, *in vitro* assays for cell invasion and cell migration are performed usually with Matrigel™. This synthetic matrix has also used for other cell culture studies including angiogenesis, 3D cell culture, neuroscientific studies, metabolic and toxicological studies.[274–276] The composition of this proprietary matrix includes laminutesin (56%), type IV collagen (31%) and entactin.[277] These three proteins are also important components of the extracellular matrix in physiological conditions.[278–280] Matrigel™ also contains several growth factors such as bFGF, EGF, IGF-1, PDGF, NGF, and TGF-β. Two- and three-dimensional arrays have been developed to investigate the invasiveness and motility of cancer cells in *in vitro* studies.[281,282]

### *Two-Dimensional Experiments*

The main objective of two-dimensional studies is to mimic three-dimensional environments that tumor cells encounter when they are attempting to invade surrounding tissues. Perhaps, this approach is one of the most used in cancer research when testing novel compounds that delay cell invasion.

Typically, *in vitro* invasion assays are carried out using Boyden chambers. The chambers consist of two parts: a polycarbonate membrane. Special features of the layer are its porosity and average diameters of its pores, which range between 1 and 8 μm. The second part of the chamber is the thick layer of a basement membrane component (i.e. collagen) or a combination of them (i.e. Matrigel™). This layer is also present as a porous material. A similar format can be found for migration assays. The main difference from migration assays is Boyden chambers do not contain a basement membrane (i.e. Matrigel™).[283] Cells are placed in the inner chamber (a top chamber incubated for an extended period of time, usually 24 hours). Cells with invasive characteristics tend to invade thorough they layer of basement membrane. Cells that pass the membrane(s) can be determined by colorimetric or fluorometric methods. A chemoattractant is also added in the bottom chamber in order to attract cells. A graphical representation of the modified Boyden chamber for invasion and migration assays can be seen in Figure 18.[284]

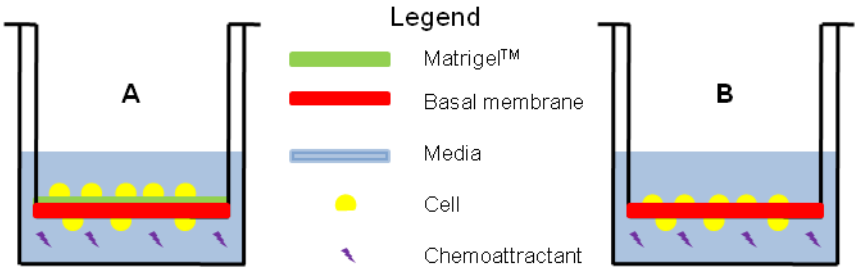



Figure 18. Graphical Representation of Boyden Chambers for Cell (**A**) Invasion; and (**B**) Migration Assays

There are also attempts to correlate cathepsin L mRNA and protein expression as a potential biomarker for the detection of breast cancer. This idea seems to be

impractical due to discrepancies in similar studies that have or have not found a direct relationship between cathepsin L and breast cancer samples.[285,286]

*Cell Invasion and Migration. Role of Proteases in ECM Degradation*

As previously stated, the metastatic cascade can be summarized in nine general steps. The mechanism is not simple due to the numerous biochemical and physiological processes that need to be carried out in order for one cancer cell to metastasize from the primary tumor. The cell then must migrate, invade and be deposited to its final destination (metastatic place)

Numerous proteases (MMPs, cysteine and serine proteases) are required for cell invasiveness. Proteolytic reactions degrade surrounding extracellular matrix and permit cancer cells to migrate from the primary site and invade surrounding organs or tissues. The degradation of the ECM occurs via two mechanisms:

1. Proteolytic degradation caused by individual cell-secreted proteases from cells.
2. Surface-related proteolytic degradation caused by membrane-bound proteases.[283]

Examples of diffuse degradation proteases are MMPs and some cathepsins. MMP-2 and MMP-14 are examples of surface-related proteases. Finally, MMPs and cathepsins are also found in this category.

*Cell Invasion and Migration. Cell Signaling between MMPs and Cathepsins*

Cancer cells are not the only cells that contribute to their invasive/motility characteristics.[287] Stromal cells, and overall, the tumor microenvironment, which enhances protease activity are included. Table 12 summarizes the proteolytic contribution of numerous tissue connective cells as part of the tumor microenviroment. The regulatory activity of all of them (excluding endothelial cells) include mutual

activation between proteases found in both types of cells (stromal and tumor); also, inhibition of proteases found in tumor cells by more than one stromal protease. Macrophages, neutrophilis, fibroblasts, and endothelial cells express more than four cathepsins. For example, macrophages produce six out of the seven cathepsins expressed in this type of stromal cell. Fibroblasts express six different metalloproteases. Table 13 also describes the distribution of other important proteases such as urokinase-type plasminogen activators (uPAs) and plasmin in cancer stromal cells. The distribution of this enzyme, along with other proteins is more selective. For example, pericytes only express MMP9 and monocytes express only cathepsin B and MMP9.

Table 12. Enzyme Distribution of MMPs and Cathepsins within Cancer Stromal Cells. (Adapted from Mason, page 230[287])

| Protease | Metalloprotease (MMP-) | | | | | | Cathepsin | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cell | 2 | 3 | 9 | 11 | 13 | 14 | B | C | D | G | H | L | S |
| Monocyte | - | - | + | - | - | - | + | - | - | - | - | - | - |
| Pericytes | - | - | + | - | - | - | - | - | - | - | - | - | - |
| Endothelial | + | + | - | - | - | + | + | - | + | - | - | + | + |
| Fibroblast | + | + | + | + | + | + | + | + | + | - | - | + | - |
| Lymphocytes | - | + | + | - | - | - | - | + | + | - | + | - | - |
| Neutrophilis | - | - | + | - | - | - | + | + | + | + | + | - | - |
| Macrophages | + | - | + | - | - | + | + | + | + | - | + | + | + |
| Mast | + | - | + | - | - | - | - | + | + | + | - | - | - |
| Mesenchymal stem | + | - | - | - | - | - | + | - | + | - | - | - | - |

Table 13. Enzyme Distribution of Other Proteases within Cancer Stromal Cells. (Adapted from Mason, page 230 [287])

| Cell | Protease | | | | | |
|---|---|---|---|---|---|---|
| | Plasmin | tPA | uPA | Elastase | Granzyme B | MMP4 |
| Endothelial | - | - | + | - | - | - |
| Fibroblast | - | - | + | - | - | - |
| Lymphocytes | - | - | - | - | + | - |
| Neutrophilis | - | - | - | + | - | - |
| Macrophages | + | + | + | - | - | - |
| Mast | - | - | - | + | + | + |

The proteolytic network is quite complex, and extensive research has done a great deal to explain possible pathways that link these different types of proteases. For example, two cathepsins (B and C) are responsible for the activation of other proteases. Cathepsin C regulates the activity of Granzyme B, Cathepsin G, and Elastase.[288] On the other hand, cathepsin B is one of the major key regulators in proteolytic cell signaling. It regulates two metalloproteases: MMP3, and MMP2.[289] Cathepsin B also regulates (and is regulated by) cathepsin D and uPA.[290] Finally, tPA regulates this cysteine protease.[291]

Additionally, cathepsin L is regulated by cathepsin D (a cathepsin B regulated protease) and regulates two more cathepsins (C and X).[292,293] As seen in these examples, the extensive network of proteases and its members is closely linked. Proteases do not work independently; however they play proteolytic and regulatory functions among other important features. Figure 19 depicts the interaction between the members of the cathepsin family. Cathepsin L regulates two cysteine proteases (cathepsins C and X) and is regulated by cathepsin D, an aspartic protease (legend: blue: cysteine protease; red: aspartic protease; green: serine protease; adapted from Mason, page 231).[287]

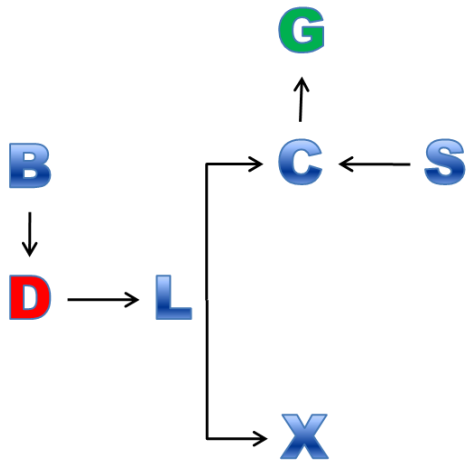



Figure 19. The Cathepsin Network

*Cell Signaling of Cell Invasion and Migration*

It has been demonstrated that both cysteine proteases and metalloproteases interact in a symbiotic fashion. However, these regulatory activities are also the result of numerous activities divided in signal transduction pathways that are specialized or related for specific functions. Cell migration and cell invasion are regulated by three important, signal transduction pathways: FAK, Integrin, and MAPK pathways.[294–296]

The regulation and activation of focal adhesion kinase (FAK) has taken a pivotal role in research due to its role in cell migration and cell invasion. Overexpression of FAK has been linked to highly invasive cancers such as squamous cell carcinomas and lung adenocarcinomas. Additionally, accumulation of FAK has been found in invadopodia, cell regions rich in integrins and MMPs.[297]

Integrins, are receptors that mediate interactions between cells and their microenvironment. In the case of cancer cells, integrins promote tumor metastasis via extracellular matrix interaction. In general, integrins partially regulate MMPs and uPA protease activities. Integrins and FAK have shown to promote cell motility in cancer cells. However, there are cases when integrins are the controllers of this characteristic.[298]

The mitogen-activated protein kinase family comprises multiple proteins that have direct effect on cell migration and invasion. Examples include p38, Erk, and JNK. The Jun N-terminus kinase promotes cell migration via phosphorylation of microtubule-associated proteins. Another member of the family, p38 regulates protein kinase 2/3, which is directly implicated in cell migration. In summary, MAPKs have proven to regulate cell migration by more than one signal transduction pathways. [299]

The basic migration mechanisms are actin and myosin II dependent. First, cell migration occurs via actin filament assembly.[299] Alternatively, myosin II performs 'contractions' in the actomyosin networks, so cells can migrate through healthy tissue. These two mechanisms are the result of the downstream product of activated (or unactivated) Rac, Rho, and Cdc42.

Perhaps, a major force of cell migration and cell invasion is the integrins.[300] These heterodimeric proteins make connections between tumor cells and ECM, so the former one can adhere to the matrix. The connections are via adaptor proteins, such as talin, vinculin, and focal adhesion kinase (FAK). Another function of the integrins is to send intercellular and extracellular signals as well. This is the moment when MMPs and cathepsins perform their hydrolytic interaction: the invasion of ECM and the implicit migration throughout the matrix. On the other hand, the cadherins are the responsibility of cancer cell invasion behavior. The list includes E-cadherin, N-cadherin, and other proteins such as conexxins and immunoglobulins. The complexity of this process can be seen in the expression of the cadherins. They both play selfregulatory activities. For instance, E-cadherin is the anti-invasion and N-cadherin gets overexpressed when the invasion process occurs.

Another class of cell migration promoter proteins is made of chemokines. The importance of these G-protein-coupled receptors lies in their ability to promote tumor invasion by using tumor macrophages and lymphocytes, important stromal cells rich in other metalloproteases and cathepsins (See Table 12).

Kinases can also promote cell invasion and migration. On one hand, migration is promoted by PI3K and Erk, two kinases that activate RTK, a third kinase. Invasion is

also promoted by the complex formed with FAK and Src. The role of FAK in cell migration and invasion is vital for tumor cells. First, the FAK-p190RohGEF complex promotes the activation of Rho, formation of actin fibers in the cytoplasm and stabilizes new adhesion points.

### *Bone Metastasis*

As previously discussed in chapter one, bone is the preferred metastatic site for breast and prostate cancers.[301] Bone metastasis can be divided into two forms. Osteoclastic metastasis occurs in male patients with prostate cancer, while osteoblastic metastasis is mainly found in breast cancer patients.[302,303] Bone metastasis is the result of what is known as the vicious cycle.[304] Tumor cells express RANKL, a ligand that binds to RANK and osteoprotegerin, a specific RANKL receptor. The ratio RANKL/OPG determines the osteoclast cycle which begins with bone resorption and growth factors. Growth factors act as tumor promoters.[305] Currently, there are compounds in early or late clinical trials. For example, imatinib is in preclinical trials. It has been proven to reduce osteolytic metastasis. Bisphosphonates have been also investigated. Other compounds include atrasentan, and denosumab in preclinical and clinical studies involving prostate cancer.[306–309]

### *Cathepsin L in Bone Metastasis, Cell Invasion and Cell Migration*

The number of research studies relating cathepsin L with bone metastasis is increasing and supports the theory of cathepsin L as one of the major macromolecules contributing to bone metastasis. Therefore, it is not surprising cathepsin L levels in breast cancer tissues and cell lines are higher than in normal tissues.[310] This

overexpression is the downstream result of some proinflammatory cytokines in the bone resorption processes (IL-1α, Il-6, TNF-α).[311,312] Specific cathepsin L inhibitors designed by Katunuma and coworkers (CLIK-148) were able to inhibit bone resorption and hypercalcemia in mice.[313] Also, *in vitro* studies osteosarcoma cells treated with cathepsin L inhibitors showed reduced levels of cathepsin L mRNA.[314] Additionally, cathepsin L is a powerful hydrolase with collagenase; elastase among other extracellular matrix components can be degraded as previously described. In depth studies have also concluded that cathepsin L participates indirectly in the bone remodeling process by activating key players in this process (MMPs, uPA and heparanase).[315] Cathepsin L might also be the responsible for the activity of E-cadherin, a cell-adhesion molecule.[255]

In general, other studies using HUVEC, U87MG (glioblastoma), and neuroblastoma cell lines also confirm the importance of cathepsin L in bone metastasis.[316–318] Cathepsin L has also shown to promote cell invasion and cell migration. Reduced cell adhesion and an increment of ECM degradation are consequences of overexpression of cathepsin L in cancer cells.[290] One of proposed mechanisms of the participation of cathepsin L in cell invasion and cell migration is the ability of the protease in cleaving E-cadherin. Lower levels of this cell-adhesion protein is downregulated in highly invasive cells.[319] The pathways for cell signaling and cathepsin L activity in cell invasion and cell migration are largely unexplored. The possibility of inhibition of cell invasion and/or cell migration via inhibition of cathepsin L has become a reference point for ongoing investigations using specific synthetic inhibitors designed for cathepsin L.

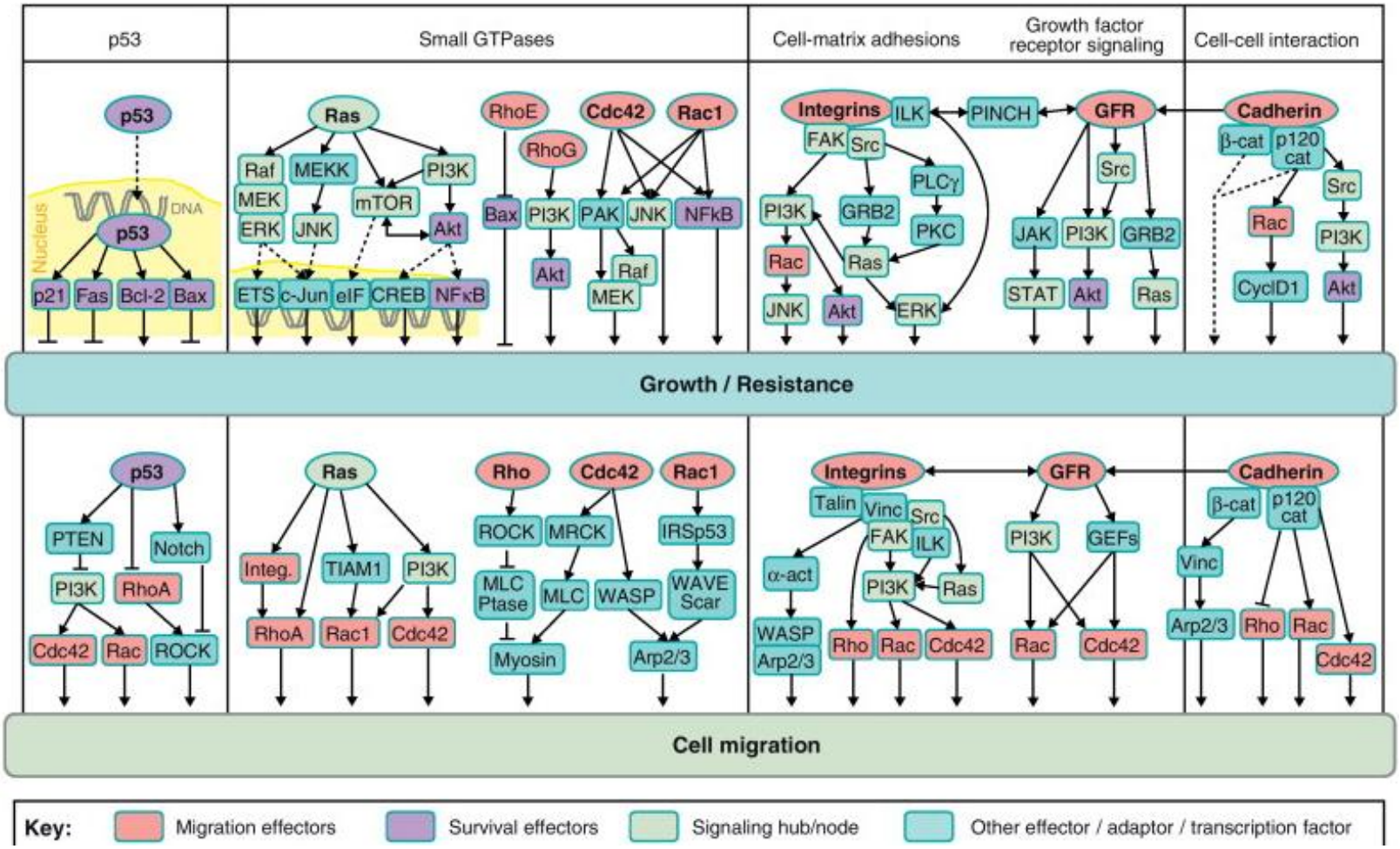



Figure 20. Signaling Pathways Controlling Tumor Cell Growth, Survival and Invasion. (reproduced from Alexander, page 17)[299]

*Material and Methods for the Biological Evaluation of Thiosemicarbazones Derivatives as Inhibitors of Cathepsin L*

*Materials and Equipment*

Ultra pure water was generated using a Barnstead Nanopure Diamond water purification system. This reagent will be called water for future references. Sodium acetate (anhydrous), ethylenediaminutesetetraacetic acid (EDTA), and dithiothreitol (DTT) were obtained from EMD. Recombinant, cathepsin L from human liver, 30 % Brij 35, and sodium dodecyl sulfate (SDS) were obtained from Sigma. Pure (99.9 %) Dimethyl sulfoxide (DMSO) and pure 7-Amino-4-methylcoumarin (AMC) were products from Acros and Anaspec. Z-Phe-Arg-AMC (7-Amino-4-methylcoumarin, N-CBZ-L-phenylalanyl-L-arginine amide, hydrochloride, Z-FR-AMC) was purchased from Bachem. Human type I collagen and gentamicin were acquired from Calbiochem. Simply Blue™ and electrophoresis reagents, including 10% Bis-Tris SDS-PAGE gels, antioxidant reagent, reducing agent, protein standard marker, SYPRO®, and loading buffer 4X were all products from Invitrogen. 3-(N-morpholino)propanesulfonic acid (MOPS) and Tris Base were purchased from BDH and Fisher Scientific. Matrigel™ invasion and migration assay kit were purchased from BD Biosciences. MDA-MB-231 cells were obtained from ATCC. DMEM media, FBS, and DIFF QUIK were products from, Cellgro, Gibco, and Imeb, respectively. Thiosemicarbazone derivatives were synthesized by members of Dr. Kevin G. Pinney' laboratory. Black 96 well flat bottom microplates (3991) and 0.2 µm filters were obtained from Corning. Ethanol was a product of Pharco. Single and twelve-channel pippeters as well pippetes tips were Eppendorf products. Range of pipettes used varied between 0.5 and 5000 µl. Finally, the

pH meter (Mettler Toledo), Fluoroskan Ascent FL (Thermo), XCell SureLock® Minutesi (Invitrogen), GE Typhoon 9400 FL, and Zeiss Axiovert 40 CFL inverted microscope were the instruments to calibrate the pH of the prepared solutions, fluorometric experiments, the separation of proteins, fluorescent imaging, and manual cell couting.

*Preparation of Buffers and Stock Solutions*

*Preparation of 1 M sodium acetate buffer, pH 5.5.* This buffer was prepared as a stock solution for further dilutions. It will be referred to as a 1 *M* buffer stock solution. One liter of this stock was prepared by weighing 69.65 g (0.85 mol) of sodium acetate anhydrous (NaOAc, MW: 82.03 g/mol) and 8.64 g (0.15 mol) of acetic acid glacial (MW: 60.05 g/mol; 1.049 g/ml) in a clean volumetric flask. The contents were dissolved with pure water. Glacial acetic acid or sodium hydroxide 5 *M* were used to adjust the pH of the buffer. The solution was stored at 4 °C .

*Preparation of 5M NaOH.* Fifty milliliters were prepared by weighing 10 g (0.25 mol) of NaOH pellets (NaOH, MW: 40 g/mol) in a clean volument flask. The contents were dissolved with pure water.

*Preparation of 400 mM sodium acetate buffer, pH 5.5*. One liter of this solution was prepared by diluting 400 ml of 1 *M* buffer stock solution to a final volume of one liter. Glacial acetic acid or sodium hydroxide 5 *M* (MW: 40 g/mol) were used to adjust the pH of the buffer. This solution was stored at 4 °C.

*Preparation of 40 mM EDTA.* A stock concentration of this solution was prepared by dissolving 14.89 mg (40 μmol) of EDTA (MW: 372.29 g/mol) in one

milliliter of water. Fifty milliliters of this solution (744.5 mg; 2 mmol) was prepared for multiple experiments. The solution of 40 m*M* EDTA was stored at 4 °C.

*Preparation of 80 mM DTT.* A freshly made stock concentration of this solution was prepared by dissolving 14 mg (80 μmol) of DTT (MW: 175 g/mol) in one milliliter of sodium acetate 400 m*M*, pH 5.5. The 80 m*M* DTT solutions were not stored for future experiments.

*Preparation of 4 mM AMC.* This stock solution was prepared by dissolving 0.91 mg (4 μmol) of AMC (MW: 229.2 g/mol) in one milliliter of pure DMSO. This solution was stable for several months when stored at low temperatures (-20 °C).

*Preparation of 29.94 mM Z-FR-AMC.* A sample of 19.43 mg (29.93 μmol) of Z-FR-AMC (MW: 649.15 g/mol) were dissolved in one milliliter of pure DMSO. This solution was stable for several months when stored at low temperatures (-20 °C).

*Recombinant cathepsin L from human liver.* Recombinant, human cathepsin L was purchased from Sigma. Concentration of the protease varied based on lot number. A typical cathepsin L concentration is 0.285 μg/μl (9.83 μ*M*; MW: 29 kDa).

*Preparation of human type I collagens solution from human skin.* Type I collagen (0.4 mg) was dissolved in 1 ml 50 m*M* acetic acid (pH 2.5). Solution was stirred for 24 hours at room temperature.

*Thiosemicarbazone derivatives stocks solutions.* More than 150 compounds synthesized in the laboratory of Dr. Kevin G. Pinney (Baylor University) were used to

make stock solutions (20 mM) in pure DMSO.[3,6,7,10–12] Usually, one milligram was weighed to give a final volume that varied between 100 and 200 µl of these stock solutions.

*Preparation of inhibitor dilutions.* Each inhibitor was serially diluted in pure DMSO to provide seven different concentrations (named A-G) that were ten-fold diluted. The concentrations of these dilutions varied between 2 m*M* (solution A) and 2 n*M* (solution G). Each dilution was diluted ten-fold in a solution that contained 35% DMSO in water (named 1-8). Concentrations of the aqueous solutions varied between 200 µi and 200 p*M*. An eighth solution was made by using solution C (20 µ*M*) to prepare an intermediate solution (1 µ*M* in 35% DMSO). Solutions made in pure DMSO were stored for two weeks at (-80 °C). Aqueous solutions were made prior to the experiment and kept at (4 °C). Figure 21. Serial Dilution Flowchart represents a graphical procedure of the preparation of the stock and inhibitor dilutions.

*Preparation of cathepsin L assay buffer.* Cathepsin L assay buffer (will be referred as assay buffer for future reference) contained 1.8 m*M* EDTA, 5.4 m*M* DTT, 0.02% Brij 35 and 120 m*M* NaOAc pH 5.5. Ten milliliters of assay buffer were made by adding 450 µl of 40 m*M* EDTA, 675 µl of 80 m*M* DTT, 6 µl of Brij 35, 3 ml of 400 m*M* NaOAc pH 5.5, and 5.87 ml of water in a plastic 50 ml conical tube. Table 14 summarizes a typical calculation sheet to make different volumes of assay buffer. The solution was stable and used for a maximum of 24 h.

*Preparation of 12% DMSO solution.* An aqueous solution of DMSO was prepared by dilution of 120 µl of DMSO with 8800 µl of water to bring a total of 1 ml.

*Preparation of 35% DMSO solution.* An aqueous solution of DMSO was prepared by dilution of 350 µl of DMSO with 650 µl of water to bring a total of 1 ml.

*Preparation of 50% DMSO solution.* An aqueous solution of DMSO was prepared by dilution of 500 µl of DMSO with 500 µl of water to bring a total of 1 ml.

*Preparation of cathepsin L stock solution.* A stock solution of cathepsin L (CL) contained 1 m*M* EDTA, 3 m*M* DTT, 0.01% Brij 35, 10 n*M* cathepsin L and 400 m*M* NaOAc pH 5.5. One milliliter of cathepsin L stock solution was made by adding 25 µl of 40 mM EDTA, 37.5 µl of 80 m*M* DTT, 0.33 µl of Brij 35, 1.02 µl of 0.285 µg/µl cathepsin L, and 936 µl of 400 m*M* NaOAc, pH 5.5, in disposable glass test tubes. Table 15 summarizes a typical calculation sheet to make different volumes of cathepsin L stock solution. The solution was made prior to every kinetic or inhibition experiment.

Table 14. Preparation Table for Cathepsin L Assay Buffer

| Assay Buffer (ml) | 40 m*M* EDTA (µl) | 80 m*M* DTT (µl) | Brij 35 (µl) | 400 m*M* NaOAc (µl) | Water (µl) |
|---|---|---|---|---|---|
| 5 | 225 | 338 | 3 | 1500 | 2935 |
| 10 | 450 | 675 | 6 | 3000 | 5870 |
| 20 | 900 | 1350 | 12 | 6000 | 11740 |
| 40 | 1800 | 2700 | 24 | 12000 | 23477 |

Table 15. Preparation Table for Cathepsin L Stock Solution

| CL (ml) | 40 m*M* EDTA (µl) | 80 m*M* DTT (µl) | Brij 35 (µl) | 400 m*M* NaOAc (µl) | Cathepsin L (µl) |
|---|---|---|---|---|---|
| 1 | 25 | 37.5 | 0.33 | 936 | 1.02 |
| 2 | 50 | 75 | 0.67 | 1872 | 2.04 |
| 3 | 75 | 112.5 | 1 | 2808 | 3.06 |

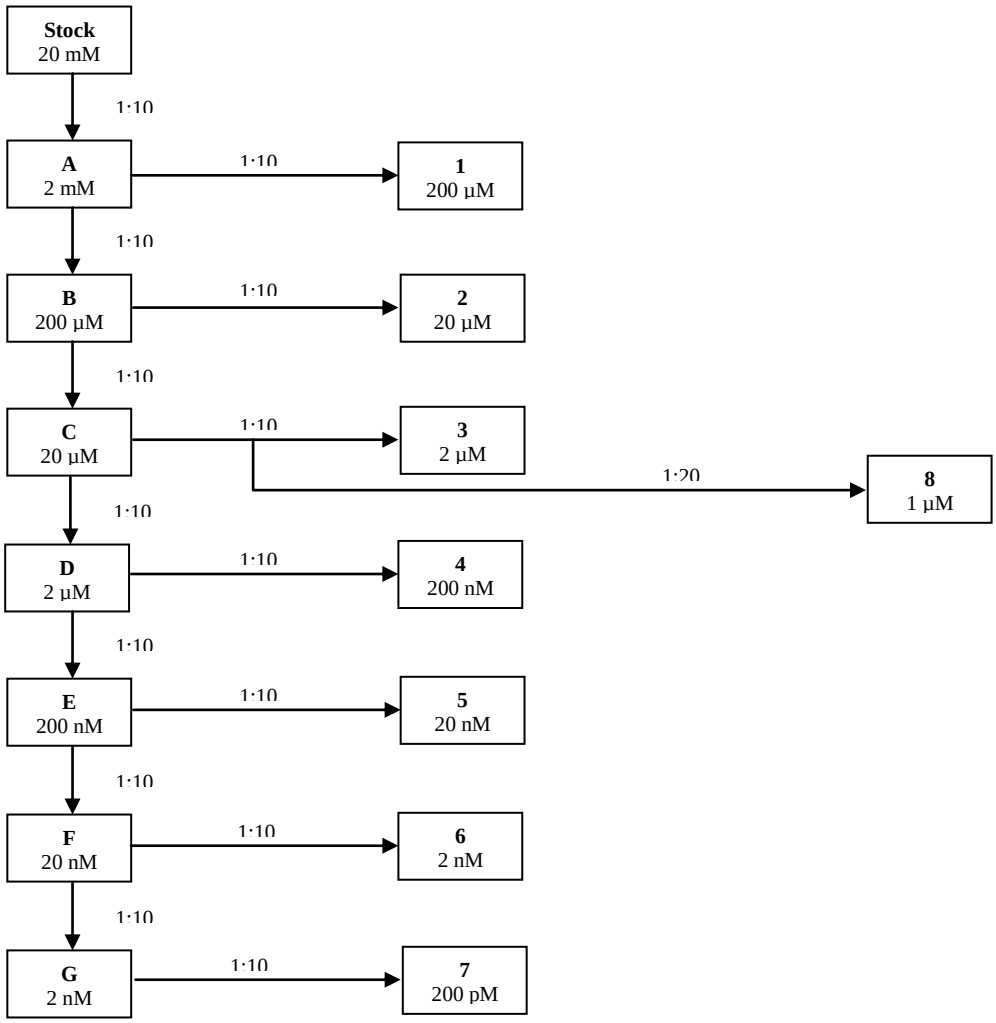



Figure 21. Serial Dilution Flowchart

*Preparation of Z-FR-AMC stock solution.* A stock solution of Z-FR-AMC (500 µM) in 2.5% DMSO was prepared by adding 20 µl of 29.94 m*M* Z-FR-AMC, 10 µl of DMSO and 1170 µl of water in order to have a final volume of 1.2 ml. The solution was made in a dark plastic 15 ml conical tube. Table 16 shows representative calculations to prepare various amounts of the substrate solution. The stock solution was stable and used for a maximum of 24 hours.

Table 16. Preparation Table for 500 µM Z-FR-AMC for Cathepsin L Assays

| 500 µM Z-FR-AMC (ml) | 29.94 mM Z-FR-AMC (µl) | DMSO (µl) | Water (µl) |
|---|---|---|---|
| 1 | 16.7 | 8.3 | 975 |
| 2.5 | 41.8 | 20.8 | 2440 |
| 5 | 83.5 | 41.5 | 4875 |
| 10 | 167.0 | 83.0 | 9750 |

*Assay buffer for reversibility studies (ABR).* Three milliliters of ABR were prepared by mixing 1200 µl of 250 m*M* NaOAc pH 5.5, 1.01 µl of Brij 30, 112.5 µl of 80 m*M* DTT, 75 µl of 40 m*M* EDTA, 55 µl of DMSO, 5 µl of 29.94 mM Z-FR-AMC and 1552 µl of water. ASR was prepared prior to the experiment in a 15 ml plastic conical tube kept in the dark. Final conditions of ASR are: 100 m*M* NaOAc pH 5.5, 1 m*M* EDTA, 80 m*M* EDTA, 0.01% Brij 30, 50.5 µ*M* Z-FR-AMC, and 2% DMSO. Table 17 summarizes these conditions.

Table 17. Preparation Table for Cathepsin L Assay Buffer for Reversibility Studies

| Volume (ml) | 40 mM EDTA (µl) | 80 mM DTT (µl) | Brij 35 (µl) | 250 mM NaOAc (µl) | 29.94 mM Z-FR-AMC (µl) | DMSO (µl) | Water (µl) |
|---|---|---|---|---|---|---|---|
| 3 | 25 | 37.5 | 0.33 | 936 | 1.02 | 55 | 1200 |

*Preparation of cathepsin L stock solution for reversibility studies (CLR).* One hundred forty-four microlites of CLR were prepared by mixing 115.2 µl of 250 m*M* NaOAc pH 5.5, 0.10 µl of Brij 30, 10.8 µl of 80 m*M* DTT, 7.2 µl of 40 m*M* EDTA, 2.93 µl of 0.285 µg/µl cathepsin L, and 7.8 µl of water in a glass test tube. Final conditions of CLR are: 200 m*M* NaOAc pH 5.5, 2 mM EDTA, 6 m*M* EDTA, 0.02% Brij 30, and 100 n*M* cathepsin L. Table 18 summarizes the preparation tables for CLR.

Table 18. Preparation Table for Cathepsin L Stock Solution for Reversibility Studies

| | Volume (ml) | 40 m*M* EDTA (µl) | 80 m*M* DTT (µl) | Brij 35 (µl) | 250 m*M* NaOAc (µl) | Cathepsin L (µl) | Water (µl) |
|---|---|---|---|---|---|---|---|
| CLR | 0.14 | 7.2 | 10.8 | 0.1 | 115.2 | 2.93 | 7.8 |

*Preparation of cathepsin L stock solution for collagenase studies (CLC).* Seventy-two microlitters of CLC were prepared by mixing 31.2 µl of 1 *M* NaOAc pH 5.5, 0.10 µl of Brij 30, 11.7 µl of 80 m*M* DTT, 7.8 µl of 40 m*M* EDTA, 8.7 µl of 0.285 µg/µl cathepsin L, and 12.5 µl of water in a glass test tube. Conditions of CLC are: 433.3 m*M* NaOAc pH 5.5, 4.3 m*M* EDTA, 13 m*M* EDTA, 0.04% Brij 30, and 1192 n*M* cathepsin L. Table 19 summarizes preparation tables for CLC.

*Dulbecco's modified Eagle's medium (DMEM) supplemented media*. A sterile bottle of DMEM media is supplemented with 50 ml of 10X FBS and 830 µl of a 3 mg/ml gentamicin stock solution. Supplemented DMEM media conditions are: 10% FBS and 50 ng/ml gentamicin. The media will be called as supplemented DMEM medium in future references.

Table 19. Preparation Table for Cathepsin L Assay Buffer for Collagenase Inhibition Studies

| CLC | Total Volume | 40 m*M* EDTA | 80 m*M* DTT | Brij 35 | 1 *M* NaOAc | CL | Water |
|---|---|---|---|---|---|---|---|
| Volume (µl) | 72 | 7.8 | 11.7 | 0.1 | 31.2 | 8.7 | 12.5 |

## *Experimental Section*

### *Kinetic Cathepsin L Assay*

Kinetic analysis of cathepsin L was carried out by using a Thermo Fluoroskan microplate reader with 3691 96-well black microplates.[5] The total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 50 µl of water, 10 of 35% DMSO solution, 20 µl cathepsin L stock solution, and 20 µl of Z-FR-AMC stock solution. A mixture containing assay buffer, 35% DMSO, was preincubated at 25 °C during 5 minutes using 96-well black microplates. The catalytic activity of cathepsin L was monitored by adding a concentration of Z-FR-AMC to every reaction was carried out in triplicates. The production of AMC was monitored for 5 minutes at 25 °C using excitation and emission reference filters of 355 and 460 nm respectively. Readings were taken every 10 seconds for five minutes. Reactions were carried out in triplicate. The final concentrations of the kinetic assay were: 100 m*M* NaOAc pH 5.5, 1 m*M* EDTA, 3 m*M* EDTA, 0.01% Brij 35, and 1 n*M* cathepsin L. Final concentrations of Z-FR-AMC varied between 0.2 and 14 µ*M*.

### *Preliminary Inhibition Studies*

Thiosemicarbazone analogs (provided by Dr. Kevin G. Pinney's laboratory, Baylor University)[3,6,7,10–12] were prescreened to determine if they have inhibitory activity against cathepsin L. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 50 µl of water, 10 µl of 35% DMSO or 10 µl of dilution "1" (See Figure 21; final concentration: 10 µM), 20 µl cathepsin L stock solutions, and 20 µl of Z-FR-AMC stock solution. The enzyme-inhibitor mixtures (180 µl) assay buffer, 35%

DMSO or inhibitor, and cathepsin L was preincubated at 25 °C during 5 minutes using 96-well black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and uninhibited samples were monitored for five minutes. Final reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and uninhibited samples were monitored for five minutes. The final concentrations of the preliminary inhibitory studies were: 100 mi NaOAc pH 5.5, 1 m*M* EDTA, 3 m*M* EDTA, 0.01% Brij 35, 1 n*M* cathepsin L, 10 µ*M* of the screened inhibitor and 50 µ*M* of Z-FR-AMC. Readings were taken every 25 seconds for five minutes and reactions were carried out in triplicate. Final concentrations of the inhibition cathepsin L assay were: 100 m*M* NaOAc pH 5.5, 1 m*M* EDTA, 3 m*M* EDTA, 0.01% Brij 35, 1 n*M* cathepsin L and 50 µ*M* of Z-FR-AMC. Final concentrations of the inhibitors varied between 10 µ*M* and 10 p*M*. Readings were taken every 25 seconds for five minutes and reactions were carried out in triplicate. Compounds that did not inhibit cathepsin L activity more than 50% at a final concentration of 10 µ*M* (i.e. $v_i/v_o \leq 0.5$) were considered as 'inactive' compounds and a general $IC_{50}$ value greater than 10000 n*M* was assumed. Compounds that inhibited cathepsin L activity more than 50% were further considered for cathepsin L inhibitory studies and an exact $IC_{50}$ value was determined.

*Cathepsin L Inhibition Assay*

Analysis of cathepsin L and its inhibitors was carried out by using a modified protocol of the kinetic cathepsin L assay, previously described. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 50 µl of water, 10 µl of 35% DMSO or 10 µl of inhibitor dilutions, 20 µl cathepsin L stock solutions, and 20 µl of Z-FR-AMC stock solution. A 180 µl mixture containing assay buffer, 35% DMSO

or inhibitor, and cathepsin L was preincubated at 25 °C during 5 minutes using 96-well black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and unhinhibited samples were monitored for five minutes. The final concentrations of the cathepsin L inhibition assay were: 100 m*M* NaOAc, pH 5.5, 1 m*M* EDTA, 3 m*M* EDTA, 0.01% Brij 35, 1 n*M* cathepsin L and 50 µ*M* of Z-FR-AMC. Final concentrations of the inhibitors varied between 10 µ*M* and 10 p*M*. Readings were taken every 25 seconds for five minutes and reactions were carried out in triplicate. TablesTable 20 andTable 21summarize final volumes and final conditions for the kinetic and cathepsin L inhibition assays.

*Construction of AMC Calibration Curve*

The cathepsin L kinetic assay was slightly modified to construct the 7-amino-4-methylcoumarin calibration curve. Final volume was 200 µl. The assay consisted of 120 µl assay buffer, 10 µl of water, 10 µl of stock solution and 20 µl of AMC dilutions made in 20% DMSO. Serial dilution was used to prepare twelve AMC samples that were 10 fold final concentrations.

Table 20. Preparation Table for Kinetic, Cathepsin L Inhibition Assays and Construction of AMC Calibration Curves

| Reagent | Kinetic assay (µl) | Inhibitory assay (µl) | AMC Curve (µl) |
|---|---|---|---|
| Assay Buffer | 100 | 100 | 100 |
| Water | 50 | 50 | 60 |
| Control | 10 | 0 | 0 |
| Inhibitor | 0 | 10 | 0 |
| CL | 20 | 20 | 20 |
| Z-FR-AMC | 20 | 20 | 20 |
| Total (µl) | 200 | 200 | 200 |

Table 21. Final Conditions for Kinetic Cathepsin L and Inhibition Cathepsin L assays

| NaOAc (m*M*) | EDTA (m*M*) | DTT (m*M*) | Brij 35 (%) | DMSO (%) | Cathepsin L (n*M*) | Z-FR-AMC (μM) |
|---|---|---|---|---|---|---|
| 100 | 1 | 3 | 0.01 | 2 | 1 | 0.1- 50 |

A 4 m*M* AMC in DMSO was used as a primary stock solution. Table 22 shows calculations for the preparation of these dilutions. Final concentration varied between 15 μM and 1 nM of AMC in solution. Final conditions were 100 m*M* NaOAc pH 5.5, 1 m*M* EDTA, 3 m*M* DTT, 0.01 % Brij, and 2 % DMSO. Assay buffer, water, cathepsin L stock solution (with no inhibitor) and AMC dilutions were added to the 96-well black microplates, incubated at 25 °C for 5 minutes. Samples were monitored 25 °C for 800 seconds at excitation and emission wavelengths of 355 nm and 460 nm respectively. Linear regression for data fitting was carried out using GraphPad 5.0, where the independent and dependent variables are time (in seconds) and relative fluorescence units. The Y- intercept (dependent variable) was plotted vs AMC concentration (in μM).

Table 22. Preparation Table for the Construction of AMC Calibration Curve

| Final Concentration | | | AMC dilutions in 20% DMSO | | | | |
|---|---|---|---|---|---|---|---|
| # | AMC (μM) | DMSO (%) | Add from | Volume ( μl) | DMSO (μL) | Water (μl) | DMSO (%) |
| 0 | 25 | 2 | Stock | 63 | 138 | 800 | 20 |
| 1 | 15 | 2 | Stock | 38 | 163 | 800 | 20 |
| 2 | 10 | 2 | Stock | 25 | 175 | 800 | 20 |
| 3 | 5 | 2 | Stock | 13 | 188 | 800 | 20 |
| 4 | 1 | 2 | # 1 | 67 | 187 | 747 | 20 |
| 5 | 0.75 | 2 | # 1 | 50 | 190 | 760 | 20 |
| 6 | 0.5 | 2 | # 1 | 33 | 193 | 773 | 20 |
| 7 | 0.3 | 2 | # 3 | 60 | 188 | 752 | 20 |
| 8 | 0.1 | 2 | # 3 | 20 | 196 | 784 | 20 |
| 9 | 0.050 | 2 | # 6 | 100 | 180 | 720 | 20 |
| 10 | 0.010 | 2 | # 6 | 20 | 196 | 784 | 20 |
| 11 | 0.005 | 2 | # 6 | 10 | 198 | 792 | 20 |
| 12 | 0.001 | 2 | # 6 | 2 | 200 | 798 | 20 |

*FC: Final concentration during assay*

*Effect of Solvent Concentration on Cathepsin L Inhibition Assays*

Preincubation studies with compounds **1, 136**, and **168** were carried out by a slight modification of the cathepsin L inhibition assay. Different sets of mixtures containing assay buffer, inhibitor and cathepsin L were preincubated at various DMSO concentrations between 0.7 and 3% (final concentration). Reactions were taken every 25 seconds for five minutes and carried out in triplicate.

*Effect of Inhibitor Concentration on Cathepsin L Progress Curves*

Final concentrations, conditions, and volumes are similar to those previously described in the cathepsin L inhibition assay. Assay buffer, inhibitors (final concentrations varied between 100 nM and 20 µM) and Z-FR-AMC (final concentration: 5 µM) were added to the 96-well black plates (volume of the substrate-inhibitor mixture: 180 µl). Then, 20 µl of cathepsin L stock solution were added immediately without preincubation time. Readings were taken every 3 seconds for fifty minutes.

*Effect of Preincubation Studies on Cathepsin L Inhibition Assays*

Preincubation studies with compounds **1** and **8** were carried out by a slight modification of the cathepsin L inhibition assay. Different sets of mixtures containing assay buffer, inhibitor and cathepsin L were preincubated at various periods of times between 0 and 240 minutes. Reactions were taken every 25 seconds for five minutes and carried out in triplicate.

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation*

Data that was obtained in the effect of preincubation studies for compounds **1** and **8** was further analyzed. The possibility of these compounds to be tight binding inhibitors

was examined by fitting the data by a nonlinear regression using Morrison's quadratic equation.

*Cathepsin L Reversibility Studies*

Three milliliters of assay buffer for reversibility studies (BSR) were prepared as previously described. Twenty five microliters of cathepsin L (100X: 100 nM) assay buffer for reversibility studies (CLR) were pre-incubated with an equal amount of a concentrated solution of the inhibitor (100X: 0.1 $IC_{50}$) at 25 °C between one and four hours. Then, two microliters of the enzyme-inhibitor mixture were rapidly mixed with 198 µl of ASR in order to start the reaction. Total reaction volume was 200 µl. Readings were taken every twenty five seconds for four hours. Final concentrations are similar as described previously. Final conditions were: 100 m*M* NaOAc, pH 5.5, 1 m*M* EDTA, 3 m*M* EDTA, 0.01% Brij 35, 1 n*M* cathepsin L and 50 µ*M* of Z-FR-AMC. Table 21 describes required volumes for this experiment.

Table 23. Preparation Table for Cathepsin L Reversibility Assay

| Reagent | Reversibility studies (µl) |
|---|---|
| Assay Buffer | 198 |
| Inhibitor/CL | 2 |
| Total | 200 |

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

The effect of [Z-FR-AMC] was studied with compounds **1** and **8**. Minor modifications of the cathepsin L inhibition assay were carried out. Different sets of mixtures containing assay, inhibitor and cathepsin L were preincubated at specific preincubation times (5 or 240 minutes). Reactions were initiated by the addition of

different concentrations of Z-FR-AMC. Reading were taken every 25 seconds for five minutes and carried out in triplicate.

*Effect of Substrate Concentration (Z-FR-AMC) on Cathepsin L Progress Curves*

Final concentrations, conditions, and volumes are similar to those previously described for the cathepsin L inhibition assay. Assay buffer, one inhibitor concentration (Final concentration for **8**: 5 µM) and Z-FR-AMC (final concentrations varied between 0.5 and 20 µM) were added to the 96-well black plates (volume of the substrate-inhibitor mixture: 180 µl). Then, 20 µl of cathepsin L stock solution was added immediately without preincubating. Readings were taken every 3 seconds for fifty minutes.

*Inhibition of Cathepsin L Collagenase Activity by Thiosemicarbazone Derivatives*

A solution of human type I collagen ([ACI]: 0.4 mg/ml) and cathepsin L stock solution (CLI) were prepared as previously described. The inhibitor stock solution was prepared by mixing 1.3 µl of 20 m*M* stock solution of the inhibitor in DMSO, 9.1 µl of DMSO and 89.6 µl of water. Conditions of this stock solution are: 10.4% DMSO and 260 µM of the inhibitor. The experiment was started by preincubating 3 µl of ACI and 2.5 µl of the inhibitor stock solution in a microcentrifuge tube at 37 °C for 2 hours. A 1X staining solution was prepared by diluting 10 µl of 5000X SYPRO® red protein gel staining dye with 49990 µl of 7.5% acetic acid. The cathepsin L-inhibitor mixtures were preincubated for 2 hours. Similarly, another group of control samples (without the inhibitor) were also set up. Then, reactions were initiated by adding 7.5 µl of type I collagen and monitored between 0 and 20 hours at 37 °C. Final conditions of the reactions were: 100 m*M* NaOAc pH 5.5, 1 m*M* EDTA, 3 m*M* DTT, 0.01% Brij 30, 275

n*M* cathepsin L, 2% DMSO, 50 μ*M* of the tested inhibitor, and 0.01 mg/ml type I collagen. Reactions were stopped by adding 2 μl of LDS NuPAGE® sample buffer, heated at 90 degrees for ten minutes and immediately stored at -80°C. Inactivated samples (15 μl) were loaded onto 4-12 % NuPAGE® Bis-Tris gels. Electrophoresis was carried out by using MOPS running buffer at 200 V for 50 minutes. Gels were stained in 50 ml of 1X SYPRO® staining solution at room temperature between one and two hours followed by a destaining process using water (1X) and 7.5% acetic acid (1X). Finally, gel imaging was performed with a GE Typhoon 9400 FL with excitation and emission wavelengths of 550 and 630 nm, respectively. Another experiment was carried out with similar conditions but cathepsin L, inhibitor and type I collagen were added at the same time with no preincubation time. Table 24 describes conditions for inhibition of collagenase activity studies.

Table 24. Preparation Table for Cathepsin L Collagenase Activity Assay

| Reagent | Control (μl) | Inhibited (μl) | Control (μl) | Inhibited (μl) |
|---|---|---|---|---|
| CLC | 3.0 | 3.0 | 3.0 | 3.0 |
| Control (Water) | 2.5 | 0 | 2.5 | 0 |
| Inhibitor | 0 | 2.5 | 0 | 2.5 |
| TIC | 7.5 | 7.5 | 7.5 | 7.5 |
| Preincubation time (h) | 2 | | 0 | |
| Total | 13 | 13 | 13 | 13 |

*Cell Culture Experiments. MDA-MB-231 Cell Subculture Maintenance*

One milliliter of passage 2, MDA-MB-231 cells that were kept under cryogenic temperatures, was thawed at 0 °C. Then, cells were cultured with 12 ml supplemented DMEM medium in T75 $cm^2$ with flasks with canted necks in incubators at 37 °C and 5%

$CO_2$/air. Supplemented DMEM medium was changed every two days with 12 ml of fresh DMEM medium. Cell subculture was carried out once cells were 80% confluent.

*Cell Subculture (passaging) Protocol*

A new generation of MDA-MB-231 cells or passaging was carried out by removing DMEM medium from confluent cells. Immediately, 4 ml of 0.25% (w/l) trypsin with EDTA were added to the cells to induce cells to detach. Flasks were incubated with trypsin in the incubator for 4 minutes. After incubation time, cells were observed under the microscope to confirm if they were completely rounded and detached from the flask surface. Eight milliliters of supplemented DMEM media 1X were added to the flasks in order to neutralize the proteolytic activity of trypsin. Trypsinized cells were used to generate a new generation of cells. Four milliliters of cells were transferred to a new T75 $cm^2$ flask with canted neck that contained 8 ml of fresh DMEM media. New passages were cultured and maintained as previously described. Also, trypsinized cells were frozen to maintain and generate new passages for future experiments.

*Cell Freezing Protocol*

Trypsinized cells were periodically frozen to keep and maintain an inventory for future experiments. Stock solutions were centrifuged at 10000 rpm for 5 minutes to form a cell pellet. Suspended media was removed and replaced with DMEM medium that is supplemented with 5% DMSO. A cell counter was used to determine cell concentration in the new stock solutions. Cell concentration varied between one and 5 million cells/ml. These cell stock solutions were transferred to cryogenic vials. Cells were kept under cryogenic conditions (liquid nitrogen).

*Cell Invasion Assays*

Invasion assays were carried out using BD Bioscience Matrigel™ invasion kit assay. Independent MDA-MB-231 passages were cultured and passaged as described. Three independent experiments were performed with passages 13, 13 and 13 cells. Cells were trypsinized when they were 80% confluent. However, trypsin neutralization was done with DMEM with no FBS. Cell density was determined with a Beckman Coulter Z-Coulter cell counter. The three stock solutions were diluted with DMEM with no FBS in order to reach a 200,000 cell/ml concentration in each case. Five thiosemicarbazone inhibitors with high cathepsin L selectivity were selected to test its efficacy by delaying cell invasion. Additionally, E-64, a known generic cysteine protease inhibitor was used as a positive control. Stock solutions (20 mM in DMSO) of the compounds and E-64 were prepared by dissolving 1 mg of each sample according to the calculations listed in Table 25.

Table 25. Preparation Table for 20 mM of Selected TSCs for Invasion Assays

| Compound | Molecular Weight (g/mol) | Sample (mg) | Sample (μmol) | Volume (μl) |
|---|---|---|---|---|
| **1** | 413.13 | 1 | 2.42 | 121.0 |
| **8** | 350.23 | 1 | 2.86 | 142.8 |
| **156** | 359.44 | 1 | 2.78 | 139.1 |
| **157** | 395.42 | 1 | 2.53 | 126.4 |
| **168** | 474.32 | 1 | 2.11 | 105.4 |
| E-64 | 357.41 | 1 | 2.80 | 139.9 |

Then, a 1:10 fold dilution in DMSO ([A]: 2 mM) of each compound was used to prepare two stock aqueous dilutions using DMEM without FBS ([Stock solutions]: 50 and 20 μ*M*). DMSO was added to each of these dilutions in order to reach a concentration of 4%. Table 26 shows the preparation of these solutions.

Table 26. Preparation Table for 50 and 20 μM for Selected TSC Inhibitors

| Concentration (μ*M*) | A (μl) | DMSO (μl) | Media (μl) | [DMSO] (%) |
|---|---|---|---|---|
| 50 | 25 | 15 | 960 | 4 |
| 20 | 10 | 30 | 960 | 4 |
| Control | 0 | 40 | 960 | 4 |

Simultaneously, several BD BioCoat Matrigel® invasion chambers, which contain an 8 micron pore size PET membrane with a layer of Matrigel basement, were prepared for the experiment. Both sides of every matrix were rehydrated using 500 μl DMEM media with no FBS for two hours in a 5% $CO_2$ environment at 37 °C. Media were removed and inserts were ready to be used to start the invasion assay. The experiment was initiated by adding 750 μl of DMEM supplemented with 10% FBS (which functions as a chemoattractant) and gentamicin to a 24-well microplate (i.e. lower chamber). Then, the inserts were carefully placed on top of the wells containing the chemoattractant, avoiding air bubbles between the bottom part of the insert and medium. Two hundred fifty microliters of the cell stock solutions (200,000 cells/ml) and 250 μl of 50 and 20 μM of each compound to be tested were added to every insert. Final conditions for treated cells were: 2% DMSO, 50,000 cells and 25 or 10 μM of the compounds. Final conditions for untreated (controls) cells are: 2% DMSO and 50,000 cells. The 24-well plates containing the invasion studies chambers were placed in an incubator with a 5% $CO_2$ environment for 24 hours at 37 °C.

Reactions were terminated by removing medium from the inserts. Each insert was cleaned twice with cotton swabs to remove cells located at the top side of the membrane. Membranes were stained with a Diff-Quik staining kit (IMEB, Inc.). Cells were fixed with 100% methanol for a minimum time of two minutes. Then, membranes

were rinsed with deionized water and consecutive staining solutions of azure and xanthenes dyes. Samples were placed in each solution for a minimum of two minutes and rinsed with water in between. Finally, samples were air-dried for a minimum for two hours in a biological safety cabinet.

Membranes were removed using sterile scalpels and placed on glass slides. Each sample was observed under a Zeiss Axiovert 40 CFL inverted microscope to perform manual cell counting. Ten fields were chosen and observed under a 40 X objective. Eight fields were located at the periphery of the circular sample and numbered clockwise. Two more fields were counted at the middle of each sample, as depicted in figure 18. Experiments were performed in triplicate. The inhibition of the invasiveness of MDA-MB-231 cells was measured by using the formula: treated cells that invaded the Matrigel™ layer/untreated cells that invaded the Matrige™ layer.

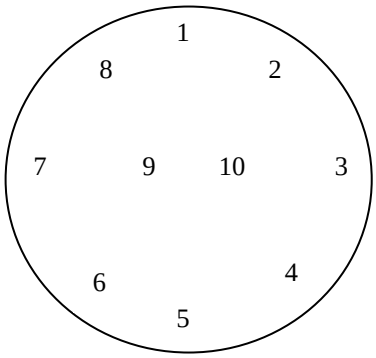



Figure 22. Invasion and Migration Assay Sample Fields

*Cell Migration Assays*

The experimental procedure for the determination of the ability of MDA-MB-231 to migrate is similar to that described for the cell invasion assays. However, the inserts with an 8 μm did not contain the Matrigel™ layer. Both, invasion and migration assays

were performed simultaneously and used the same MDA-MB-231 cell stock solutions. The inhibition of the motility of MDA-MB-231 cells was measured by using the formula: treated cells that migrated to the 8 μm membrane/untreated cells that migrated to the 8 μm layer.

*Results and Discussion*

Through a continuing collaboration with Dr. Kevin G. Pinney and his group, our results have indicated that some thiosemicarbazone derivatives are lead compounds as cathepsin L inhibitors.[3,6,7,10–12] Thiosemicarbazones were first discovered as inhibitors of cruzain, a cathepsin L-like enzyme in a high-through put screening process.[8] Synthesis of a small library of thiosemicarbazones led to the discovery of a number of thiosemicarbazone derivatives as inhibitors of cruzain. This library was tested toward cathepsin L as well showing promising results. The library of compounds required a complete screening to verify the potency of novel analogs. The Structure-Relationship Activity (SAR) was characterized after completing *in vitro* testing against cathepsin L ($IC_{50}$ values).

The characterization of compounds that could be used as potential therapeutic agents involved numerous evaluations. Understanding the kinetic effect of these inhibitors helped to identify their mechanism of action. Fluorometric based assays were utilized to study various assay parameters in inhibitory activities, determination of $K_I$, and reversibility of thiosemicarbazone inhibitors. Two dimensional cell based assays (colorimetric cell invasion and cell migration assays) were used to determine the inhibitory ability of these compounds within the extracellular matrices of metastatic breast cancer cell lines.

To understand interactions between inhibitors and cathepsin L, molecular modeling of thiosemicarbazones was employed. These studies will also aid in the understanding of developing inhibitory moieties that could be explored in combination with existing inhibitors.

Cathepsin L studies were carried out using a 96-well microplate fluorometric based assay. The major advantage of this type of assay is the duration of the assay without compromising the reproducibility of the results. Microplate-based assays are commonly used in drug discovery; furthermore, high throughput screening is also available where hundreds or thousands of samples can be evaluated simultaneously. In general, cysteine proteases, including cathepsins L, K, and cruzain, have the highest catalytic activity under acidic conditions. Thus, NaOAc was selected due to its stability, effective and optimal pH range between 3.7 and 5.6.[5,320] EDTA functions as a chelator and prevents any interaction of possible metal ion traces found in aqueous solution such as mercury with the target protein. DTT is a powerful reducing agent. It is frequently used in enzyme kinetics because it has the ability to protect thiol groups that are essential in cysteine proteases. It can also reduce disulfide bonds by undergoing self oxidation processes by difulfide exchange. Brij 35 is a nonionic detergent that is used for its ability to prevent possible protein aggregation and adherence to containers. DMSO is a universal solvent used to solubilize reagents with low aqueous solubility such as the substrate and thiosemicarbazone derivatives. Z-FR-AMC is a fluorogenic substrate that has been used with serine and cysteine proteases. The lists of enzymes that can cleave this substrate include serine proteases, cathepsins, kallikrein, and plasmin.[321,322] The basis of the fluorescent assay of cysteine proteases is their ability to release AMC from

Z-FR-AMC, a nonfluorescent substrate. AMC is a compound that fluorescences when it is excited at 354 nm (excitation: 354 nm; emission: 442 nm).[323] Final conditions for the cathepsin L fluorometric assay were: 100 mM NaOAc pH 5.5, 1 mM EDTA, 3 mM DTT, 0.01% Brij 35, 0.7 - 2% DMSO, 1 nM cathepsin L, and 0.5 - 50 μM Z-FR-AMC.

*Assay Optimization. Effect of DMSO on Cathepsin L Inhibition Assays*

Research related to cathepsin L has increased significantly over time. Numerous experimental conditions are described and all of them are customized based on the main objectives of the research. An important factor to consider is the effect of DMSO during enzymatic catalysis. Reports of cathepsin L assays range from 0.7% to 10% final concentration.[324,325] Thus, it was necessary to analyze the effect of DMSO concentration in our samples. Experiments were conducted using analogs **1**, and **156**, to study the effect of DMSO concentration (See Table 27)

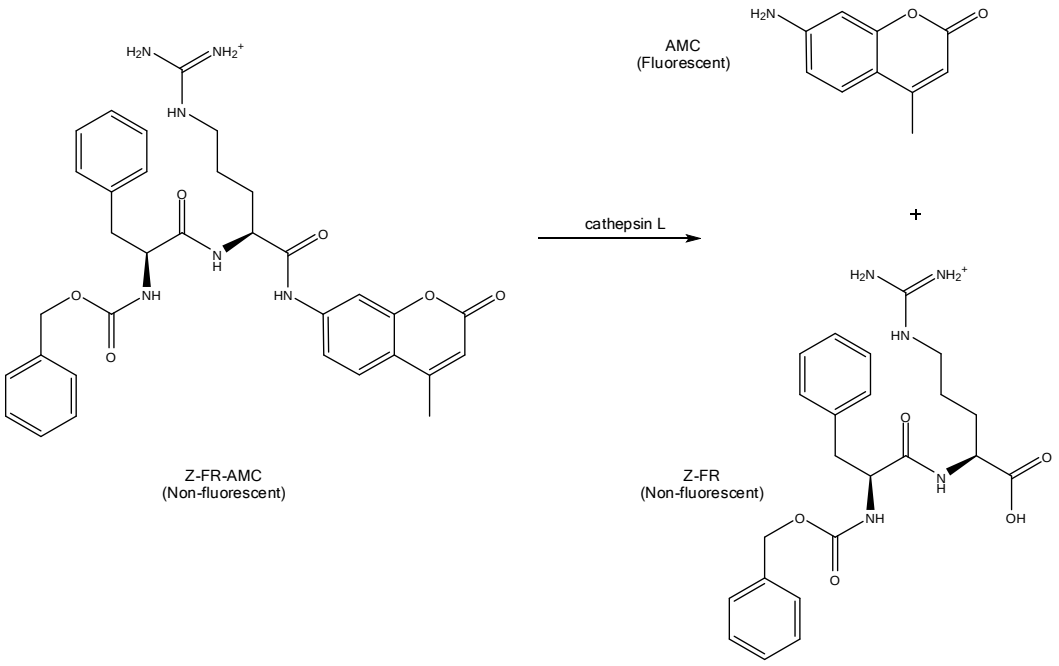
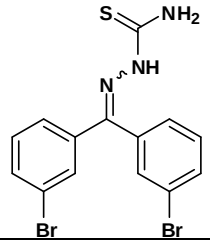
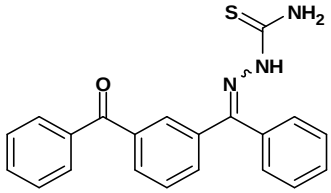



Figure 23. Hydrolysis of the Z-FR-AMC using cathepsin L

Table 27. Effect of DMSO Concentration on Inhibition Studies

| DMSO (%) |  |  |
|---|---|---|
| | **1** | **156** |
| | IC$_{50}$ (nM) | |
| 0.7 | 70 | ND |
| 2 | 16.7 | 9.85 |
| 3 | 6.64 | 7.7 |

[Z-FR-AMC]: 50 µM; preincubation time: 5 min

The obtained results showed the effect of DMSO on IC$_{50}$ values of cathepsin L inhibitors. Solubility of the lead compounds (**1** and **156**) was increased upon addition of DMSO and their inhibitory activity was enhanced accordingly. The efficiency of **1** is 6 times better when comparing its IC$_{50}$ at 0.7 and 2%, an overall of 2.85 fold increment of DMSO in the enzymatic assay. (IC$_{50}$= 70 and 11.8 nM respectively). On the other hand, the effect of the solvent is not as evident when comparing IC$_{50}$ values for **156**. (IC$_{50}$= 9.85 and 7.7 nM). It was also found the stability of the enzyme activity was also increased with higher concentrations of DMSO.

*Construction of AMC Calibrations Curves*

Fluorometric based assays were used to study the potency, mechanism and mode of inhibition of thiosemicarbazones as cathepsin L inhibitors. The principle of the assay relies on the ability of cathepsin L to cleave AMC, a fluorophore, from Z-FR-AMC, a nonfluorescent synthetic substrate that is extensively used in research involving cysteine cathepsins such as cathepsin L and K.

The construction of AMC curves was necessary in order to convert the relative fluorescence units into the concentration of the fluorophore. The curve allowed the validation of enzymatic activities, kinetic and inhibition assays. Also, kinetic parameters, velocities and rates could be compared to standard concentration units that can be found in the literature. Figure 24 and Figure 25 show the results of the construction of AMC curves. Results of the second linear regression give an equation that can correlate RFU and AMC concentration. The equation is RFU= 188.6 AMC (in μM) + 30.4 and the coefficient of determination ($r^2$) were also calculated (0.9931).

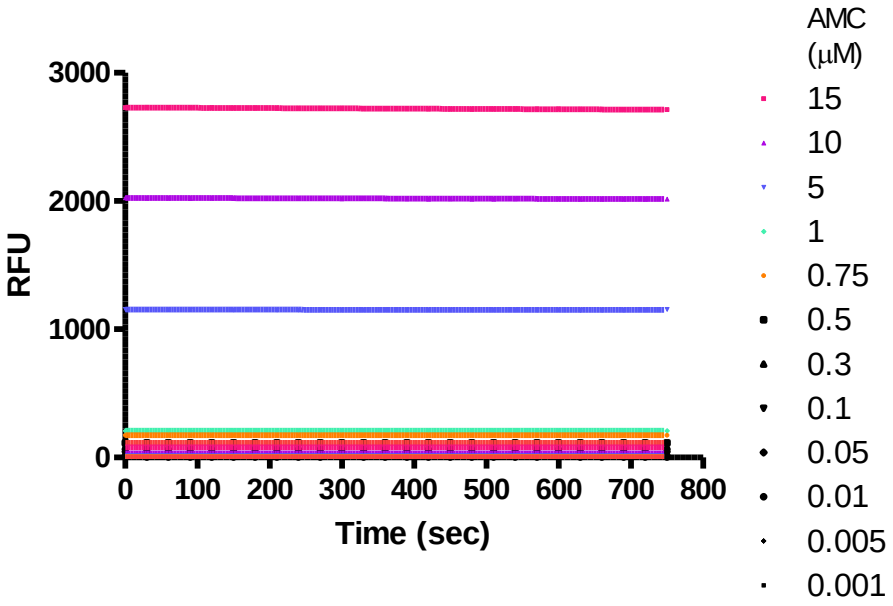



Figure 24. Fluorescence Response AMC vs. Time

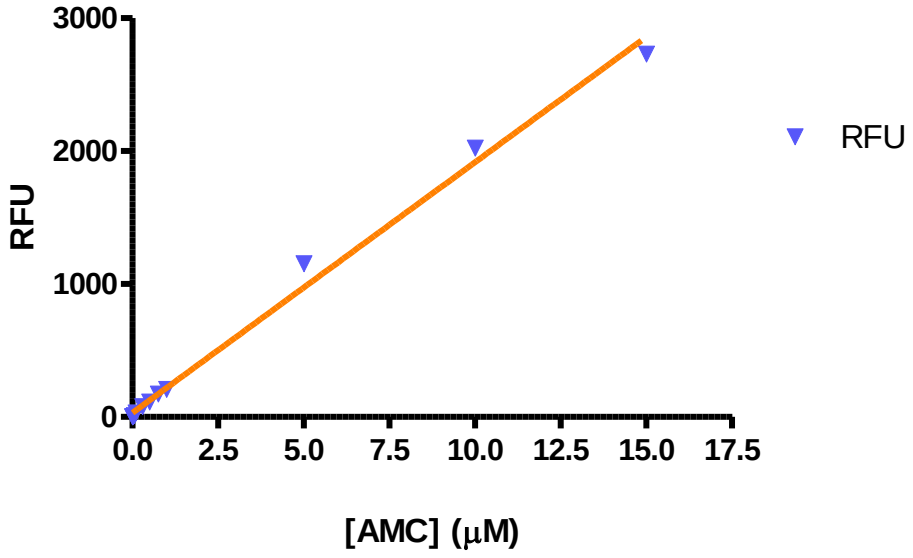



Figure 25. AMC Calibration Curve

*Determination of $K_M$, $V_{MAX}$ and $k_{CAT}$*

Determination of kinetic constants was based on the steady-state analysis assumption. A large concentration of the substrate (up to 15-fold $K_M$ value) was used for studies related to cathepsin L. Cathepsins L catalytic activity showed a linear behavior even at low substrate concentrations (0.2 μM). The cleavage of the substrate (Z-FR-AMC was monitored for at least 5 minutes. Complete details of this experiment can be found in the Material and Methods Section of this chapter. The determination of $K_M$, $V_{MAX}$ and $k_{CAT}$ was possible with experiments that observe the catalytic activity of a fixed concentration of cathepsin L (1 nM) but vary the concentration of the substrate Z-FR-AMC (0.2 -14 μM). Experiments were carried out in triplicate. Catalytic rates were calculated by applying linear regression of the data. ([AMC] is the dependent variable and time (seconds) is the independent variable). Then, a nonlinear regression analysis of the Michaelis-Menten equation (Eq. 1.1) was performed to calculate $K_M$ and $V_{MAX}$ values with the aid of commercially available software (GraphPad 5.0). The $k_{CAT}$ constant value

was determined using equation 2.1. The velocity $v_o$ is the initial rate velocity at a specific substrate concentration. The $V_{MAX}$ is the maximum velocity, $K_M$ is the Michaelis-Menten constant, [S] is the substrate concentration, and $k_{CAT}$ is the catalytic rate constant of the reaction.

$$v_o = \frac{V_{MAX}[S]}{K_M + [S]} \qquad (1.1) \qquad\qquad k_{CAT} = \frac{V_{MAX}}{[cathepsin\ L]} \qquad (2.1)$$

$K_M$, $V_{MAX}$ and $k_{CAT}$ values were found to be 1.1 ± 0.17 μM, 1.56 ± 0.04 AMC nM/s and 1.56 $s^{-1}$, respectively.[326]

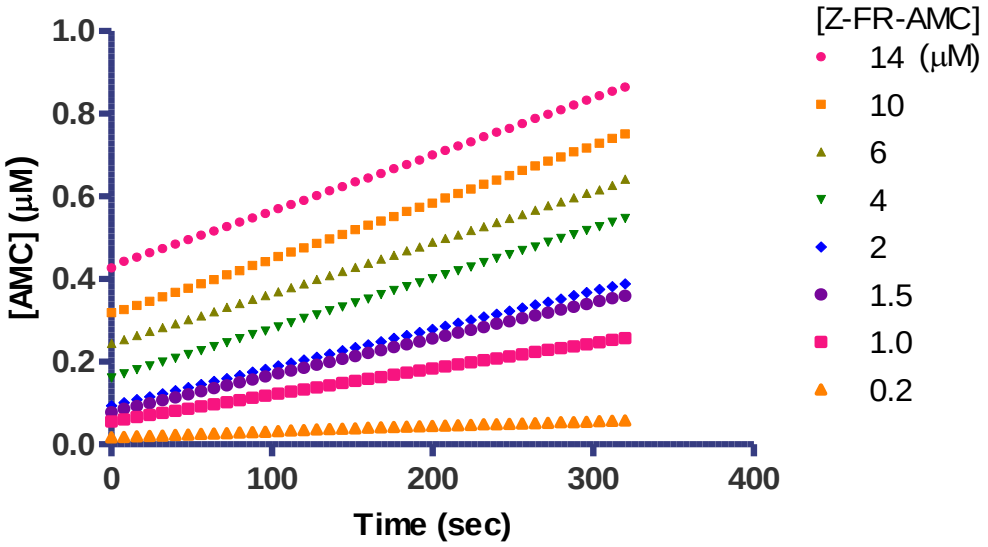



Figure 26. Catalytic Activity of Cathepsin L Using Z-FR-AMC as the Fluorogenic Substrate

*Determination of Inhibitory Efficacy of Thiosemicarbazone Analogs at 10 μM*

The first set of experiments was performed at a set concentration (10 micromolar) of individual compounds to verify if they were cathepsin L inhibitors. If the catalytic activity of the cysteine protease was inhibited by 50% or more by a fixed concentration (final concentration: 10 μM) of the potential inhibitor. Three independent sets of

experiments of untreated ([I]: 0 μM) and treated samples ([I]:10 μM) were preincubated with 1 nM cathepsin L for 5 minutes at 25 °C.

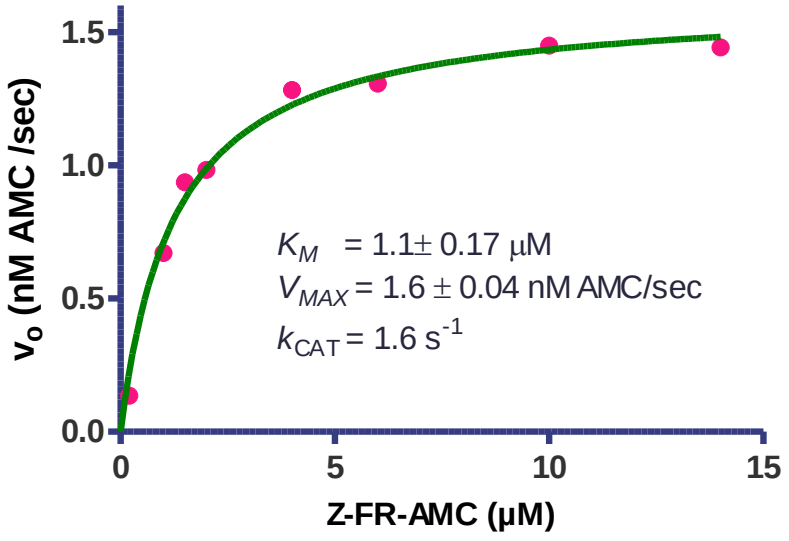



Figure 27. Determination of $K_M$ and $V_{MAX}$ for Human Cathepsin L Using Z-FR-AMC as a Fluorogenic Substrate

Inhibitory activities were monitored when reactions were started by adding 50 μM Z-FR-AMC as a fluorogenic substrate. Reactions demonstrated a linear behavior for the first five minutes. Catalytic rates of uninhibited and inhibited samples were calculated by linear regression of the data ([AMC]: dependent variable and time (seconds): independent variable). Compounds that exhibited inhibition of more than 50% cathepsin L activity at 10 μM in solution were further analyzed to determine their exact $IC_{50}$ values. If the ratio $v_i/v_o \leq 0.5$, then the compounds were not considered potential inhibitors and an approximated $IC_{50}$ value ≤ 10000 nM was assigned to them.

*Determination of $IC_{50}$ Values*

A library of more than one hundred fifty thiosemicarbazone analogs was analyzed to verify their efficacy to inhibit the catalytic activity of cathepsin L. The synthetic

compounds were synthesized under the guidance of Dr. Kevin G. Pinney by several members of his research group. The synthesis of these compounds has been previously published.[3,6,7,10–12] A 96-well microplate fluorometric based assay was utilized to determine the inhibitory activity of each of these inhibitors. Uninhibited cathepsin L catalytic activity showed linear behavior when 50 μM Z-FR-AMC was used for reactions times that were 5 minutes long. The determination of the $IC_{50}$ values was carried out with experiments that observed the inhibitory capacity of the synthetic compounds when a fixed concentration of cathepsin L (1 nM) was preincubated for 5 minutes at 25 °C. The final concentration of each compound varied between 10 pM and 10 μM. Experiments were carried out in triplicate. Catalytic rates of uninhibited and inhibited samples were calculated by linear regression of the data. ([AMC]: dependent variable and time (seconds): the independent variable). Data followed a typical sigmoidal dose response, and therefore, a nonlinear regression of the equation 1.2 was performed to calculate $IC_{50}$ values with the aid of commercially available software (GraphPad 5.0). The value Y represents the inhibited activity (normalized relative to control), X is log([inhibitor]) in *M*. The velocities $v_{min}$ and $v_{max}$ represent cathepsin L preincubation with the highest and lowest inhibitor concentrations respectively (10 μM and 0 pM or control). The Hillslope value is the slope of the sigmoidal curve. $IC_{50} \pm$ S.E. values represent the average and standard errors of at least three experiments. Structure-activity relationship (SAR) is shown in Tables 28 to 45 that group these inhibitors by functional groups or families.

$$Y = \frac{v_{MIN} + (v_{MAX} - v_{MIN})}{1 + 10^{(\log(IC_{50} - X) * Hillslope)}} \tag{1.2}$$

Table 28. Inhibition of Human Cathepsin L by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12]

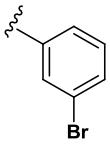
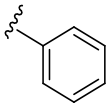
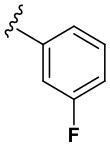
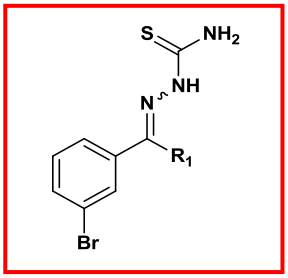
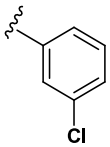
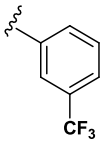
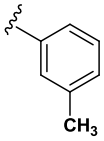
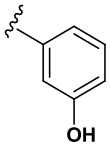
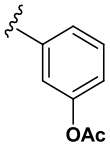
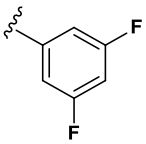
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E, (nM)* |
|---|---|---|---|
| **1** | |  | 16.7 ± 1.0 |
| **2** | |  | ≥ 10000 |
| **3**[8] | |  | 832.9 ± 16.9 |
| **4** | |  | 250.3 ± 18.9 |
| **5** |  |  | 131.4 ± 4.0 |
| **6** | |  | 46.5 ± 2.8 |
| **7** | |  | 224.4 ± 2.8 |
| **8** | |  | 188.7 |
| **9** | |  | 150.8 ± 16.0 |
| **10** | |  | 59.4 ± 0.7 |

Table 28. Inhibition of Human Cathepsin L by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

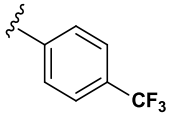
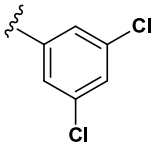
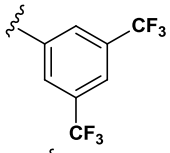
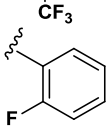
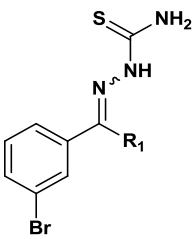
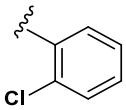
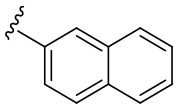
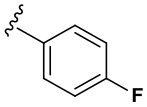
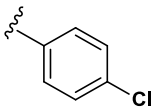
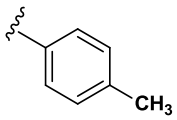
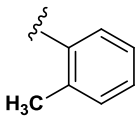
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **11** | |  | 520.9 ± 24.6 |
| **12** | |  | 415.2 ± 35.6 |
| **13** | |  | 96.0 ± 4.0 |
| **14** | |  | 23.83 ± 0.9 |
| **15** |  |  | 1610.0 ± 150.0 |
| **16** | |  | 1200.0 ± 130.0 |
| **17** | |  | 79.6 ± 2.7 |
| **18** | |  | 327.1 ± 13.8 |
| **19** | |  | 2156.0 ± 70 |
| **20** | |  | ≥ 10000 |

Table 28. Inhibition of Human Cathepsin L by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

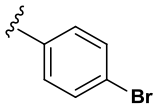
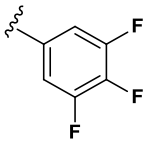
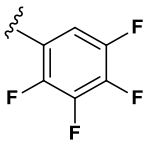
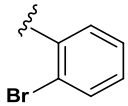
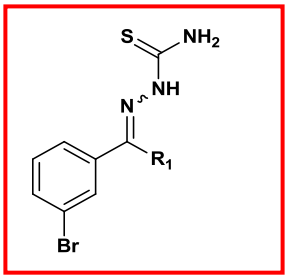
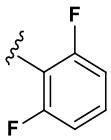
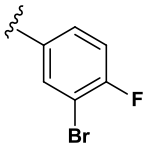
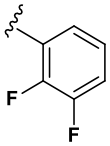
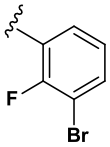
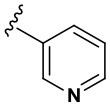
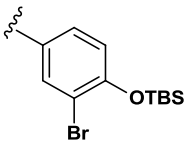
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.), ( nM)* |
|---|---|---|---|
| **21** | |  | ≥ 10000 |
| **22** | |  | 118.0 ± 4.0 |
| **23** | |  | 63.2 ± 4.6 |
| **24** | |  | 2600.0 ± 546 |
| **25** |  |  | 609.7 ± 48 |
| **26** | |  | 113.6 ± 10.8 |
| **27** | |  | 83.3 ± 3.3 |
| **28** | |  | 233.3 ± 29.0 |
| **29** | |  | 1000 ± 90 |
| **30** | |  | >10000 |

Table 28. Inhibition of Human Cathepsin L by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

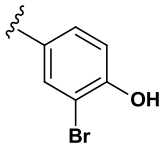
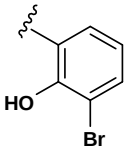
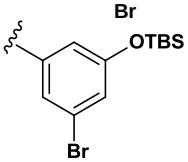
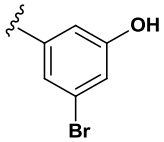
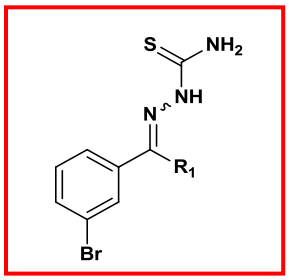
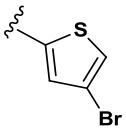
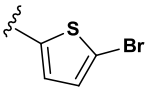
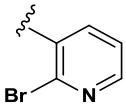
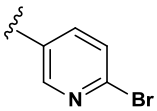
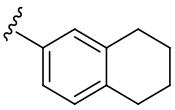
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **31** | |  | 126.1 ± 0.6 |
| **32** | |  | ≥ 10000 |
| **33** | |  | ≥ 10000 |
| **34** | |  | 232.4 ± 19.3 |
| **35** |  |  | 814.5 ± 78.9 |
| **36** | |  | 931.3 ± 124.5 |
| **37** | |  | 369.7 ± 7.8 |
| **38** | |  | ≥ 10000 |
| **39** | |  | ≥ 10000 |
| **40** | |  | 9650 ± 1.65 |

Table 28. Inhibition of Human Cathepsin L by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

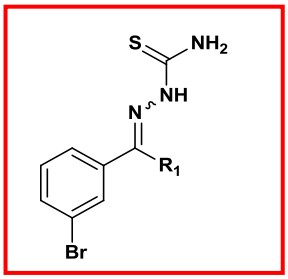
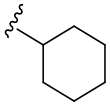
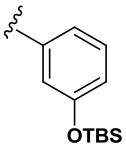
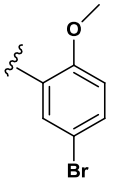
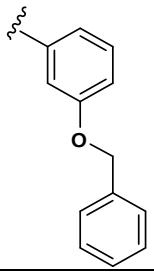
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **41** |  |  | ≥ 10000 |
| **42** | |  | ≥ 10000 |
| **43** | |  | ≥ 10000 |
| **44** | |  | ≥ 10000 |

Table 29. Inhibition of Human Cathepsin L by *meta*-Substituted Propanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

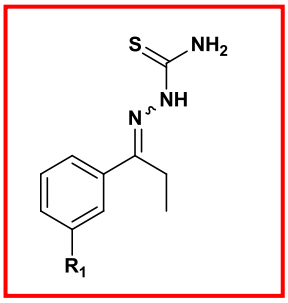
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **45** |  | $NO_2$ | 1905 ± 206 |
| **46** | | $NH_2$ | ≥ 10000 |

Table 30. Inhibition of Human Cathepsin L by *para*-Bromo Functionalized Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

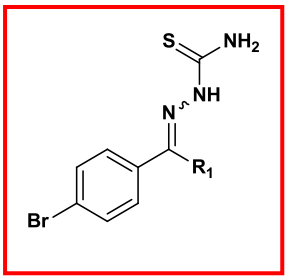
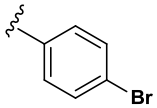
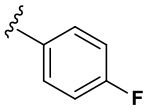
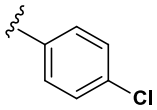
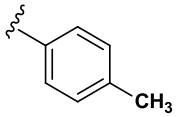
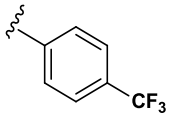
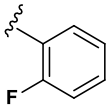
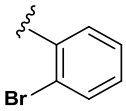
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **47** |  |  | ≥ 10000 |
| **48** | |  | 3320 ± 260 |
| **49** | |  | ≥ 10000 |
| **50** | |  | 2828 ± 92 |
| **51** | |  | ≥ 10000 |
| **52** | |  | 2423 ± 576 |
| **53** | |  | ≥ 10000 |

Table 31. Inhibition of Human Cathepsin L by Dihalogen-substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

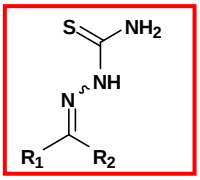
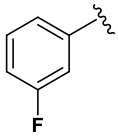
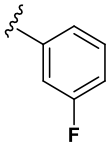
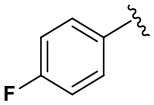
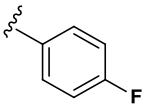
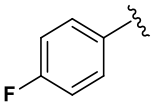
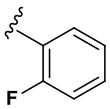
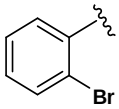
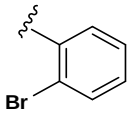
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **54** |  |  | 4871 ± 89 |
| **55** |  |  | ≥ 10000 |
| **56** |  |  | 2458 |
| **57** |  |  | ND |

Table 32. Inhibition of Human Cathepsin L by 3,3'-Dibromo-*N*-Substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

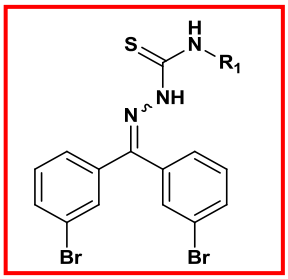
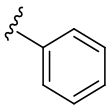
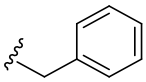
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **58** |  |  | ≥ 10000 |
| **59** | |  | ≥ 10000 |
| **60** | |  | ≥ 10000 |

Table 33. Inhibition of Human Cathepsin L by Thiosemicarbazones Containing a Phenyl Group. For Synthesis of Compounds: [3,6,7,10–12]

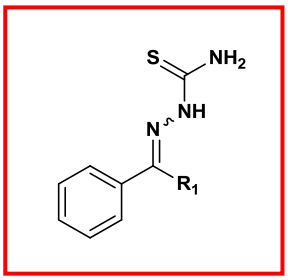
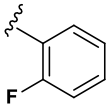
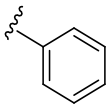
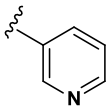
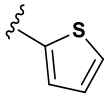
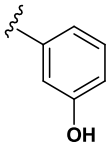
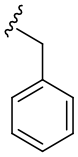
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **61** |  |  | ≥ 10000 |
| **62** | |  | ≥ 10000 |
| **63** | |  | ≥ 10000 |
| **64** | |  | ≥ 10000 |
| **65** | |  | ≥ 10000 |
| **66** | |  | 5410 ± |

Table 34. Inhibition of Human Cathepsin L by Di-halogenated or Monohalogenated Ketones. For Synthesis of Compounds: [3,6,7,10–12]

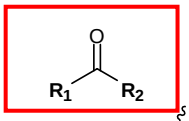
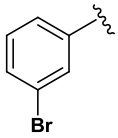
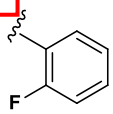
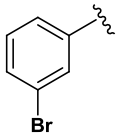
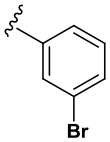
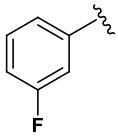
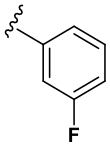
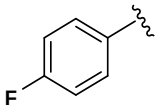
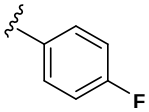
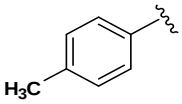
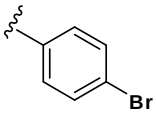
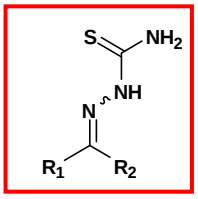
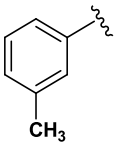
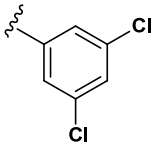
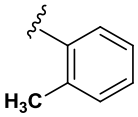
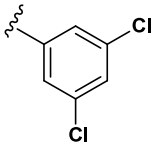
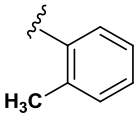
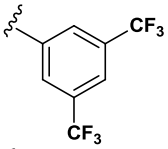
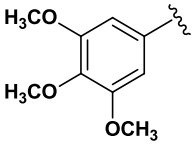
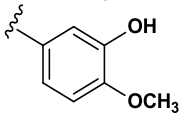
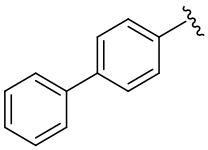
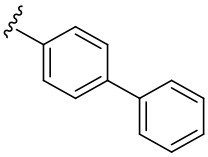
| *Compound* | *$X_1$* | *$X_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **67** |  |  | ≥ 10000 |
| **68** |  |  | ≥ 10000 |
| **69** |  |  | ≥ 10000 |
| **70** |  |  | ≥ 10000 |
| **71** |  |  | ≥ 10000 |

Table 35. Inhibition of Human Cathepsin L by Substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **72** |  |  | ≥ 10000 |
| **73** |  |  | 86.4 ± 6.0 |
| **74**[327] |  |  | 101.4 ± 4.8 |
| **75** |  |  | 3600 ± 250 |
| **76** |  |  | ND (Insoluble) |

*ND: Not Determined

Table 36. Inhibition of Human Cathepsin L by Functionalized Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

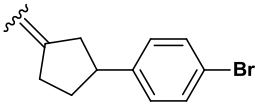
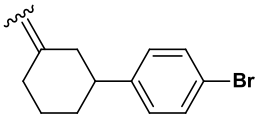
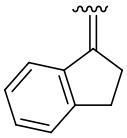
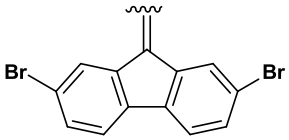
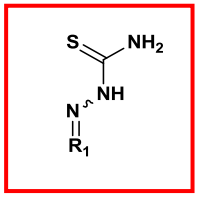
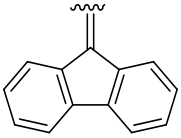
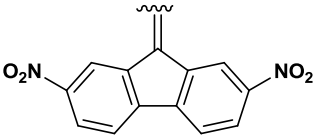
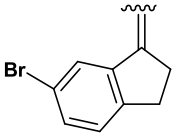
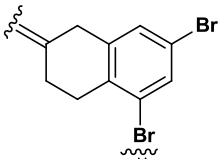
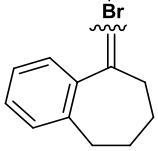
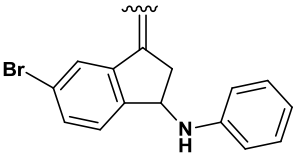
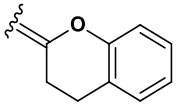
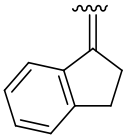
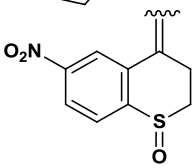
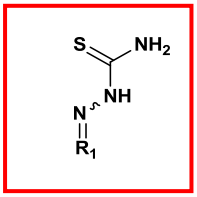
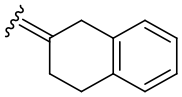
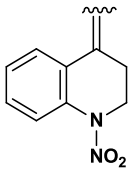
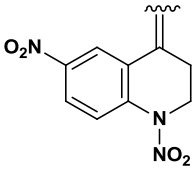
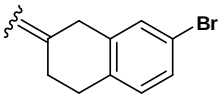
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **77** | |  | ≥ 10000* |
| **78** | |  | ≥ 10000* |
| **79** | |  | ≥ 10000* |
| **80** | |  | ≥ 10000 |
| **81** |  |  | ≥ 10000 |
| **82** | |  | ≥ 10000 |
| **83** | |  | ≥ 10000 |
| **84** | |  | ≥ 10000* |
| **85** | |  | ≥ 10000 |
| **86** | |  | ≥ 10000 |

Table 36. Inhibition of Human Cathepsin L by Functionalized Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12](Continued)

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **87** | |  | ≥ 10000* |
| **88** | |  | ≥ 10000* |
| **89** | |  | ≥ 1740 ± 50 |
| **90** |  |  | ≥ 10000 |
| **91** | |  | ND |
| **92** | |  | 228.4 ± 7.2 |
| **93** | |  | ≥ 10000* |

* [Z-FR-AMC]: 5 μM; ND: not determined

Table 37. Inhibition of Human Cathepsin L by Functionalized Annulone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

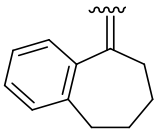
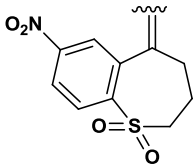
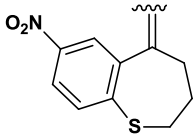
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **94** |  |  | ≥ 10000 |
| **95** | |  | ≥ 10000 |
| **96** | |  | 658.2 ± 114.0 |

Table 38. Inhibition of Human Cathepsin L by Substituted Quinolone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

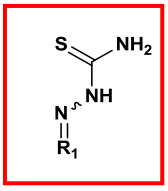
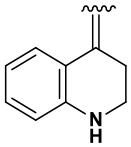
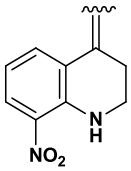
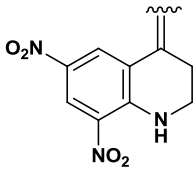
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **97** |  |  | ≥ 10000 |
| **98** | |  | ≥ 10000 |
| **99** | |  | 337.9 ± 9.6 |

Table 38. Inhibition of Human Cathepsin L by Substituted Quinolone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

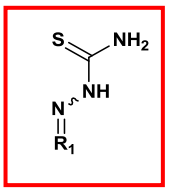
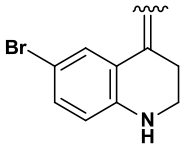
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **100** |  |  | 163.8 ± 16.4 |

Table 39. Inhibition of Human Cathepsin L by Non-Thiosemicarbazone Based Analogs. For Synthesis of Compounds: [3,6,7,10–12]

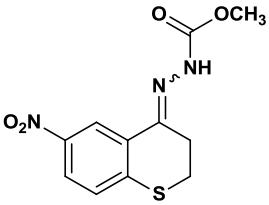
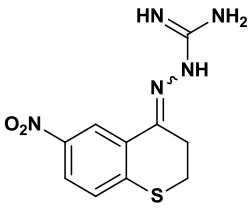
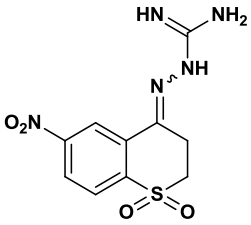
| *Compound* | *Structure* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|
| **101** |  | ≥ 10000 |
| **102** |  | ≥ 10000 |
| **103** |  | ≥ 10000 |

Table 39. Inhibition of Human Cathepsin L by Non-Thiosemicarbazone Based Analogs. For Synthesis of Compounds: [3,6,7,10–12](Continued)

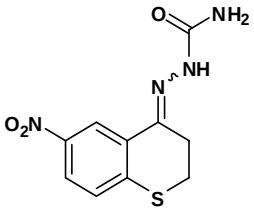
| *Compound* | *Structure* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|
| **104** |  | ≥ 10000 |

Table 40. Inhibition of Human Cathepsin L by Substituted Tetralone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

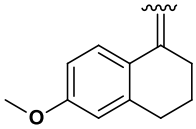
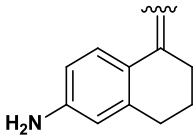
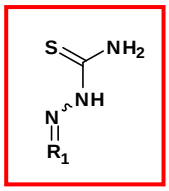
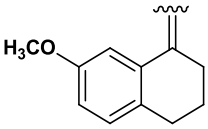
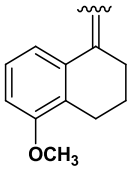
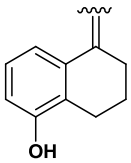
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **105** | |  | ≥ 10000 |
| **106** | |  | ≥ 10000 |
| **107** |  |  | ≥ 10000 |
| **108** | |  | ≥ 10000 |
| **109** | |  | ≥ 10000 |

Table 41. Inhibition of Human Cathepsin L by Functionalized Chromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

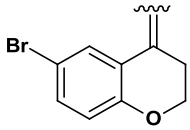
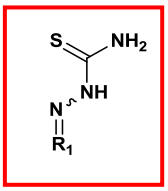
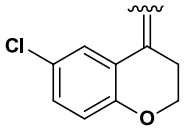
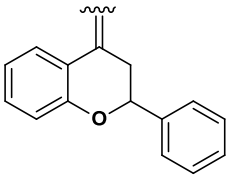
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **110** | |  | 303.0 ± 24.7 |
| **111** |  |  | 369.4 ± 4.5 |
| **112** | |  | ≥ 10000 |

Table 42. Inhibition of Human Cathepsin L by Substituted Thiochromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

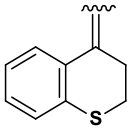
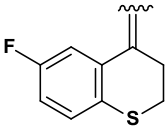
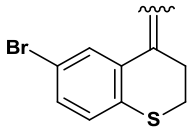
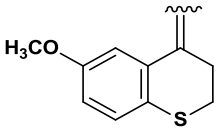
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **113** | |  | ≥ 10000* |
| **114** |  |  | 741.8 ± 23.3 |
| **115** | |  | 152.3 ± 2.7 |
| **116** | |  | ≥ 10000 |

Table 42. Inhibition of Human Cathepsin L by Substituted Thiochromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12](Continued)

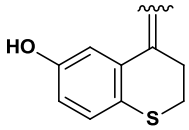
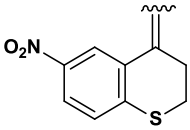
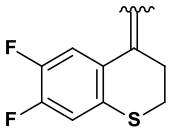
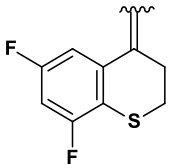
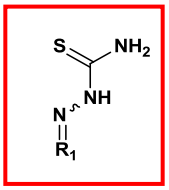
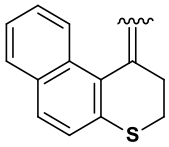
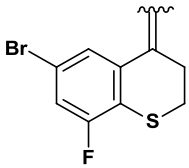
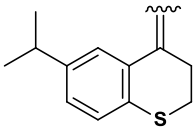
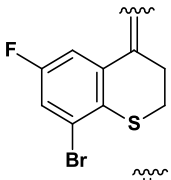
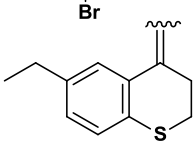
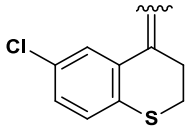
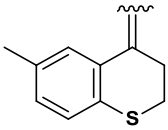
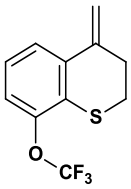
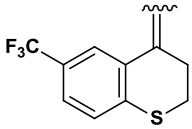
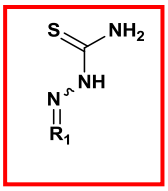
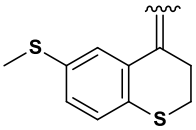
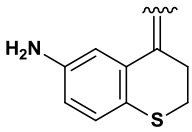
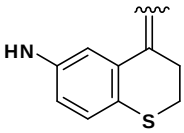
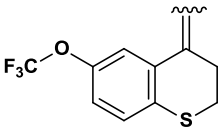
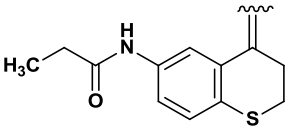
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **117** | |  | ≥ 10000 |
| **118** | |  | 67.85 ± 4.5 |
| **119** | |  | 54.0 ± 7.6 |
| **120** | |  | 1500 ± 100.1 |
| **121** |  |  | ≥ 10000 |
| **122** | |  | 434.2 ± 14.8 |
| **123** | |  | ≥ 10000 |
| **124** | |  | ≥ 10000 |
| **125** | |  | 2720 ± 240 |

Table 42. Inhibition of Human Cathepsin L by Substituted Thiochromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12](Continued)

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **126** | |  | 228.4 ± 11.6 |
| **127** | |  | 213.8 ± 16.4 |
| **128** | |  | ≥ 10000 |
| **129** | |  | 284.1 ± 10.7 |
| **130** |  |  | ≥ 10000 |
| **131** | |  | ≥ 10000 |
| **132** | |  | ≥ 10000 |
| **133** | |  | 255.5 ± 26.2 |
| **134** | |  | ≥ 10000 |

*[Z-FR-AMC]: 5 μM

Table 43. Inhibition of Human Cathepsin L by Substituted Sulfone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

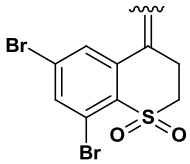
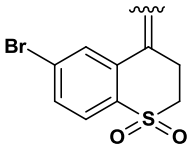
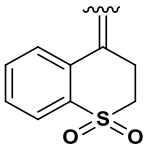
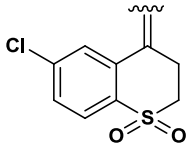
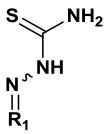
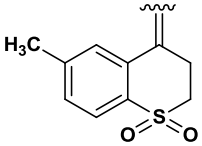
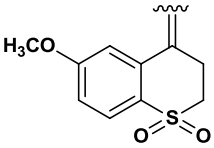
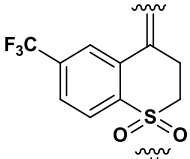
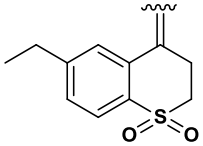
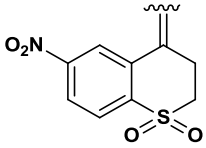
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **135** | |  | 355.9 ± 11.8 |
| **136** | |  | 573.9 ± 37.4 |
| **137** | |  | ≥ 10000* |
| **138** | |  | ND |
| **139** |  |  | ≥ 10000 |
| **140** | |  | ≥ 10000 |
| **141** | |  | 259.8 ± 34.4 |
| **142** | |  | 6521 ± 490 |
| **143** | |  | 112 |

Table 43. Inhibition of Human Cathepsin L by Substituted Sulfone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

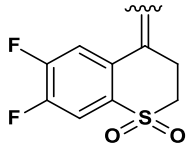
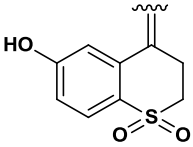
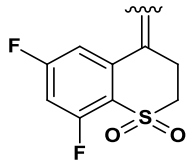
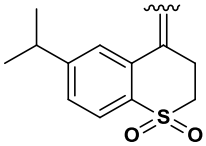
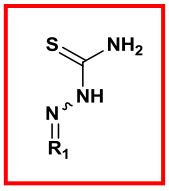
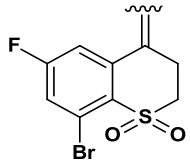
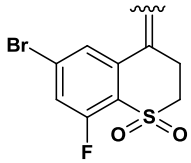
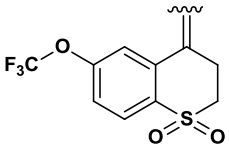
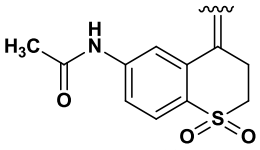
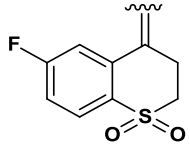
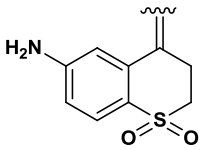
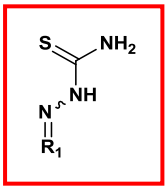
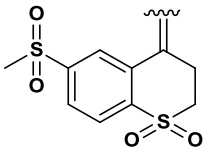
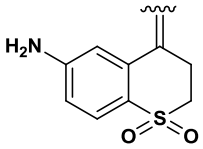
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **144** | |  | 3650 ± 378 |
| **145** | |  | ≥ 10000 |
| **146** | |  | ≥ 10000 |
| **147** | |  | ≥ 10000 |
| **148** |  |  | ≥ 10000 |
| **149** | |  | 1117 ± 102.7 |
| **150** | |  | 3960 ± 40 |
| **151** | |  | ≥ 10000 |
| **152** | |  | ≥ 10000 |

Table 43. Inhibition of Human Cathepsin L by Substituted Sulfone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **153** | |  | ≥ 10000 |
| **154** |  |  | ≥ 10000 |
| **155** | |  | ≥ 10000 |

*ND: 5 μM

Table 44. Inhibition of Human Cathepsin L by FunctionalizedBenzoyl-Benzophenone Thiosemicarbazones*

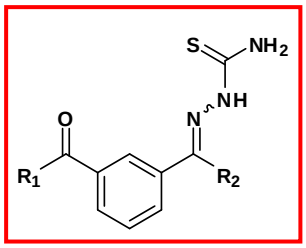
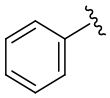
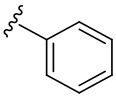
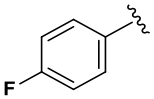
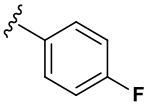
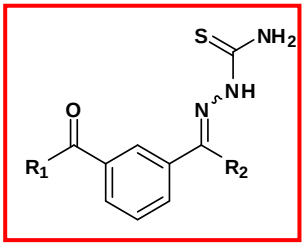
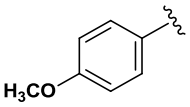
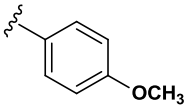
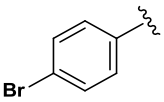
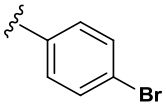
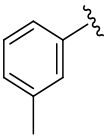
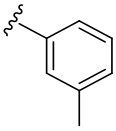
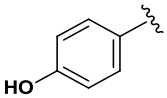
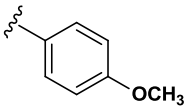
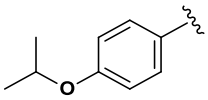
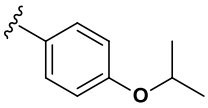
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **156** |  |  | 9.85 ± 0.60 |
| **157** |  |  | 24.3 ± 1.2 |

Table 44. Inhibition of Human Cathepsin L by FunctionalizedBenzoyl-Benzophenone Thiosemicarbazones* (Continued)

| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **158** |  |  | 586.9 ± 46.3 |
| **159** |  |  | ≥ 10000 |
| **160** |  |  | 7823 ± 736 |
| **161** |  |  | ≥ 10000 |
| **162** |  |  | ≥ 10000 |

*Compounds **156-162** have been synthesized by members of the Kevin G. Pinney laboratory (Baylor University)

Table 45. Inhibition of Human Cathepsin L by Substituted-Benzoyl-Benzophenone Thiosemicarbazones*

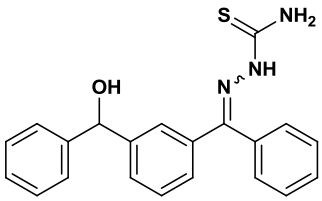
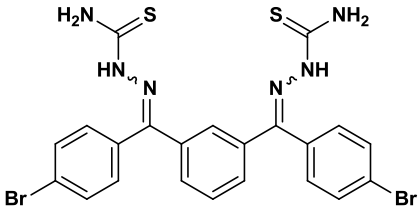
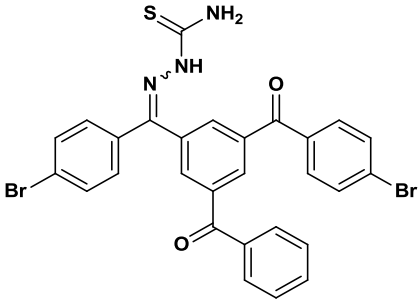
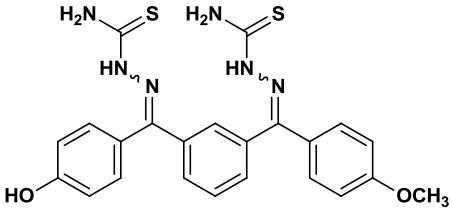
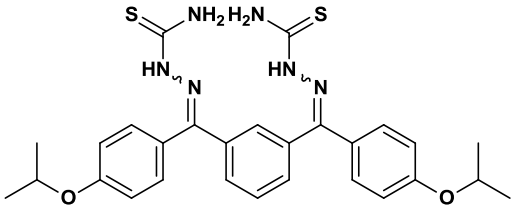
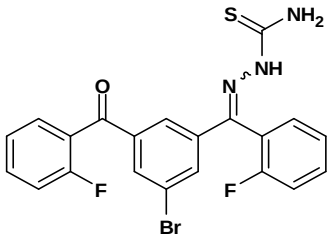
| *Compound* | *Structure* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|
| **163** |  | 23.8 ± 0.8 |
| **164** |  | ≥ 10000 |
| **165** |  | 25.0 ± 0.2.3 |
| **166** |  | ≥ 10000 |
| **167** |  | ≥ 10000 |

Table 45. Inhibition of Human Cathepsin L by Substituted-Benzoyl-Benzophenone Thiosemicarbazones* (Continued)

| *Compound* | | *$IC_{50}$ (± S.E.), nM* |
|---|---|---|
| **168** |  | 8.12 ± 0.5 |

*Compounds **163-168** have been synthesized by members of the Kevin G. Pinney laboratory (Baylor University)

*Structure-Activity Relationship (SAR) of Thiosemicarbazones as Cathepsin L Inhibitors*

More than 150 thiosemicarbazone (TSC) and ketone-derived synthetic compounds were analyzed during the preliminary screening in order to determinate the potency of these compounds. Approximately half of the samples showed a significant inhibitory potency, and $IC_{50}$ values were calculated. A selected group of eight different compounds were excellent cathepsin L inhibitors with $IC_{50}$ values between 8 and 55 nM. Eleven compounds demonstrated to be good inhibitors ($50 \leq IC_{50} \leq 100$ nM). One fifth of the library (36 analogs) had a moderate activity. $IC_{50}$ values of these synthetic compounds were ranged between 100 and 1000 nM.

*Potent cathepsin L inhibitors*. Figure 28 summarizes the structures of the compounds with $IC_{50}$ values less than 55 nM. Three analogs belong to the subfamily of *meta*-brominated benzophenone TSC analogs. Fluorine, bromine, and trifluoromethyl, are all halogenated moieties and the substituents in this subfamily. The second subgroup consists of four unsubstituted and substituted benzoyl benzophenone thiosemicarbazones.

Interestingly, the best two compounds found in the family of TSC are benzoyl-benzophenone TSC with $IC_{50}$ values less than 10 nM. The list of this series is completed with a 6,7-*di*fluoro thiochromanone and a benzoyl benzhydrol thiosemicarbazone.

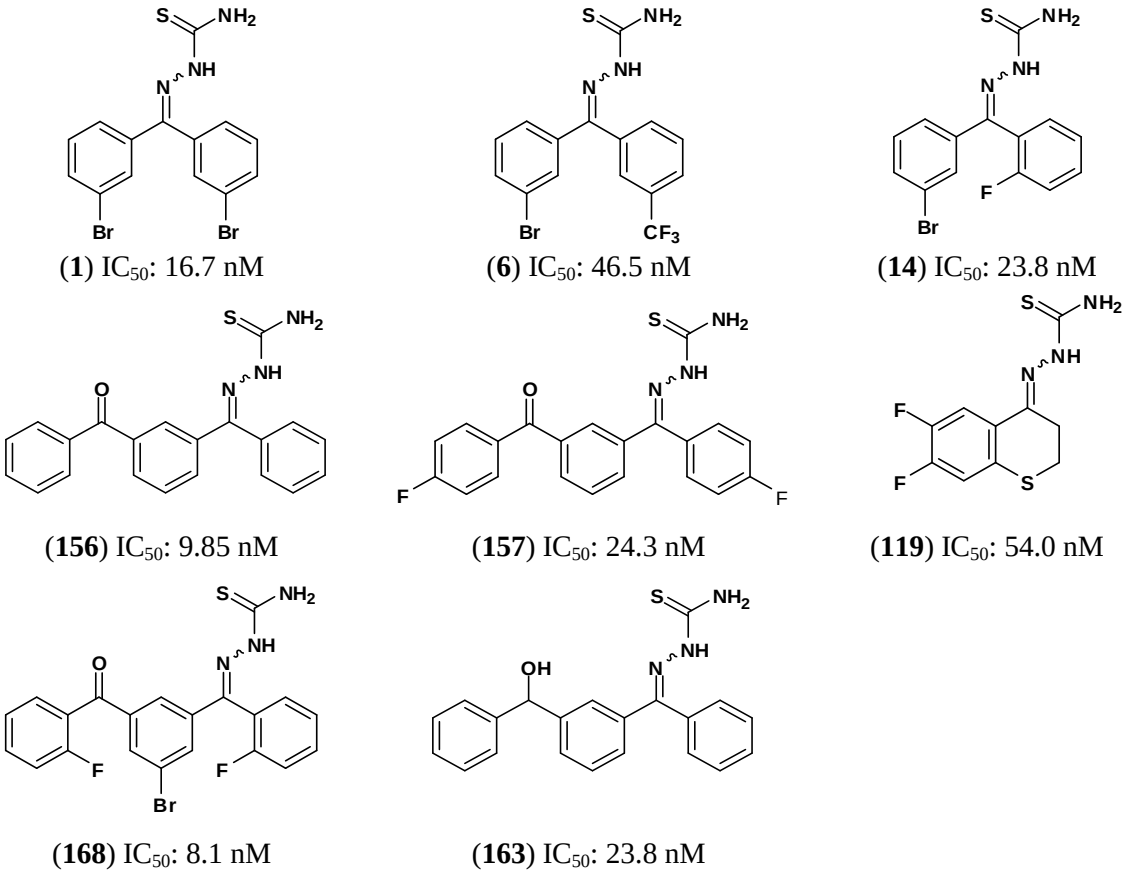



Figure 28. Thiosemicarbazone analogs with Potent Inhibitory Activity against Human Cathepsin L[9–12]

*Compounds with $IC_{50}$ less than 100 nM*. Eleven compounds proved to have good inhibitory activity (less than 100 nM). In addition to the eight compounds in Figure 28, these compounds included: 3-bromo-3',5'-*di*fluorobenzophenone TSC, 3-bromo-3',5'-*di*trifluoromethylbenzophenone TSC, 3-bromo-3'-fluorobenzophenone TSC, 3-bromo-

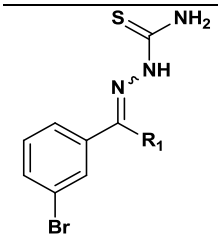
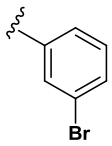
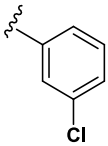
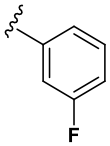
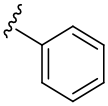
3',4',5',6'-fluorobenzophenone TSC, 2'-methyl-3',5'-difluoro benzophenone TSC, 6'-nitrothiochromanone TSC, 6'-nitrosulfone TSC, and four others.

*Compounds with moderate activity*. Thirty six TSC analogs showed $IC_{50}$ values between 100 and 1000 nM. The list included, but was not limited to: *meta*-brominated benzophenone TSCs, substituted tetrahydroquinoline TSCs halogenated chromanone TSCs, some naphthalene, sulfides, and sulfone TSCs and some substituted benzoyl benzophenone TSCs.

*General Remarks of the Structure-Activity Relationship*

Further analysis of the structure-activity relationship revealed that thiosemicarbazone inhibitory activity can be enhanced or reduced dramatically by moiety substitution. The series of substituted *meta*-brominated benzophenone thiosemicarbazones is the largest subgroup among the compounds that were tested. It consists of forty-four different analogs with inhibitory activities that varied between 16.7 and ≥ 10000 nM. Halogenated substituents greatly enhanced the activity of the thiosemicarbazones. Compounds **1**, **2**, **4**, **5** (Table 46) are thiosemicarbazones sharing the main moiety (monobromobenzophenone thiosemicarbazone) but varying the *meta*-halogenated substituent. Compound **2** is used as a comparison to complete the series. Analog **2** does not possess any inhibitory activity towards cathepsin L. The activity is inversely proportional to the atomic size of the substituent. The brominated analog is 15-fold more active than its fluorinated analog.

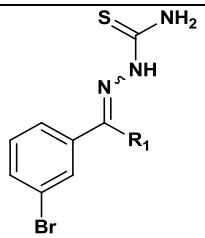
Table 46. Inhibition of Human Cathepsin L by 3-Bromo-3'-Halogen Benzophenone Thiosemicarbazones

| | *R₁* | | | |
|---|---|---|---|---|
|  |  | <  | <  | <  |
| | **1** | **5** | **4** | **2** |
| $IC_{50}$ (nM) | 16.7 | 131.4 | 250.3 | ≥ 10000 |

Five fluorinated-substituted benzophenone TSC analogs (**4**, **10**, **22**, **23**, and **25**; Table 47) were also compared. Polyfluorination of the benzyl rings indicates that cathepsin L-inhibitory activity of the compounds can be increased with the addition of the electronegative fluoro-substituents. However, a deeper comparison between the two difluoro benzophenone TSCs also reveals that position of the substitutions plays a key role in the activity of these analogs. The *ortho*-difluoro analog is 10 times less active than the di*meta* counterpart.

The effect of heteroatomic substituents was also evaluated (Table 48). Various groups were substituted at the *meta*- position of one of the phenyl rings. The group includes trifluoromethyl, methyl, hydroxyl, and acetate moieties. Three compounds showed modest activities compared to the trifluoromethyl analog, which is almost five times more potent than the methylated analog. Notably, compound **8** has improved solubility characteristics and was further investigated.

Table 47. Inhibition of Human Cathepsin L by 3-Bromo-Fluorinated-benzophenone Thiosemicarbazones



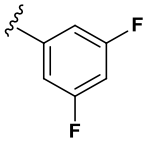
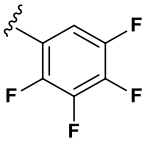
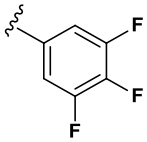
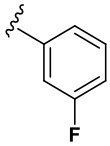
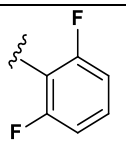
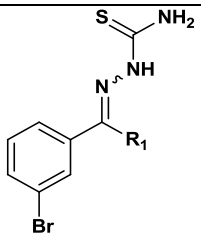
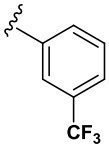
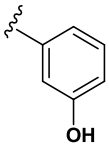
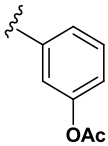
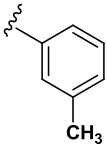
| | | | | | |
|---|---|---|---|---|---|
| *$R_1$* |  | >  | >  | >  | >  |
| | **10** | **23** | **22** | **4** | **25** |
| $IC_{50}$ (nM) | 59.4 | 63.2 | 118.1 | 250.3 | 609.7 |

Table 48. Inhibition of Human Cathepsin L by 3-Bromo-3'-Heteroatomic groups Benzophenone Thiosemicarbazones

| | *$R_1$* | | | |
|---|---|---|---|---|
|  |  < |  < |  < |  |
| | **6** | **8** | **9** | **7** |
| $IC_{50}$ (nM) | 46.5 | 131.4 | 150.8 | 224.4 |

A small trifluoromethyl series of three compounds (**6**, **11**, and **13**; Table 49) also revealed that substituted compounds often show better activity against cathepsin L when compared to monosubstituted. *Meta*-compounds are also better cathepsin L inhibitors than *para*- compounds. Additionally, the position of the substituent enhances or decreases the activity of these synthetic compounds.

Table 49. Inhibition of Human Cathepsin L by 3-Bromo-Trifluoromethyl Benzophenone Thiosemicarbazones

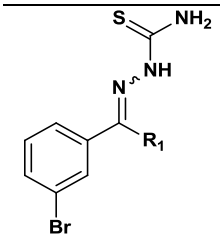
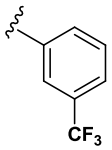
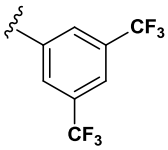
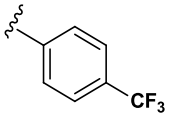
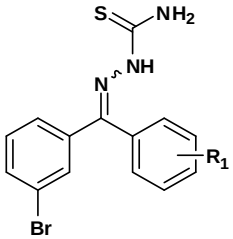
| | *R1* | | |
|---|---|---|---|
|  |  < |  < |  |
| | **6** | **13** | **11** |
| $IC_{50}$ (nM) | 46.5 | 96.0 | 520.9 |

Table 50. Effect of the Position of Substituents in the Inhibitory Activity of 3-Bromo-Benzophenone Thiosemicarbazones



| *R1* | *ortho* | *meta* | *para* |
|---|---|---|---|
| -F | (**14**) 23.9 | (**4**) 250.3 | (**17**) 79.6 |
| -Cl | (**15**) 1610.0 | (**5**) 131.4 | (**18**) 327.1 |
| -Br | (**24**) 2600 | (**1**) 16.7 | (**21**) ≥ 10000 |
| $CH_3$ | (**20**) ≥ 10000 | (**7**) 224.4 | (**19**) 2156.0 |

However, the effect of position of the substituent can be analyzed in depth with the series of twelve compounds that are monosubstituted with halogen and aliphatic groups (F, Cl, Br, and $CH_3$). Table 50 compares $IC_{50}$ for these compounds. The inhibitory activity of the compounds increases with more electronegative substituents if they are positioned in the *ortho* position. On the other hand, the trend is the opposite for halogen based the substituents are found in the *meta* positions. Furthermore, the *meta*-brominated analog is the third most potent inhibitor found in this library. In general, *meta*-substituted TSCs analogs are better inhibitors than *ortho*- and *para*-substituted TSCs with the exception of fluoro-substituted benzophenone thiosemicarbazones.

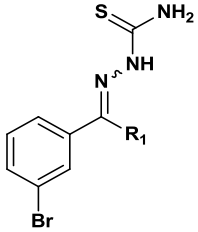
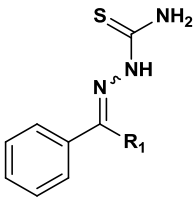
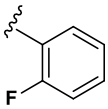
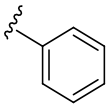
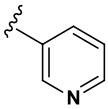
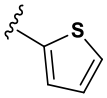
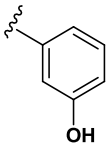
Compounds **61-65** also demonstrated the importance of the presence of bromine in one of the phenyl rings in order to enhance the potency of thiosemicarbazones. A comparison of unsubstituted and bromo-substituted analogs unsubstituted and halogenated benzophenones, phenols, pyridines, and thiophenes; revealed than none of the unsubstituted compounds showed significant inhibitory activity. Similar trends were observed with substituted hydroquinolines, and naphthalenes (Table 51).

The activity of TSCs was greatly enhanced by the bromination of the benzophenone in the *meta* position (up to 41 times more potent in the case of the fluorinated analogs) with the exception of the chlorinated derivative which showed no activity in either case (Table 52).

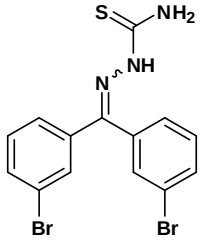
Bromination of the benzophenone group is also critical for the activity of TSC as potential cathepsin L inhibitors. A closer examination of halogenated compounds with similar chemical structures reveals that a different halogen substituent, such as fluorine, greatly reduces the activity of the compounds. The 3,3'-difluorobenzophenone TSC (**54**) is almost 300 times less active than 3,3'-dibromobenzophenone TSC (**1**). Furthermore, both 4,4'-difluorobenzophenone TSC (**54**) and 4,4'-dibromobenzophenone TSC (**47**) showed no activity against cathepsin L, showing that *para*-disubstitution with halogen substituents cannot be used as a possible route for the design of similar compounds.

Compounds **58**, **59**, and **60** (Table 32) demonstrated that modification in the thiosemicarbazone moiety was detrimental in the activity of the synthetic compounds. None of them showed significant activity towards cathepsin L.

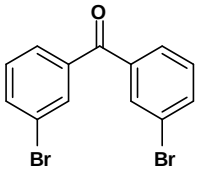
Table 51. Comparison between Brominated and Unbrominated Benzophenone, Thiophene and Pyridine Thiosemicarbazones for Cathepsin L

| $R_1$ |  |  |
|---|---|---|
|  | (**14**) 23.8 | (**61**) ≥ 10000 |
|  | (**2**) ≥ 10000 | (**62**) ≥ 10000 |
|  | (**29**) 1000 | (**63**) ≥ 10000 |
|  | ND | (**64**) ≥ 10000 |
|  | (**8**) 188.7 | (**65**) ≥ 10000 |

Screening of building blocks also demonstrated that TSC is one of the key features for a successful inhibition of cathepsin L. Compound **1**, the reference compound, is a potent inhibitor with an $IC_{50}$ value of 16.7 nM. On the other hand, the dibromosubstituted ketone did not show any activity under the same conditions (Figure 29). A similar situation was observed with 3-bromo-2'-fluoro TSC and its ketone counterpart..



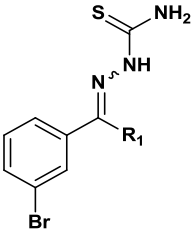
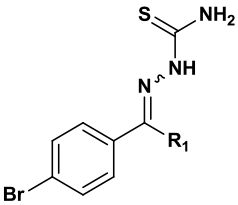
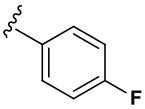
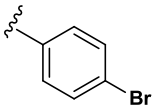
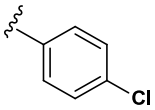
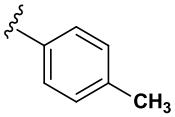

(**1**) $IC_{50}$: 16.7 nM




(**68**) ≥ 10000

Figure 29. Comparison between Inhibitory activities of **1** and **68**

Table 52. Comparison of the Inhibitory Activity between *meta*- and *para*-Bromine-substituted Benzophenone Thiosemicarbazones

| *R1* |  |  |
|---|---|---|
|  | (**17**) 79.6 | (**48**) 3320 |
|  | (**21**) 327.1 | (**47**) ≥ 10000 |
|  | (**18**) ≥ 10000 | (**49**) ≥ 10000 |
|  | (**19**) 2156.0 | (**50**) 4570 |

A series of non benzophenone thiosemicarbazone analogs was explored (Table 36). The series includes fluorenes, naphthalenes, indenes, annulenes, chromanones, aminoquinolines, and others. The majority of these derivatives were inactive compounds

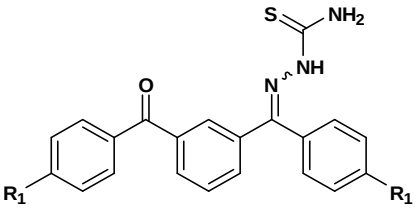
and the $IC_{50}$ values were not be determined. Only **88** and **92** were capable to inhibit cathepsin L.

Thiochromanones (Table 42) and sulfones (Table 43) are important subclasses of thiosemicarbazones that were screened in this library. Their $IC_{50}$ values ranged between 46.5 and 10000 nM. In general, sulfides showed better inhibitory activities when compared with sulfones. Small size substitution in the thiochromanone moiety in various positions series reveals that thiochromanone TSCs become better inhibitors with nitro-aromatic or *poly*-fluorinated groups. *Mono*-halogenation, on the other hand, does not improve their activity yielding modest activities. Once again, the position of the substituent plays a critical role in the activity of thiosemicarbazones. A comparison of two sets of halogenated sulfones and thiochromanones were compared. First, 6,7-difluorothiochromanone TSC (**119,** $IC_{50}$: 54 nM) is 200 times more efficient that its analog, a 6,8-*di*fluorothiochromanone TSC (**120** $IC_{50}$: 1500 nM). Additionally, 6,7-*di*fluorosulfone TSC TSC (**144,** $IC_{50}$: 3650 nM) is 3 times more potent than 6,8-*di*fluorosulfone TSC (**146,** $IC_{50} \geq$ 10000 nM). These observations also supported the effect of the substituent's position. Similarly, 6-fluoro-8-bromothiochromanone TSC (**124**) and 6-bromo-8-fluorothiochromanone TSC (**122**), two isomers, showed inhibitory activities that were ≥ 10000 and 434.2 nM, respectively.

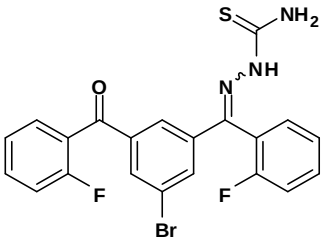
Lastly, the recent discovery of a new potent thiosemicarbazone with an $IC_{50}$ value of 9.85 nM (compound **156**), led to the synthesis of a new generation of thiosemicarbazones (unpublished results). Chemically, compound **156** is a benzoyl benzophenone thiosemicarbazone. The small subclass of TSCs that was screened is composed of thirteen analogs with similar structures. Unsubstituted and *para*-substituted

analogs were also examined (Table 53). The unsubstituted analog showed an extraordinary potency with an $IC_{50}$ less than 10 nM, followed by a difluorobenzoyl benzophenone TSC and diacetate analog (24.3 and 587 nM). Interestingly, the dibromobenzoyl benzophenone analog did not prove to be a good inhibitor for cathepsin L. The presence of two thiosemicarbazones moieties did not prove to be beneficial for the activity of these synthetic compounds.

Table 53. Inhibition of Human Cathepsin L by Difunctionalized Benzoyl Benzophenone Thiosemicarbazones

| | *$R_1$* | | | | |
|---|---|---|---|---|---|
|  | H **(156)** | F **(157)** | $OCH_3$ **(158)** | $OCH(CH_3)_2$ **(162)** | Br **(159)** |
| $IC_{50}$ (nM) | 9.9 | 24.3 | 587 | ≥ 10000 | |

Finally, a polyhalogenated benzoyl benzophenone thiosemicarbazone was one of the last compounds to be screened. Compound **168** showed an extraordinary potency of 8.1 nM, becoming the most potent cathepsin L inhibitor found in this library of thiosemicarbazone analogs.

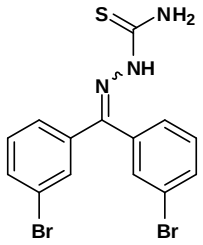


$IC_{50}$: 8.1 nM

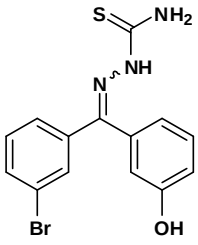
Figure 30. Chemical Structure of **168**

*Advanced Kinetic Studies*

Two compounds, **1** and **8** proved to be potent and good cathepsin L inhibitors ($IC_{50}$ values are 16.7 and 188.7 nM respectively) and were further characterized in order to understand their mechanism of action. Additionally, **8** offers the possibility of a derivatization that could increase the solubility of the thiosemicarbazone. Both compounds are structurally related. Their scaffold consists of a *para* brominated benzophenone thiosemicarbazone. However, the major difference between **1** and **8** is the presence of different *para*substituents (**1**:-Br and **8**: -OH) in the second (B) phenyl ring. A more detailed analysis was carried out using the phenolic analog (**8**).




**1**: 16.7 nM




**8:** 188.7 nM

Figure 31. Chemical Structures of **1** and **8**

The characterization of compounds that could be used as potential therapeutic agents involved numerous evaluations. Understanding the kinetic effect of these inhibitors helped to identify their mechanism of actions. Fluorometric based assays were utilized to study various assay parameters inhibition activities, determination of $K_I$, and reversibility of thiosemicarbazone inhibitors, mechanism of inhibition and mode of inhibition.

*Kinetic Analysis of 3-Bromo-3'-Hydroxybenzophenone Thiosemicarbazone (**8**) as a Cathepsin L Inhibitor*

*Effect of Inhibitor Concentration on Cathepsin L Progress Curves*

Cathepsin L (1 nM) was added to six different concentration curves ranging from 0 to 20 μM of compound **8**. Reactions were initiated by the rapid addition of Z-FR-AMC (Final concentrations between 1 and 50 μM). The release of AMC from the nonfluorescent substrate was monitored every 3 seconds for fifty minutes. Figure 32-Figure 36 show the effect of inhibitor concentration on cathepsin L progress curves at different substrate concentrations. Uninhibited and inhibited reactions were monitored for 3000 seconds. The first conclusion, based on visual observations, is the clear inhibition dependence with time. Then, data were fitted to equation 1.7, by nonlinear regression analysis using GraphPad 5.0. P is the concentration of product (μM), $v_o$ and $v_s$ are the initial and steady-state velocities (μM/s), t is the time in seconds and $k_{obs}$ the rate constant for conversion of the initial velocity $v_o$ to the steady state velocity $v_s$. The rate constant ($k_{obs}$) units are given in $s^{-1}$. The equation was entered into the computational software, knowing that P and t are the dependent and independent variables, while keeping the velocities and the rate constant as unknowns. For each case, the constraints for their calculation were to give positive values (i.e. $k_{obs} \geq 0$). It is also worth noting that equation 1.7 is only valid when substrate depletion is insignificant.[188] Therefore, some points were excluded in every case for data fitting. Velocities, rates, $r^2$ and points analyzed for each substrate concentration are shown in Tables 54-58.

$$P = v_s t \frac{(v_o - v_{s)}}{k_{obs}} \left(1 - e^{-k_{obs} t}\right) \tag{1.7}$$

Table 54. Calculated Kinetic Parameters from Eq.1.7 for Cathepsin L Progress Curves with 50 μM Z-FR-AMC

| [I] (μM) | 20 | 10 | 5 | 1 | 0.5 |
|---|---|---|---|---|---|
| $v_s$ (μM/s) | 0.0007067 | 0.001048 | 0.001150 | 0.001340 | 0.001376 |
| $v_o$ (μM/s) | 0.4482 | ~ 12.38 | ~ 6.261 | ~ 5.213 | ~ 5.182 |
| $k_{obs}$ ($s^{-1}$) | 0.5569 | ~ 17.64 | ~ 9.394 | ~ 8.232 | ~ 8.597 |
| $r^2$ | 0.9904 | 0.9988 | 0.9990 | 0.9988 | 0.9987 |
| Points analyzed | 1000 | 1000 | 1000 | 1000 | 1000 |

Table 55. Calculated Kinetic Parameters from Eq.1.7 for Cathepsin L Progress Curves with 25 μM Z-FR-AMC

| [I] (μM) | 20 | 10 | 5 | 1 | 0.5 |
|---|---|---|---|---|---|
| $v_s$ (μM/s) | 0.0003545 | 0.0008308 | 0.0009205 | 0.001256 | 0.001283 |
| $v_o$ (μM/s) | 0.002994 | 0.2028 | ~ 0.2786 | ~ 2.064 | ~ 2.577 |
| $k_{obs}$ ($s^{-1}$) | 0.003522 | 0.4363 | ~ 0.6231 | ~ 6.293 | ~ 7.384 |
| $r^2$ | 0.9734 | 0.9926 | 0.9953 | 0.9994 | 0.9993 |
| Points analyzed | 1000 | 1000 | 1000 | 1000 | 1000 |

Table 56. Calculated Kinetic Parameters from Eq.1.7 for Cathepsin L Progress Curves with 10 μM Z-FR-AMC

| [I] (μM) | 20 | 10 | 5 | 1 | 0.5 |
|---|---|---|---|---|---|
| $v_s$ (μM/s) | 8.407e-005 | 0.0001831 | 0.0002328 | 2.674e-014 | 0.001182 |
| $v_o$ (μM/s) | 0.001703 | 0.001567 | 0.001577 | 0.001591 | ~ 0.3222 |
| $k_{obs}$ ($s^{-1}$) | 0.002457 | 0.001048 | 0.0008123 | 0.0002494 | ~ 1.800 |
| $r^2$ | 0.9793 | 0.9956 | 0.9967 | 0.9982 | 0.9972 |
| Points analyzed | 1000 | 1000 | 1000 | 1000 | 1000 |

Table 57. Calculated Kinetic Parameters from Eq.1.7 for Cathepsin L Progress Purves with 5 μM Z-FR-AMC

| [I] (μM) | 20 | 10 | 5 | 1 | 0.5 |
|---|---|---|---|---|---|
| $v_s$ (μM/s) | 2.732e-005 | 8.012e-006 | 4.627e-005 | ~ 1.888e-16 | 2.040e-15 |
| $v_o$ (μM/s) | 0.001111 | 0.001048 | 0.0009878 | 0.001053 | 0.001114 |
| $k_{obs}$ ($s^{-1}$) | 0.003824 | 0.001566 | 0.0009347 | 0.0004653 | 0.0003643 |
| $r^2$ | 0.9617 | 0.9898 | 0.9953 | 0.9976 | 0.9989 |
| Points analyzed | 667 | 667 | 667 | 667 | 1000 |

Table 58. Calculated Kinetic Parameters from Eq.1.7 for Cathepsin L Progress Curves with 1 μM Z-FR-AMC

| [I] (μM) | 20 | 10 | 5 | 1 | 0.5 |
|---|---|---|---|---|---|
| $v_s$ (μM/s) | 1.327e-005 | 2.622e-005 | 4.346e-005 | 8.961e-5 | 3.944e-5 |
| $v_o$ (μM/s) | 0.0006334 | 0.0004182 | 0.0004304 | 0.0004607 | 0.0003939 |
| $k_{obs}$ ($s^{-1}$) | 0.01499 | 0.005331 | 0.003741 | 0.002667 | 0.001218 |
| $r^2$ | 0.9273 | 0.9816 | 0.9907 | 0.9965 | 0.9974 |
| Points analyzed | 402 | 402 | 402 | 402 | 402 |

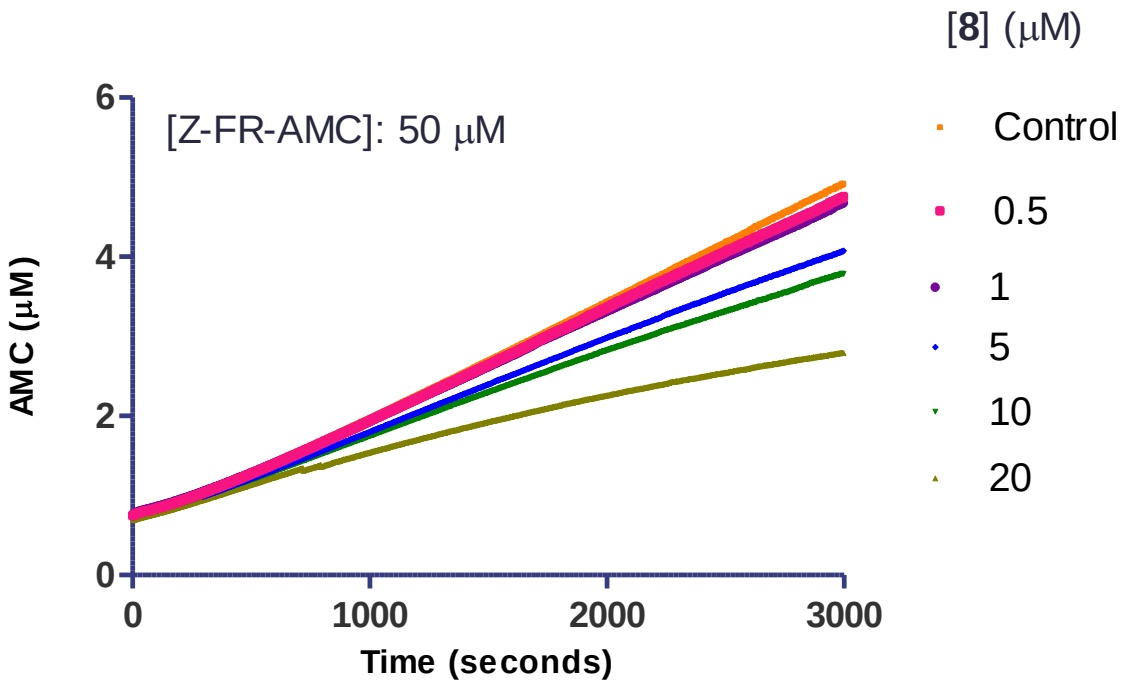



Figure 32. Cathepsin L Progress Curves with **8** Using 50 μM Z-FR-AMC

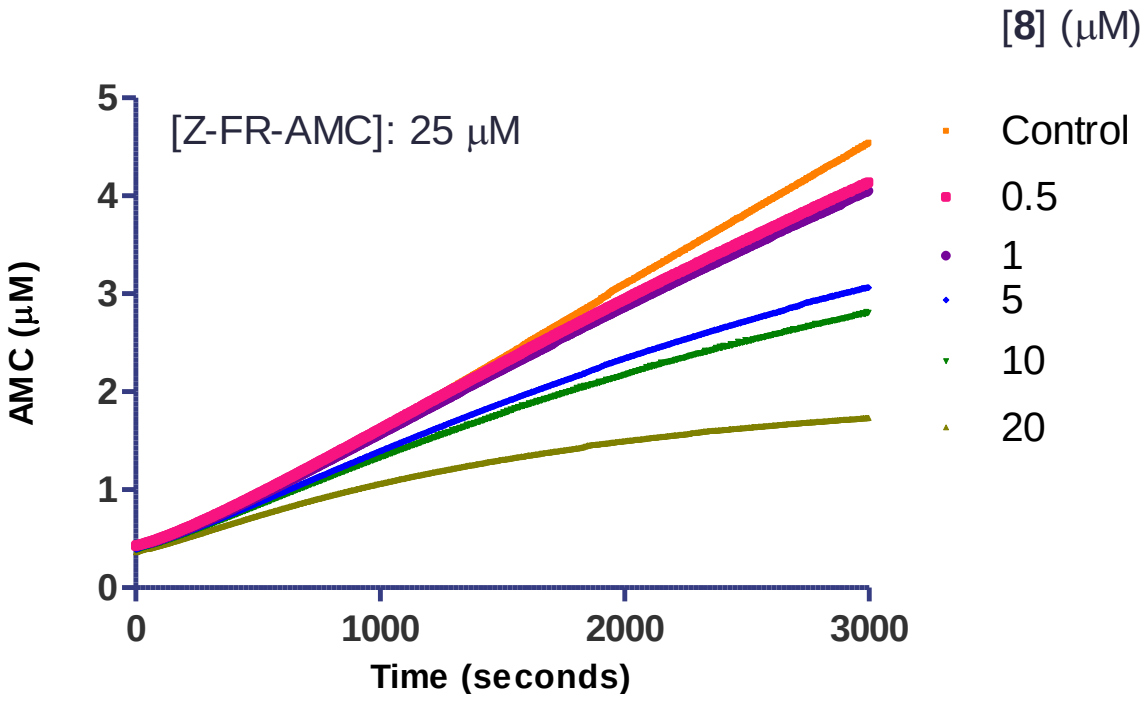



Figure 33. Cathepsin L Progress Curves with **8** Using 25 μM Z-FR-AMC

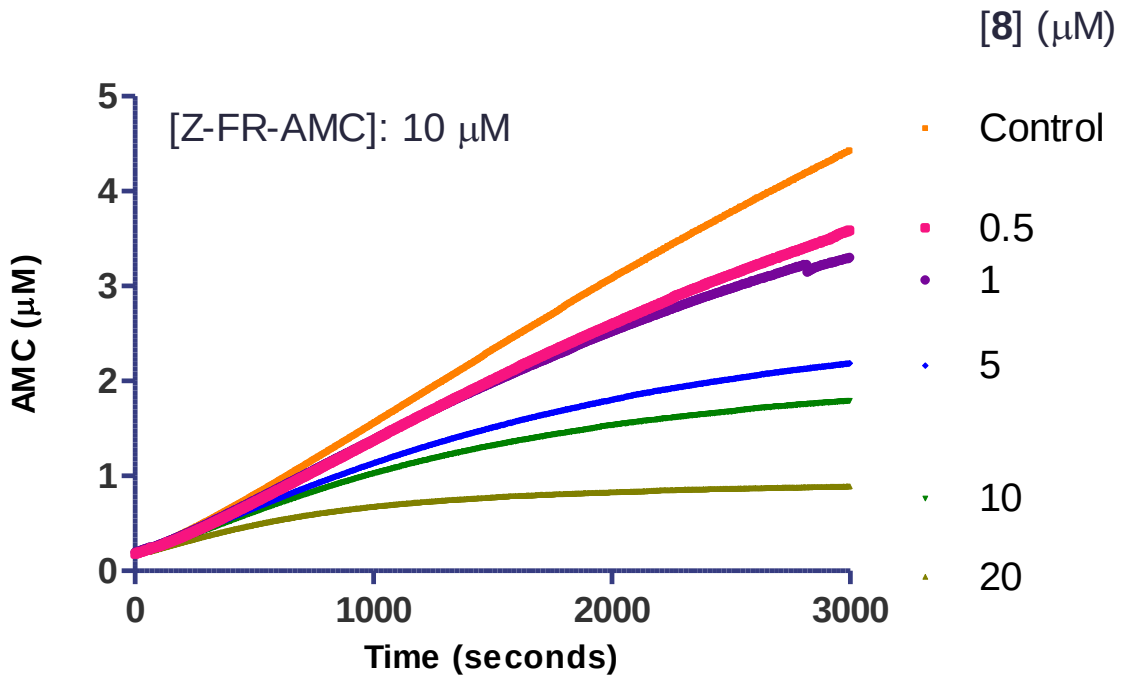



Figure 34. Cathepsin L Progress Curves with **8** Using 10 μM Z-FR-AMC

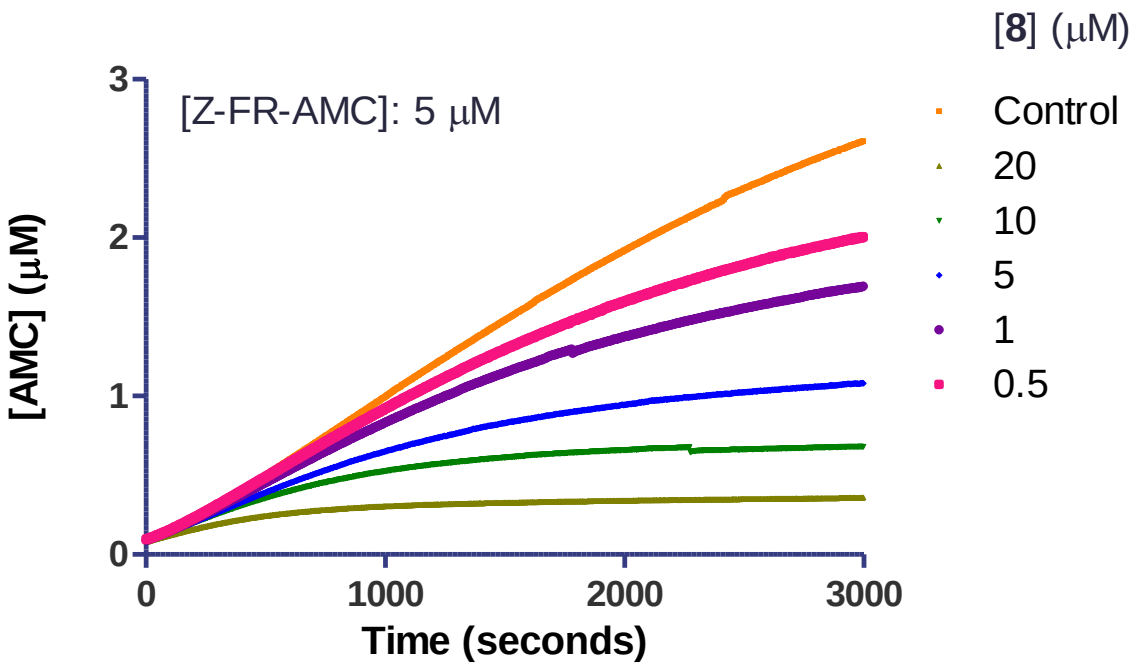



Figure 35. Cathepsin L Progress Curves with **8** Using 5 μM Z-FR-AMC

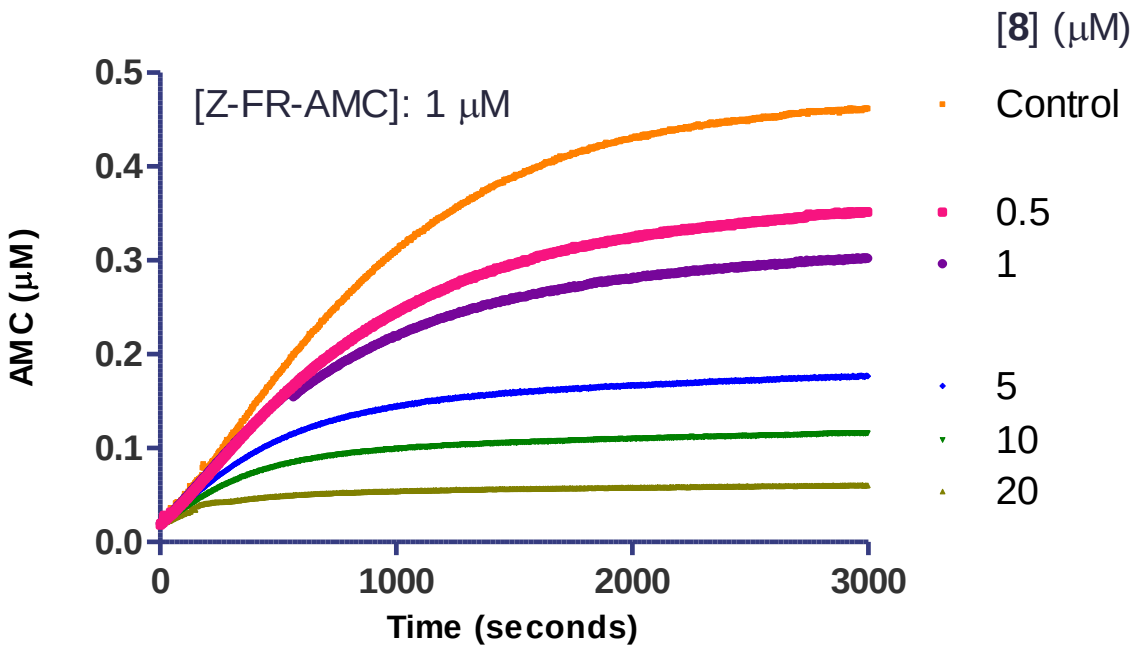



Figure 36. Cathepsin L Progress Curves with **8** Using 1 μM Z-FR-AMC

Approximate values for the rate constants, and velocities were obtained for every substrate concentration (25 and 50 μM). Data were fitted into Equation 1.7, but due to the elevated amount of substrate present in solution, compared to the concentration of the phenolic analog, the results were not reasonable. A better fit was acquired with lower concentrations of Z-FR-AMC (1-10 μM). We decided to investigate the mechanism of inhibition of **8** as a slow binding inhibitor of cathepsin L, due to the clear time-dependence of the progress curves. There are two possible mechanisms that are known for slow binding inhibitors. Figure 37 shows the two proposed mechanisms that could be observed using the phenolic bromobenzophenone thiosemicarbazone (3,3'-Br-PhOH-TSC) as a slow-binding cathepsin L inhibitor. Mechanism A summarizes a simple reversible inhibition with $k_{on}$ and $k_{off}$ values relatively small. The parameters $k_{on}$ and $k_{off}$ are the rate constants for the formation and dissociation of the cathepsin L-3,3'-Br-PhOH-TSC complex. Mechanism B offers an extra step which is more complicated to monitor or verify. This is a more general approach where it is assumed the cathepsin L -3,3'-Br-PhOH-TSC complex undergoes an auto-isomerization, or a possible covalent modification of the enzyme due to the presence of the inhibitor.

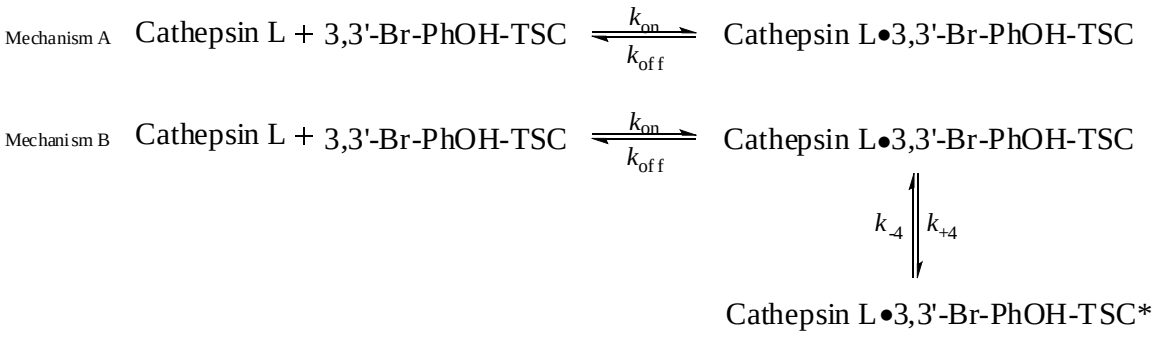



Figure 37. Possible Mechanisms of Inhibition of Human Cathepsin L by **8**

Initial velocities and $k_{obs}$ values were plotted versus inhibitor concentration. The first plot ($v_o$ vs [I]) is utilized to determine the mechanism of inhibition that compound **8** follows under *in vitro* conditions. Figure 38 and Figure 39 show the $v_o$ vs [I] and $k_{obs}$ vs [I] plots. Initial velocities of the reactions with **8** at different substrate concentrations (1-10 μM) are the same as the reactions without **8**. The independence of $v_o$ with respect to the inhibitor concentration is a clear indication that a simple reversible inhibition (Mechanism A, Figure 37) is carried out for slow-binding inhibitors.

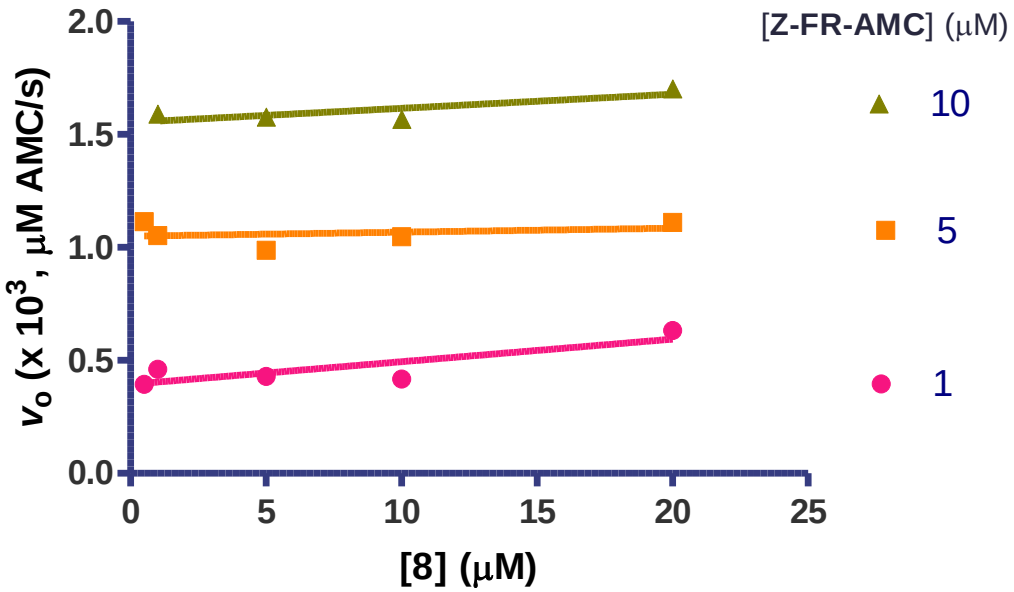



Figure 38. Calculated $v_o$ from Eq. 1.7 Cathepsin L Progress Curves with **8**

Figure 39 shows the linear trend between $k_{obs}$ and [I]. This figure would confirm that a simple reversible mechanism can be observed between cathepsin L and **8**. Interestingly, many groups have reported similar plots, but their final remarks suggest a covalent modification.[328,329] For example, Hang investigated the behavior of non-nucleoside inhibitors of hepatitis C virus (HCV) polymerase. Their results also confirmed the behavior of these compounds as slow binding inhibitors.

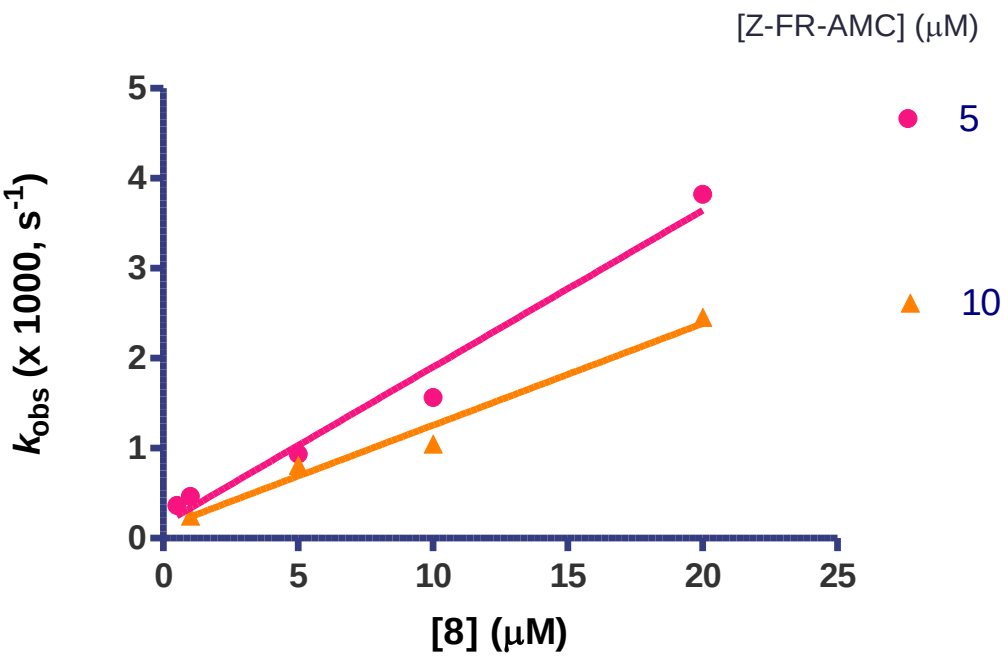



Figure 39. Calculated $k_{obs}$ from Eq. 1.7 Cathepsin L Progress Curves with **8**

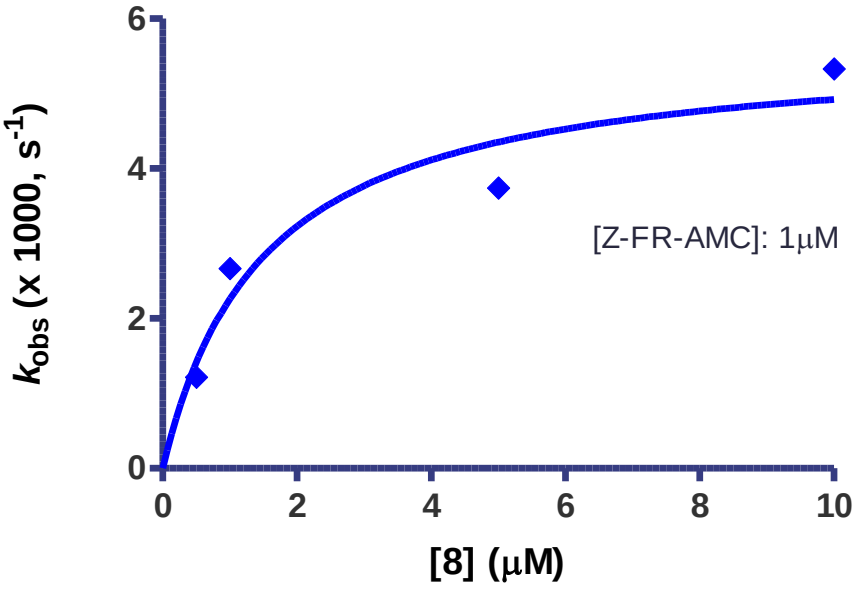



Figure 40. Calculated $k_{obs}$ from Eq. 1.7 Cathepsin L Progress Curves with **8** using 1 μM Z-FR-AMC

Linear regression analysis was performed using GraphPad 5.0 in order to analyze the data (5 and 10 μM Z-FR-AMC). Results can be seen in Table 55. The $k_{on}$ and $k_{off}$

values can be then calculated using equations 2.1 and 2.2.[328] However, a closer examination of the data showed that lower concentrations ([Z-FR-AMC]: 1 μM) there was a hyperbolic relation between $k_{obs}$ and [I] (Figure 40). These results indicated that, instead, a possible covalent modification might occur, but further experiments will need to be performed to validate this statement. Non-linear regression analysis was performed using GraphPad 5.0 in order to analyze the data (1 μM Z-FR-AMC). These results indicate that the mechanism of inhibition for slow binding inhibitors (i.e. simple reversibility, isomerization, or covalent modification) is substrate dependent. Higher concentrations of substrate force the inhibition to a simple mechanism. In contrast, the inhibitor might form a complex species with the enzyme when lower concentrations of substrate are present.

$$k_{obs} = k_{on}^{app}[I] + k_{off} \tag{2.1}$$

$$k_{obs} = \frac{k_{on}[I]}{(1 + \frac{[S]}{K_M})} + k_{off} \tag{2.2}$$

Table 59. Calculated $k_{on}$ and $k_{off}$ for Selected Substrate Concentrations

| [Z-FR-AMC] (μM) | 10 | 5 |
|---|---|---|
| $k_{on}^{app}$ (μM·s)$^{-1}$ | 0.17 | 0.11 |
| $k_{on}$ (μM·s)$^{-1}$ | 0.96 | 1.14 |
| $k_{off}$ (s$^{-1}$) | 0.16 | 0.12 |
| $r^2$ | 0.9770 | 0.9761 |

The results that are summarized in table 55 also confirm the slow release of the inhibitor from the cathepsin L·3,3'-Br-PhOH-TSC complex. $k_{on}$ rates are 6 fold higher

than $k_{off}$ kinetic constants. The $k_{on}$ and $k_{off}$ values indicate the dissociation of the enzyme-inhibitor complex happens at a much slower rate than the formation of the species.

*Effect of Preincubation Studies on Cathepsin L Inhibition Assays Using **8***

Progress curves provided strong evidence that **8** is a slow-binding inhibitor. Therefore, we explored the influence of preincubation time on the inhibitory potency of thiosemicarbazones and their $IC_{50}$ values. It is often found that reported inhibitory assays differ in this specific parameter in the literature literature. Therefore, it is quite difficult to compare $IC_{50}$ values because it is an assay dependent result. Compound **8** $IC_{50}$ values were determined at seven different preincubation times ranging between 0 and 240 minutes. Inhibitor final concentration varied between 0 and 10 μM. Figure 41 shows the results of the studies.

Table 60. Effect of Preincubation Time on $IC_{50}$ Values Using **8**

| *Pre-incubation times (*minutes*)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 240 | 46.97 ± 1.99 |
| 120 | 79.72 ± 2.03 |
| 60 | 96.84 ± 4.17 |
| 30 | 131.2 ± 2.05 |
| 5 | 188.7 ± 18.7 |
| 1 | 6081 ± 334.3 |
| 0 | ≥ 10,000 |

The effect of preincubation on the potency of **8** as seen in Table 60 is quite evident. The phenolic bromobenzophenone thiosemicarbazone (**8**) showed no inhibitory activity when cathepsin L was not in preincubated with the compound. Compound **8** has poor activity with one minute preincubation time. However, the trend changed dramatically when the preincubation time was increased to 5 minutes. The inhibitory

activity of **8** increased 32-fold with a remarkable $IC_{50}$ less than 200 nM (188.7 nM). The potency of **8** modestly increased with longer preincubation times. Finally, the best potency activity was found when **8** was preincubated for four hours with cathepsin L, the activity of the phenolic analog increased 4-fold to give a value of 47 nM. These results confirmed the strong dependence of $IC_{50}$ value determination with respect to the preincubation time parameter. Controls (i.e. uninhibited cathepsin L) were monitored at every preincubation time. There was no significant loss of catalytic activity at longer preincubation times.

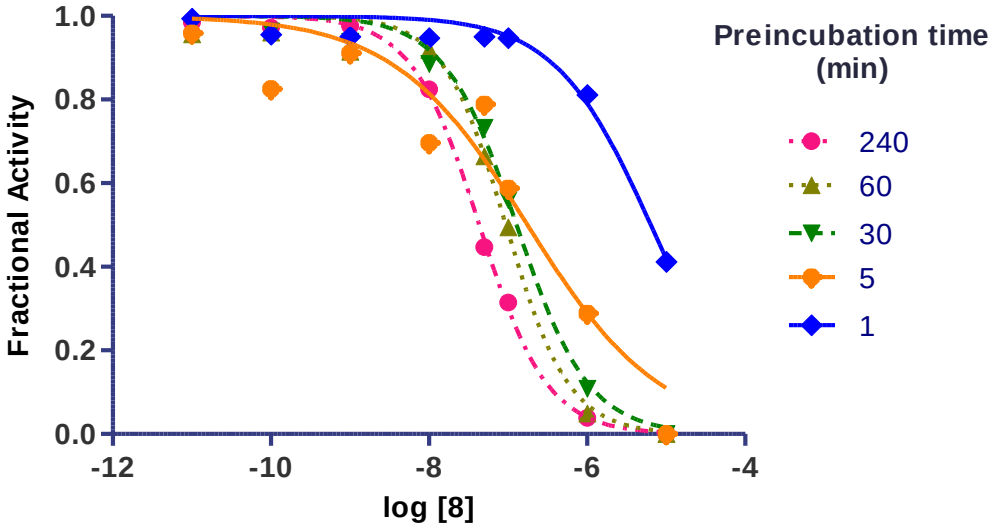



Figure 41. Effect of Preincubation Time on $IC_{50}$ Values of **8** against Cathepsin L

*Determination of $K_i^{app}$ Using Morrison's Quadratic Equation. Effect of Preincubation Time Using **8**.*

The data obtained from the preincubation was further analyzed. The possibility that **8** was a tight-binding inhibitor was analyzed with Morrison's quadratic equation. (See equation 1.11).

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \qquad (1.11)$$

The rates $v_i$ and $v_o$ are the inhibited and uninhibited cathepsin L velocities (RFU/s); $[\mathrm{E}]_\mathrm{T}$ (nM) is, the total concentration of enzyme found in solution (free enzyme and inhibitor-enzyme complex); $[\mathrm{I}]_\mathrm{T}$ (nM) is the total concentration of inhibitor present in solution (free inhibitor and inhibitor-enzyme complex); and $K_\mathrm{I}$ (nM) is the inhibition constant, often referred as the dissociation constant. The equation may be solved to give two possible answers. However, the equation is written so that there is only one possible answer that fits physiological conditions (i.e. $K_\mathrm{I}^{\mathrm{app}} > 0$). GraphPad 5.0 was used to fit the data after data manipulation for every preincubation time point. Inhibitor, substrate, and enzyme concentrations, as well as $K_\mathrm{M}$, $K_\mathrm{I}^{\mathrm{app}}$, were all in micromolar units (μM). The residual activity (or $v_i$/vo) was normalized to 1 (i.e. vo: 1 and $0 \leq v_i \leq 1$). Normalized residual activity and [I] were defined as the dependent and independent variables respectively. Nonlinear regression was applied using the following conditions: [S]: 50 μM, $K_\mathrm{M}$: 1.1 μM, $[\mathrm{E}]_\mathrm{T}$: 0.001 μM and vo: 1. Normalization and nonlinear conditions were used to three or four independent experiments per preincubation time. Average and standard errors can be seen in Table 57 and Figure 57.

Table 61. Effect of Preincubation Time in $K_I^{app}$ Values of **8** against Cathepsin L

| *Pre-incubation times (*minutes*)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 1 | 121.5 ±1 8.1 |
| 5 | 3.7 ± 0.3 |
| 30 | 2.7 ± 0.2 |
| 60 | 2.0 ± 0.1 |
| 120 | 1.6 ± 0.1 |
| 240 | 1.1 ± 0.07 |

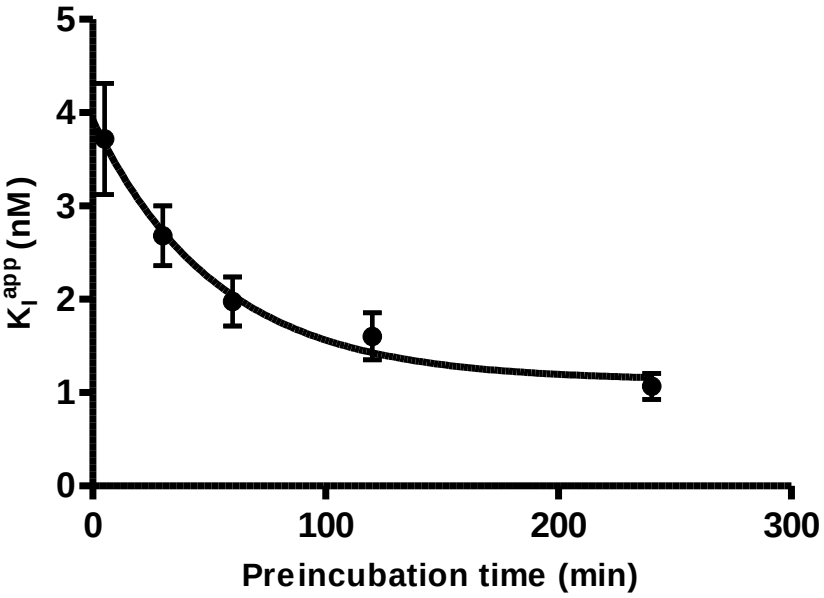



Figure 42. Effect of Preincubation Time in $K_I^{app}$ Values of **8** against Cathepsin L

The $K_I^{app}$ values are shown in Table 61. It is apparent that the inhibition constants are also time-dependent. One minute of preincubation shows a good inhibition constant with a value of 121.5 nM. Furthermore, the potency is extremely enhanced (more than 30-fold) for the calculated value at standard conditions (preincubation time: 5 minutes). The best inhibition constant value was obtained at the longest preincubation time (4 hours). The apparent $K_I$ was 120-fold better when compared with the corresponding value at one minute preincubation time ($K_I^{app}$ = 1.1 nM)

*Cathepsin L Reversibility Studies*

Compound **8** was found to be a time-dependent inhibitor in the preincubation studies. Therefore, we decided to explore if this specific compound was a reversible inhibitor of cathepsin L. A mixture containing 100 X cathepsin L and 10 X $IC_{50(\text{preincubation time: 4 hours})}$, which are 100 and 470 nM respectively, were incubated at 25 °C for four hours. The inhibition of cathepsin L by **8** was able to be monitored by the rapid dilution of the mixture (100-fold) with assay buffer containing Z-FR-AMC. Experiment was set up in order to monitor these reactions almost immediately after adding cathepsin L to the assay. Final conditions are 1 nM cathepsin L, 4.7 nM of the phenolic bromobenzophenone thiosemicarbazone (**8**), and 50 μM. Additionally, a control experiment (cathepsin L with DMSO as control vehicle) was also carried out. Figure 38 shows the release of AMC for the first 5000 seconds. However, uninhibited and inhibited reactions were followed for a total time of four hours. Cathepsin L was able to recover its catalytic activity after the rapid dilution with assay buffer containing Z-FR-AMC. Apparent substrate depletion can be observed in the uninhibited reaction after 5000 seconds. Thus, linear regression was applied to the first 5000 seconds of the reactions to determine cathepsin L activity. Cathepsin L activity for both reactions were found to be 1.62 and 0.44 nM AMC/s for control and inhibited reactions. The results showed cathepsin L was recovered up to 30% after 5000 seconds of reaction**.** Furthermore, a closer inspection of the early points showed that cathepsin L recovered its activity within the first 150 seconds of the reaction up to a 17%. Cathepsin L activity for both reactions were 1.76 and 0.31 nM AMC/s. (See Figure 43 and Figure 44).

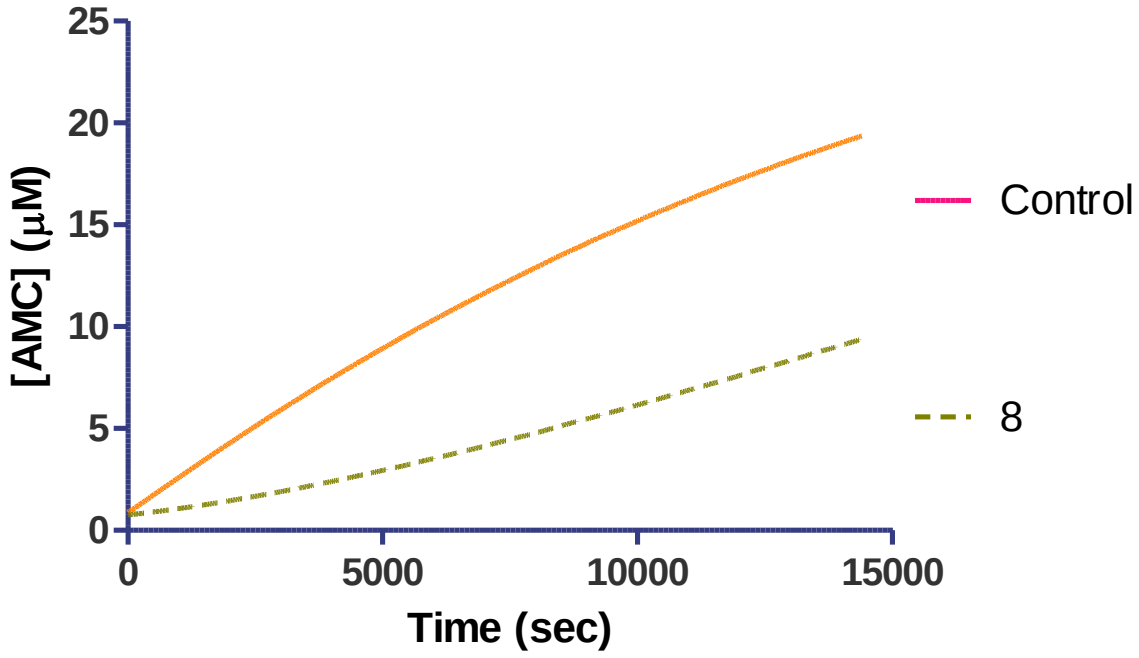



Figure 43. Cathepsin L Reversibility Studies with **8** Using 50 μM Z-FR-AMC

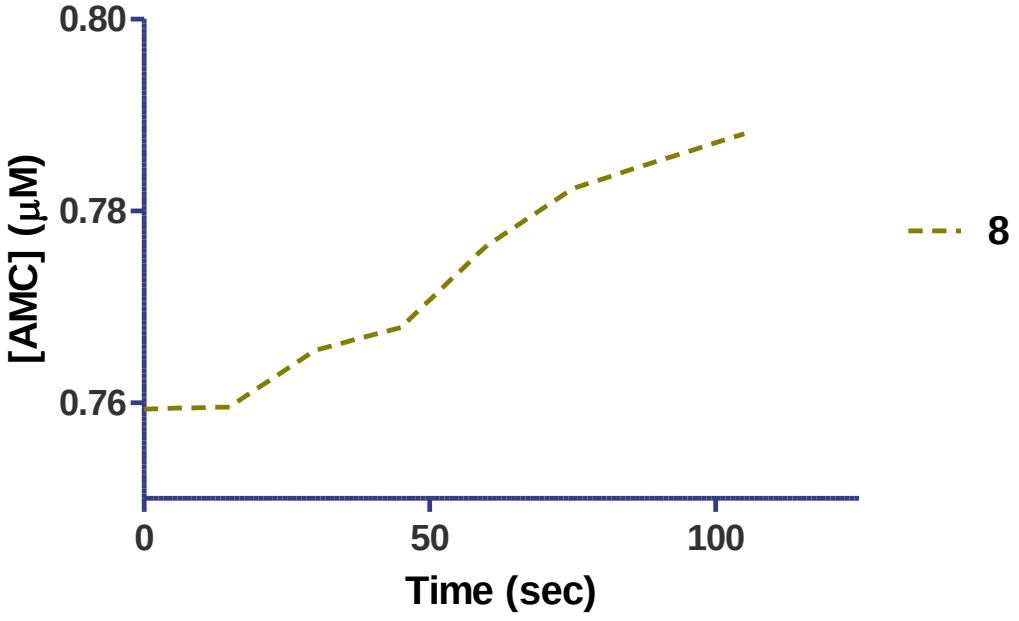



Figure 44. Cathepsin L Reversibility Studies with **8** Using 50 μM Z-FR-AMC

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

We finally tried to determine the mode of inhibition of thiosemicarbazones as slow-binding inhibitors of cathepsin L. Therefore, we investigated the effect of substrate concentration on $IC_{50}$ values.[177] Classical methods for the determination of mode of inhibition (steady-state kinetics) fail when studying a slow, tight-binding inhibitor. In these cases, Copeland offers a suitable alternative for the investigation. $IC_{50}$ values decrease hyperbolically if the compound is an uncompetitive inhibitor. In the case of competitive inhibitors, $IC_{50}$ values increase in a linear trend with higher substrate concentrations. Three substrate concentrations (Final concentrations: 50, 10, and 5 μM) were used to determine $IC_{50}$ values for compound **8.** Cathepsin L and the series of inhibitors were incubated for four hours. Results can be found in table 58 and figure 40.

Table 62. Effect of Substrate Concentration in $IC_{50}$ Values of **8** against Cathepsin L

| *[Z-FR-AMC] (μM)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 50 | 46.97 ± 1.99 |
| 10 | 31.0 ± 2.1 |
| 5 | 28.4 ± 1.1 |

The effect of substrate concentration was investigated using **8** as a lead compound of the series of thiosemicarbazones in order to investigate its mode of inhibition. A linear behavior can be observed with the values. Lowering the amount of substrate in solution gives a subtle, yet evident increase in the potency of the phenolic analog. According to Copeland, a positive linear behavior (i.e. $IC_{50}$ α [S]) is an indication the compound acts as a competitive inhibitor; that is, both substrate and the compound compete for the cathepsin L active site. However, the results cannot be taken as conclusive and more experiments were necessary to confirm these findings.

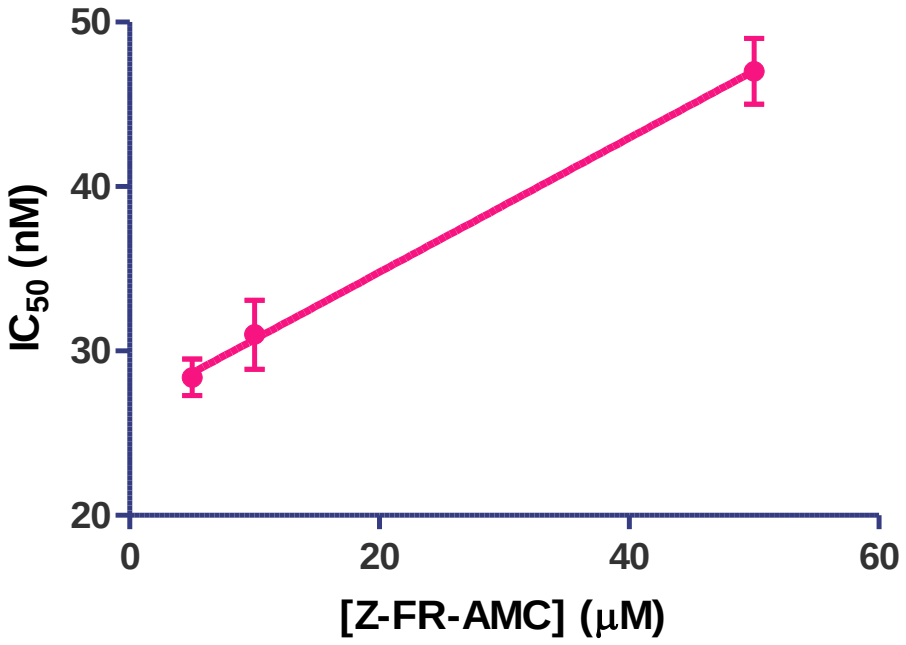



Figure 45. Effect of Substrate Concentration in $IC_{50}$ Values of **8** against Cathepsin L

*Effect of Substrate Concentration (Z-FR-AMC) on Cathepsin L Progress Curves*

The effect of substrate concentration in $IC_{50}$ values suggested the phenolic bromobenzophenone thiosemicarbazone (**8**) was a competitive inhibitor of cathepsin L. Thus, we decided to explore other alternatives to validate previous experiments. A fixed concentration of **8** was used to monitor the effect of this compound on cathepsin L progress curves when using different substrate concentrations. Then, by using non-linear regression analysis (equation 1.7), the parameters $k_{obs}$, $v_o$, and $v_s$ were determined. Compound **8** (0.5 μM) and seven Z-FR-AMC concentrations (ranging from 0.5 and 20 μM) were set up in order to verify the effect of substrate on inhibited cathepsin L progress curves. Reactions were started by the addition of cathepsin L (final concentration: 1 nM). The release of AMC was monitored every five seconds for sixty minutes. Figure 46 shows a typical result of this experiment. The results once again

showed the strong time dependence of 8 when inhibiting cathepsin L. Substrate depletion is negligible for every substrate concentration as seen in Figure 46. Data were fitted into equation 1.7, (see Effect of inhibitor concentration in cathepsin L progress curves). P is the concentration of product (μM), vi and vs are the initial and steady-state velocities (μM/s), t is the time in seconds, and $k_{obs}$ is the rate constant for conversion of the initial velocity vo to the steady state velocity $v_s$. The kobs units are given in $s^{-1}$. Similar constraints were also given (i.e. $k_{obs} \geq 0$). Velocities, rates, $r^2$ and points analyzed for each substrate concentration are shown in tables Table 63,Table 64 and Table 65.

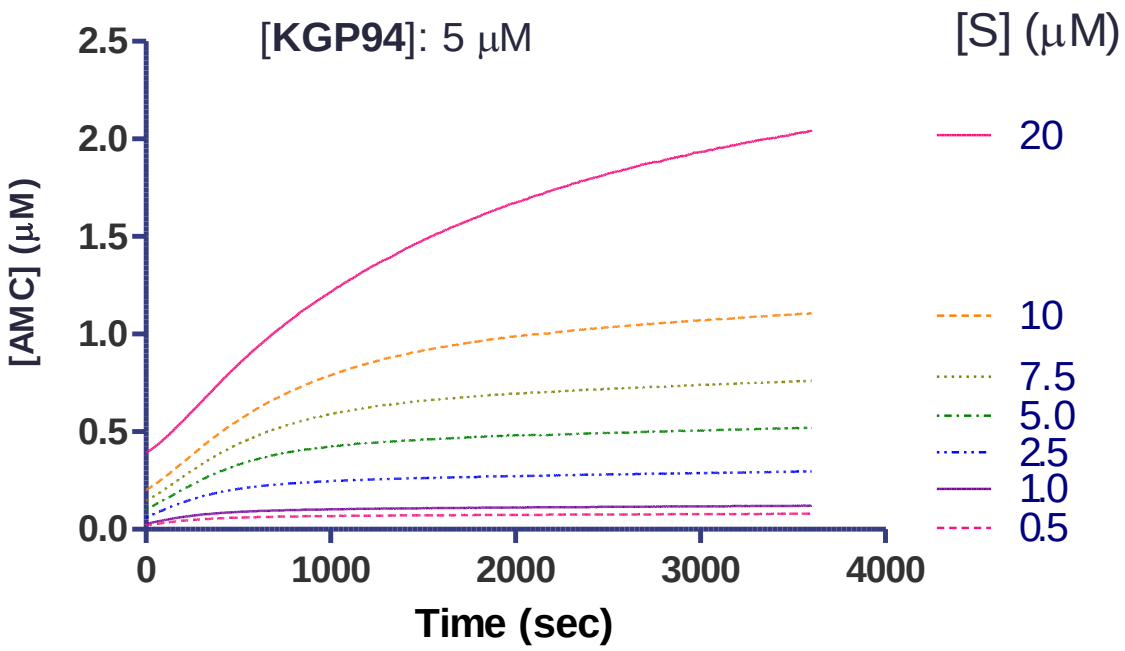



Figure 46. Cathepsin L Progress Curves with 5 μM **8** Using Z-FR-AMC

Table 63 shows the average and standard error of calculated $k_{obs}$ as a function of [S]. Data were graphed as a function of the unitless substrate/Michaelis-Menten constant $[S]/K_M$, where $K_M$: 1.1 μM as previously determined. The rate constant $k_{obs}$ follows an

inversely hyperbolic trend. The rate constant values increased as the concentration of Z-FR-AMC is decreased in solution. This is a clear indication that **8** is a cathepsin L competitive inhibitor with respect to the fluorogenic substrate. A positive hyperbolic trend can be seen for the initial velocities with higher substrate concentrations. (See Table 63 and Figure 47).

Table 63. Effect of [Z-FR-AMC] in $k_{obs}$ Values When Using **8** against Cathepsin L

| [Z-FR-AMC] (μM) | $k_{obs}$ ± *Standard Error (x $10^3$, $s^{-1}$)* |
|---|---|
| 20 | 2.58 ± 0.31 |
| 10 | 2.15 ± 0.04 |
| 7.5 | 2.50 ± 0.05 |
| 5 | 3.03 ± 0.01 |
| 2.5 | 4.22 ± 0.09 |
| 1 | 5.72 ± 0.18 |
| 0.5 | 6.63 ± 0.37 |

Table 64. Effect of [Z-FR-AMC] in $v_o$ Values When Using **8** against Cathepsin L

| [Z-FR-AMC] (μM) | $v_o$ ± *Standard Error (x $10^3$, μM/s)* |
|---|---|
| 20 | 2.81 ± 0.13 |
| 10 | 1.96 ± 0.08 |
| 7.5 | 1.77 ± 0.12 |
| 5 | 1.44 ± 0.07 |
| 2.5 | 1.04 ± 0.02 |
| 1 | 0.65 ± 0.09 |
| 0.5 | 0.46 ± 0.07 |

Table 65. Effect of [Z-FR-AMC] in $v_s$ Values When Using **8** against Cathepsin L

| [Z-FR-AMC] (μM) | $v_s$ ± *Standard Error (x $10^4$, μM/s)* |
|---|---|
| 20 | 2.91 ± 0.11 |
| 10 | 0.90 ± 0.04 |
| 7.5 | 0.60 ± 0.09 |
| 5 | 0.42 ± 0.09 |
| 2.5 | 0.20 ± 0.002 |
| 1 | 0.10 ± 0.01 |
| 0.5 | 0.06 ± 0.008 |

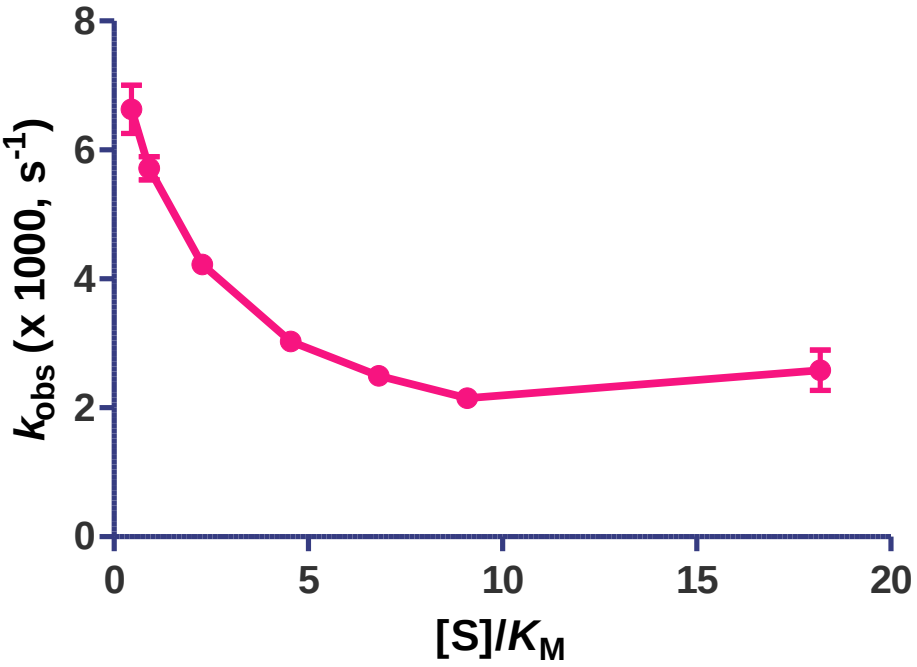



Figure 47. Effect of [Z-FR-AMC] in $k_{obs}$ Values When Using **8** against Cathepsin L

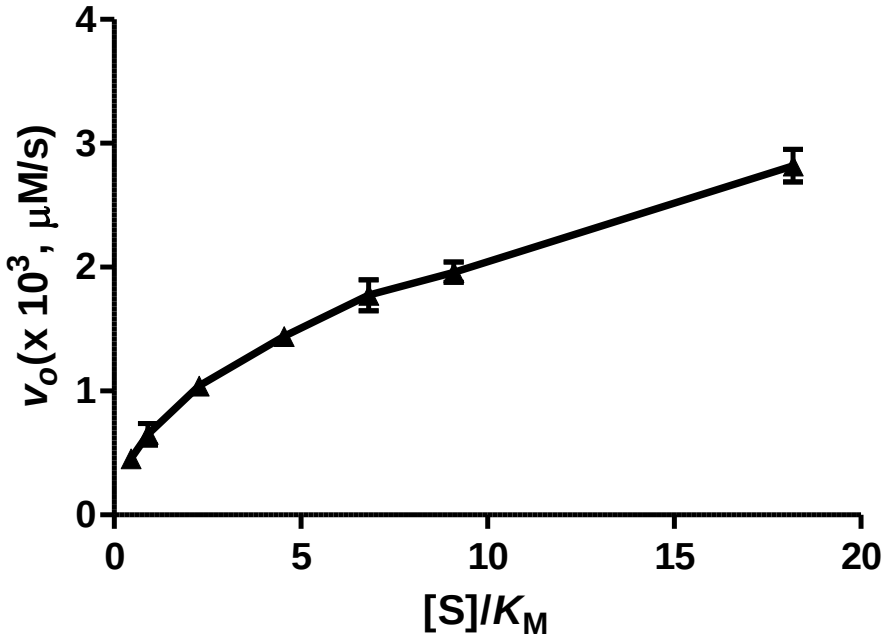



Figure 48. Effect of [Z-FR-AMC] on $v_o$ Values When Using **8** against Cathepsin L

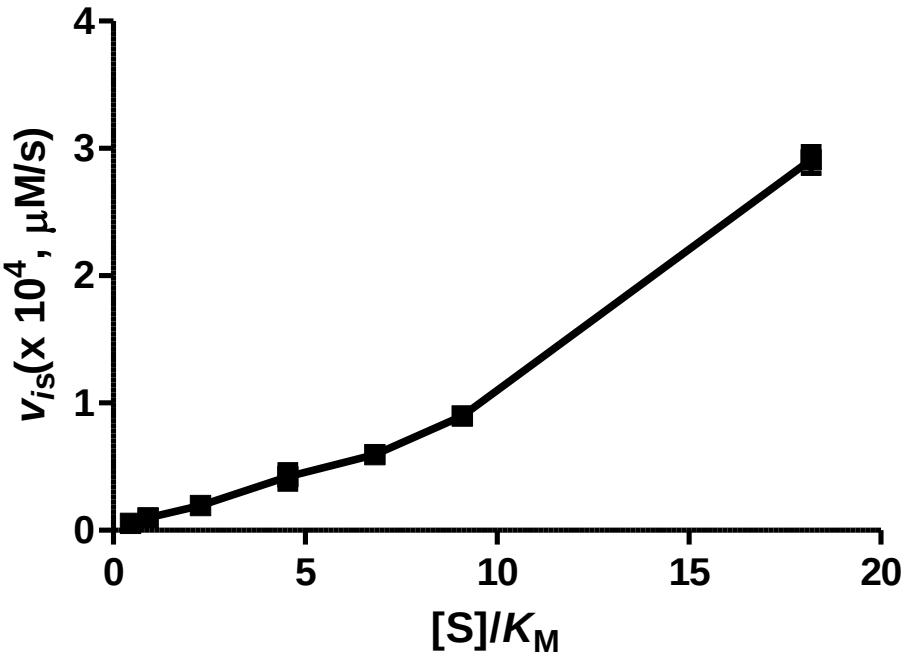



Figure 49. Effect of [Z-FR-AMC] on $v_s$ Values When Using **8** against Cathepsin L

*Inhibition of Cathepsin L Collagenase Activity by Thiosemicarbazone Derivatives*

Cathepsin L is an extremely powerful hydrolase. It is considered one of the most powerful cathepsin among the cathepsin family which also includes, cathepsin K, B, S, and V. It has also been shown that cathepsin L is also capable of degrading high molecular weight proteins. Cathepsin L plays a vital role in the degradation of the extracellular matrix under normal and pathological conditions. Collagen is a major component of the ECM and a substrate for cathepsin L. Thus, we explored the catalytic activity of cathepsin L by using a natural substrate, type I collage from human skin. We also tested the ability of one of the lead compounds (**8,** $IC_{50}$: 188.7 nM), to inhibit the collagenase catalytic activity of cathepsin L. Untreated and treated samples were incubated at 37 °C under acidic conditions (pH 5.5). Four sets of samples were prepared for this experiment. Two sets were not treated with **8** (DMSO was used as control vehicle). The third and fourth sets were treated with 50 μM of compound **8** in 2%

DMSO. The effect of preincubation activity on the catalytic activity of cathepsin L and the inhibitory potency of compound **8** was also tested**.** Thus, two fixed preincubation times (0 and 120 minutes) were chosen to carry out these experiments.

*Inhibition of Cathepsin L Collagenase Activity by Thiosemicarbazone Derivative **8**. (Preincubation Time: 2 Hours).*

A sample containing three microliters of 1.2 μM cathepsin L in CLI and 2.5 μl of 260 μM compound **8** were preincubated as previously described for two hours. Individual samples were prepared to complete a series of six samples. Then, 7.5μl of 0.4 mg/ml type I collagen in acetic acid were added to the cathepsin L-inhibitor mixture. Every sample was carefully mixed and placed in a 37°C water bath. Reactions were monitored between 0 and 20 hours, stopped with 2 μl LDS sample buffer and heated at 90 °C for every time point. Inactivated samples were immediately stored at -80 °C to preserve them.

Untreated and treated samples were loaded onto 4-12% Bis-Tris SDS gel. A sample of molecular weight standards, one sample with type I collagen only, and one sample of cathepsin L only were also loaded as reference controls. These last two samples were inactivated after two hours of preincubation time. Electrophoresis was performed at 200 V for 60 minutes. The gel was rinsed off with water and placed into a 1X SYPRO® solution, a fluorescent protein staining dye. The gel was stained for a minimum of one hour at room temperature. The destaining process was made by washing the gel once with water and twice with 7.5% acetic acid. Finally, a digital image of the gel was obtained by using a fluorescence scanner imager.

*Inhibition of Cathepsin L Collagenase Activity by Thiosemicarbazone Derivative **8**. (Preincubation Time: 0 hours).*

A sample containing three microliters of 1.2 μM cathepsin L in CLI , 2.5 μl of 260 μM compound **8** and 7.5 μl of 0.4 mg/ml type I collagen were added to start the reaction. Figure 50 shows a picture of the collagenase inhibition of cathepsin L by **8** when they were preincubated for two hours. Similarly, Figure 51 shows the conditions for a similar experiment when there was no preincubation time.

Figure 50 shows the results of degradation of type I collagen from human skin using human cathepsin L soluble in acidic solutions.[330] The activity of the protease was stopped at different time points (0-20 hours). The figure shows the progress of the reaction for the first six hours after collagen I was added to the cathepsin L-inhibitor mixture that was previously incubated for two hours. The first lane shows the molecular marker (MM), which shows ten different bands ranging from 12 to 225 kDa. The second lane is type I collagen that was not treated with cathepsin L after two hours of preincubation time in the vehicle control (2% DMSO). Reported literature establishes that collagen is large protein that is divided into three defined chains: α, β, and γ heavy chains and is crosslinked and polymerizes into fibrils. The molecular weights of the heavy chains are approximately $80 \leq MW_{\alpha} \leq 125$ kDa for α chain ; $160 \leq MW_{\beta} \leq 250$ kDa for β chain; and $240 \leq MW_{\gamma} \leq 375$ kDa for γ chain.[331,332] Purity and integrity of the sample are some of the biggest obstacles prior the development of the experiment. Collagen I extractions usually require the use of potent proteolytic enzymes, such as pepsin under acidic conditions, in order to increase the solubility of the macromolecule.[333] Due to its size (MW: 300 kDa), type I collagen solubility is very poor in polar solvents, such as water.[334] However, solubility of the protein is enhanced if

acidic solvents are used such as acetic acid. Results show the presence of three major bands with high molecular weights that are consistent with the reported values for α, β, and γ heavy chains in the literature. Approximate molecular weights are: α ≈ 102 kDa, β ≈ 150 kDa and γ ≈ 200 kDa (lane 2). The third and fourth bands of Figure 50 represent untreated and treated samples that were stopped after 20 minutes of the natural substrate. Treated samples contained type I collagen, and cathepsin L in DMSO, while treated samples consisted of type I collagen, cathepsin L, and compound **8** in DMSO. The untreated sample showed a slight degradation process even as little as 20 minutes. The α, β, and γ bands look more defined than their respective control. The degradation is more evident in the case of the α bands (lane 3). The degradation process is evident after 90 and 180 minutes. All of the bands α, β, and γ look well defined and there is evidence that heavy bands compared to untreated collagen at time 0. (lane 2). A comparison between treated (lanes 5 and 7) and untreated (lanes 6 and 8) samples reveals that compound **8** was able to inhibit the collagenase activity of cathepsin L. Longer preincubation times (t ≥ 6 hours) indicate that uninhibited cathepsin is able to completely degrade 3 μg of type I collagen. However, that is not the case for the treated samples with 50 μM of compound **8**. These results indicate that the phenolic brominated benzophenone thiosemicarbazone is capable to inhibit cathepsin L collagenase activity up to a fifty percent.

Additionally, Figure 51 shows the results of samples containing cathepsin L with **8** (50 μM) and type I collagen that was immediately added without preincubation time. Results are consistent with previous observations where the protease was preincubated with **8** for two hours. Cathepsin L was able to degrade the fibrillar protein in a lesser extent. However, untreated samples were not able to completely degrade three

micrograms at longer preincubation times. A small sample of α and β chains can be seen after 20 hours of reactions, but γ chains were completely degraded. Inhibition can be observed because the protease was not capable to degrade the γ chain. Finally, cathepsin L samples (275 nM) can be seen at the bottom of both gels. The bands showed an apparent molecular weight between 24 and 31 kDa, which is consistent with reported values (MW: 29 kDa).

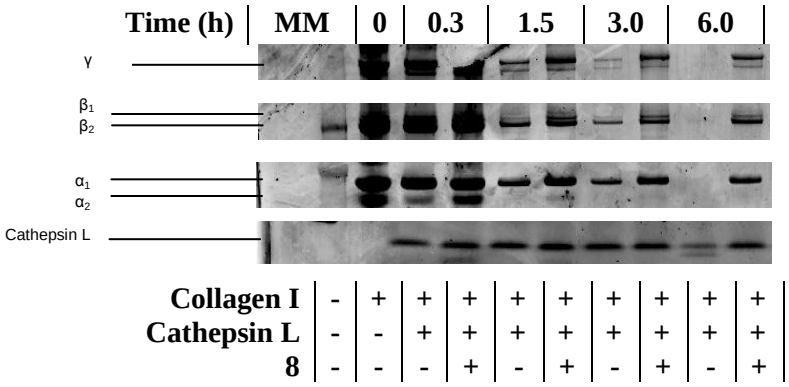



Figure 50. Inhibition of Collagenase Activity of Cathepsin L by **8**, Preincubation Time: 2 Hours.

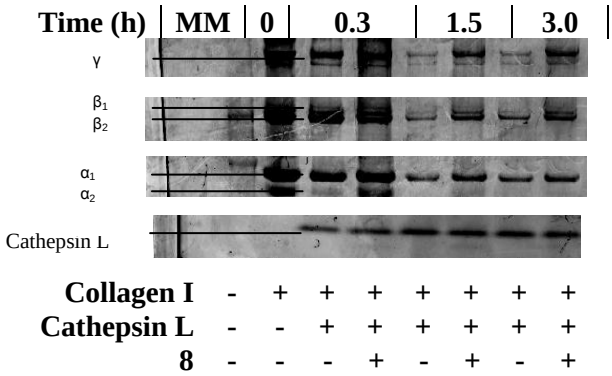



Figure 51. Inhibition of Collagenase Activity of Cathepsin L by **8**, No Preincubation time

*Molecular Docking of **8** with Cathepsin L*

In a collaborative work with Dr. Shen-En Chen, compound **8** was docked in in cathepsin L using computational software (Insight II, 2005). The crystal structure was originally reported by Guncar and coworkers[228] and the crystal structure was deposited in the protein data bank (PDB ID1ICF).[335] Briefly, the macromolecule was minimized by deleting the original ligand (thyroglobulin type-1). Also, the hydrogen atoms were added to the structure and were set to pH 5.5. The Cys25-His163 thiolate-imidazolium ion pair was generated by changing the setting the hybridization state of the species to $sp^2$ . Finally, water molecules were added in order to calculate electronic potentials. Compound **8** was also prepared by changing the hybridization states of its nitrogen atoms ($sp^2$); solvated, and minimized.[4]

Overall, two important hydrogen bonds between the inhibitor and cathepsin L were found. The first hydrogen bond was between the NH terminal of **8** and the carboxamide chain residue of Gln19. The second hydrogen bond was between the side chain of Asp162 and the phenolic moiety of **8**. Gln19 is part of the oxyanion hole, while Asp162 forms a wall on the right side of the $S_2$ pocket (See Crystal Structure of Cathepsin L, page 48). These two amino acid residues help placing **8** in close proximity to the active site of cathepsin L. Figure 52 shows the orientation of cathepsin L S1 and S1' pockets and the relative position of **8** with respect of cathepsin L. Specifically, the S1' (facing down) was found to interact with the bromophenyl group of **8.** The molecular docking suggested the formation of a transient tetrahedral intermediate (i.e. covalent modification).

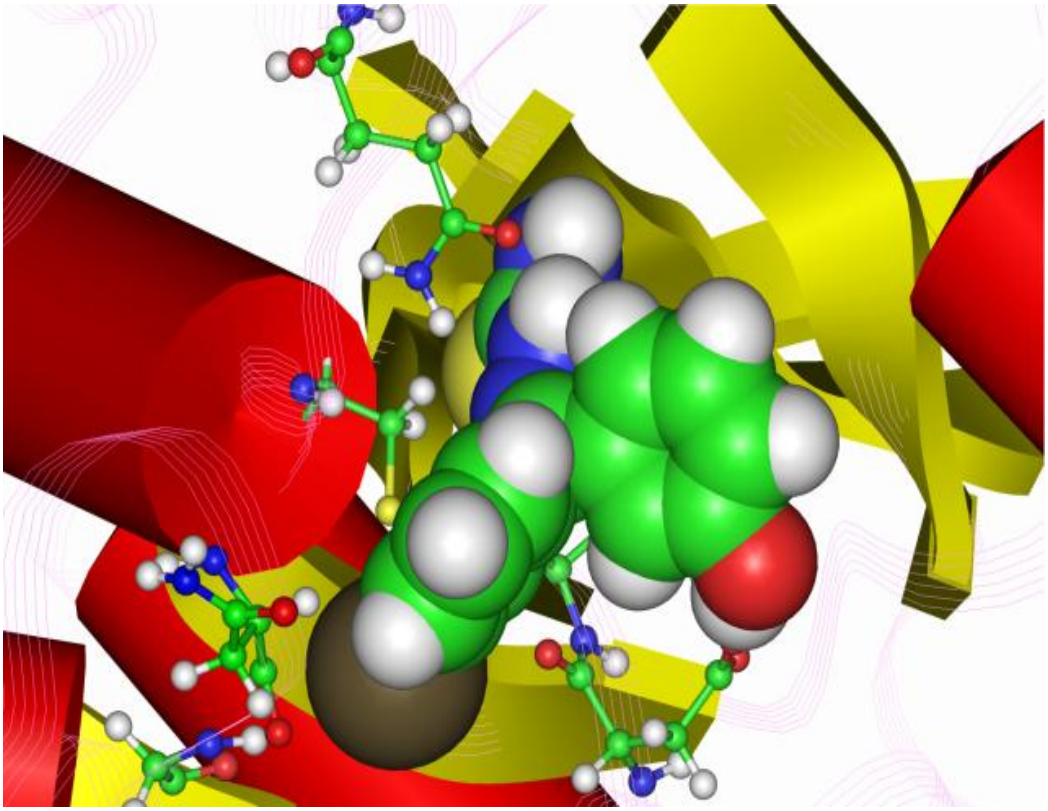


Figure 52. Molecular Modeling of Cathepsin L with **8** (Image courtesy of Dr. Shen-En Chen, Baylor University)

*Proposed Mechanism of Inhibition of Cathepsin L by **8***

Advanced kinetic analysis and molecular docking of **8** as good cathepsin L inhibitors were used to propose a mechanism of inhibition of the protease by this thiosemicarbazone analog. Figure 53 shows a two dimensional representation of the most important hydrogen bonds between the inhibitor and two residues of the protease (Gln19 and Asp162) and a covalent bond between Cys25 and thiocarbonyl carbon of **8**. The possibility of formation of a thiocarbamoylated enzyme cannot be ruled out.

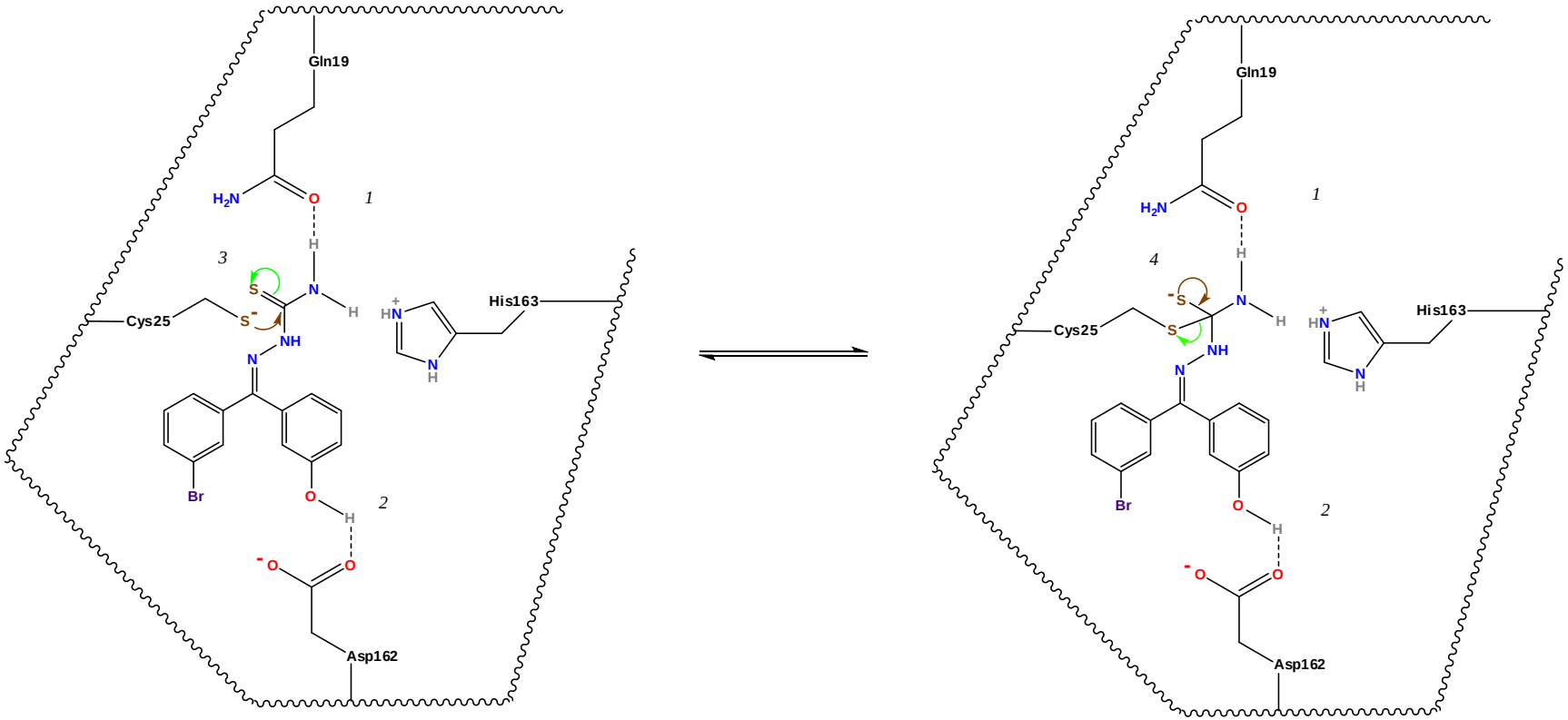



Figure 53. Two-Dimensional Representation of the Proposed Mechanism of Inhibition of Cathepsin L by **8.** *1*. Hydrogen Bond between Carbonyl Oxygen of Gln19 Side Chain and KPG94 NH Moiety. *2*. Hydrogen Bond between **8** Phenolic Group and Asp162 Carboxylic Acid Side-Chain. *3*. The Cys25 Thiolate is in Position to Attack the **8** Thiocarbonyl Carbon Forming a Reversible Transient Covalent Bond (*4*) and Orients **8** at the Active Site.

*Kinetic Analysis of 3,3'-Dibromobenzophenone Thiosemicarbazone (**1**) as a Cathepsin L Inhibitor*

*Effect of Inhibitor Concentration on Cathepsin L Progress Curves*

Cathepsin L (1 nM) was added to six different concentration curves ranging from 0 to 10 μM of **1**. Reactions were initiated by the rapid addition of Z-FR-AMC (final concentration: 5 μM). The release of AMC from the nonfluorescent substrate was monitored every 3 seconds for fifty minutes. Figure 54 shows the effect of inhibitor concentration on cathepsin L progress curves. Based on visual observations, the inhibition of cathepsin L by **1** is time-dependent. Data were fitted to equation 1.7, by nonlinear regression analysis using GraphPad 5.0. P is the concentration of product (μM), $v_i$ and $v_s$ are the initial and steady-state velocities (μM/s), t is the time in seconds and $k_{obs}$ the apparent first-order rate constant for the steady state that is formed between all the species present in the reaction (enzyme, inhibitor, enzyme-inhibitor complex, and substrate). The $k_{obs}$ units are given in $s^{-1}$. The equation was entered into the computational software, knowing that P and t are the dependent and independent variables, while keeping the velocities and the rate constant as unknown. For each case, the constraints for their calculation were set to give a positive value (i.e. $k_{obs} \geq 0$). It is also worth noting that equation 1.7 is only valid when substrate depletion is insignificant. Therefore, some points were excluded in every case for data fitting. Velocities, rates, $r^2$, and points analyzed for each substrate concentration are shown in Table 66 and Figure 54.

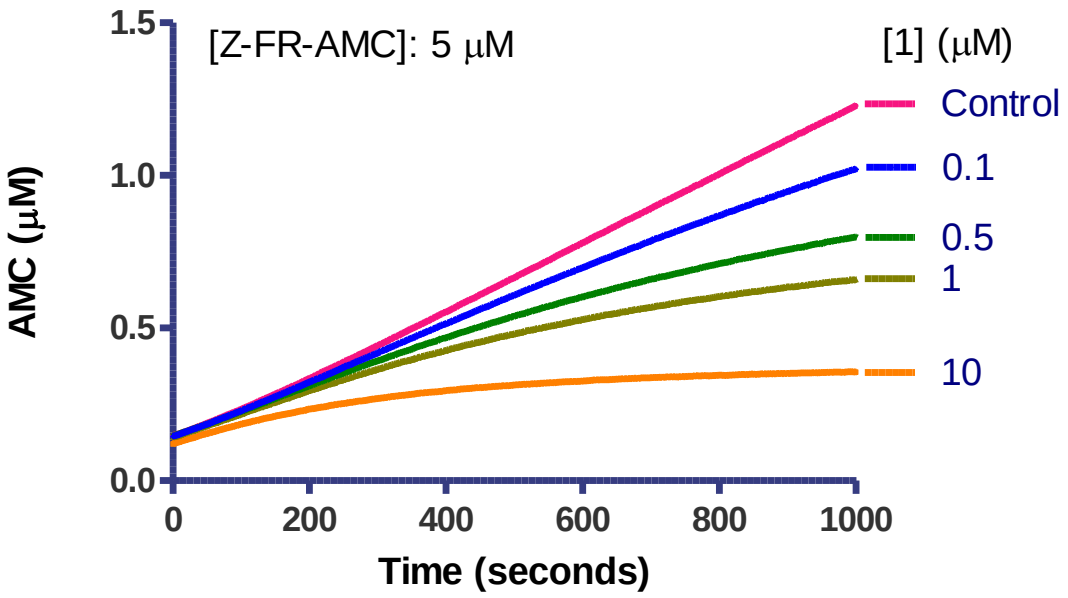



Figure 54. Cathepsin L Progress Curves with **1** using 5 μM Z-FR-AMC

Data sets were fitted into equation 1.7. Ambiguous results were obtained for reactions low inhibitor concentration or uninhibited sets ([I] ≤ 0.1 μM). Only three concentrations (10, 1, and 0.5 μM) were able to be resolved. Further investigation needs to be done to make stronger conclusions about the mechanism of inhibition of analog **1.**

*Effect of Preincubation Studies on Cathepsin L Inhibition Assay Using **1***

Progress curves provided strong evidence that **1** is a slow-binding inhibitor. Therefore, we explored the influence of preincubation time on the inhibitory potency of thiosemicarbazones and their $IC_{50}$ values. Compound **1** $IC_{50}$ values were determined at 7 different preincubation times ranging between 0 and 240 minutes. Substrate concentration was 5 μM. Inhibitor final concentration varied between 0 and 10 μM. Table 66 and Figure 55 show the results of the studies.

Table 66. Calculated Kinetic Parameters from Eq.1.7 for Inhibited Cathepsin L Progress Curves with **1** Using 5 μM Z-FR-AMC

| [I] (μM) | 10 | 5 | 0.5 |
|---|---|---|---|
| $v_s$ (μM/s) | 0.0001502 | 0.0004666 | 0.0006583 |
| $v_i$ (μM/s) | 0.004320 | 0.005051 | 0.06016 |
| $k_{obs}$ ($s^{-1}$) | 0.01850 | 0.02006 | 0.3207 |
| $r^2$ | 0.9117 | 0.9736 | 0.9867 |
| Points analyzed | 335 | 335 | 335 |

Table 67. Effect of Preincubation Times on $IC_{50}$ Values of **1** against Cathepsin L

| *Pre-incubation times (*minutes*)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 0 | 752 ± 17.77 |
| 1 | 155.9 ± 8.44 |
| 5 | 8.26 ± 0.8 |
| 15 | 7.504 ± 0.8 |
| 30 | 6.98 ± 0.7 |
| 60 | 3.08 ± 0.8 |
| 120 | 2.75 ± 0.3 |
| 240 | 2.48 ± 0.1 |

The effect of preincubation in the potency of **1** as seen in Table 67 is quite evident. The dibromobenzophenone thiosemicarbazone showed a potent activity when cathepsin L was not preincubation with the compound (752 nM). Compound **1** improved its activity five times with only one minute preincubation time. The trend changed dramatically when the preincubation time was increased to 5 minutes. The inhibitory activity of **1** increased 19-fold with a remarkable $IC_{50}$ less than 10 nM (8.3 nM). The potency of **1** was modestly increasing with longer preincubation times. Finally, the best potency was found when **1** was preincubated for four hours with cathepsin L. The activity of the brominated analog increased 4-fold to give a value of 2.5 nM. These results confirmed the strong dependence of $IC_{50}$ value determination with respect to the preincubation time parameter for **1**. Controls (i.e. uninhibited cathepsin L) were

monitored at every preincubation time. There was no significant loss of catalytic activity at longer preincubation times.

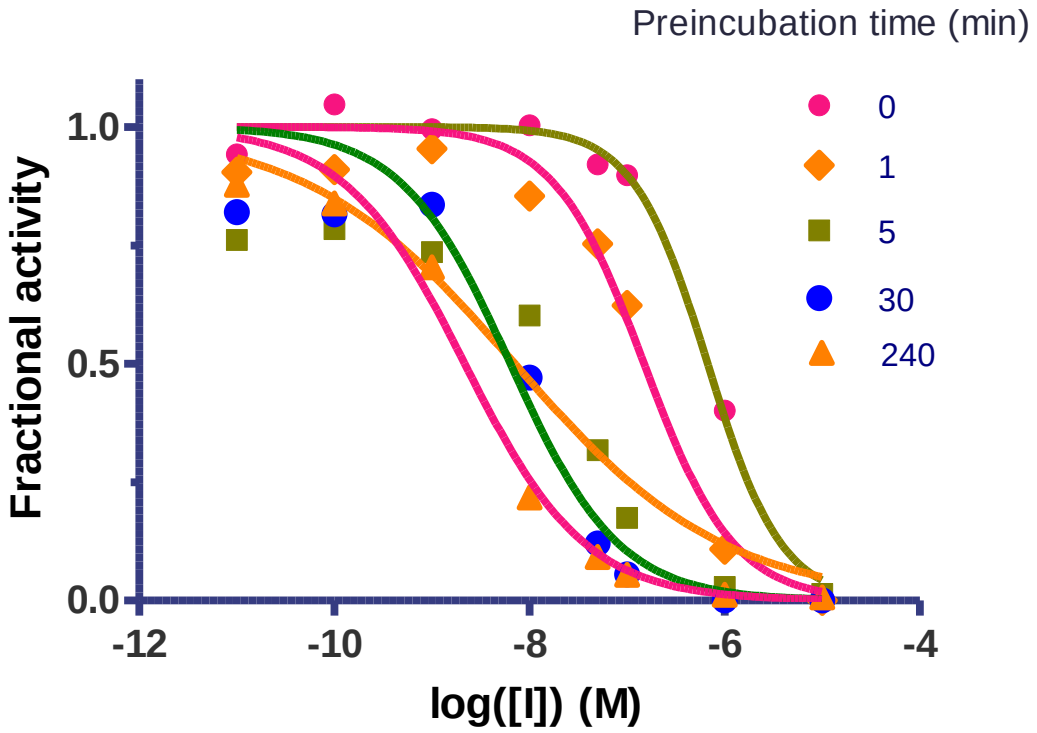



Figure 55. Effect of Preincubation Time on $IC_{50}$ Values of **1** against Cathepsin L

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation. Effect of Preincubation Time Using **1**.*

The data obtained from the preincubation time was further analyzed. The possibility that **1** was a tight-binding inhibitor was analyzed with the Morrison's quadratic equation. (See equation 2.3). The velocities $v_i$ and $v_o$ are the inhibited and uninhibited cathepsin L velocities (RFU/s), $[E]_T$ (nM) is the total concentration of enzyme found in solution (free enzyme and inhibitor-enzyme complex), $[I]_T$ (nM) is the total concentration of inhibitor present in solution (free inhibitor and inhibitor-enzyme complex), and $K_I$ (nM) is the inhibition constant, often referred as the dissociation

constant. The equation may be solved to give two possible answers. However, the equation is written so that only there is only one possible answer that fits the physiological conditions (i.e. $K_I^{app} > 0$). GraphPad 5.0 was used to fit the data for every preincubation time point. Inhibitor, substrate, and enzyme concentrations, as well as $K_M$, $K_I^{app}$, were all in micromolar units (μM). The residual activity (or $v_i/v_o$) was normalized to 1 (i.e. $v_o$: 1 and $0 \leq v_i \leq 1$). Normalized residual activity and [I] were defined as the dependent and independent variables, respectively. Nonlinear regression was applied using the following conditions: [S]: 50 μM, $K_M$: 1.1 μM, $[E]_T$: 0.001 μM and $v_o$: 1. Normalization and nonlinear conditions were used to three or four independent experiments per preincubation time. Average and standard errors can be seen in Table 64 and Figure 49. The $K_I^{app}$ values are shown in Table 68 and Figure 56. It can also be seen that the apparent inhibition constants are also time-dependent. No preincubation results in a good inhibition constant with a value of 138 nM. Furthermore, the potency is extremely enhanced (more than 38-fold) for the calculated value at a long preincubation time (1 hour). The apparent $K_I$ was 125-fold better when compared with the corresponding value at no preincubation time ($K_I^{app}$ : 1.1 nM).

Table 68. Effect of Preincubation Time in $K_I^{app}$ values of **1** against Cathepsin L

| *Pre-incubation times (*minutes*)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 0 | 138 ± 13.1 |
| 1 | 24.7 ± 1.2 |
| 5 | 3.6 ± 0.6 |
| 15 | 2.1 ± 0.2 |
| 30 | 1.5 ± 0.1 |
| 60 | 1.1 ± 0.1 |
| 120 | ND |
| 240 | ND |

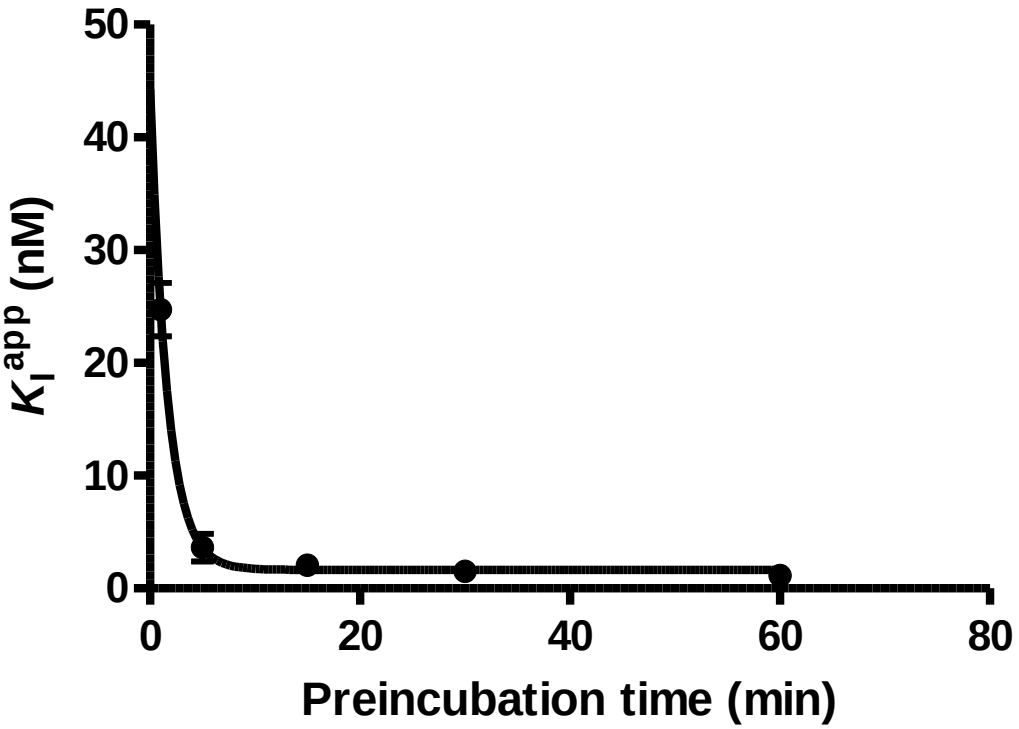



Figure 56. Effect of Preincubation Time in $K_I^{app}$ Values of **1** against Cathepsin L

*Cathepsin L Reversibility Studies*

Compound **1** was found to be a time-dependent inhibitor in the preincubation studies. However, we decided to explore if this specific compound was a reversible inhibitor of cathepsin L. A mixture containing 100 X cathepsin L and 10 X $IC_{50(\text{preincubation time: 5 minutes})}$, which are 100 and 170 nM, were incubated at 25 °C between 0 and 60 minutes. The inhibition of cathepsin L by **1** was able to be monitored by the rapid dilution of the mixture (100-fold) with assay buffer containing Z-FR-AMC. Final conditions are 1 nM cathepsin L, 1.7 nM of dibromobenzophenone thiosemicarbazone (**1**), and 50 μM. Additionally, a control experiment (cathepsin L with DMSO as control vehicle) was also set up. Complete experimental details can be found in the material and methods seconds. Figure 57 shows the release of AMC for the first 3000 seconds. Uninhibited and inhibited reactions were followed for a total time of four hours.

Cathepsin L was able to recover its catalytic activity after the rapid dilution with assay buffer containing Z-FR-AMC. Apparent substrate depletion can be observed in the uninhibited reaction after 5000 seconds. A linear regression was applied to the first 3000 seconds of the reactions to determinutese cathepsin L activity. Cathepsin L activity for the untreated reaction was found to be 0.71 nM AMC/s. Similarly, the inhibited reaction when the system was preincubated for one hour was 0.20 nM AMC/s. The results showed cathepsin L was inhibited more than 80% with **1.** Enzymatic activity can be seen within the first 300 seconds (Figure 51).

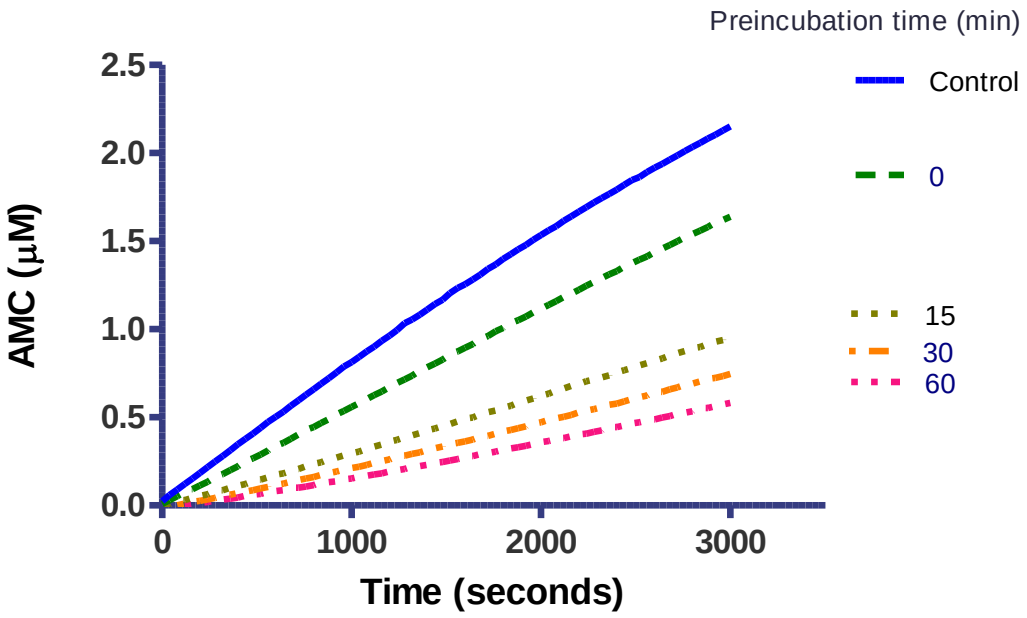



Figure 57. Cathepsin L Reversibility Studies with **1** Using 50 μM Z-FR-AMC

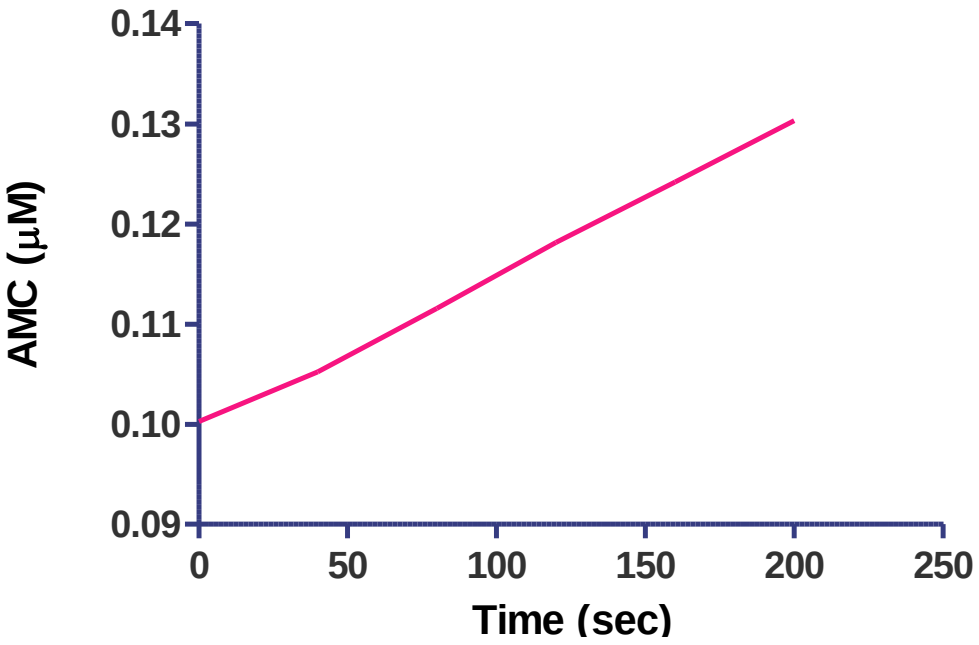



Figure 58. Cathepsin L Reversibility Studies with **1** Using 50 μM Z-FR-AMC. Preincubation time: 1 hour.

*Effect of Substrate Concentration (Z-FR-AMC) in $IC_{50}$ Values Using **1***

We finally tried to determine the mode of inhibition of the one of the most potent of the thiosemicarbazone inhibitors of cathepsin L. Therefore, we investigated the effect of substrate concentration to verify if these synthetic compounds are competitive, noncompetitive, or uncompetitive inhibitors. Three substrate concentrations (Final concentrations: 50, 20, and 5 μM) were used to determine $IC_{50}$ values for compound **1.** Cathepsin L and the series of inhibitors were incubated for five minutes. Results can be found in Table 69 and Figure 59.

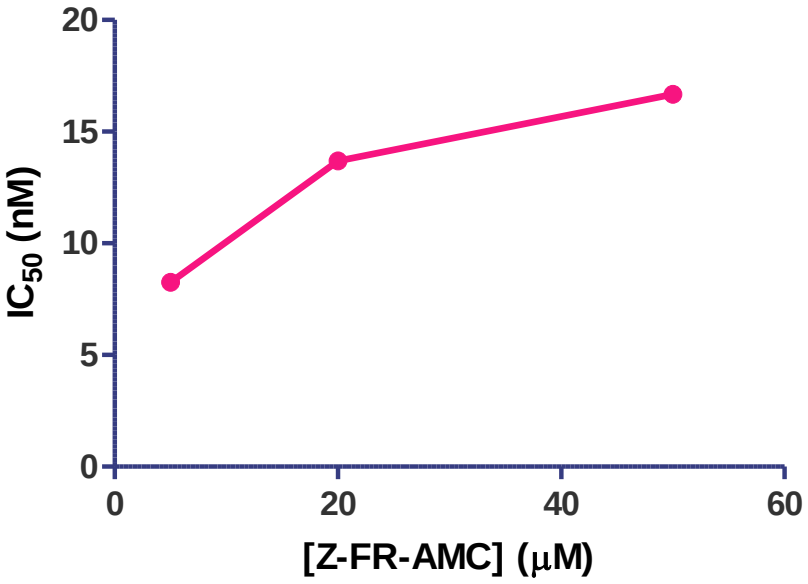



Figure 59. . Effect of Substrate Concentration in $IC_{50}$ values of **1** against cathepsin L

Table 69. Effect of Substrate Concentration on $IC_{50}$ Values of **1** against Cathepsin L

| *[Z-FR-AMC] (μM)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 50 | 16.7 ± 1.0 |
| 20 | 13.7 ± 1.7 |
| 5 | 8.3 ± 0.8 |

The effect of substrate concentration was investigated using **1** as the lead compound of the series of thiosemicarbazones. A pseudo linear behavior can be observed with the values. Lowering the amount of substrate in solution gives a subtle, yet evident increase in the potency of the dibrominated analog. According to literature, a positive linear behavior (i.e. $IC_{50}$ α [S]) is an indication of the compound acts as competitive inhibitor; that is, both substrate and the compound compete for cathepsin L active site. However, the results cannot be taken as conclusive and more experiments were necessary to carry out to confirm these findings.

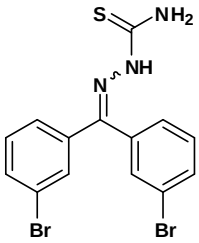
*Cell Culture Experiments*

Preliminary screening of a library of more than 150 thiosemicarbazones provided interesting and promising results. More than 6 compounds were found to potent cathepsin L inhibitors, (See Table 28-Table 45). It was also shown that one of the compounds, the phenolic brominated benzophenone thiosemicarbazone was capable to inhibit the collagenase activity of cathepsin L *in vitro.* Thus, we decided to explore the potency of some of the most promising compounds to arrest the invasion and motility of MDA-MB-231 cells, a human breast cancer cell line that has been previously investigated and showed high levels of cathepsin L, invasiveness and motility properties. Figure 60 shows five of the most promising compounds that were used to perform cell culture experiments. Two analogs belong to the *para*brominated benzophonenone thiosemicarbazone series (Compounds **1** and **8**), and the remaining three compounds are part of the benzoylbenzophenone subfamily. E-64 (Table 8, page 31) , a nonspecific cysteine protease irreversible inhibitor was used as a positive control to verify the activity of the compounds. E-64 was selected due to being highly selective towards cysteine proteases and not reacting with other enzymes. E-64 has also been used in cell studies due its permeability characteristics.[336]

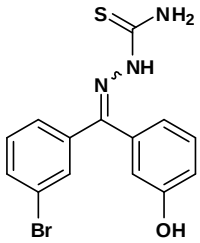
Invasion and migration assays were carried out using MDA-MB-231 cells that have been harvested with DMEM media supplemented with 10% FBS. Maintenance and subculture procedures have been described in the material and procedure secondstion of this chapter. Three separate T75 flasks with MDA-MB-231 cells (passages 13, 13, and 13) were used to carry out invasion and migration experiments. Briefly, stock solutions

of the compounds and the positive control were made according to Table 25 and Table 26. Stock solution concentrations for each compound were 50 and 20 μM in 4% DMSO.

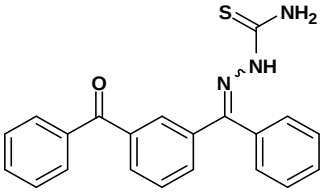
MDA-MB-231 cells were trypsinized when they were 90% confluent. DMEM media with low FBS was used to neutralize the enzymatic activity of the solution. Cell counting was used to calculate the cell density of each stock solution (cells/ml). Then, DMEM without FBS was added to each stock solution to make a dilution with a concentration of 200,000 cells/ml. Invasion inserts (inserts with a Matrigel® layer) and migration inserts, (inserts without a Matrigel® layer with an 8 μm pore) were rehydrated with no-FBS DMEM media at 37 °C in 5% $CO_2$/air environment. Media was removed and discarded. Immediately, inserts were placed into 24-well companion plates that contained DMEM media supplemented with 10% FBS. Every insert (for migration and invasion assays) contained a mixture with cells and two concentrations for every compound. Final concentrations were 50,000 cells, 2% DSMO, and 0, 10 (or 25) μM of compounds **1**, **8**, **156**, **157**, **168** and E-64. Samples were incubated for 24 hours at 37 °C in 5% $CO_2$/air environment. After incubation time, media was removed, and the inner sides of the membranes were cleaned to remove cells that did not invade or migrate. Then, membranes were stained with DiffQuik staining kit with azure and xanthenes dyes. Membranes were air-dried and placed on slides. Manual cell counting was performed using a Zeiss inverted microscope. Slides were placed onto the instrument, and were observed under a 40X objective. Cell that migrated the 8 μm membrane or cells that invaded the Matrigel® layer have a dark purple shade with a round shape. The purple coloration is due that cytoplasmic matter and nuclei are sensitive to both dyes. The dark color is also easier to follow and distinguish from other particles.



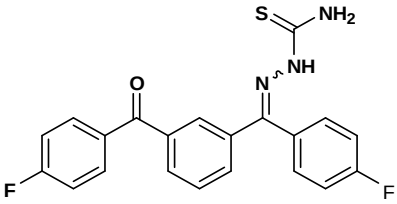

$IC_{50}$: 16.7 nM



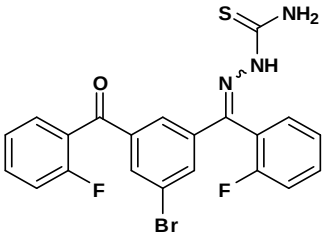

$IC_{50}$: 188.7 nM



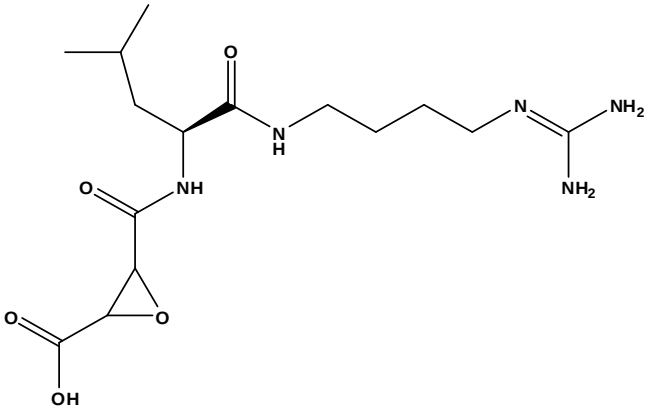

$IC_{50}$: 9.85 nM




$IC_{50}$: 24.3 nM




$IC_{50}$: 8.1 nM




$IC_{50}$: 2.5 nM[156]

Figure 60. Chemical Structures of **1**, **8**, **156**, **157**, **168** and E-64.

Figure 61. Invasion and Migration Assay Samples using MDA-D231 cells treated with **1**, **8**, **156**, **157**, **168** and E-64

Figure 61 shows the samples that were preserved in glass slides for further investigation. Manual cell counting was performed using a Zeiss inverted microscope. Samples were placed onto the instrument, and were observed under a 40X objective. Cells that migrated the 8 μm membrane or cell that invaded the Matrigel® layer have a dark purple shade with a round shape. The purple coloration is due that cytoplasmic matter and nuclei that are sensitive to both dyes. The dark color is also easier to follow and distinguish from other particles. On the other hand, pores look clear with a round shape, similar to a 'bubble trapped in the membrane'. Only purple dots were manually counted, but there were some cells that did not migrate/invade entirely and have a dark

shadow semi-round shape. These cells were also taken into account for the manual cell counting process.

Ten representative areas were taken for every sample. Eight areas located in the outer periphery of the sample were counted. Also, two more areas located at the center of the membrane were counted. Figure 18 shows a scheme that describes approximate positions for every counted area. Eighty-four samples were observed and analyzed. The list includes triplicates of controls with 2% DMSO (control vehicle), controls with no DMSO, and 10 or 25 μM of compounds **1**, **8**, **156**, **157**, **168** and E-64 (positive control). Figure 62, Figure 63, Figure 64, Figure 65,Figure 66, Figure 67 show panels of representative experiments of some of the samples that were analyzed.

Results were tabulated and total numbers of samples were calculated. Table 70, Table 71, Table 72, Table 73 show the final results of manual cell counting. Furthermore, the percentages of cell invasion and migration that were inhibited by the several compounds are defined in Equations 2.4, 2.5, 2.6 and 2.7.

$$\%\,Invasion = \frac{Cells\ that\ invaded\ \ the\ Matrigel(R)\ in\ the\ presence\ of\ compound}{Cells\ that\ invaded\ the\ Matrigel\ (R) with\ no\ compound} \tag{2.4}$$

$$\%\,Migration = \frac{Cells\ that\ migrated\ the\ Matrigel(R)\ in\ the\ presence\ of\ compound}{Cells\ that\ migrated\ the\ Matrigel\ (R) with\ no\ compound} \tag{2.5}$$

$$\%\,Inhibition\ (invasion)\text{: 100 - \% Invasion} \tag{2.6}$$

$$\%\,Inhibition\ (migration)\text{: 100 - \% Migration} \tag{2.7}$$

Table 70. Cell Migration (25 μM) Results

| 25 μM | Vehicle (2% DMSO) | No vehicle | # 1 | #8 | #156 | #157 | #168 | E-64 |
|---|---|---|---|---|---|---|---|---|
| | | | | | **Cell Count** | | | |
| E1 | 1002 | 1004 | 29 | 119 | 136 | 328 | 336 | 603 |
| E2 | 807 | 727 | 76 | 168 | 339 | 367 | 532 | 541 |
| E3 | 1101 | 766 | 67 | 137 | 121 | 635 | 332 | 657 |
| **Average** | 970 | 832.3 | 57.3 | 141.3 | 198.7 | 443.7 | 400.0 | 600.3 |
| **S.E.** | 86.4 | 86.6 | 14.4 | 14.3 | 70.3 | 96.3 | 66.0 | 33.5 |
| **% ± S.E.** | *100 ± 8.9* | *85.8 ± 8.9* | *5.9 ± 1.5* | *14.6± 1.5* | *20.4 ± 7.2* | *45.74± 9.9* | *41.2 ± 6.8* | *61.9± 3.5* |

Table 71. Cell Migration (10 μM) Results

| 10 μM | Vehicle (2% DMSO) | No vehicle | # 1 | #8 | #156 | #157 | #168 | E-64 |
|---|---|---|---|---|---|---|---|---|
| E1 | 1002 | 1004 | 96 | 222 | 576 | 360 | 750 | 904 |
| E2 | 807 | 727 | 286 | 350 | 578 | 668 | 727 | 871 |
| E3 | 1101 | 766 | 87 | 252 | 492 | 640 | 653 | 754 |
| **Average** | 970 | 832.3 | 156.3 | 274.7 | 548.7 | 547.7 | 710 | 843 |
| **S.E.** | 86.4 | 86.6 | 64.9 | 38.7 | 28.3 | 10.3 | 29.3 | 45.5 |
| **% ± S.E.** | *100 ± 8.9* | *85.8 ± 8.9* | *16.1 ± 6.7* | *28.3 ± 4.0* | *56.6 ± 2.9* | *66.8 ± 1.1* | *73.2 ± 3.0* | *86.9 ± 4.7* |

Table 72. Cell Invasion (25 uM) Results

| 25 μM | Vehicle (2% DMSO) | No vehicle | # 1 | #8 | #156 | #157 | #168 | E-64 |
|---|---|---|---|---|---|---|---|---|
| E1 | 714 | 728 | 70 | 86 | 583 | 361 | 402 | 209 |
| E2 | 889 | 819 | 102 | 69 | 535 | 270 | 316 | 167 |
| E3 | 699 | 632 | 89 | 114 | 476 | 237 | 90 | 160 |
| **Average** | 767.3 | 726.3 | 87.0 | 89.7 | 531.3 | 289.3 | 269.3 | 178.7 |
| **S.E.** | 61.0 | 54.0 | 9.3 | 13.1 | 30.9 | 37.1 | 93.0 | 15.3 |
| **% ± S.E.** | *100 ± 7.9* | *94.7 ± 7.0* | *11.3 ± 1.2* | *11.7 ± 1.7* | *69.3 ± 4.0* | *37.7 ± 4.8* | *35.1 ±12.1* | *23.3 ± 2.0* |

Table 73. Cell Invasion (10 μM) Results

| 10 μM | **Vehicle (2% DMSO)** | **No vehicle** | **# 1** | **#8** | **#207** | **#244** | **E#312** | **E-64** |
|---|---|---|---|---|---|---|---|---|
| E1 | 714 | 728 | 128 | 279 | 586 | 735 | 219 | 347 |
| E2 | 889 | 819 | 155 | 139 | 644 | 690 | 231 | 213 |
| E3 | 699 | 632 | 222 | 248 | 636 | 540 | 249 | 367 |
| **Average** | 767.3 | 726.3 | 168.3 | 222.0 | 622.0 | 655.0 | 233.0 | 309.0 |
| **S.E.** | 61.0 | 54.0 | 27.9 | 42.5 | 18.2 | 59.0 | 8.7 | 48.4 |
| **% ± S.E.** | *100 ± 7.9* | *94.7 ± 7.0* | *21.9 ± 3.6* | *28.9 ± 5.5* | *81.1 ± 2.4* | *85.4 ± 7.7* | *30.4 ± 1.1* | *40.3 ± 6.3* |

Six different compounds were tested to see their ability to inhibit cell invasion and migration using a well established colorimetric method to quantify these two characteristics. Five of these compounds are thiosemicarbazones analogs, and E-64 is an non-specific irreversible cysteine inhibitor that is used as a positive control. MDA-MB-231 cells are known to show high levels of invasiveness and motility.[337]

Cell migration was achieved with the set of six compounds that were tested. In general, substittuded *para*-brominated benzophenone thiosemicarbazones were the best compounds to delay cell migration. Compounds **1** and **8** showed an extraordinary capacity to delay cell migration. More than 85% of the cells (compared to control) did not migrate the 8 μM membranes. Furthermore, the benzoylbenzophenone subfamily was also able to significantly arrest cell migration. The unsubstituted analog (compound **156**) inhibited almost 80% of cell migration. Overall, this group of selected thiosemicarbazones demonstrated to have a better inhibitory activity compared to the positive control.

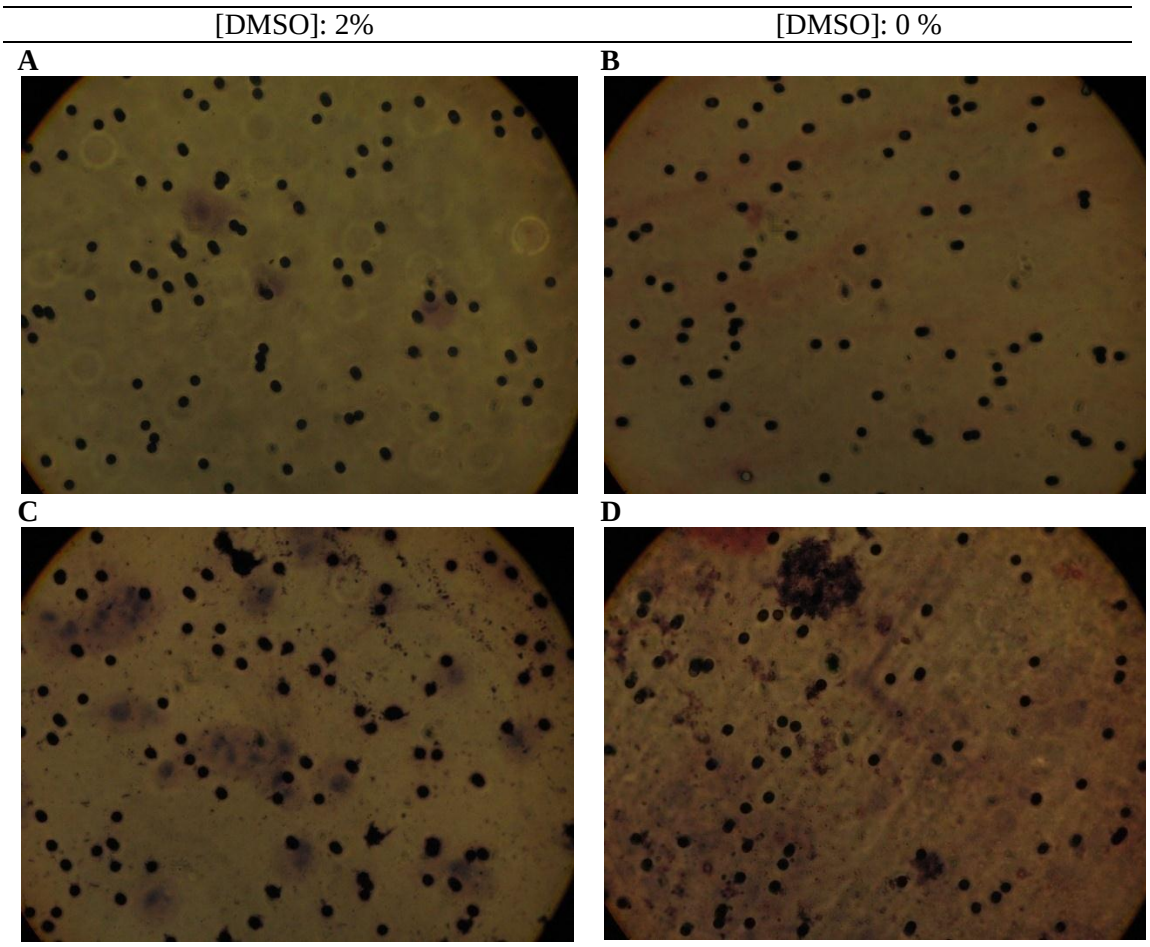
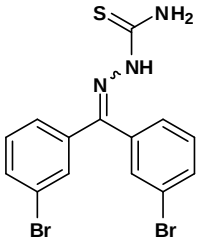



Figure 62. Representative Samples of Cell Invasion (A and B) and Cell Migration (C and D) Assays Using Non-Treated MDA-MB-231 Breast Cancer Cells. **A**. Invasion Assay, Control 2% DMSO. **B**. Invasion Assay, Control 0% DMSO. **C**. Migration Assay, Control 2% DMSO. **D**. Migration Assay, Control 0% DMSO**.**




| [**1**]: 25 μM | [**1**]: 10 μM |
|---|---|

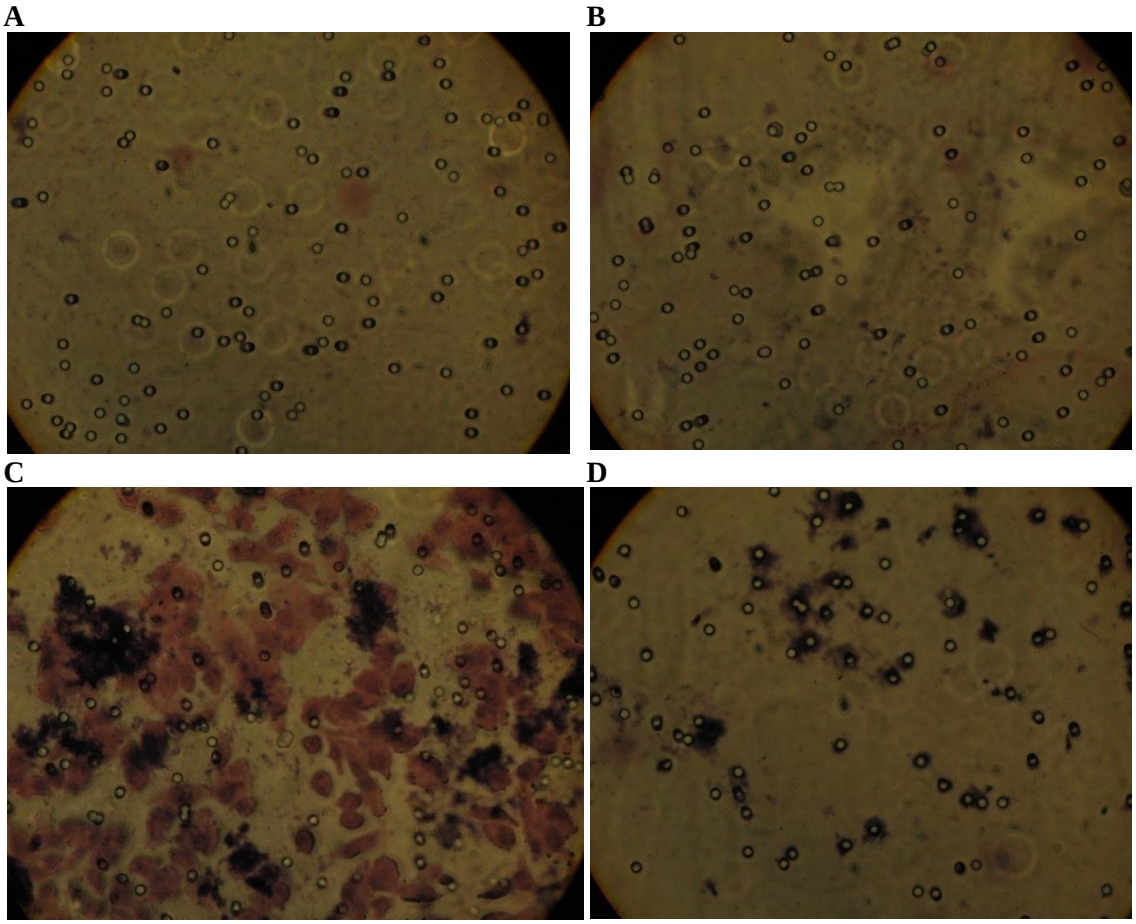
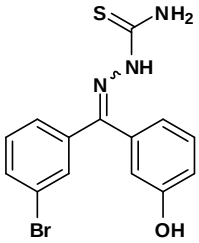



Figure 63. Representative Samples of Cell Invasion (A and B) and Cell Migration (C and D) Assays using MDA-MB-231 Breast Cancer Cells treated with 1. A. Invasion Assay, 25 μM of **1**. B. Invasion Assay, 10 μM of **1**. C. Migration Assay, 25 μM of **1**. D. Migration Assay, 10 μM of **1**




| [**8**]: 25 μM | [**8**]: 10 μM |
|---|---|

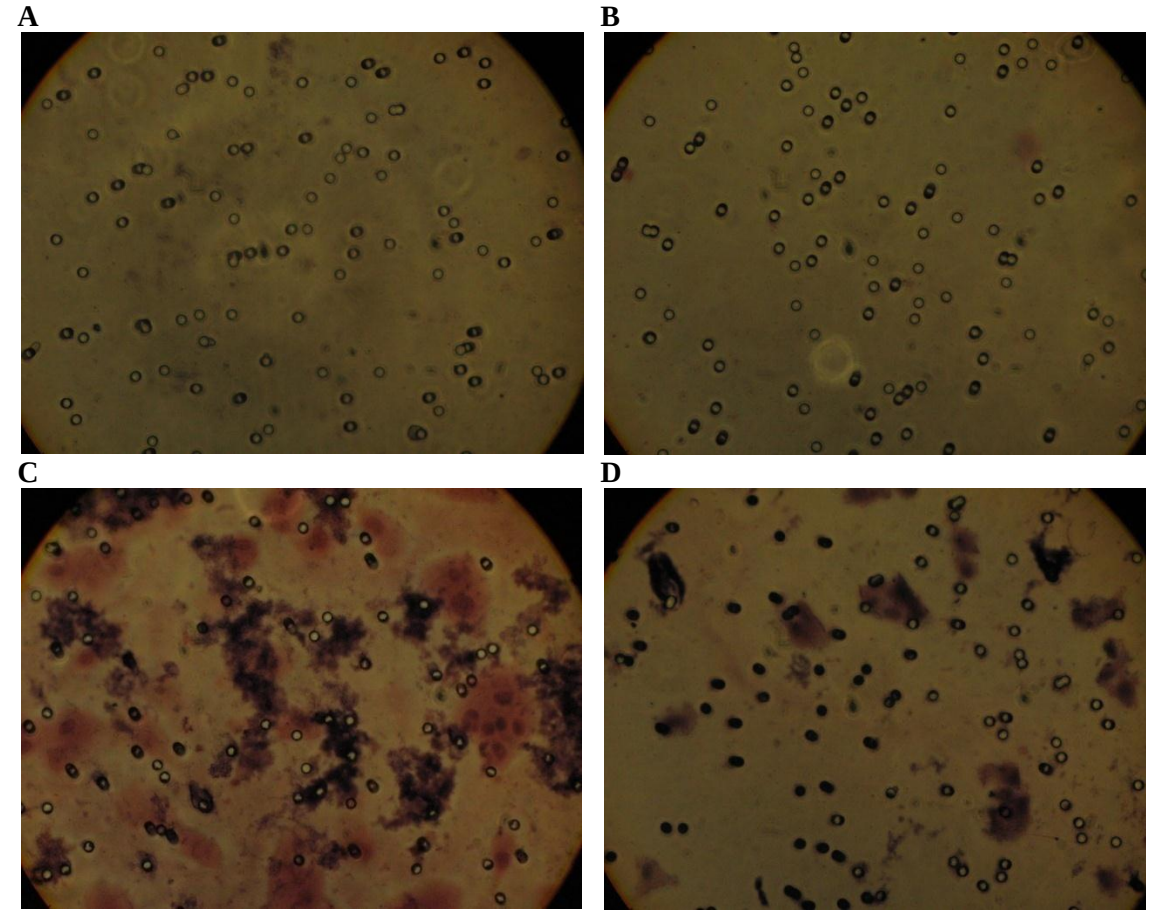
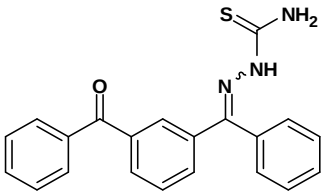



Figure 64. Representative Samples of Cell Invasion (A and B) and Cell Migration (C and D) Assays using MDA-MB-231 Breast Cancer Cells treated with **8**. A. Invasion Assay, 25 μM of **8**. B. Invasion Assay, 10 μM of **8**. C. Migration Assay, 25 μM of **8**. D. Migration, 10 μM of **8**



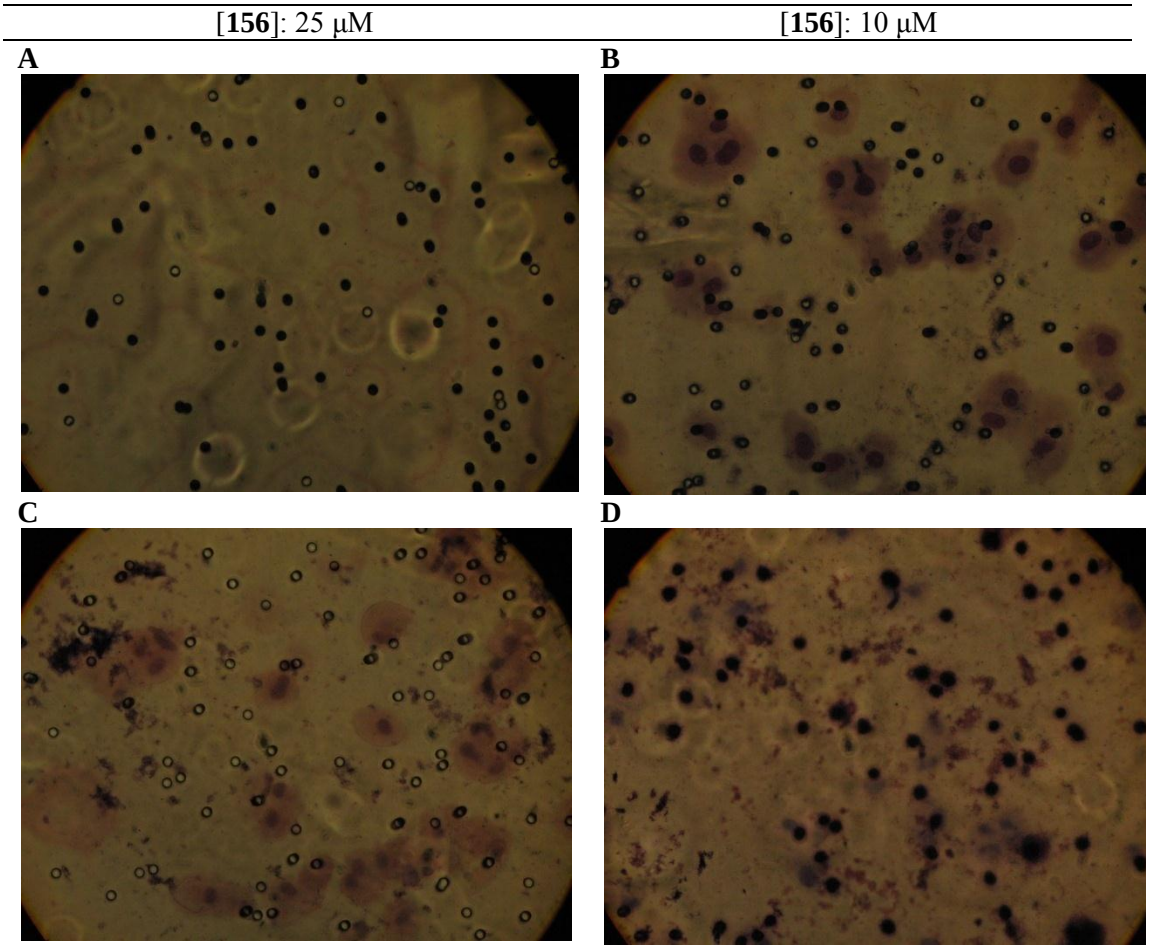
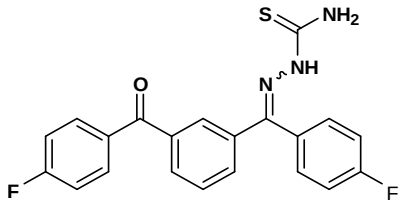




Figure 65. Representative Samples of Cell Invasion (A and B) and Cell Migration (C and D) Assays using MDA-MB-231 Breast Cancer Cells treated with **156**. A. Invasion Assay, 25 μM of **156**. B. Invasion Assay, 10 μM of **156**. C. Migration Assay, 25 μM of **156**. D. Migration Assay, 10 μM of **156**



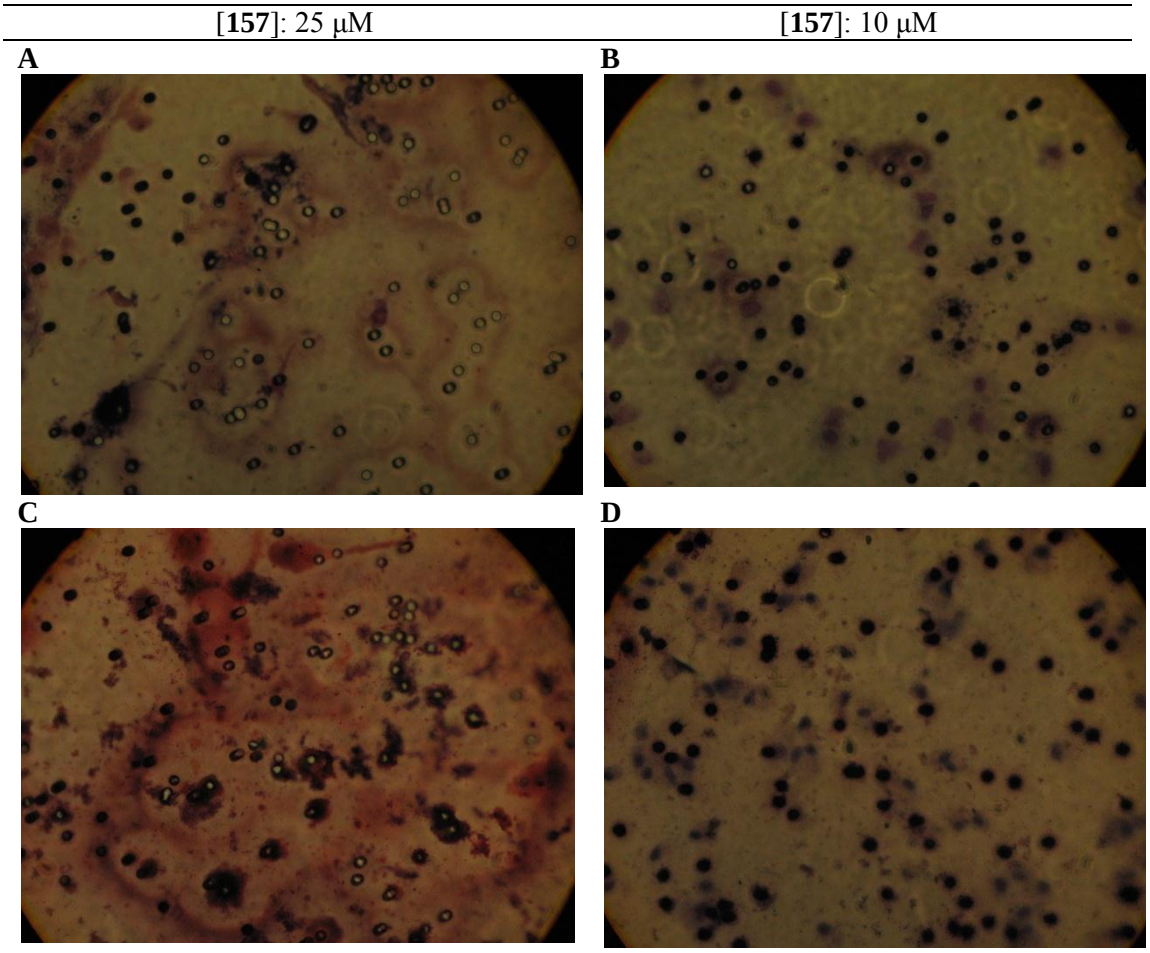
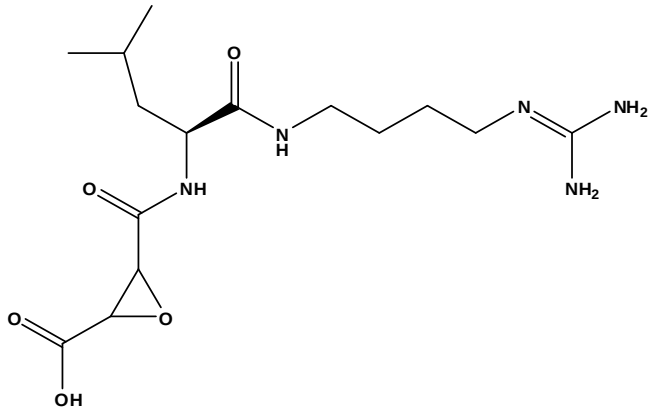




Figure 66. Representative Samples of Cell Invasion (A and B) and Cell Migration (C and D) Assays using MDA-MB-231 Breast Cancer Cells treated with **157**. A. Invasion Assay, 25 μM of **157**. B. Invasion Assay, 10 μM of **157**. C. Migration Assay, 25 μM of **157**. D. Migration Assay, 10 μM of **157**




| [**E-64**]: 25 μM | [**E-64**]: 10 μM |
|---|---|

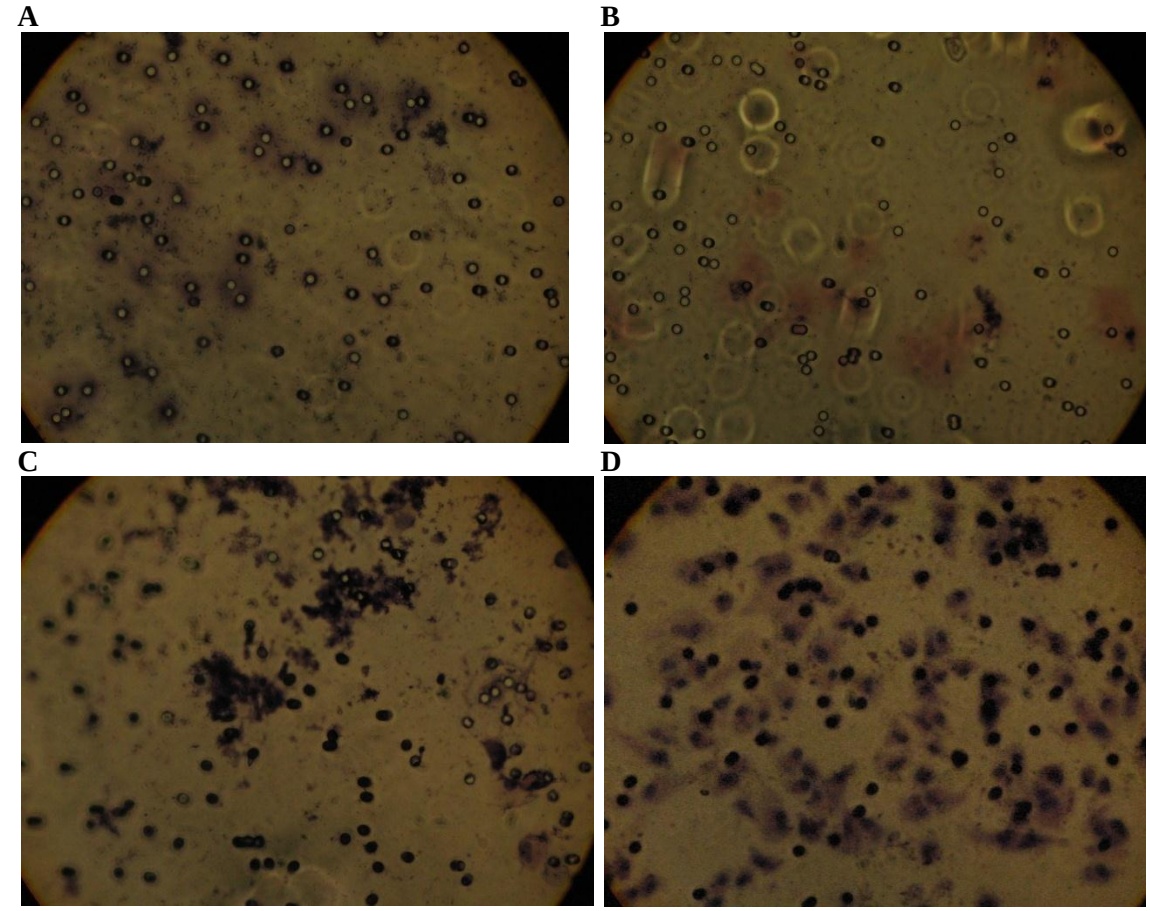



Figure 67. Representative Samples of Cell Invasion (A and B) and Cell Migration (C and D) Assays using MDA-MB-231 Breast Cancer Cells treated with E-64. A. Invasion Assay, 25 μM of E-64. B. Invasion Assay, 10 μM of E-64. C. Migration Assay, 25 μM of E-64. D. Migration Assay, 10 μM of E-64

Compound **157**, the difluorobenzoyl benzophenone thiosemicarbazone is almost 1.5 times more potent that E-64. Results may also indicate that thiosemicarbazones have a good permeability when a proper solvent vehicle concentration is used (2% DMSO).

Compounds **1** and **8** also showed to be the best compounds of this series of analogs (more than 88% inhibition). Furthermore, the hetero halogenated analog of the benzoylbenzophenone series was able to significantly inhibit cell invasion. However, the concentration effect does not seem to enhance the potency of this compound. It was the only one of the subfamily that proved to be a better cell invasion agent when compared to E-64 results. Finally, the unsubstituted benzoyl benzophenone did not show a significant activity when inhibiting MDA-MB-231 cell invasion.

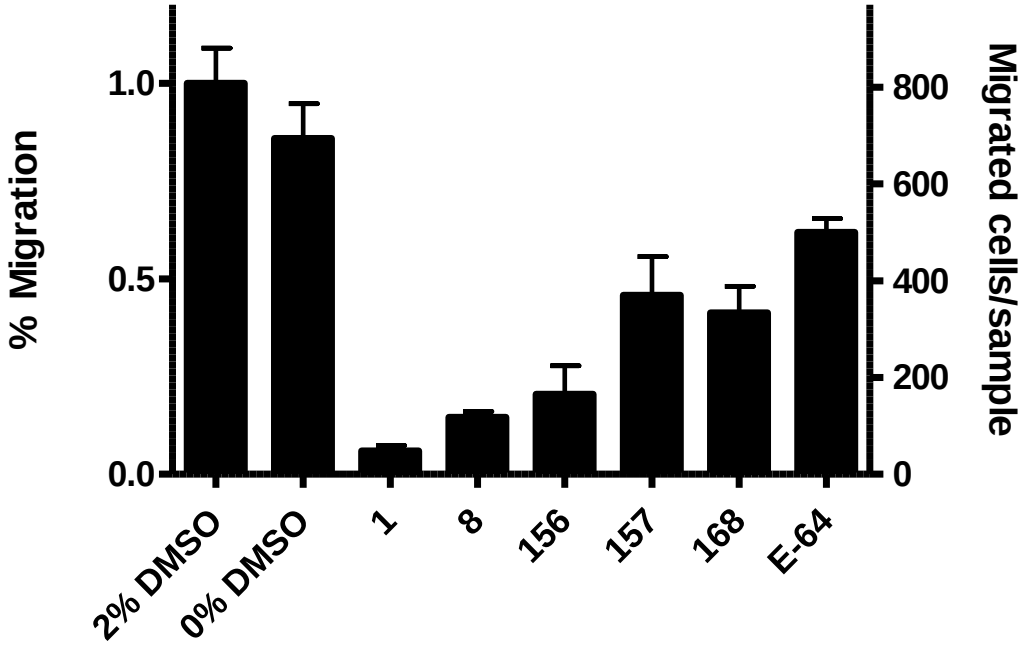



Figure 68. Cell Migration (25 μM) Results

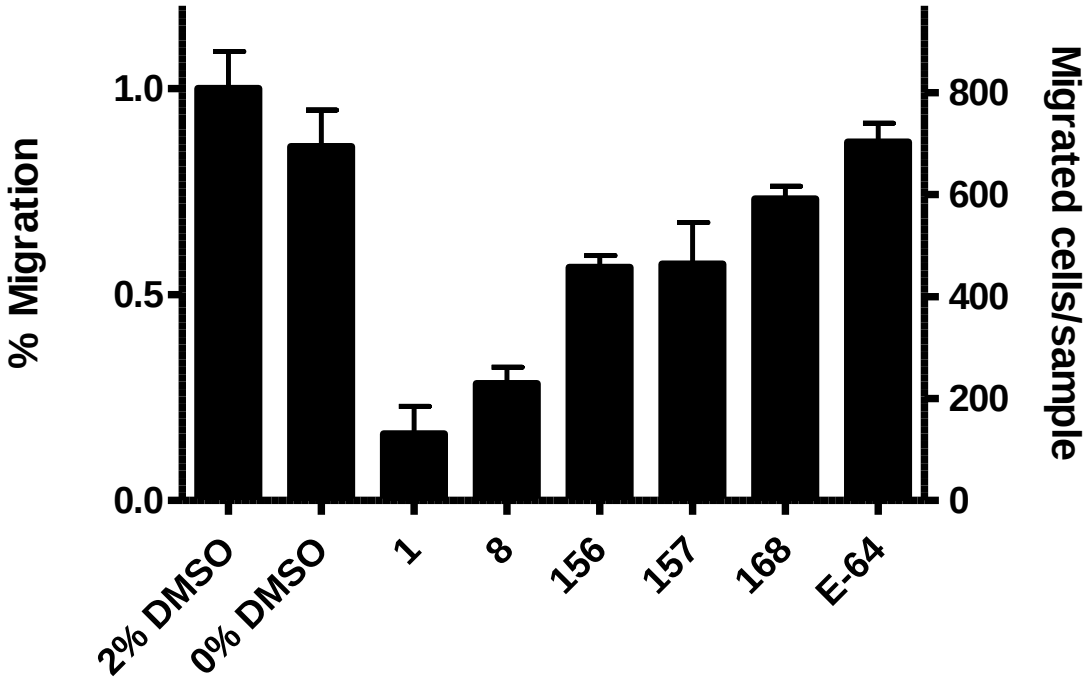



Figure 69. Cell Migration (10 μM) Results

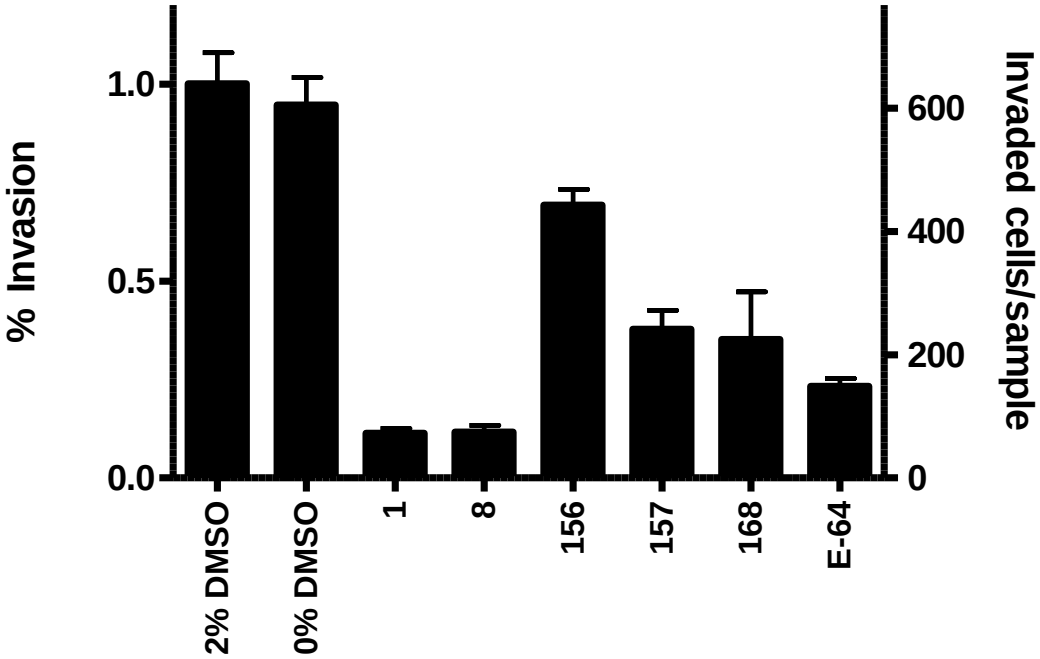



Figure 70. Cell Invasion (25 μM) Results

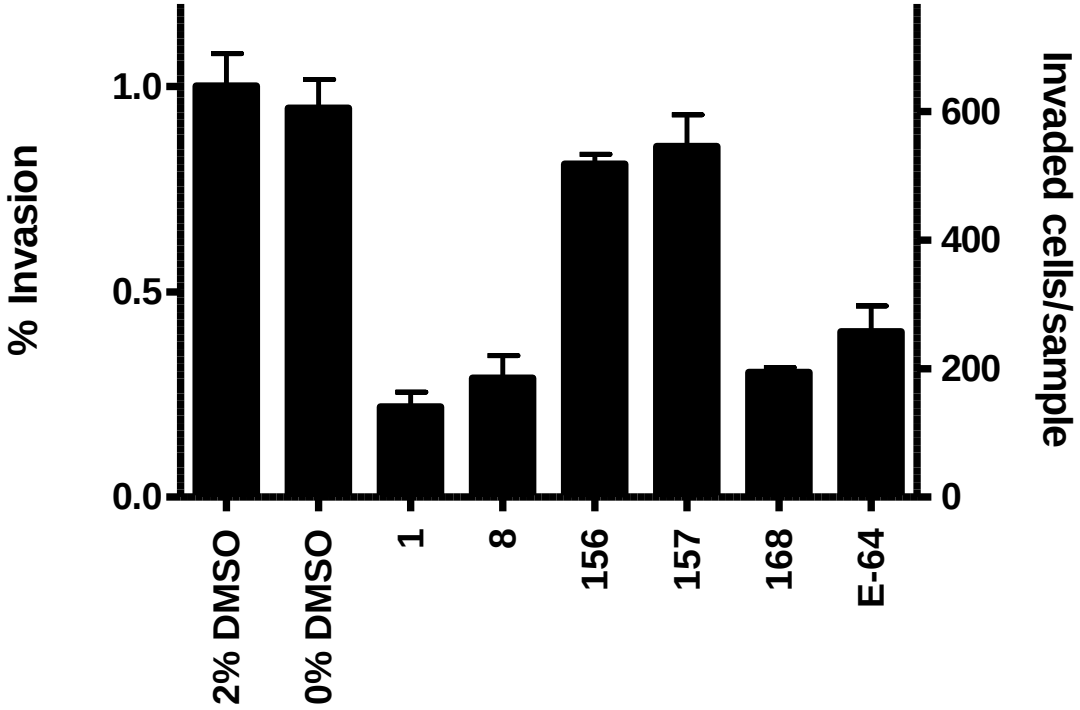
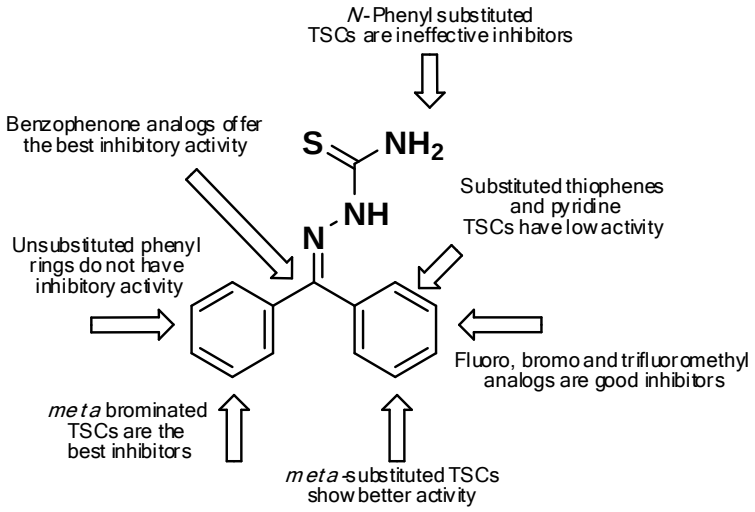



Figure 71. Cell Invasion (10 μM) Results

*Conclusions*

A medium size library of 168 thiosemicarbazones analogs were evaluated as inhibitors of recombinant, cathepsin L from human liver. Eight analogs arose as excellent cathepsin L inhibitors ($IC_{50} \leq 50$ nM). The examined library included functionalized benzophenone, thiochromanone, sulfones, and benzoyl benzophenone thiosemicarbazones. Figure 72 shows the Structure-Activity Relationship of thiosemicarbazones based on their chemical structures.

**A.**




**C.**

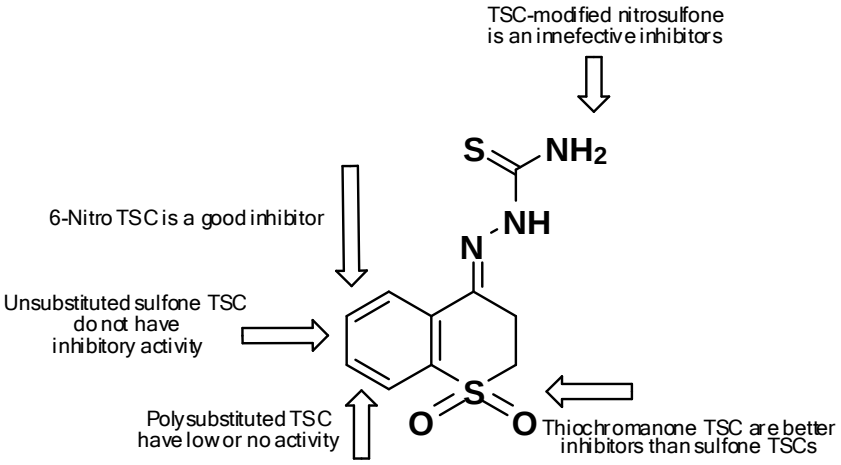
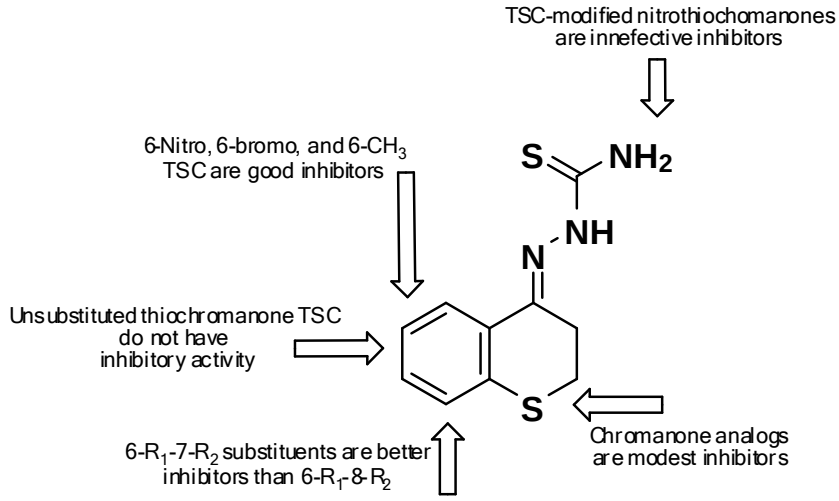



Figure 72. Structure-Activity Relationship for TSCs as Cathepsin L Inhibitors. **A**. Benzophenone TSCs. **B**. SulfoneTSCs

**A.**




**B.**

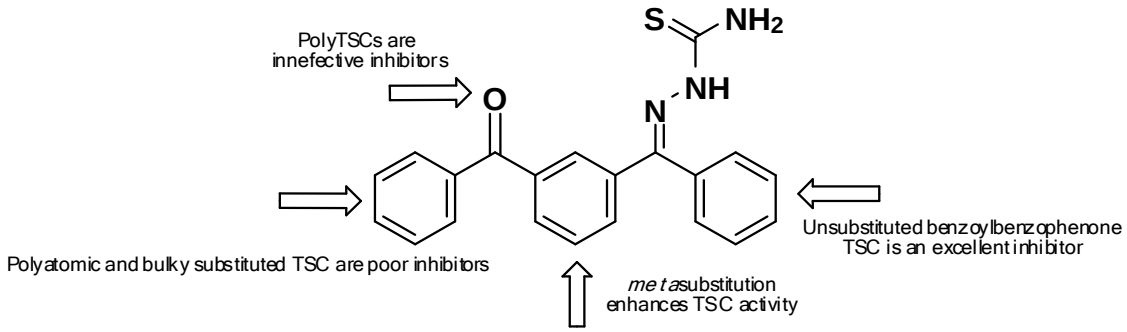



Figure 73. Structure-Activity Relationship for TSCs as Cathepsin L Inhibitors. **A**. Thiochromanone TSCs. **B**. Benzoylbenzophenone TSCs

Two lead compounds, **1** and **8** were characterized by using advanced kinetic studies to determine kinetic parameters, mechanisms and mode inhibition of these non-peptidic inhibitors. Both compounds were determined to be slow, time-dependent, reversible, competitive inhibitors of the fluorogenic substrate Z-FR-AMC. Kinetic studies with **8**, are consistent the phenolic brominated benzophenone thiosemicarbazone inhibits cathepsin L by simple reversible mechanism. However, a two-step mechanism is suggested at lower substrate concentrations. The inhibition of cathepsin L collagenase activity was also studied and inhibited when **8** was used in *in vitro* studies using type I

collagen from human skin as a natural substrate. Compound **8** was able to inhibit the catalytic activity of cathepsin L more than 50% after 6 hours of reaction.

Finally, five compounds (**1**, **8**, **156**, **157**, **168**) were used with MDA-MB-231 cells, (adenocarcinoma breast cancer), to test their ability to inhibit cell invasion and cell migration of this cell line. Functionalized benzophenone thiosemicarbazones showed the most potent activities among the selected group by inhibiting more than 85% of cells invading Matrigel®, a mixture containing collagen, fibronectin and laminin, which resembles the complex extracellular matrix found in human tissues.

# CHAPTER THREE

## Evaluation of Thiosemicarbazones as Cathepsin K Inhibitors

### *Nomenclature, Classification and Historical Background*

Human cathepsin K (EC 3.4.22.38) is a hydrolase and belongs to the subclass of peptidases, enzymes that act on peptidic bonds.[338] Similarly to cathepsin L, cathepsin K is also classified as an endopeptidase, that is, a hydrolase that only acts on peptidic bonds located within amino acid sequences.[339] Basically, cathepsin K is a cathepsin L-like enzyme as well.[340]

The official name of the protein is cathepsin K (CK). CK is also known as cathepsin O or cathepsin O2.[193] The discovery of cathepsin K is more recent that its cathepsin L analogue. reports of the presence of cathepsin K were reported by Tezuka and coworkers in 1994 when they cloned this protease from rabbit osteoclasts.[341] Cathepsin K is so far, one of the most studied cysteine proteases in mammalian organisms due to its extraordinary versatility and hydrolytic capacities. These aspects will be discussed in this chapter.

### *Localization of cathepsin K*

Interestingly, cathepsin K is a lysosomal cathepsin L-like protease and its distribution in human tissues is more limited compared to cathepsin L. This might be due to its highly capacity to degrade macromolecules among tissues and organs. Specifically, expression of cathepsin K has been found in osteoclasts (these cells are found in bone matrix and are responsible of the degradation of the proteins found in bone tissue),

macrophages, epithelial cells of several systems (respiratory, and gastrointestinal), urinary tract of human fetus, skin, ear, lung, ovary, testis, and colon.[105,342–347]

*Natural Substrates of Cathepsin K*

Cathepsin K (CK) is one of the most potent of the cysteine cathepsins. It plays a critical role in bone resorption as it will be discussed later in this chapter. The activity of this cysteine protease has been well investigated in numerous in *in vitro* studies.

Research has proved that extracellular matrix components are the preferred substrates for cathepsin K.[348] Cathepsin K is capable of degrading aggrecan, elastin, and osteonectin among other proteins.[342,349,350]

Perhaps collagen is cathepsin K's special substrate. Cathepsin K collagenase activity has been verified by using multiple collagen substrates including types I, II, III, and IV collagen.[351–354] Types I and II collagen are very resistant to degradation.[355] Few enzymes are able to utilize them as substrates. The selected group includes MM1, MMP8, MMP13, MMP14, and cathepsin K.[356] The uniqueness of cathepsin K among the cathepsins resides in its extraordinary capacity for degrading both the telopeptides-globular structures located at the ends of native molecules of collagen- and the triple helix chains.[222,351,357,358] Other cysteine cathepsins fail in this task. Cysteine cathepsins have telopeptidase activity, while MMPs are only able to degrade the triple helices of collagens. The difference between CK and the rest of the papain family is found in its catalytic active site. Cathepsins L, K, V, and S belong to the cathepsin L-like subfamily. They all share similar structural characteristics, but there are differences in their substrate specificity. Cathepsin L tends to favor aromatic hydrophobic residues for peptidic cleavage at the P2 position (i.e. N-Phe-X-C); on the other hand, cathepsins S and K

might cleave any aliphatic residues (i.e. N-Leu-X-C). The main difference between cathepsin L and K is their specificity at the P2 position. Cathepsin L can accept both aliphatic and aromatic residues, while cathepsin K's position only favors aliphatic amino acids residues in the P2 position (i.e. N-Leu-X-C or N-Ile-X-C. Furthermore, cathepsin K is able to cleave peptides with proline and glycine at the P2 and P3 positions, which are unusual residues for substrates of cysteine cathepsins.[225,359] This explains why proline-rich macromolecules, such as type I and II collagens, are excellent cathepsin K substrates.[360]

### *Biological Roles of Cathepsin K*

Cathepsin K distribution is selective in human tissues. Therefore, its physiological function is limited to a specific number of processes.

The most important, and perhaps the best known, process is the participation of cathepsin K in bone resorption. Type I collagen is one of the most abundant fibrillar proteins in the bone matrix. This fibrillar protein has a molecular size of almost 300 kDa. Type I collagen (TIC) is composed of three helical chains, and telopeptides, the end of these chains.[361] MMPs with collagenase activity lack telopeptidase activity. This is not the case for cathepsin K, which has both activities combined (collagenase, telopeptidase).[351,362,363]

### *The Collagen Family: Type I and Type IV Collagens*

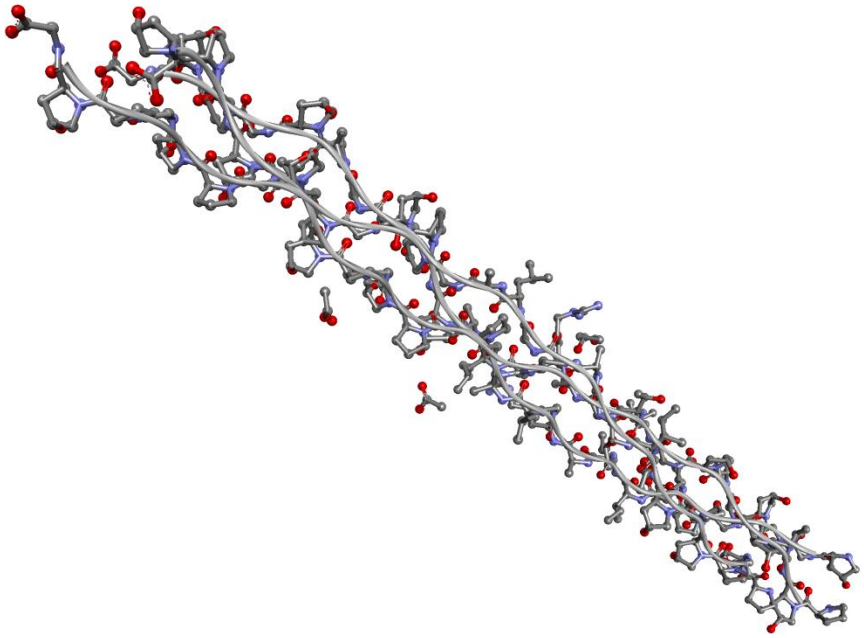
The collagen family comprises, to date, twenty-eight types of this fibrillar protein.[333] However, the collagens share twenty-five α-chains. There are several characteristics that make collagens one of the most studied proteins. Even though the collagen are practically insoluble, their versatility is very well known and are used in

cosmetics, biomaterials, tumorigenic studies, drug delivery, ulcer treatment among other applications.[364,365] Collagens' amino acid sequence is quite simple, yet efficient: Gly-X-Y. X and Y are usually proline and hydroproline (Hyp) residues. The roles of each amino acid, glycine, proline and hydroxyproline are well known. Glycine side chain, a hydrogen atom gives the characteristic compactness to the molecule. The side chain is able to fit into the tight conformation formed by three helical chains. Also, hydrogen bonds between NH and C=O groups of two adjacents helical chains backbone are formed. Both, proline and hydroxyproline provide the characteristics kinks found in the molecule, due to the fixed angle formed in the amino acid structure (C-N).[224]

Tropocollagen, or simply collagen is made of three left handed α-chains that form a right-handed macromolecule. The main function of the collagens is to provide support and anchor surrounding cells within the body, such as connective tissues, bones, and extracellular matrix. It was believed that fibroblasts, connective tissues cells, were the only ones responsible for the expression of collagen. Nevertheless, additional research has demonstrated that other epithelial cells can also express several types of collagens. Figure 74 shows the crystal structure of type I collagen (PDB ID: 1CAG)[366] and the non collagen domain of type IV collagen (PDB ID: 1T61)[367]

Collagens can be classified by their α-chain composition (homotrimeric or heterotrimeric), and quaternary structures (or form of polymerization): fibrillar, fibril associated with interrupted triple-helices (FACIT), short chain, basement membrane, multiplexins, membrane associated with interrupted triple-helices (MACIT).[333,368]

**A.**



**B.**

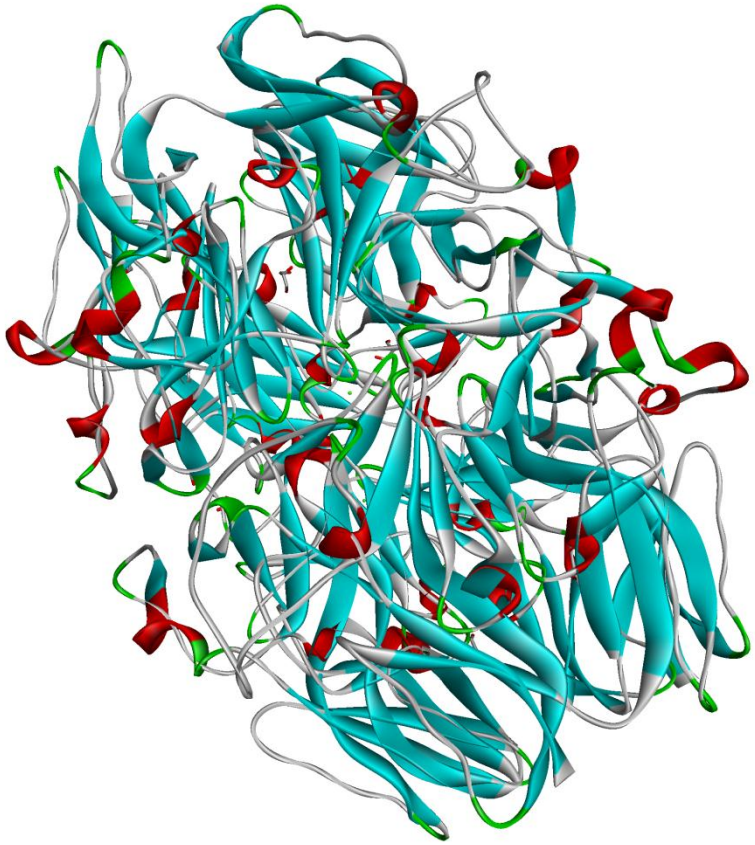


Figure 74. Crystal Structure of Collagen. **A** Type I Collagen Triple Helix Structure (PDB ID: 1CAG). **B.** Type IV Collagen NC1 Domain from Human Placenta (PDB ID: 1T61).

Type I collagen (TIC) is considered a heterotrimeric, fibril-forming macromolecule with a molecular formula of $[\alpha1(I)]_2\ \alpha2(I)$. It is the most abundant member of this family (up to 90%) and is widely found in extracellular matrix, bone, skin, tendons, ligaments, cornea, etc. The average length of each α-chain is approximately 1400 residues (approximately 300 nm, and 0.5- 3 μm diameter in their aggregated form).[368]

Type IV collagen (TIVC) is also a heterotrimeric, network forming macromolecule with a molecular formula of $[\alpha1(IV)]_2\ \alpha2(IV)$. TIVC is only found in the basal lamina, which is a layer of the extracellular matrix.

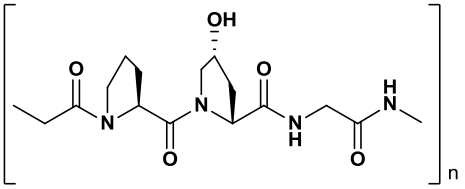



Figure 75. Representative Amino Sequence of the Collagens[369]. The sequence represents Pro-Hyp-Gly

Similar to type I collagen, TIVC's three α-chains contain 1400 residue per chain. Each one of them is composed of three different domains: an amino-terminal, the Gly-Pro-Hyp main chain and the C-terminal which contains approximately 230 residues that have non-collagen characteristics. The N-terminal domain is rich in cysteine and lysine residues. Both amino acids are responsible for the interchain crosslinking between four triple helical molecules through disulfide bonds and lysine-hydroxylysine bonds. The strong interactions between the triple helices provide the macromolecule high stability and low susceptibility to proteolytic degradation. Perhaps the unique feature of TIVC is the presence of a pattern of interruptions that occur in the main chain (~ 21-26). The

interruptions provide the macromolecule the ability to form two-dimensional networks rather than usual 3-D structures that are more common to find within other members of the collagen family, such as TIC. Khoshnoodi offered an extensive review about mammalian collagen IV.[370]

*The Effect of Chondroitin 4-Sulfate on Cathepsin K Collagenase Activity*

The collagenase activity of cathepsin K is greatly enhanced by the presence of chondroitin 4-sulfate (C4-S), a glycosaminoglycan.[371] In constrast, the macromolecule acts as a inhibitor of other member of the cathepsin family, like cathepsin L and S, which share structural similarities with cathepsin K.[372] The crystal structure of cathepsin K complexed with E-64 and C4-S was solved by Li and coworkers (Figure 76).[373] Their findings can be summarized as follows:

1. E-64 binds to Cys25 via covalent modification. (Cys25 thiol group – E64 epoxide ring. E-64 occupies cathepsin K's S2 and S3 pockets
2. C4-S binds to the R-domain of cathepsin K which is located on the opposite side of the protease's active site.
3. Hydrogen bonds and ion-pairs are the responsible of stabilizing the complex formed by cathepsin K and C4-S.
4. The backbones and/or side chains of D6, R8, K9, K10, I171, E172 and N190 interact with C4-S moieties.
5. The minimal ratio between C4-S and cathepsin K was calculated to be 1:3.

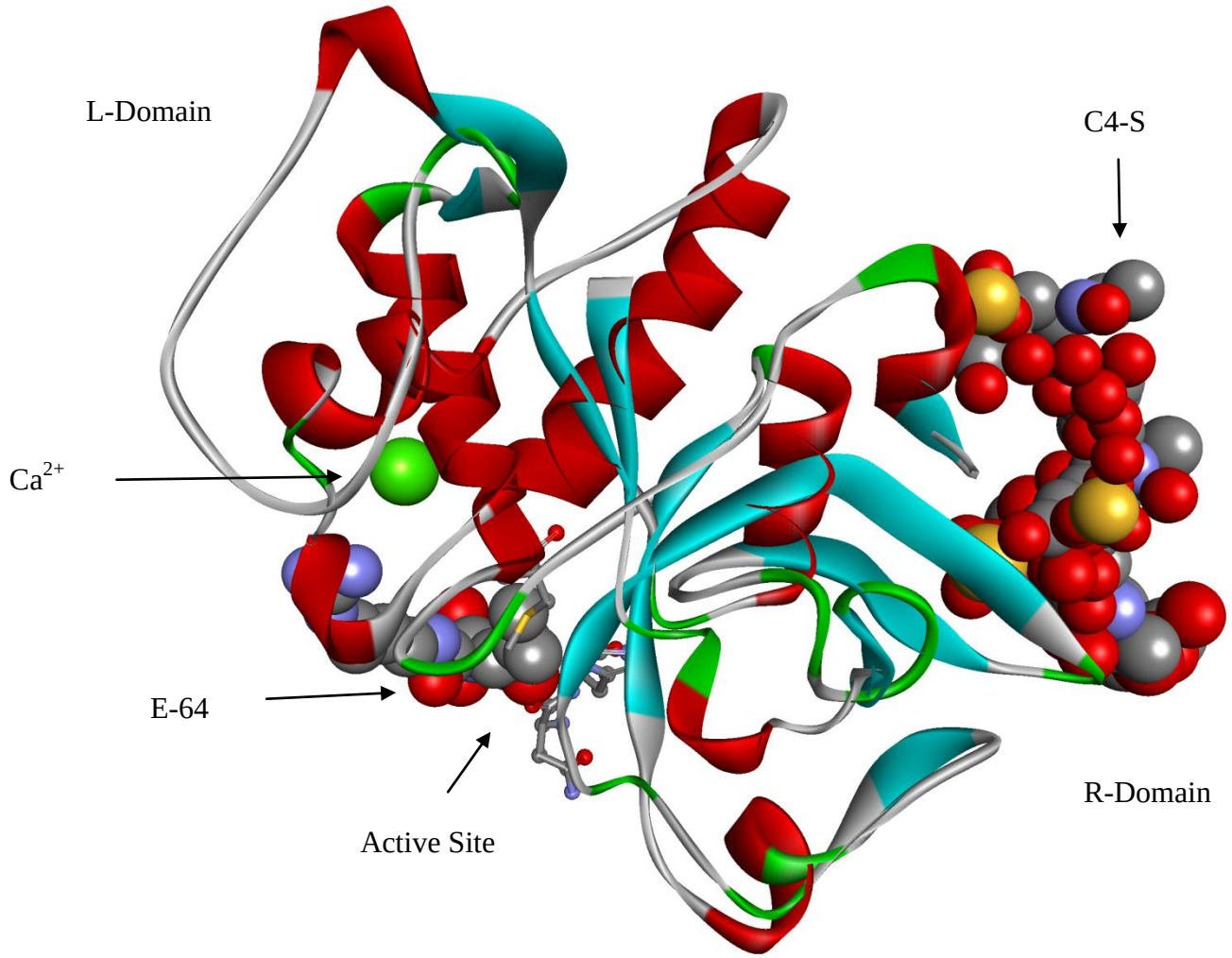



Figure 76. Crystal Structure of Cathepsin K Complexed with E-64 and Chondroitin 4-Sulfate (PDB ID: 3C9E)

*Catalytic Activity of Cathepsin K and Substrate Specificity*

Cathepsin K possesses similar characteristics to cathepsin L. Their amino acid sequence is 51% identical.[135] Thus, it is not surprising that share similar substrate specificities. However, there are subtle differences as pointed earlier in this chapter. Cathepsin K prefers substrates with amino acid residues containing aliphatic and hydrophobic (i.e. Ala, Ile, Leu) residues at the second position prior to the peptidic cleavage. Furthermore, proline is a very unusual residue for peptidic cleavage at the P2 position. Few enzymes are able to cleave these residues.[374] Cathepsin K is the only cysteine cathepsin that has a preference for proline at the aforementioned position (X-Pro---X).[359,375]

*Structure of Cathepsin K*

Cathepsin K is a homodimer containing 215 residues per monomer with a molecular weight between 25 and 29 kDa.[193] Amino acid sequences of both cysteine cathepsins (L and K) have been identified and confirmed. Literature reports identities between 50 and 60 percent.[135] Figure 77shows the amino acid sequence of mature cathepsin K containing 215 residues.[376] The three most abundant residues are glycine, asparagine and arginine (24, 19, and 19 residues respectively). On the other hand, the least abundant residues are histidine (2 residues), tryptophan and methionine (four residues each). Similar to cathepsin L, hydrophobic residues account for almost one third of the total composition of cathepsin K. Also, 44 percent of the composition is entirely made of polar, uncharged residues. The secondary structure also reveals that eight alpha helices accounting for 27% of the structure. However, 23% of the sequence forms beta sheets distributed in seventeen strands.[377]

```
  1  APDSVDYRKKGYVTPVKNQGQCGSCWAFSSVGALEGQLKKKTGKLLNLSP
 51  QNLVDCVSENDGCGGGYMTNAFQYVQKNRGIDSEDAYPYVGQEESCMYNP
101  TGKAAKCRGYREIPEGNEKALKRAVARVGPVSVAIDASLTSFQFYSKGVY
151  YDESCNSDNLNHAVLAVGYGIQKGNKHWIIKNSWGENWGNKGYILMARNK
201  NNACGIANLASFPKM
```

Figure 77. Amino Acid Sequence of Cathepsin K (PDB ID: 3KX1).[376]

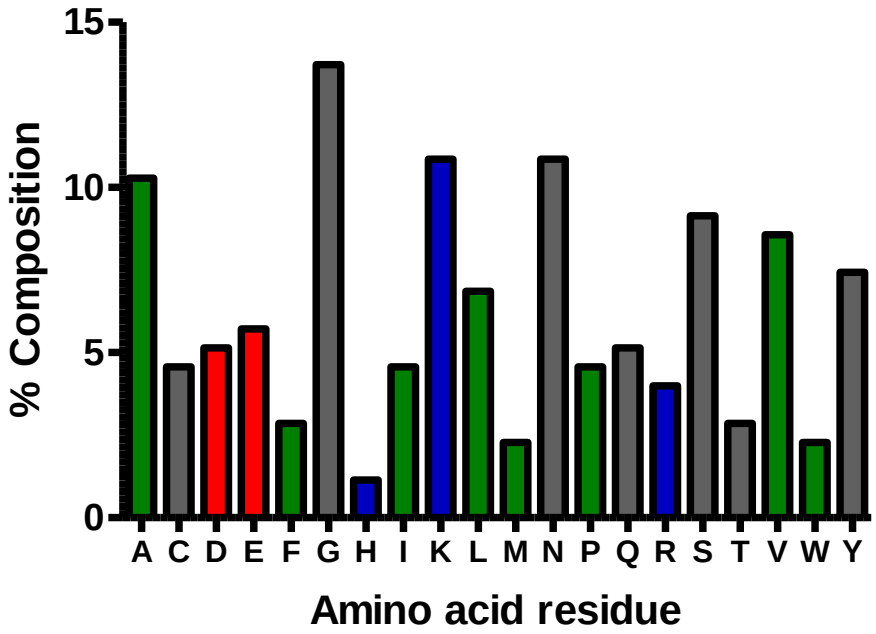



Figure 78. Composition of Human Cathepsin K. Legend= Blue: Basic Residues; Red: Acidic Residues; Green: Hydrophobic (Non Polar) Residues; Gray: Neutral, Polar Residues

Table 70. Amino acid Composition of Cathepsin K.

| AA | # | % | AA | # | % | AA | # | % |
|---|---|---|---|---|---|---|---|---|
| Ala | 18 | 8.37 | Ile | 8 | 3.72 | Arg | 7 | 3.26 |
| Cys | 8 | 3.72 | Lys | 19 | 8.84 | Ser | 16 | 7.44 |
| Asp | 9 | 4.19 | Leu | 12 | 5.58 | Thr | 5 | 2.33 |
| Glu | 10 | 4.65 | Met | 4 | 1.86 | Val | 15 | 6.98 |
| Phe | 5 | 2.33 | Asn | 19 | 8.84 | Trp | 4 | 1.86 |
| Gly | 24 | 11.16 | Pro | 8 | 3.72 | Tyr | 13 | 6.05 |
| His | 2 | 0.93 | Gln | 9 | 4.19 | Total | 215 | 100 |

*Crystal Structure of Cathepsin K*

One of the first crystal structures of human cathepsin K was resolved by McGrath and coworkers in 1996, two years after the protease was discovered in rabbits.[124] The macromolecule was crystallized with an inhibitor, and important interactions were determined between cathepsin K and the inhibitor. Mature cathepsin K is a macromolecule with 215 amino acids. The approximate size of the crystal was determined as 55 X 35 X 37 Å. In the structure, Cys25 and His159, two of the members of the catalytic triad appeared in the form of an ion pair (thiolate-imidazolium pair). The third residue, Asn175 was also found but protected by Trp77 (papain numbering). The overall structure contains several pockets; some of them are well defined while others were poorly formed. Seven of these pockets interact with potential cysteine protease inhibitors.

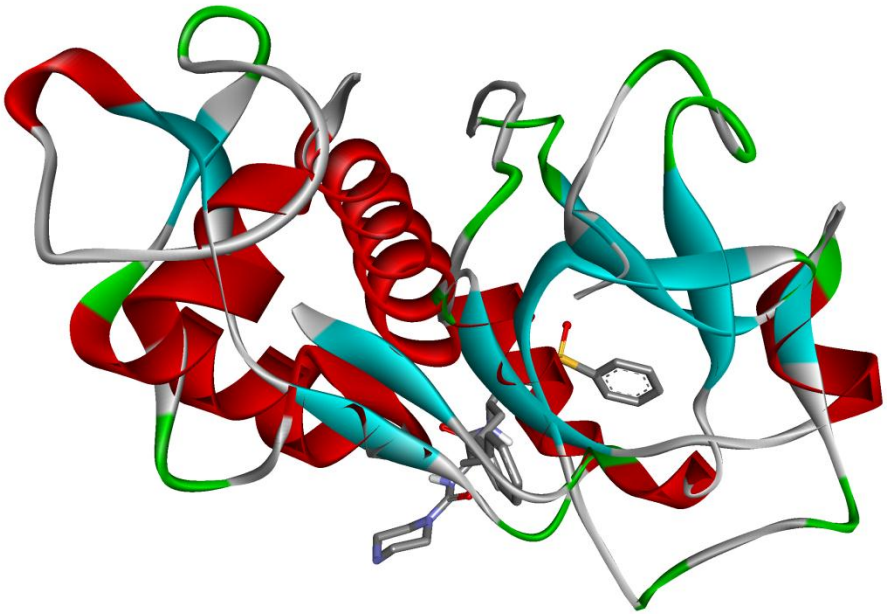
Eight amino acid residues located on the surface of the molecule could not be fully characterized. These residues are Ser4, Lys41, Lys44, Glu59, Lys77, Lys119, Arg127, and Lys147. However, comparisons with similar cysteine cathepsin structures revealed little deviation between them. Important interactions were found at the active site of cathepsin K. For example, one hydrogen bond is formed between the carboxyl oxygen of Asn175 and one of the nitrogens of the imidazole ring of His159. The crystal structure also reveals a close proximity between one of the His159 nitrogens and the thiol group of Cys25 (*N-S:* 3.65 Å). The importance of this inhibitor-cathepsin K complex resides in the interaction of the inhibitor (vinyl sulfone) that binds to the macromolecule as a substrate. The Cys25 thiol moiety attacks a vinyl carbon of the carbon that mimics the substrate carbonyl carbon in a possible substrate. The crystal structure of cathepsin K

inhibited by the vinyl sulfone is shown in Figure 79. Also, Figure 80 shows the crystal structure of uninhibited cathepsin K.[376,377] Three hydrogen bonds are formed between inhibitor and cathepsin K Gly65, Gly66, and Asn158 residues. Additionally, van der Waal contacts were detected with Gly64, Gly65, Cys63, and Gly23. The almost-hydrophobic S2 subsite is formed by six residues, (Tyr67, Met, 68, Ala133, Leu157, Ala160, and Leu205). The presence of Leu205 makes the S2 pocket be extremely shallow for bulky residues, such as phenylalanine. On the other hand, the S3 pocket is made only with two residues: Tyr67 and Asp61. The P1´ pocket is made with side chains of Asn158, Trp177, and Ala136. Trp177, the protective group of Asn175, appears to be in close proximity to the substrate as well. Finally, the key signature of cysteine proteases, a disulfide bond between two cysteine residues was found in cathepsin K (Cys153 and Cys200). Figure 80 shows the crystal structure of human cathepsin K. Three active residues (Cys25, His159, and Asn175) are shown in ball and stick form. The structure is colored according to the following: α-helices are red, β-sheets are cyan, turns are green, and coils are white PDB: 3KX1.[376]

*General Considerations of Procathepsin K*

Lalonde and coworkers elucidated the crystal structure of human procathepsin K.[138] Again, the macromolecule is formed by a 99-residue (the propeptide region) and the mature form of 215 residues). The mature form was found to have similar characteristics without major conformational changes when the propeptide is present. The complete amino acid sequence is presented in Figure 82. The proregion has a globular-shaped form and is made with four secondary structures: 3 α-helices and 1 β-sheet. The proregion is located above the β-sheet domain of the mature cathepsin K.

**A**



**B**

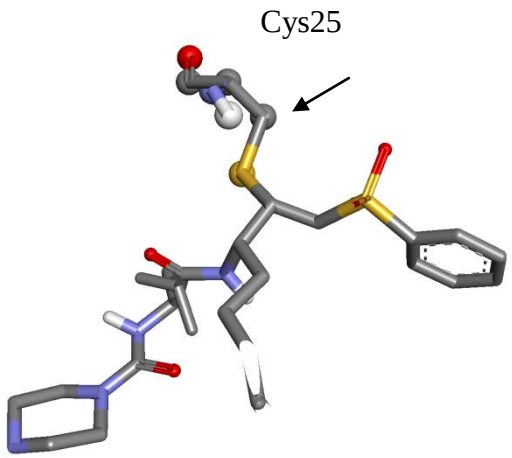



Figure 79. Crystal Structure of Human Cathepsin K with Z-Phe-Tyr(OBut)-COCHO, a Reversible, Slow, Tight Binding Inhibitor. **A**. Overall Structure. **B**. Covalent Modification of Cys25 by the Inhibitor (PDB ID: 1MEM)

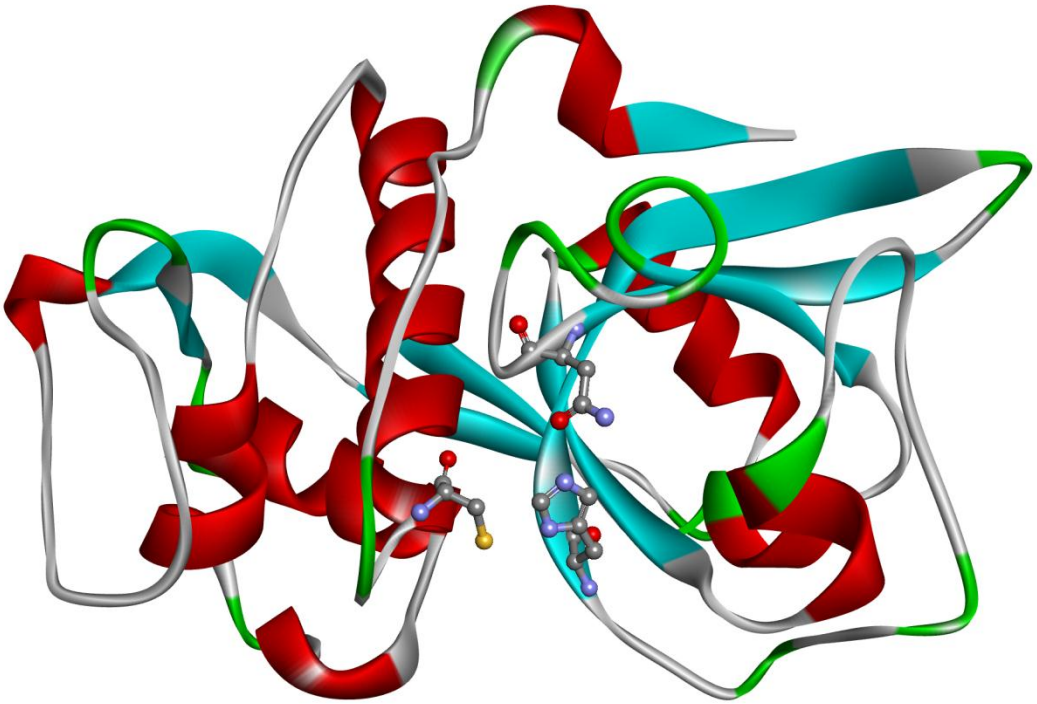


Figure 80. Crystal Structure of Human Cathepsin K (PDB: 3KX1)

The first helix of the propeptide region is an 11-residue almost found to be orthogonal compared to the second helix (27-residues). This change in direction forces the tripeptidic β-sheet to be in close to cathepsin K's β-sheet domain forming a 2-stranded β-sheet. Finally, the fourth helix (9 residues) is found right next to the active site forcing the cleft to appear stretched. The investigation also found that the propetide can be divided into three major groups: the globular domain (69 residues), active site-binding segment (8 residues) and the C-terminal segment with 18 segments.

The globular domain of the propeptide shows hydrophobic and hydrogen bond interactions with mature cathepsin K. For example, the hydrophobic sides of helices 2 and 3 make contact with the nonpolar cathepsin K surface (residues 236-251 in the mature form). The responsible of the active site stretching in the mature form is called

the propeptide binding loops and is a 6-residue α-helix. They also found six hydrogen bonds between main chain and side-chains. Phe243 from the mature form also shows van der Waals interactions between Asn38, Ile42, Leu58, Asn61, Leu63 and Gly64 (all of them form part of the propeptide). Conversely, the propeptidic Tyr56, a hydrophobic residue, interacts with 6 residues found in the mature form (Tyr244, Gly247, Val248, Tyr249, Tyr250 and Asp251). Figure 82 shows the crystal structure of human procathepsin K. Three active residues (Cys25, His159, and Asn175) are shown in ball and stick form. The structure is colored according to the following: α-Helices are red, β-sheets are cyan, turns are green, and coils are white (PDB ID: 1BY8).[138]

Eight residues of the propeptide (p74-p81, p indicates is part of the propeptide) are located close to the active site. This is the region where several electrostatic and hydrogen bonds are found. pThr76 and pMet75 occupy S' subsites, above the catalytic Cys25. On the other hand, pGly77, Leu78 and Lys79 occupy the S subsites. pLys74 also forms hydrogen bonds with Cys20 and Glu19 (mature form numbering). pMet75 interacts with two residues of the mature form: Gln18and Trp182. Lastly, another hydrogen bond is made between Thr76 and Gln17 of the mature form. A full description of the structure can found in their presented work.

```
  1  LYPEEILDTHWELWKKTHRKQYNNKVDEISRRLIWEKNLKYISIHNLEAS
 51  LGVHTYELAMNHLGDMTSEEVVQKMTGLKVPLSHSRSNDTLYIPEWEGRA
101  PDSVDYRKKGYVTPVKNQGQCGSCWAFSSVGALEGQLKKKTGKLLNLSPQ
151  NLVDCVSENDGCGGGYMTNAFQYVQKNRGIDSEDAYPYVGQEESCMYNPT
201  GKAAKCRGYREIPEGNEKALKRAVARVGPVSVAIDASLTSFQFYSKGVYY
251  DESCNSDNLNHAVLAVGYGIQKGNKHWIIKNSWGENWGNKGYILMARNKN
301  NACGIANLASFPKM
```

Figure 81. Amino Acid Sequence of Procathepsin K (PDB ID: 3KX1).[138] The Propeptide, a 99-Residue Polypeptide, is Shown in Red.

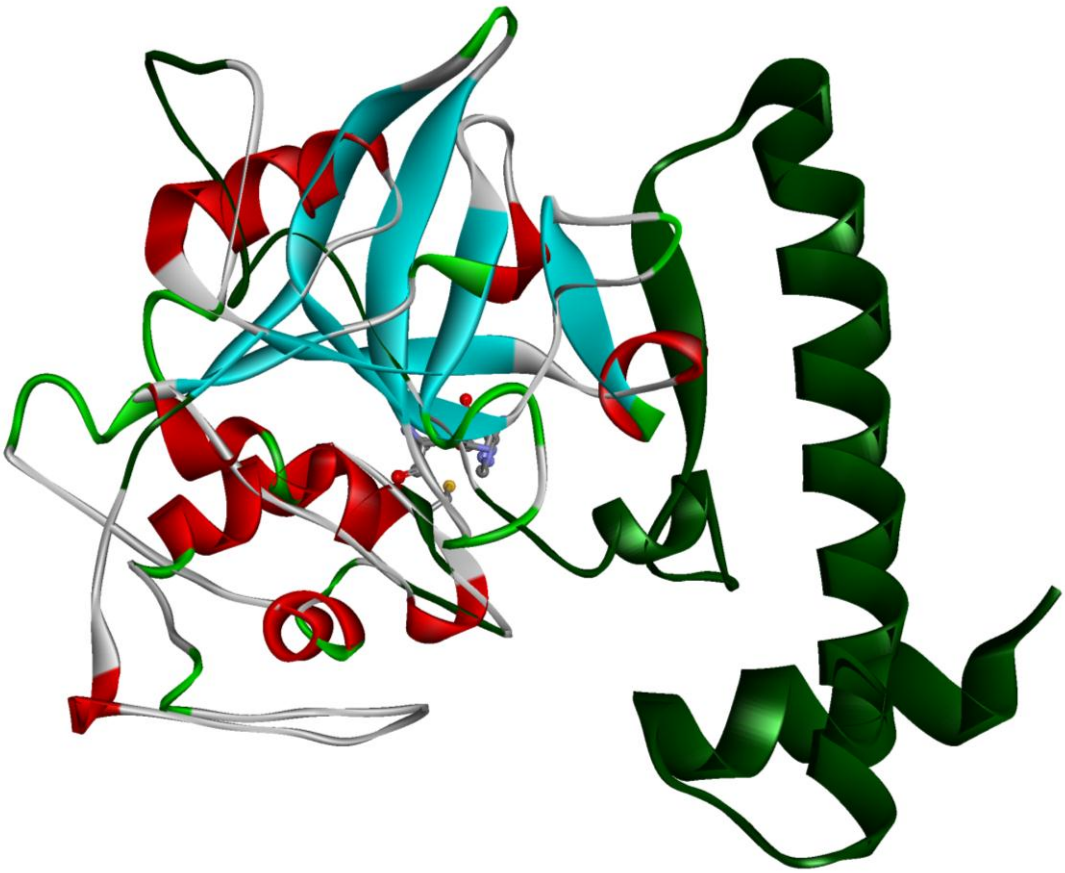


Figure 82. Crystal Structure of Human Procathepsin K (PDB ID: 1BY8)

*Inhibitors of Cathepsin K*

Research related to cathepsin K has found numerous inhibitors, natural and synthetic. To date, there are two compounds that have been successfully tried in early clinical phases, and they still are under investigation.

*Natural Inhibitors of Cathepsin K*

Cathepsin K propeptide is perhaps one of the best inhibitors of this enzyme showing a $K_I$ value of 4 nM. However, it does not show selectivity towards this protease because it also inhibits cathepsins L and S.[135] Mammalian organisms express glycosaminoglycans in bone tissues. These macromolecules dramatically increase the

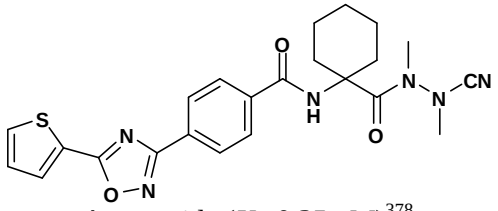
collagenase activity of cathepsin K by forming complexes with the protease. Interestingly, these molecules do not have any effect on the collagenase activity of cathepsins L and S.[357] Cystatins, stefins, kininogens, thyropins, and serpins are powerful cathepsin K inhibitors. *In vitro* experiments were carried out demonstrating the inhibitory efficacy of these molecules. For example, the $K_I$ values for Stefin B, Cystain C, L-kininogen and H-kininogen varied between 1 and 15 pM. Finally, α-macroglobulin, a well known inhibitor of other cysteine proteases (cathepsins B, H and L) has not been tested to verify its inhibitory activity towards cathepsin K.

*Synthetic Inhibitors of Cathepsin K in Preclinical Trials*

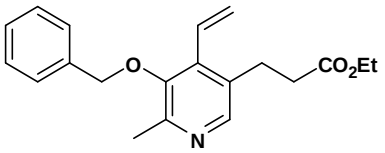
The importance of cathepsin in bone resorption has propelled the search for synthetic inhibitors of this cysteine protease. A selected group of these compounds have reached clinical phases in the treatment of osteoporosis. *In vitro* studies show derivatives of cyclohexanones, dihydrofuranones, ketoamides, sulfonamidoketone, ketooxadiazole, and peptidic nitriles are excellent cathepsin K inhibitors that exhibit showing relative selectivity towards the protease.[386] Figure 83 presents selected compounds that showed inhibitory activity towards this protease.

*Synthetic Inhibitors of Cathepsin K in Clinical Trials*

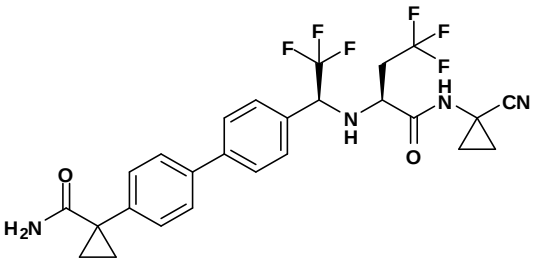
The number of synthetic inhibitors in clinical trials is outstanding. Four compounds have reached clinical trials and some of them have shown promising results.[158,387] Balicatib is a nitrile-derived compound developed by Novartis. *In vitro* studies showed this compound is a selective inhibitor of cathepsin K.



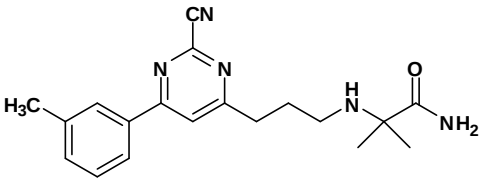
Azopeptide ($K_I$: 0.35 nM)[378]



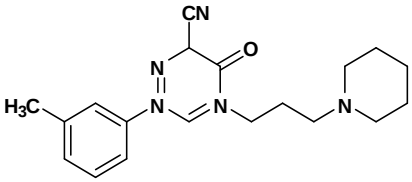
CLIK-166 (Aromatic Vinyl analog) (inhibits ~ $10^{-5}$)[155]



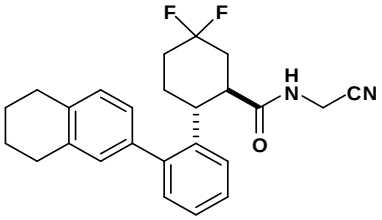
Ditrifluoroethylamines (Odanacatib derivative)[379] $IC_{50}$: 0.3 nM



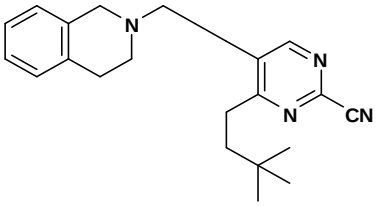
2-Cyano-pyrimidine ($IC_{50}$: 0.3 nM)[380]



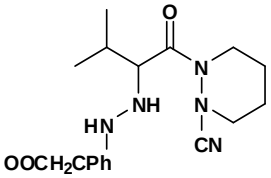
Dioxo-triazine ($IC_{50}$: 17 nM)[381]



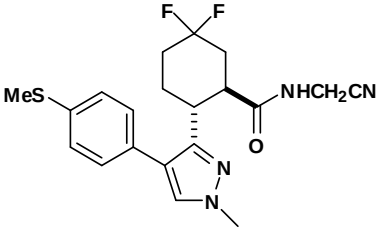
Cyclohexylcarboxamide ($IC_{50}$:0.28 nM)[382]



Aromatic acetylene base ($IC_{50}$: 3.2 nM)[383]



Cyano-pyridazine ($IC_{50}$: 1.0 nM)[384]



Cyclohexamide pyrazole ($IC_{50}$: 0.4 nM)[385]

Figure 83. Selected Cathepsin K Inhibitors Reported between 2007 and 2012

Interestingly, the compound loses its selectivity in cell culture studies by inhibiting other cysteine cathepsins. Phase II studies are completed with female patients with osteoarthritis showing promising results in a one year study where bone resorption marker levels were decreased. Clinical trials were discontinued due to adverse side effects seen in patients' skin.[388]

Glaxo Smith Kline had a cathepsin K inhibitor in its clinical trials (Relacatib, an α-heteroatom cyclic ketone). This inhibitor shows no selectivity against cathepsin L-like proteases (cathepsins L, K, V). Studies were stopped during phase I trials due to possible interactions with acetaminophen, ibuprofen, and atorvastatin. These compounds are usually prescribed to people with osteoporosis and osteoarthritis).[389]

Ono Pharmaceutical has ONO-5334, a compound with modest results in phase II trials. Their results showed a dramatic decrease in bone resorption markers only in patients that were treated with high doses of the compound. However, reports of side effects were less severe in treated patients.[390]

The most promising inhibitor of cathepsin K is Odanacatib. The nitrile-based analog is one of Merck's most studied compounds. The compound shows a high selectivity against cathepsin K. The efficacy of this compound is due to the reversible thioimidate formed between odanacatib's nitrile moiety and cathepsin K Cys25 thiol group.[391,392] The advantages of this compound are that patients tend to tolerate Odinacatib quite well; with minor side effects, and have a long half life, between 3-4 days. Daily doses (2.5 mg/patient/day) showed a significant decrease in bone resorption markers. Odanacatib is considered an excellent candidate for osteoarthritis patients because it does not interfere with the bone formation processes, a situation that has been

found with other candidates such as alendronate and denosumab. Currently, Odanacatib is under investigation in Phase III clinical trials.[393–396]

*Importance of Cathepsin K in Medicine*

Cathepsin K is one of the most studied cysteine proteases along with Cathepsin L and B. Its powerful hydrolytic activity that degrades components of the ECM is unique and fascinating. Cathepsin K is mainly involved in bone remodeling and resorption. However, it is well known that overexpression of this protease in specific tissues and processes promote the development of serious pathological diseases (See Table 9). In summary, Cathepsin K is involved in thyroidal and bone-related conditions. Bone related diseases have major implications in health issues across the country. Prostate, breast, and skin cancer tend to metastasize to the bone and patient survival probabilities tend to decrease because current treatments fail to delay or arrest bone metastasis. Additionally, overexpression of proteases is an indication of excessive osteoclast activity. This proteolytic activity led to pathological conditions such as osteoporosis and osteoarthritis. Osteoporosis is characterized for excessive fragility in bone tissue and it is observed more often in female patients that are 55 years old. Overall, an alarming number of 10 million people older than 50 years will develop osteoporosis. Indirect and direct costs of people with osteoporosis represent more than 19 billion of dollars.

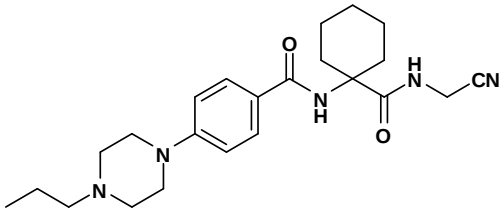
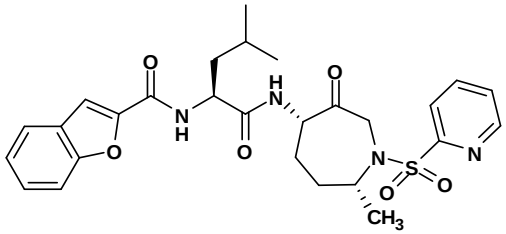
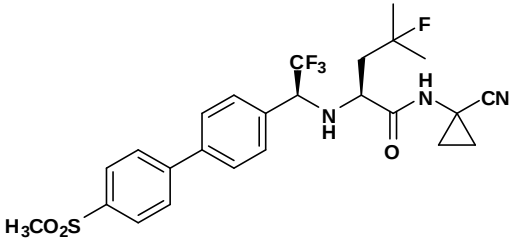
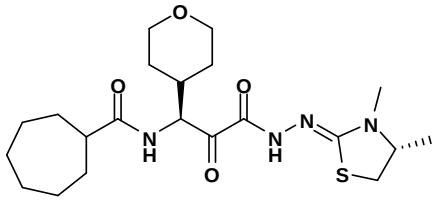
| Compound | Name | Phase | CathK | CathB | CathL | CathS | CathV | CathF |
|---|---|---|---|---|---|---|---|---|
|  | Balicatib | II | $IC_{50}$: 1.4 | $IC_{50}$: 4800 | $IC_{50}$: 503 | $IC_{50}$: 65000 | Not determined | Not determined |
|  | Relacatib | I | $K_I$: 0.041 | $K_I$: 13 | $K_I$: 0.068 | $K_I$: 1.6 | $K_I$: 0.063 | Not determined |
|  | Odanacatib | III | $IC_{50}$: 0.2 | $IC_{50}$: 1034 | $IC_{50}$: 2295 | $IC_{50}$: 60 | $IC_{50}$: 762 | $IC_{50}$: 795 |
|  | ONO-5334 | I | $K_I$: 0.1 | $K_I$: 32 | $K_I$: 1.7 | Not determined | Not determined | Not determined |

Table 74. $IC_{50}$ and $K_I$ Values of Cathepsin K Inhibitors in Clinical Trials.[397,398] Values are in nM

### *Regulation of Cathepsin K in Osteoclasts*

Little is known about cathepsin K and its implication in signal transduction and its participation in cellular processes. Most of the work done related to the protease involves its proteolytic activity in osteoclasts by degrading numerous fibrillar proteins that form part of the extracellular matrix. However, the expression of cathepsin K gene in osteoclasts is tightly regulated by numerous activators and inhibitors. Troen offers a comprehensive explanation of the cell signaling involving activators or inhibitors of the expression of the enzyme.[399]

RANKL is the main regulator of cathepsin K due to its participation in the differentiation and activity of the osteoblasts. Other stimulators include vitamin D, glucocorticoids, interleukins, prostaglanding, histamine, etc. Therefore, it is not surprising that RANKL inhibitors also suppress the activity and expression of cathepsin K.[400]

The receptor activator of NF-κB ligand stimulates cathepsin K mRNA expression in human osteoclasts. Additionally, studies have demonstrated the direct inhibition of the expression of cathepsin K. Calcitonin, IFN-γ, estradiol, and calcium are some examples that have been studied.

### *Cathepsin K and Thyroid-Related Diseases*

The thyroid is one of the most important organs regulating hormonal processes. It requires specialized proteins for hormone processing. Cathepsin K is one of the members of a selected group with thyroglobulinase activity. High concentrations of cathepsin K have been found in patients with thyroid cancer when compared to a sample of healthy

people. Animal studies also revealed that Cathepsin K deficiency is translated into longer small intestines.[111,401]

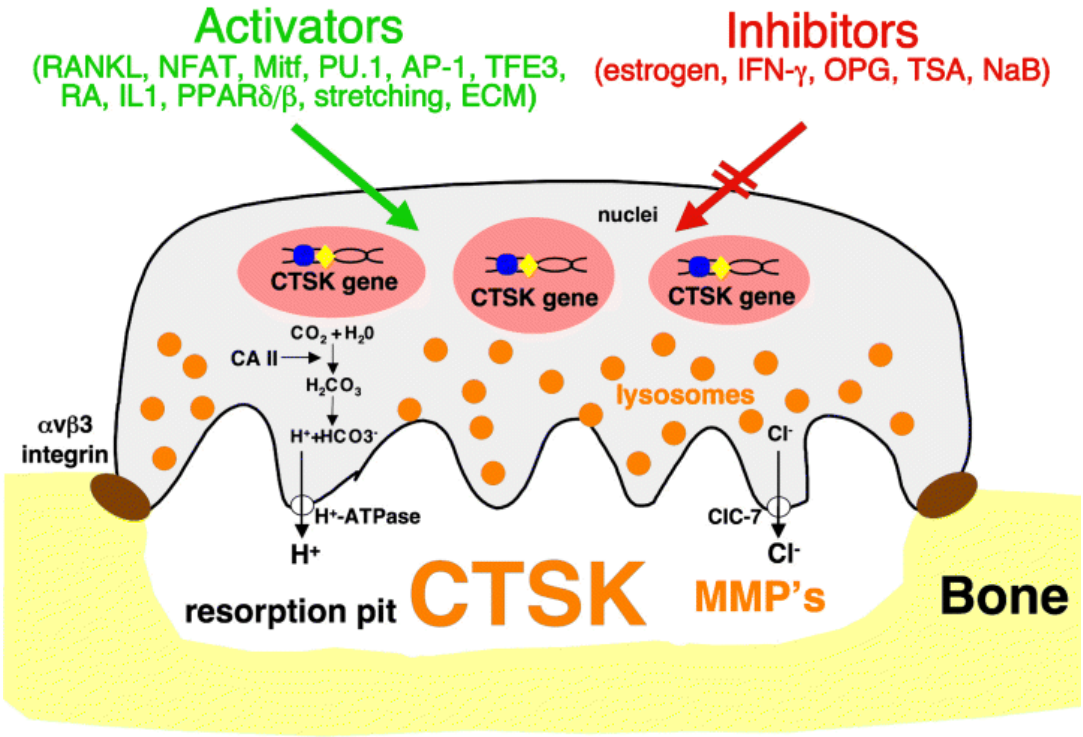



Figure 84. Regulation of Cathepsin K (CTSK) Gene Expression in Osteoclasts (Reproduced from Troen, page 169)

*Cathepsin K and Bone Resorption*

Bone remodeling is intrinsically related to bone resorption, a function that is tightly regulated by osteoclasts. It is also proven that abnormal osteoclast activity is responsible for pathological conditions such as osteoporosis, osteoarthrititis, and bone metastasis. In summary, the role of these specialized cells is the degradation of organic bone material and minerals found in bone tissues.

*The Resorption Cycle*

The osteoclastic activity can be summarized in what is called the resorption cycle. Kalervo and coworkers explained in recent reviews the steps of this cycle.[402] Full reviews can also be found elsewhere.[403]

1. Osteoclasts migrate to the 'resorption cycle' and attach to the bone tissue.
2. Osteoclasts undergo a self-polarization and form new membrane domains
3. Osteoclasts dissolve hydroxyapatite, the mineral component of bones
4. Osteoclasts degrade the organic bone material
5. Cells perform a housekeeping process by removing the degraded material
6. Cells undergo apoptosis or migrate to a non-resorbing stage

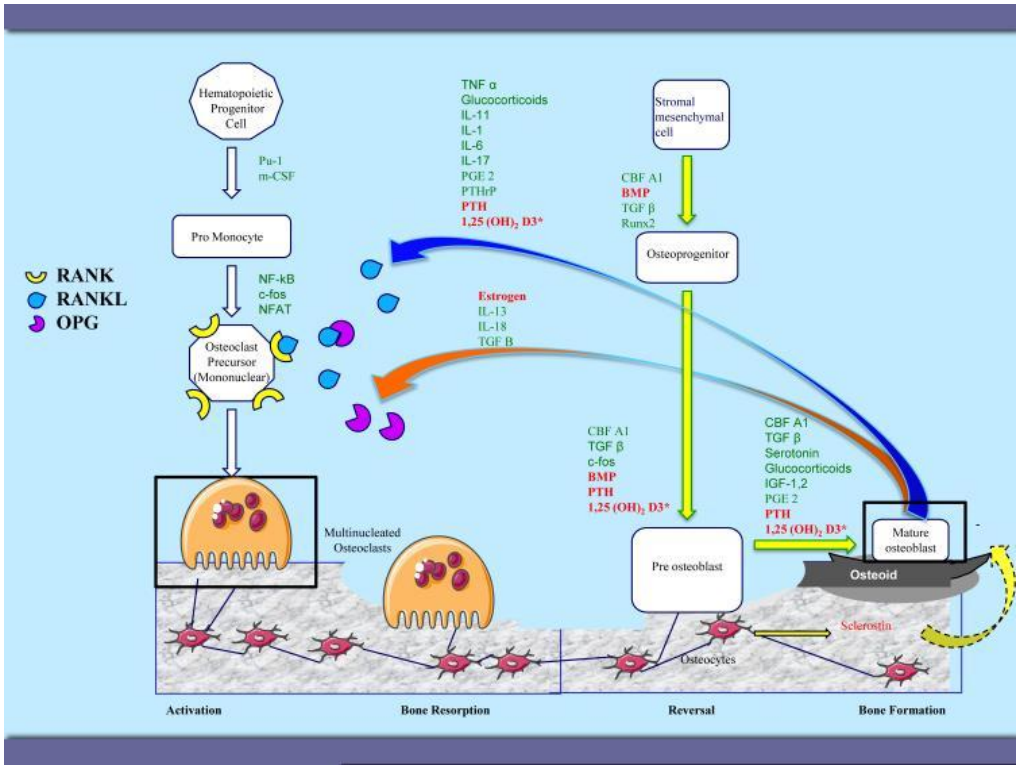



Figure 85. Bone Remodeling Showing the Various Stages and the Factors Involved. Also Shown is the Development of Osteoblasts and Osteoclasts from Precursors. (reproduced from Gallaguer, page 303 [404])

It has been showed that osteoclast expresses significant levels of cathepsin K in normal conditions. The activity of cathepsin K and other important metalloproteases is required in the fourth step. These enzymes are responsible for the proteolytic degradation of the organic bone material. Specifically, MMP-9 and cathepsin K are the main players in the resorption cycle. Furthermore, the expression of cathepsin K accounts for 98% of the cysteine proteases found in osteoclasts. Its activity has been found in the lysosomes, ruffled border and resorption lacuna (See The resorption Cycle, page 210).[405]

### *Cathepsin K and Bone Remodeling*

Bone remodeling comprises a complex cycle where bone tissues are 'formed' and 'degraded'. Both processes are necessary in healthy bone tissues and help in bone growth, bone fracture, etc. This is a lengthy process taking approximately six months to be completed. Proteases, such a metalloproteases and cysteine proteases are important players in the normal development of bone processes. Overexpression of cathepsin K in animal studies revealed an accelerated degradation of the organic bone material, a clear indication of osteoporosis.[406] On the other hand, pycnodysosis is a pathological disease characterized by short stature and osteopretosis. Pycnodysosis is the result of a mutation in the genetic code that produces sub-expression of Cathepsin K.[407]

### *Role of Cathepsin K in Osteoporosis*

Extensive research has been carried out in the search for selective and efficient cathepsin K inhibitors. Examples of these compounds in clinical trials have been discussed early in this chapter. All of these compounds were developed as potential therapeutic agents against osteoporosis. Osteoclasts are the main cells responsible for the development of osteoporosis in patients suffering this condition. Currently, the reatment

of osteoporosis is limited to antiresorptive and anabolic agents. The list of antiresorptive agents (agents that inhibit the formation of new bone tissue) includes estrogen-based therapy, selective estrogen receptor modulators, calcitonin treatment, and bisphosphonates. In addition, anabolic agents include parathyroid hormone and strontium ranelate, a treatment where the agent is given to patients with a positive outcome in phase III clinical trials.[404]

### *Role of Cathepsin K in Arthritis*

It is not surprising that cathepsin K is also implicated in the development of rheumatoid arthritis in human and animal studies. Two types of stromal cells, fibroblast and chondrocytes, are cells that are strongly linked to this disease. Table 12 shows the expression of cysteine proteases in some of these stromal cells. Furthermore, the expression of cathepsin K has also been found in fibroblasts and chondrocytes.[160,398,408]

### *Cathepsin K and Atherosclerosis*

One of the least investigated diseases where the cathepsin K role is crucial is atherosclerosis. The role of cysteine cathepsins in cardiac research has been recently reviewed due to their importance in cardio-vascular conditions.[409] Investigators have demonstrated that cathepsin K is involved in the deterioration of blood vessels walls and the destabilization of plaques.[410,411]

### *Bone Metastasis*

Bones are one of the most common sites affected by metastasis. Thus, research has also been devoted to study the biology of bone metastasis and its implication in cancer development.[300,412] Breast and prostate cancer, the leading types of cancer in

female and male patients, usually metastasize to bone as their primary metastatic site. Once again, osteoclasts are mainly involved in the metastatic colonization of the bone. However, metastasis in female and male patients with breast or prostate cancer is different. For example, bone metastasis in female cancer is predominantly osteolytic, while osteogenic lesions are found in male patients. Cysteine cathepsins and metalloproteases are also responsible for the facilitation of bone metastasis. These two types of proteases greatly influence and modify the tumor microenviroment promoting the dispersion and generation of new metastatic lesions. Specifically, overexpression of cathepsin K in osteoclast cells in breast cancer samples facilitates cell migration, cell invasion, adhesion to a vessel, and extravasation to the final destination (metastatic organ). The search for excellent antimetastatic agents is not exclusively against cathepsin K. In depth investigations are also being carried out for agents inhibiting the activity of Src, TFGβ, MMPs, RANKL ET-1, uPA, DKK-1, WNT, etc. Figure 86 shows some of the most important targets in bone metastasis research that are currently in extensive studies and their most important inhibitors.

### *Cathepsin K and Skin Cancer*

The evidence of cathepsin K expression in normal skin tissue has led to the evaluation of cathepsin K in melanoma and other skin cancer diseases. This form of cancer is extremely aggressive and also metastasizes to bone as well. Recent studies showed higher levels of cathepsin K in squamous cell carcinoma in more than 100 samples of benign and malignant skin cancer tumors. The invasiveness of this type of cancer was also investigated and concluded that the high aggressiveness that

characterizes this type of tumor is partially due to the overexpression of cathepsin K in fibroblast cells.[350,354,413]

### *Implications of Cathepsin K in Prostate Cancer*

Cathepsin K activity was detected by several techniques, including reverse-transcription –polymerase chain reaction, western blotting, and immunochemistry in prostate cancer cell lines and tumors from patients. However, there are no recent studies linking the proteolytic activity of cathepsin K to prostate cancers.[414] Treatments with denosumab, a RANKL inhibitor, and bisphosphonates have shown modest results in patients with metastasized prostate cancer.[415] In addition, SRC, cABL kinases, and cathepsin K inhibitors have been under investigation to verify their potency against prostate cancer.

### *Cathepsin K and Breast Cancer*

The importance of cathepsin K in breast cancer progression is more evident. There is strong evidence that numerous breast carcinomas samples express cathepsin K. Inhibition studies with animals showed that cathepsin K played an important role in bone metastasis. BT474 cells (invasive ductal breast carcinoma) were implanted in mice. A cathepsin K inhibitor significantly reduced osteolytic lesions when a cathepsin K inhibitor was administered in clinical doses.[416] Additional investigation links the activity of cathepsin K in angiogenesis and expression of pro-inflammatory cytokines as well.[417]

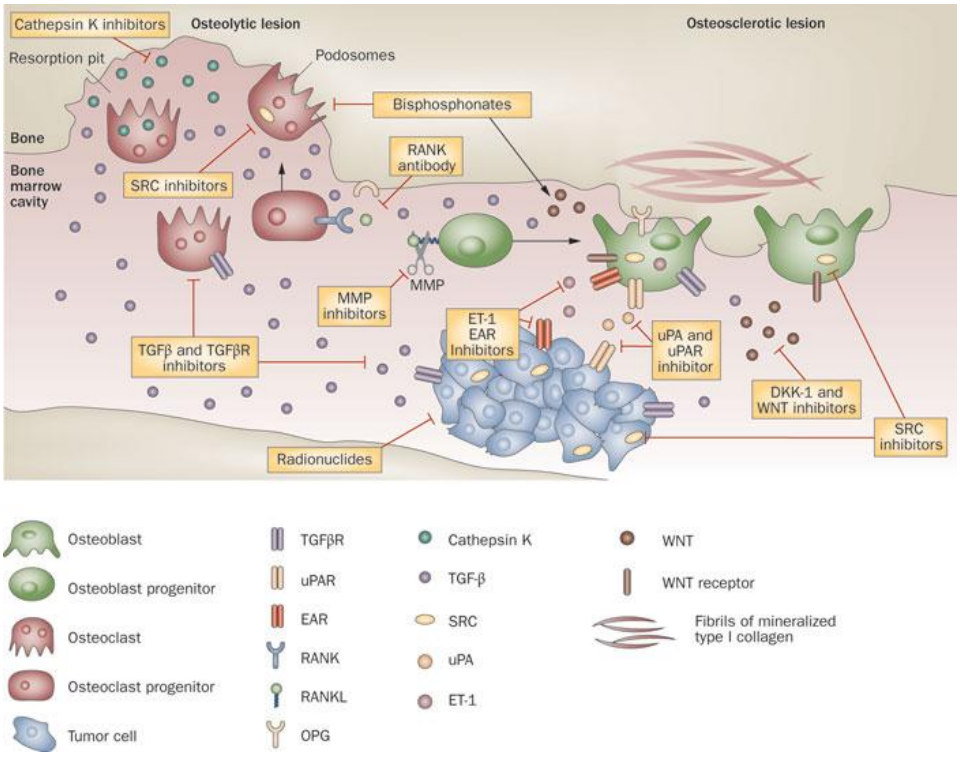



Figure 86. Bone-targeted therapy in Metastatic Lesions. (reproduced from Sturge, page 359)

*Role of Collagens in Cell Signaling and Cancer*

It has been demonstrated that both type I and type IV collagens participate in numerous cell-related processes including cell adhesion, migration, survival proliferation and differentiation.[370] Type IV is one of the main substrates of numerous types of cells (i.e. hepatocytes, keratinocytes, and tumor cells such as breast and prostate carcinoma, melanoma, fibrosarcoma and gliomas).

Cell attachment to the collagens is via integrins, which act as mediators between cells and their surrounding microenvironment. In general, integrins can bind to specific peptide sequenes found in ECM proteins. The most important integrin-binding sites found within the collagens are located in the main triple-helical chains and the non-collagen domains (NC1D).

The two most important integrin receptors of type I and type IV collagen triple helical chains are $\alpha_1\beta_1$ and $\alpha_2\beta_1$. One of the most important integrins binding recognition sites is the Arg-Gly-Asp peptide and is also found in collagens. However, cell adhesion in type IV is independent because integrins cannot reach that site found deep in the triple helical chain of the macromolecule. Both integrins belong to the β1 subfamily. The first integrin, $\alpha_1\beta_1$ shows more affinity toward type IV collagen, while the latter one prefers type I collagen. Research showed that deletion of the binding group reduced fibroblastic cell adhesion using type IV collagen as a basement. Other integrins that interact with the main chain are $\alpha_3\beta_1$, $\alpha_{11}\beta_1$, and $\alpha_{10}\beta_1$.

The non collagen domain (NC1D) also increases cell adhesion in collagen. It might also inhibit angiogenesis.

Additionally, collagen degradation is crucial in the diagnosis of diseases. The presence of bone turnover markers in serum and urine has lead to promising investigation and early treatment of diseases including osteoporosis, bone turnover, bone loss, renal failure and Paget's disease.[418]

Degradation of type I collagen, the major component of the extracellular matrix produces more than ten biomarkers. Three of them are the aminoterminal (NTX) and the carboxyterminal (CTX-I, and ICTP) cross linked telopeptides of type I collagen. Specific bone resorption markers are especially investigated due to their presence in samples of patients with bone metastasis.[419] Patients with breast or prostate cancer have showed elevated amounts of CTX-I or NTX-I in their samples.[420,421]

*Material and Methods for the Biological Evaluation of Thiosemicarbazones Derivatives as Inhibitors of Cathepsin K*

*Materials*

The comprehensive list of materials and equipment is listed in chapter two under the Materials and Equipment section. Recombinant, human procathepsin K (0.9 μg/μl) was purchased from Enzo. The proenzyme was expressed in insect cells according to the manufacturer. Type IV collagen (0.327 mg/ml) from human placenta was a product from BD Biosciences. Type IV collagen was extracted and purified using proteolytic enzymes. Chondroitin Sulfate A (≥ 60 % pure) sodium salt from bovine trachea is a product from Sigma. Tubulin from calf brain (1.25 mg/ml) was purified by members of the Trawick Group under the guidance and supervision of Dr. Mary-Lynn Trawick.

NuPAGE® MES SDS Running Buffer (20X) and 4-12% Bis-Tris SDS-PAGE are products from Invitrogen. Western blotting detection kit was purchased from Thermo Scientifics. The detection kits includes: wash buffer (30X), blocking buffer (1X), PVDF membranes, molecular marker and two secondary antibodies (goat anti-mouse Cy-3 labeled and goat anti-rabbit Cy-5 labeled). Mouse anti-human cathepsin K that can recognize both procathepsin and active cathepsin K was obtained from Sigma. Trans-Blot® SD Semi-Dry Electrophoretic Transfer Cell was from BioRad.

*Preparation of Buffers and Stock Solutions*

*Stock Solutions.* Protocol details about 1 M sodium acetate buffer, pH 5.5, 40 mM EDTA, 80 mM DTT, 4 mM AMC, 29.94 mM Z-FR-AMC, thiosemicarbazone derivatives stock solutions, preparation of inhibitor dilutions, and TSC derivatives stock

solutions can be found in chapter two in the preparation of buffers and stock solutions. Synthesis of thiosemicarbazones have been previously published.[3,6,7,10–12]

*Preparation of 180 mM sodium acetate buffer, pH 5.5*. One liter of this solution was prepared by diluting 180 ml of 1 M buffer stock solution with 820 ml of water. Acetic acid glacial and/or sodium hydroxide 5 M (MW: 40 g/mol) was used to adjust the pH of the buffer. This solution was stored at 4 °C to increase its shelf life.

*Preparation of 32.5 mM sodium acetate buffer, pH 3.5.* A fresh solution of this buffer was prepared by diluting 3.5 mg of solid sodium acetate anhydrous and 43.9 µl of acetic acid glacial in a 25 ml volumetric flask. Acetic acid glacial or sodium hydroxide 5 M (MW: 40 g/mol) were used to adjust the pH of the buffer. This buffer is also referred as solution B.

*Procathepsin K.* Ten micrograms of human procathepsin K (0.9 µg/µl) were kept at -80 °C to preserve the integrity of the protein.

*Preparation of inhibitors dilutions*. Each inhibitor was serially diluted in pure DMSO to provide seven different concentrations (named A-G) that were ten-fold diluted. The concentrations of these dilutions varied between 2 mM (solution A) and 2 nM (solution G). Each dilution was diluted ten-fold in a solution that contained 50% DMSO in water (named 1-8). Concentrations of the aqueous solutions varied between 200 µl and 200 pM. An eighth solution was made by using solution C (20 µM) to prepare an intermediate solution (1 µM in 50% DMSO). Solutions made in pure DMSO were stored

for two weeks at (-80 °C). Aqueous solutions were made prior to the experiment and kept at (4 °C).

*Preparation of 1 M NaCl.* A sample of 584.4 mg of sodium chloride was dissolved in 10 ml of water. This solution was stored at 4 °C to increase its shelf life.

*Preparation of 10 mM EDTA.* A sample of 37.2 mg of EDTA was dissolved in ten milliliters of water. Alternatively, 2.5 ml of 40 mM EDTA were diluted with 7.5 ml of water. The dilution was stored at 4 °C to increase its shelf life.

*Preparation of 60 mM EDTA.* A stock concentration of this solution was prepared by dissolving 22.34 mg (60 μmol) of EDTA (MW: 372.29 g/mol) in one milliliter of water. Fifty milliliters of this solution (744.5 mg; 2 mmol) was prepared for multiple experiments. The solution of 40 mM EDTA was stored at 4 °C.

*Preparation of 1.9 % chondroitin sulfate A.* The solution was prepared by dissolving 0.19 grams of Chondroiting Sulfate A (C4-S) with 10 milliliters of 2 M sodium acetate buffer. The solution was vigorously mixed at room temperature to dissolve the contents. The solution was made prior the experiments.

*Preparation of 3% chondroitin 4-sulfate in 50% DMSO*. Five milliliters of this solution were prepared by dissolving 0.15 g of C4-S with 2.5 ml of pure DMSO and 2.5 ml of water. The solution was prepared prior to the experiment. Final conditions were 3% C4-S and 50% DMSO.

*Preparation of inhibitor dilutions with C4-S.* Each inhibitor was serially diluted in pure DMSO to provide seven different concentrations (named A-G) that were ten-fold diluted. The concentrations of these dilutions varied between 2 mM (solution A) and 2 n*M* (solution G). Each dilution was diluted ten-fold in a solution that contained 50% DMSO in water with C4-S (named 1-8), by adding 90 µl of 3% Chondroitin 4-Sulfate in 50% DMSO and 10 µl of stock solution (A-G). Concentrations of the aqueous solutions varied between 200 µl and 200 p*M*. An eighth solution was made by using solution C (20 µ*M*) to prepare an intermediate solution (1 µM in 50% DMSO). Solutions made in pure DMSO were stored for two weeks at (-80 °C). Aqueous solutions were made prior to the experiment and kept at (4 °C). Final concentrations are 55% DMSO and 2.7% C4-S.

*Solution A.* Five milliliters of this solution were made by mixing 2.5 ml of 1 M NaCl, 0.5 ml 10 mM EDTA and 2 ml of water. Conditions of this solution are: 0.5 M NaCl and 1 mM EDTA. The dilution was stored at 4 °C to increase its shelf life.

*Activation of procathepsin K under acidic conditions at room temperature.* The proenzyme was activated by mixing 1 µl of 0.9 µg/µl (33.3 µM) procathepsin K, 1 µl 32.5 mM NaOAc pH 3.5 (solution B) and 4 µl of solution A. Total volume is 6 µl. Conditions for activation of the enzyme are: 5.6 µM of mature cathepsin K, 5.4 mM NaOAc pH 3.5, 0.7 mM EDTA and 333.3 mM NaCl. The mixture was incubated in a siliconized tube at room temperature (24 °C) between 30 and 180 minutes. Activation time varied depending of the proenzyme batch. Activation of the proenzyme was done prior to any experimentation with the mature enzyme.

*Preparation of cathepsin K assay buffer*. Cathepsin K assay buffer (will be referred as assay buffer for future references in this chapter) contained 3 mM EDTA, 3 mM DTT, 0.01% Brij 35 and 180 mM NaOAc pH 5.5. Ten milliliters of assay buffer were made by adding 750 µl of 40 mM EDTA, 375 µl of 80 mM DTT, 4.33 µl of Brij 35 and 8.9 ml of 180 mM NaOAc pH 5.5 in a plastic 15 ml conical tube. Table 75 summarizes a typical calculation sheet to make different volumes of the assay buffer. The solution was stable and used for a maximum of 24 h.

Table 75. Preparation Table for Cathepsin K Assay Buffer

| Assay Buffer (ml) | 40 mM EDTA (µl) | 80 mM DTT (µl) | Brij 35 (µl) | 180 mM NaOAc (µl) |
|---|---|---|---|---|
| 5 | 375 | 187.5 | 2.2 | 4436 |
| 10 | 750 | 375 | 4.3 | 8871 |
| 20 | 1500 | 750 | 8.7 | 17743 |
| 40 | 3000 | 1500 | 17.3 | 35482 |

*Preparation of cathepsin K stock solution*. A stock solution of cathepsin K (CK) contained 2.5 mM EDTA, 2.5 mM DTT, 1.5 nM cathepsin K and 150 mM NaOAc pH 5.5. One milliliter of cathepsin K stock solution was made by adding 62.5 µl of 40 mM EDTA, 31.2 µl of 80 mM DTT, 2.7 µl of 5.6 µM activated cathepsin K, 833 µl of 180 mM NaOAc pH 5.5 and 70.2 µl of water in disposable glass test tubes. Table 76 summarizes a typical calculation sheet to make different volumes of cathepsin K stock solution. The solution was made prior to every kinetic or inhibition experiment.

Table 76. Preparation Table for Cathepsin K Stock Solution

| CK (ml) | 40 mM EDTA (µl) | 80 mM DTT (µl) | Brij 35 (µl) | 180 mM NaOAc (µl) | Cathepsin K (µl) | Water (µl) |
|---|---|---|---|---|---|---|
| 1 | 62.5 | 31.2 | 0 | 833 | 2.7 | 70.2 |
| 2 | 125 | 62.5 | 0 | 1666 | 5.4 | 140.4 |
| 4 | 250 | 125 | 0 | 3332 | 10.8 | 280.8 |

*Preparation of Z-FR-AMC stock solution.* A stock solution of Z-FR-AMC (500 µM) in 15 % DMSO was prepared by adding 20 µl of 29.94 mM Z-FR-AMC, 160 µl of DMSO, and 1020 µl of water in order to have a final volume of 1200 µl. The solution was made in a plastic dark 15 ml conical tube. Table 77 shows representative calculations to prepare various amounts of the substrate solution. The stock solution was stable and used for a maximum of 24 hours.

Table 77. Preparation Table for 500 µM Z-FR-AMC for Cathepsin K Assays

| 500 µM Z-FR-AMC (ml) | 29.94 mM Z-FR-AMC (µl) | DMSO (µl) | Water (µ) |
|---|---|---|---|
| 1 | 16.7 | 133.3 | 850.1 |
| 2.5 | 41.8 | 333.3 | 2125.2 |
| 5 | 83.5 | 666.6 | 4250.4 |
| 10 | 167.1 | 1333.1 | 8500.7 |

*Assay buffer for reversibility studies (AKR).* Three milliliters of AKR were prepared by mixing 900 µl of 500 mM NaOAc pH 5.5, 1.0 µl of Brij 30, 93.8 µl of 80 mM DTT, 187.5 µl of 40 mM EDTA, 115 µl of DMSO, 5.1 µl of 29.94 mM Z-FR-AMC and 1698 µl of water. AKR was prepared prior to the experiment in a 15 ml plastic conical tube kept in the dark. Final conditions of AKR are: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 mM EDTA, 0.01% Brij 30, 50.5 µM Z-FR-AMC, and 4% DMSO. Table 78 summarizes these conditions.

Table 78. Preparation Table for Cathepsin K Assay Buffer for Reversibility Studies

| Volume (ml) | 40 mM EDTA (µl) | 80 mM DTT (µl) | Brij 35 (µl) | 500 mM NaOAc (µl) | 29.94 mM Z-FR-AMC (µl) | DMSO (µl) | Water (µl) |
|---|---|---|---|---|---|---|---|
| 3 | 187.5 | 93.8 | 1 | 900 | 5.1 | 115 | 1698 |

*Preparation of cathepsin K stock solution for reversibility studies (CKR).* One hundred and forty-four microliters of CKR were prepared by mixing 86.4 µl of 500 mM NaOAc pH 5.5, 9.0 µl of 80 mM DTT, 18 µl of 40 mM EDTA, 7.8 µl of 5.6 µM activated cathepsin K, and 22.8 µl of water in a glass test tube. Final conditions of CSR are: 300 mM NaOAc pH 5.5, 5 mM EDTA, 5 mM EDTA, and 300 nM cathepsin K. Table 79 summarizes preparation tables for CKR.

Table 79. Preparation Table for Cathepsin K Stock Solution for Reversibility Studies

| | Volume (ml) | 40 mM EDTA (µl) | 80 mM DTT (µl) | Brij 35 (µl) | 500 mM NaOAc (µl) | Cathepsin K (µl) | Water (µl) |
|---|---|---|---|---|---|---|---|
| CKR | 0.144 | 18 | 9 | 0 | 86.4 | 7.8 | 22.8 |

*Cathode buffer*. One liter of cathode buffer (25 mM Tris base; 40 mM glycine; 0.1% SDS) was prepared by dissolving 3.03 grams of Tris base, 3 grams of glycine and 1 gram of SDS with water. The solution was then filtered using a 0.2 µm filter and stored at 4 ºC. The recorded pH of the solution was 8.8.

*Anode I buffer*. Half liter (500 ml) of anode buffer (300 mM Tris base; 15 % methanol) was prepared by dissolving 18.2 grams of Tris base, and 75 ml of HPLC grade methanol with water. The solution was filtered with a 0.2 µm filter and stored at 4 ºC. The recorded pH of the solution was 10.4.

*Anode II buffer*. Half liter (500 ml) of anode buffer (25 mM Tris base; 15 % methanol) was prepared by dissolving 1.5 grams of Tris base, and 75 ml of HPLC grade methanol with water. The solution was filtered with a 0.2 µm filter and stored at 4 ºC.

The recorded pH of the solution was 9.7. Table 80 and Table 81 summarize conditions and procedures for cathode, anode I and anode II buffers.

Table 80. Concentrations of Cathode, Anode I and Anode II Buffers

| Buffer | Tris Base (mM) | Gly (mM) | Methanol (%) | SDS (%) | Total (ml) |
|---|---|---|---|---|---|
| Cathode | 25 | 40 | 0 | 0.1 | 1000 |
| Anode I | 300 | 0 | 15 | 0 | 500 |
| Anode II | 25 | 0 | 15 | 0 | 500 |

Table 81. Preparation Table for Cathode, Anode I and Anode II Buffers

| Buffer | Tris Base (g) | Gly (g) | Methanol (ml) | SDS (g) | Total (ml) |
|---|---|---|---|---|---|
| Cathode | 3.03 | 3.00 | 0 | 1 | 1000 |
| Anode I | 18.2 | 0 | 75 | 0 | 500 |
| Anode II | 1.51 | 0 | 75 | 0 | 500 |

*Preparation of NuPAGE® MES SDS running buffer (1X).* Fifty milliliters of the NuPAGE® MES SDS Running Buffer (20X) were diluted with 950 milliliters of ultrapure water in a volumetric flask. Final volume of the solution was one liter. Final conditions of the buffer (1X) were: 50 mM MES, 50 mM Tris Base, 0.1% SDS, and 1 mM EDTA. The pH of the buffer was 7.3 and was stable for up to 6 months when stored at 4 °C.

*Preparation of cathepsin K calibration curves for detection limits by using western blotting.* Four stock dilutions of inactive cathepsin K (pCKSS) ranging between 10 and 150 ng/µl were prepared by diluting procathepsin K stock solution (0.9 µg/µl ≡ 900 ng/µl) with water as described in Table 82. Then, fourteen samples ranging between 5 and 200 ng were set up by using the former stock solutions. Samples were brought up

to a final volume of 10 µl by adding 2.5 µl of loading buffer 4X  Dilution volumes, and masses are shown in Table 83.

Table 82.  Preparation Table of Procathepsin K Stock Solutions Detection Limit by Western Blotting Studies

| pCKSS | Procathepsin K (ng/µl) | Dilution Factor | From S (µl) | Water (µl) |
|---|---|---|---|---|
| S | 900 | (Stock) | 1 | 0 |
| S1 | 150 | 1:6 | 1 | 5 |
| S2 | 50 | 1:18 | 1 | 17 |
| S3 | 25 | 1:36 | 1 | 35 |
| S4 | 10 | 1:90 | 1 | 89 |

Table 83.  Preparation Table of Procathepsin K Samples Ranging Between 5 and 200 ng

| Sample | Procathepsin K (ng) | From pCKSS | From CKSS (µl) | Loading Buffer (µl) | Water (µl) |
|---|---|---|---|---|---|
| 1 | 5 | S4 | 0.5 | 2.5 | 7 |
| 2 | 10 | S4 | 1 | 2.5 | 6.5 |
| 3 | 15 | S4 | 1.5 | 2.5 | 6 |
| 4 | 20 | S4 | 2 | 2.5 | 5.5 |
| 5 | 25 | S3 | 1 | 2.5 | 6.5 |
| 6 | 30 | S3 | 1.2 | 2.5 | 6.3 |
| 7 | 35 | S3 | 1.4 | 2.5 | 6.1 |
| 8 | 40 | S3 | 1.6 | 2.5 | 5.9 |
| 9 | 45 | S2 | 1.8 | 2.5 | 5.7 |
| 10 | 55 | S2 | 1.1 | 2.5 | 6.4 |
| 11 | 65 | S2 | 1.3 | 2.5 | 6.2 |
| 12 | 100 | S1 | 2 | 2.5 | 5.5 |
| 13 | 150 | S1 | 1 | 2.5 | 6.5 |
| 14 | 200 | S1 | 1.33 | 2.5 | 7 |

*Preparation of 12 mM Sodium Acetate Buffer, pH 3.5.*  A fresh solution of this buffer was prepared by diluting 369 µl of 32.5 mM sodium acetate buffer, pH 3.5 to a final volume of 1000 µl by adding 631 µl of water.  Glacial Acetic acid or sodium hydroxide 5 M (MW: 40 g/mol) were used to adjust the pH of the buffer.  This buffer is also referred as modified solution B.

*Solution A with DMSO (SA-C).* Five milliliters of this solution were made by mixing 3.21 ml of 1 M NaCl, 0.67 ml 10 mM EDTA, 385 µl DMSO and 738 µl of water. Conditions of this solution are: 0.64 M NaCl, 1.34 mM EDTA, and 7.68 % DMSO.

*Solution A with **1** in DMSO (SA-I).* Five milliliters of this solution were made by mixing 3.21 ml of 1 M NaCl, 0.67 ml 10 mM EDTA, 375 µl DMSO, 9.61 µl of 20 mM **1** diluted in DMSO and 738 µl of water. Conditions of this solution are: 0.64 M NaCl, 1.34 mM EDTA, 7.68 % DMSO and 38.5 µM **1.**

*Activation of procathepsin K with DMSO as vehicle control.* The proenzyme was activated by mixing 3 µl of 0.9 µg/µl (33.3 µM) procathepsin K, 45 µl 12 mM NaOAc pH 3.5 (modified solution B) and 52 µl of solution *SA-C*. Total volume is 100 µl. Conditions for activation of the enzyme are: 1 µM of mature cathepsin K, 5.4 mM NaOAc pH 3.5, 0.7 mM EDTA, 333.3 mM NaCl and 4% DMSO. The mixture was incubated in a siliconized tube at 3 °C between zero and seven hours.

*Activation of procathepsin K with **1** and DMSO as vehicle control.* The proenzyme was activated by mixing 3 µl of 0.9 µg/µl (33.3 µM) procathepsin K, 45 µl 12 mM NaOAc pH 3.5 (modified solution B) and 52 µl of solution *SA-I).* Reaction volume is 48 µl. Conditions for activation of the enzyme are: 5.6 µM of mature cathepsin K, 5.4 mM NaOAc pH 3.5, 0.7 mM EDTA, 333.3 mM NaCl, 20 µM **1**, and 4% DMSO. The mixture was incubated in a siliconized tube at 3 °C between zero and seven hours.

*Inhibition of cathepsin K proteolytic activity by thiosemicarbazone derivatives.* Three stock solutions of human type I collagen ([ACI]: 0.4 mg/ml), human type IV collagen([ACIV]: 0.327 mg/ml) and Tubulin ([TUB]: 1.25 mg/ml) were obtained or

prepared as recently described in the materials and methods of chapters two and three. The three macromolecules were used as cathepsin K substrates. Two inhibitor stock solutions with final concentrations of 20 μM were prepared by mixing 6.25 μl 20 mM stock solution of the inhibitor in DMSO, 243.75 μl of DMSO and 250 μl of water. The experiment was performed by the addition of mature cathepsin K (activated under standard procedures) to a solution containing sodium acetate buffer pH 5.5, EDTA, DTT, C4-S, inhibitor, DMSO and substrate. Final volumes were 12.5 microliters in every case. Similarly, another group of control samples (no inhibitor) were also set up. The degradation of type I collagen, type IV collagen and tubulin were performed at 37 °C and monitored between 0 and 6 hours. Volumes, concentrations of stock solutions and final conditions of the reactions are summarized in Tables 83, 84 and 85. Inactivation of untreated and treated samples, electrophoresis and staining procedures using SYPRO® were described in chapter two in the section material and methods.

Table 84. Assay Conditions for the Degradation of Type IV Collagen by Cathepsin K

| Reagent | Stock Solution | Volume (μl) | Final Conditions |
|---|---|---|---|
| EDTA | 60 mM | 0.54 | 2.6 mM |
| DTT | 80 mM | 0.41 | 2.6 mM |
| NaOAc | 1 M | 1.00 | 80 mM |
| C4-S | 1.9% | 1.00 | 0.15% |
| Inhibitor (in DMSO) | 250 μM | 1.00 | 20 μM |
| DMSO | 50 % | | 4 % |
| Type IV Collagen | 0.327 μg/ μl | 8.00 | 2.6 μg (mass) |
| Cat K | 1728 nM | 0.50 | 70 nM |
| | Total | 12.5 | |

Table 85. Assay Conditions for the Degradation of Tubulin by Cathepsin K

| Reagent | Stock Solution | Volume (µl) | Final Conditions |
|---|---|---|---|
| EDTA | 60 mM | 0.54 | 2.6 mM |
| DTT | 80 mM | 0.41 | 2.6 mM |
| NaOAc | 1 M | 1.00 | 80 mM |
| C4-S | 1.9% | 1.00 | 0.15% |
| Inhibitor (in DMSO) | 250 µM | 1.00 | 20 µM |
| DMSO | 50 % | | 4 % |
| Tubulin | 1.25 µg/ µl | 2.4 | 3 µg (mass) |
| Cat K | 1728 nM | 0.50 | 70 nM |
| Water | | 5.60 | - |
| | Total | 12.5 | |

Table 86. Assay Conditions for the Degradation of Type I Collagen by Cathepsin K

| Reagent | Stock Solution | Volume (µl) | Final Conditions |
|---|---|---|---|
| EDTA | 60 mM | 0.54 | 2.6 mM |
| DTT | 80 mM | 0.41 | 2.6 mM |
| NaOAc | 1 M | 1.00 | 80 mM |
| C4-S | 1.9% | 1.00 | 0.15% |
| Inhibitor (in DMSO) | 250 µM | 1.00 | 20 µM |
| DMSO | 50 % | | 4 % |
| Type I Collagen | 0.4 µg/ µl | 7.5 | 3 µg (mass) |
| Cat K | 1728 nM | 0.50 | 70 nM |
| Water | | 5.60 | - |
| | Total | 12.5 | |

## *Experimental Section*

### *Activation of Procathepsin K*

The protocol for the activation of procathepsin K was slightly modified as suggested from the manufacturer. Activation of the proenzyme needed to be carried out under acidic conditions. Several parameters were optimized prior to the beginning of the screening of thiosemicarbazones as potential cathepsin K inhibitors. The parameters that

were optimized were: incubation time, enzyme concentration, substrate effect, stability of the mature enzyme after activation, and assay condition effects.

*Assay Optimization*

Five different cathepsin K assays were compared in order to determine the optimal condition to obtain the highest cathepsin K catalytic activity. Table 87 summarizes final conditions that were studied and compared. Assays A and B compare the effect of DTT, a reducing agent, in cathepsin K activity. Assays A and C compare the effect of Brij 30, a detergent, in cathepsin K activity. Finally, assays A and D compare the effect of DMSO, an aprotic solvent, in cathepsin K activity, without the presence of Brij 30. Four proenzyme activation times were arbitrarily set between 50 and 200 minutes for these studies.

Table 87. Cathepsin K Assay Buffer Conditions

| Assay | EDTA mM | DTT mM | BRIJ % | NaOAc mM | DMSO % | Cathepsin K nM | Z-FR-AMC μM |
|---|---|---|---|---|---|---|---|
| A | 2.5 | 2.5 | 0 | 150 | 2 | 3.6 | 50 |
| B | 2.5 | 20 | 0 | 150 | 2 | 3.6 | 50 |
| C | 2.5 | 2.5 | 0.01 | 150 | 2 | 3.6 | 50 |
| D | 2.5 | 2.5 | 0 | 150 | 4 | 3.6 | 50 |
| E | 2.5 | 2.5 | 0.01 | 150 | 4 | 3.6 | 50 |

*Stability of Activated Cathepsin K*

Literature also reports activated cathepsin K is extremely unstable and loses its enzymatic activity fairly quickly. Therefore, we tested the stability of cathepsin K after the protease was fully activated. Stability studies were done for twelve time periods between 0 and 55 minutes after activation time.

*Enzyme Concentration Effect*

Four different cathepsin K concentrations were also evaluated to determine an optimal catalytic activity. Activated cathepsin K final concentrations were 1, 1.5, 2 and 2.5 nM using the conditions that were previously described. Additionally, every cathepsin K concentration was tested at seven different activation times (between 1 and 4 hours).

*Kinetic Cathepsin K Assay*

Kinetic analysis of cathepsin K was carried out by using a Thermo Fluoroskan microplate reader, 3691 96-well black microplates. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 10 µl of 50% DMSO solution, 20 µl cathepsin K stock solutions, and 20 µl of Z-FR-AMC stock solution. A mixture containing assay buffer, and 50% DMSO, was preincubated at 25 °C during 5 minutes using 96-black microplates. Triplicate sets of the catalytic activity of cathepsin K were monitored by adding a concentration of Z-FR-AMC to every reaction. The production of AMC was monitored for 5 minutes at 25 °C using excitation and emission references of 355 and 460 nm respectively. Readings were taken every 8 seconds for five minutes. The final concentrations of the kinetic assay are: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, and 1.5 nM activated cathepsin K. Final concentrations of Z-FR-AMC varied between 1 and 75 µM.

*Preliminary Inhibition Studies*

Thiosemicarbazone analogs (provided by Dr. Kevin G. Pinney's laboratory) [3,6,7,10–12] were prescreened to determine if they have inhibitory activity against cathepsin K. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay

buffer, 50 µl of water, 10 µl of 50% DMSO or 10 µl of dilution "1" (final concentration: 10 µM), 20 µl cathepsin K stock solutions, and 20 µl of Z-FR-AMC stock solution. The enzyme-inhibitor mixtures (180 µ) assay buffer, 50% DMSO or inhibitor, and cathepsin K was preincubated at 25 °C during 5 minutes using 96-black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and uninhibited samples were monitored for five minutes. Final reactions were started by adding 20 µl of Z-FR-AMC. The final concentrations of the preliminary inhibitory studies are: 150 mM NaOAc pH 5.5, 1 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 1.5 nM cathepsin K, 10 µM of the screened inhibitor and 50 µM of Z-FR-AMC. Readings were taken every 25 seconds for five minutes and reactions were carried out in triplicate. Compounds that did not have cathepsin K inhibitory activity more than 50% (i.e. $v_i/v_o \leq 0.5$) were considered as 'inactive' compounds and a general $IC_{50}$ value greater than 10000 nM was assumed. Compounds that inhibited cathepsin K inhibitory activity more than 50% were further considered for cathepsin K inhibitory studies and an exact $IC_{50}$was determined.

*Cathepsin K Inhibition Assay*

Analysis of cathepsin K and its inhibitors was carried out by using a modified protocol of the kinetic cathepsin K assay. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 50 µl of water, 10 µl of 35% DMSO or 10 µl of inhibitor dilutions, 20 µl cathepsin K stock solutions, and 20 µl of Z-FR-AMC stock solution. An 180 µl mixture containing assay buffer, 50% DMSO or inhibitor, and cathepsin K was preincubated at 25 °C during 5 minutes using 96-black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited

and uninhibited samples were monitored for five min. The final concentrations of the inhibitory cathepsin K assay are: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 1.5 nM cathepsin K and 50 µM of Z-FR-AMC. Final concentrations of the inhibitors varied between 10 µM and 10 pM. Readings were taken every 25 seconds for five minutes and reations were carried out in triplicate. Table 89 summarizes final volume sand final conditions for the kinetic and inhibitory cathepsin K assays.

*Construction of AMC Calibration Curve*

The calibration curve obtained for cathepsin L experiments (Chapter 2) was used for similar cathepsin K experiments. Complete details and results of these experiments can be found in chapter 2 under the subsection "Construction of AMC calibration curve".

Table 88. Final Conditions for Cathepsin K Kinetic and Inhibitory Assays

| NaOAc (mM) | EDTA (mM) | DTT (mM) | Brij 35 (%) | DMSO (%) | Cathepsin K (nM) | Z-FR-AMC (µM) |
|---|---|---|---|---|---|---|
| 150 | 2.5 | 2.5 | 0.01 | 4 | 1.5 | 1 - 75 |

Table 89. Preparation Table for Kinetic, Inhibitory Cathepsin K Assays and Construction of AMC Calibration Curves

| Reagent (µl) | Kinetic assay | Inhibitory assay | AMC Curve |
|---|---|---|---|
| Assay Buffer | 150 | 150 | 150 |
| Control | 10 | 0 | 10 |
| Inhibitor | 0 | 10 | 0 |
| CK | 20 | 20 | 20 |
| Z-FR-AMC/AMC | 20 | 20 | 20 |
| Total | 200 | 200 | 200 |

*Effect of Inhibitor Concentration on Cathepsin K Progress Curves*

Final concentrations, conditions, and volumes are similar as previously described in the cathepsin K inhibition assay. Assay buffer, inhibitors (Final concentrations varied

between 100 nM and 10 µM of **1**) and Z-FR-AMC (final concentration: 50 µM) were added to the 96-well black plates (volume of the substrate-inhibitor mixture: 180 µl). Then, 20 µl of cathepsin K stock solution were added immediately without preincubation time. Readings were taken every 3 seconds for fifty minutes.

*Effect of Inhibitory Activity of Cathepsin K Inhibitors on Preincubation Studies*

Preincubation studies with compound **1** were carried out by a slight modification of the cathepsin K inhibition assay. Different sets of mixtures containing assay buffer, inhibitor and cathepsin K were preincubated at various periods of times between 0 and 180 minutes. Reactions were taken every 25 seconds for five minutes and carried out in triplicate.

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation*

Data that were obtained in the effect of preincubation studies was further analyzed. The possibility of **1** to be a tight binding inhibitor was examined by fitting the data by a nonlinear regression using Morrison's quadratic equation.

*Cathepsin K Reversibility Studies*

Three milliliters of assay buffer for reversibility studies (AKR) were prepared (see section Materials and methods). Twenty five microliters of cathepsin K (100X: 150 nM) assay buffer for reversibility studies (ASL) were pre-incubated with an equal amount of a concentrated solution of the inhibitor (100X: 0.1 $IC_{50}$) at 25 °C for sixty minutes. Then, two microliters of the enzyme-inhibitor mixture were rapidly mixed with 198 µl of AKR in order to start the reaction. Total reaction volume was 200 µl. Readings were

taken every twenty five seconds for four hours. Final concentrations are similar as described previously. Table 90 describes required volumes for this experiment.

Table 90. Preparation Table for Cathepsin K Reversibility Assay

| Reagent (µl) | Reversibility studies |
|---|---|
| Assay Buffer | 198 |
| Inhibitor/CK | 2 |
| Total | 200 |

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

The effect of [Z-FR-AMC] was studied with compound **1**. Minor modifications of the cathepsin K inhibition assay were carried out. Different sets of mixtures containing assay, inhibitor and cathepsin K were preincubated at a specific preincubation time (5 minutes). Reactions were initiated with the addition of different concentrations of Z-FR-AMC. Readings were taken every 25 sec for five minutes and carried out in triplicate.

*Detection Limits of Procathepsin K by Fluorescent Western Blotting*

Fourteen samples of human cathepsin K ranging between 5 and 200 ng were prepared as previously described. Samples also contained sample buffer (1X) and water for a final volume of 10 µl each.

*Selection of the Antibody.* The selection of the antibody was extremely important for the immunodetection of procathepsin K and mature cathepsin K. In summary, the ideal antibody could be able to detect both special simultaneously. Monoclonal human anti-cathepsin K (host: mouse) was chosen due to its ability to recognize specifically mature cathepsin K. The epitope of the antibody recognizes the 14-residue peptide with

the sequence: RGYREIPEGNEKAL (residues 222-235 in mature cathepsin K). This antibody can also recognize procathepsin K since both species share that specific peptide in their amino acid sequences. Figure 87 shows the polypeptide sequence that can be recognized by the antibody.

```
  1  LYPEEILDTHWELWKKTHRKQYNNKVDEISRRLIWEKNLKYISIHNLEAS
 51  LGVHTYELAMNHLGDMTSEEVVQKMTGLKVPLSHSRSNDTLYIPEWEGRA
101  PDSVDYRKKGYVTPVKNQGQCGSCWAFSSVGALEGQLKKKTGKLLNLSPQ
151  NLVDCVSENDGCGGGYMTNAFQYVQKNRGIDSEDAYPYVGQEESCMYNPT
201  GKAAKCRGYREIPEGNEKALKRAVARVGPVSVAIDASLTSFQFYSKGVYY
251  DESCNSDNLNHAVLAVGYGIQKGNKHWIIKNSWGENWGNKGYILMARNKN
301  NACGIANLASFPKM
```

Figure 87. Amino Acid Sequence of Procathepsin K (PDB ID: 3KX1).[138] Legend: Red: Propeptide; Green: Mature Cathepsin K; Purple: Antibody Recognition Site

*Electrophoresis.* The set of samples were heated at 90 ºC for ten minutes and loaded into 4-12% Bis-Tris SDS-PAGE gels. Electrophoresis of the gel containing was performed using MES running buffer (1X) containing SDS at 200 V for 35 minutes. A fluorescent marker (2.0 µl) was also loaded as a reference.

*Equilibration of the Blotting Sandwich.* Six pieces of blotting paper, low fluorescent PVDF membrane and gel were equilibrated in cathode, anode I and anode II buffers. In summary, the SDS-PAGE gel was equilibrated in cathode buffer for 30 minutes; three pieces of blotting paper in cathode buffer for 30 minutes; two pieces of blotting paper in anode I buffer for 30 minutes; one piece of blotting paper in anode II buffer for 30 minutes. Finally, the membrane was equilibrated in pure methanol for 2 minutes and then equilibrated in anode II buffer for 40 minutes.

Table 91 summarizes conditions for equilibration of the blotting sandwich components. The usage of methanol and SDS has been explained by the manufacturer

(Biorad). Discontinuous systems, that is, the use of three buffers in semi-dry protein transfers systems require the optimization and empirical determination of optimal conditions when transferring proteins from gels to membranes.

Methanol is the preferred solvent in protein transfer for numerous reasons. First, methanol can have the ability to remove SDS excess that is bound to the target protein. This removal promotes a higher binding affinity between the macromolecule and the membrane. One the major disadvantages of methanol is the reduction of pore size in membrane causing poor transfer efficiencies if the concentration of methanol is relative high. It is usually recommended to keep higher concentration of methanol if the target protein to transfer is relative small of medium in size. Lower concentrations of methanol may improve transfer efficiencies when transferring larger proteins (i.e. MW ≥ 100 kDa).

On the other hand, lower amounts of SDS may improve the solubility of the proteins, thus facilitating the transfer of the former ones from the gel to the membrane. The most important disadvantage of SDS in discontinuous systems is the poor efficiency of protein binding to the membrane.

Table 91. Equilibration Times for the Components of the Blotting Sandwich

| Buffer | Blotting paper | Gel | PVDV membrane |
|---|---|---|---|
| Methanol | | | 2 min |
| Cathode | 3 pc (30 min) | 30 min | |
| Anode I | 2 pc (30 min) | | |
| Anode II | 1 pc (30 min) | | 40 min |

*Assembly of the transfer stack and protein transfer*. The blotting sandwich was prepared as described. All of the components were placed in the center of the anode plate as described: two pieces equilibrated in anode buffer I; one piece equilibrated in anode buffer I; PVDF membrane equilibrated with methanol and anode II buffer; SDS- PAGE

gel equilibrated in cathode buffer; and three pieces of blotting paper equilibrated in cathode buffer. Air trapped between each layer was carefully removed. Finally, the cathode electrode plate was firmly placed on top of the stack and secured. Figure 74 shows a schematic representation of the transfer stack in a semi-dry protein transfer system. Protein transfer was run at 20 V for 30 minutes.

*Protein blocking and blotting*. The membrane was blocked with a blocker solution containing 1% BSA at room temperature for one hour. Then, the blocked membrane was blotted in a 50 ml solution containing 1.0 µg/ml of mouse anti-human cathepsin K ([Stock solution]: 2 mg/ml; dilution factor: 1:2000) in blocking solution (3% BSA). The blotting was incubated at 4 °C for 12 hours.

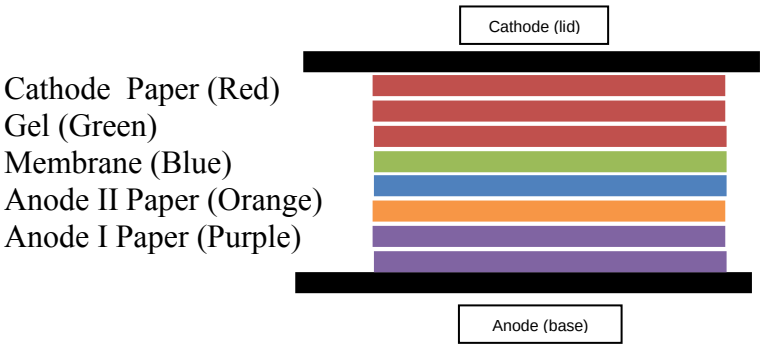



Figure 88. Schematic Representation of the Blotting Sandwich using a Semi-Dry Transfer Device

*Protein detection*. Unbound antibody was washed with wash buffer (PBS-Tween) three times, 10 minutes each. Then, the washed membrane was incubated using wash buffer with DyLight 649 goat anti-mouse IgG (H + L) secondary antibody (1: 2500) for one hour at room temperature. Finally, the excess of antibody was removed by washing the membrane with wash buffer six times, 5 minutes each. Fluorescent detection was

performed by using a GE typhoon FLA 9000 using an excitation wavelength of 635 nm. Conditions of the scanning were: PMT: 100 V, method: Cy5, excitation wavelength: 635 nm, pixel size 100 µm.

*Inhibition of the Activation of Human Procathepsin K*

Activation of human procathepsin K was slightly modified in order to extend activation time of procathepsin K. Two sets of samples (untreated and treated with compound **1**) were prepared as previously described. Three micrograms of procathepsin K (3 µl; stock solution: 0.9 µg/µl) were incubated at 3 degrees activation buffer with DMSO (control). Similarly, three micrograms of procathepsin K (3 µl; stock solution: 0.9 µg/µl) were incubated at 3 degrees activation buffer with **1** and DMSO (treated samples). Final conditions were: 1 µM (27 ng/µl) active cathepsin K, 0.7 mM EDTA, 333.3 mM NaCl, 5.4 mM NaOAc pH 3.5, 4% DMSO, and 20 µM of **1** (for the treated samples). Then, samples consisting of 200 ng (7.5 µl of activated cathepsin K) of protein were taken every 45 minutes and inactivated using 1% SDS (1.5 µl) and heated at high temperatures (100 ºC) in order to stop the activation process. Final volumes of the samples for western blotting were 9.5 µl. Immediately, heated samples were stored at -80 ºC. Also, aliquots (2.5 µl) were also taken to measure cathepsin K catalytic activity using fluorometric techniques. Final concentrations in fluorometric assays were: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 DTT, 0.1 % Brij, 50 µM Z-FR-AMC, 4% DMSO, 1.5 nM cathepsin K and 30 nM **1** (for treated ones)**.** Activation of procathepsin K was followed for a total time of 6.75 hours.

*Electrophoresis and Western Blotting.* Electrophoresis was carried out as previously described in this chapter under the section "Detection limits of procathepsin K by Fluorescent Western Blotting" with minor modifications. Briefly, 9.5 µl of inactivated cathepsin K (200 ng) were mixed with 2.5 µl of loading sample buffer (4X) containing LDS and 1 µl of reducing agent. The set of mixtures (treated and untreated) were heated at 90 ºC for ten minutes and loaded into 4-12% Bis-Tris SDS-PAGE gels. Electrophoresis of two gels containing treated and untreated samples separately was performed using MES running buffer containing SDS at 200 V for 35 minutes. A fluorescent marker (2.0 µl) was also loaded as a control. Equilibration of the blotting sandwich, Assembly of the transfer stack and protein transfer, Protein blocking and blotting and Protein detection were carried out as previously described in this chapter.

*Effect of Chondroitin 4-Sulfate on Cathepsin K Inhibition Assay*

Analysis of cathepsin K and **1** was carried out by using a modified protocol of the Cathepsin K Inhibition Assay. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 50 µl of water, 10 µl of 35% DMSO or 10 µl of inhibitor dilutions, containing 2.7 % C4-S, 20 µl cathepsin K stock solutions, and 20 µl of Z-FR-AMC stock solution. An 180 µl mixture containing assay buffer, 50% DMSO or inhibitor, and cathepsin K was preincubated at 25 °C during 5 minutes using 96-black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and uninhibited samples were monitored for five min. The final concentrations of the inhibitory cathepsin K assay are: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 1.5 nM cathepsin K, 0.14 % C4-S and 50 µM of Z-FR-AMC. Final concentrations of the inhibitors varied between 10 µM and 10 pM.

Readings were taken every 25 seconds for five minutes and reations were carried out in triplicate.

## *Results and Discussion*

Previous results obtained in collaboration results between Trawick and Pinney groups indicated that thiosemicarbazones are lead compounds as cathepsin L inhibitors[10–12]. Therefore, the library of compounds required a complete screening to verify the potency of novel analogs against human cathepsin K and potential lead compounds that could show selectivity when compared to cathepsin L's case (Chapter Two). The Structure-Relationship Activity (SAR) was characterized after completing *in vitro* testing against cathepsin K ($IC_{50}$ values).

Fluorometric based assays were utilized to study various assay parameters in inhibition activities, determination of $K_I$, and reversibility of thiosemicarbazone inhibitors. Finally, the inhibition of the activation of procathepsin K under acidic conditions by using a thiosemicarbazone as a lead inhibitor was followed by using western blotting and fluorometric techniques.

### *Assay Optimization*

To optimize the assay a variety of reaction conditions and activation times were explored. The catalytic activity of activated cathepsin K was measured by using Z-FR-AMC, a fluorogenic substrate at specific conditions and activation times. Triplicates or more for each condition were carried out. Averages and standard errors are presented in Table 92 and Figure 89.

Five different enzymatic conditions were tested to verify the activity of cathepsin K over time. The lowest activity was obtained when final conditions contained relatively lower concentrations of the solvent vehicle (DMSO, 2%). Another factor that enhanced the activity was the presence of detergent (Brij 30). Cathepsin K activity at longer activation times (220 min) was 1.3 fold greater when a non-ionic detergent was present. A similar trend could be observed when DMSO was 4% (Conditions D and E, Table 87).

Table 87. Cathepsin K Assay Buffer Conditions

| Assay | EDTA mM | DTT mM | BRIJ % | NaOAc mM | DMSO % | Cathepsin K nM | Z-FR-AMC µM |
|---|---|---|---|---|---|---|---|
| A | 2.5 | 2.5 | 0 | 150 | 2 | 3.6 | 50 |
| B | 2.5 | 20 | 0 | 150 | 2 | 3.6 | 50 |
| C | 2.5 | 2.5 | 0.01 | 150 | 2 | 3.6 | 50 |
| D | 2.5 | 2.5 | 0 | 150 | 4 | 3.6 | 50 |
| E | 2.5 | 2.5 | 0.01 | 150 | 4 | 3.6 | 50 |

Table 92. Comparison of Cathepsin K Activity by Varying Assay Conditions

| Activation time (min) | Condition (see Table 87, page 237, for complete details) | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | E |
| | (Cathepsin K Activity, nM AMC/s) | | | | |
| 50 | 0.17 | 0.23 | 0.42 | 0.51 | 0.60 |
| 90 | 0.55 | 0.63 | 1.01 | 1.38 | 1.51 |
| 150 | 1.33 | 1.65 | 2.62 | 3.39 | 4.13 |
| 200 | 1.36 | 2.11 | 2.75 | 3.68 | 4.60 |

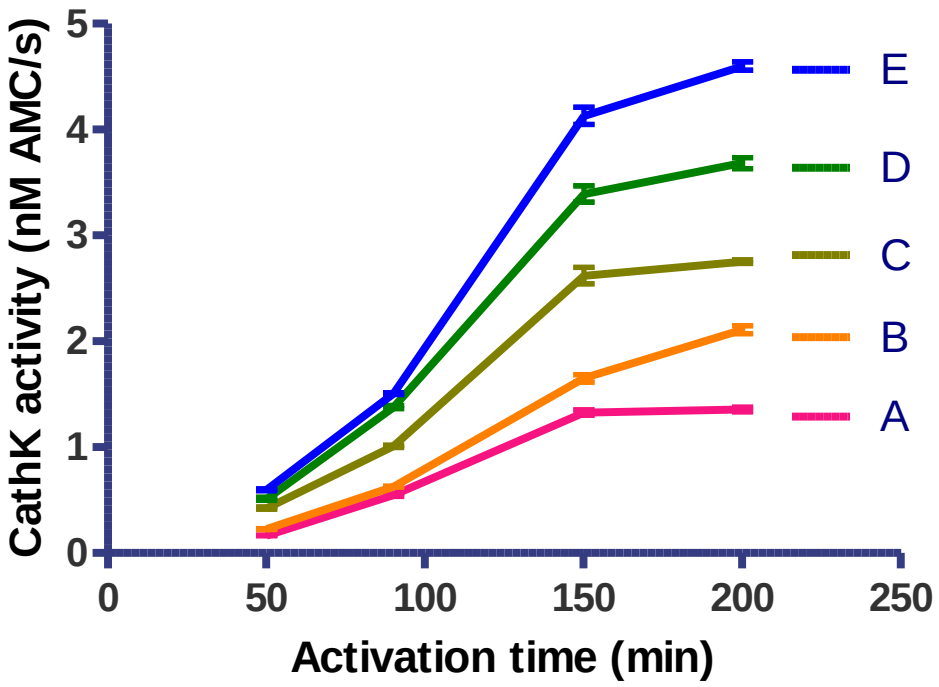



Figure 89. Comparison of Cathepsin K Activity by Varying Assay Conditions. (See Table 87 and Table 92)

Based on experimental results, condition E (150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 mM DTT, 0.01% Brij, 4% DMSO, 50 µM Z-FR-AMC) was chosen as the preferred condition for future cathepsin K experiments. It was also established that activation times can vary and they can be quite lengthy. However, activation time was determined empirically for every batch of procathepsin K that was obtained. Maximum activity was obtained with condition E when procathepsin K was activated for 220 minutes (Activity: 4.6 nM AMC/s).

*Stability of Activated Cathepsin K*

Stability studies were performed within the first hour after cathepsin K was fully activated. Results are the average of four independent experiments per stability time and can be seen in Figure 90.

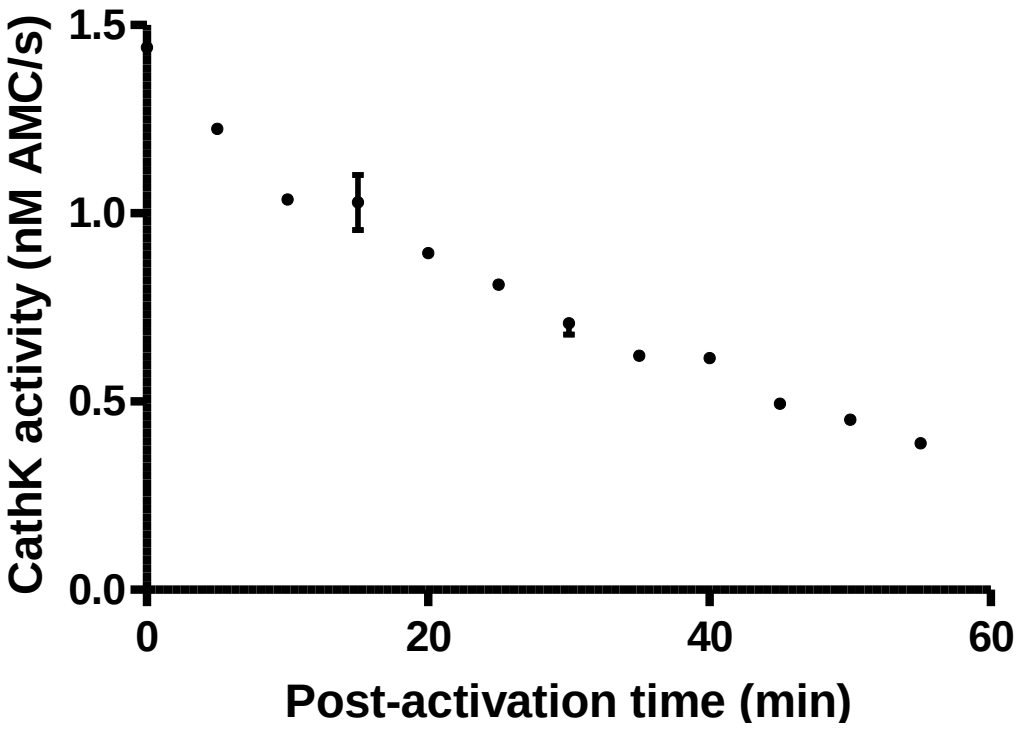



Figure 90. Stability of Activated Cathepsin K Curve

As seen in Figure 90, cathepsin K is extremely unstable under this specific activation conditions (pH: 3.5) and loses its activity at a very fast pace. Nevertheless, linear regression of this data (not shown, $r^2$: 0.9612) reveals cathepsin K loses 1.3% its catalytic activity per minute. For example, cathepsin K lost 30% of its activity after 10 minutes it was activated. Therefore, it was determined that cathepsin K experiments should be carried out within the first ten minutes after activation time was reached.

*Enzyme Concentration Effect*

Finally, the minimum enzyme concentration was also determined to verify if a proper detection limit was obtained. Results are the average of three independent results and can be seen in Table 93 and Figure 91.

Table 93. Cathepsin K Activity vs Enzyme Concentration

| Activation time (h) | Cathepsin K (nM) 1 | 1.5 | 2 | 2.5 |
|---|---|---|---|---|
| | (nM AMC/s) | | | |
| 1.00 | 0.07 | 0.25 | 0.60 | 0.75 |
| 1.50 | 0.24 | 0.48 | 0.72 | 1.42 |
| 2.00 | 0.32 | 0.67 | 0.97 | 1.35 |
| 2.50 | 0.39 | 0.79 | 1.33 | 1.65 |
| 3.00 | 0.46 | 0.49 | 0.71 | 1.20 |
| 3.50 | 0.14 | 0.46 | 0.66 | 1.15 |
| 4.00 | 0.19 | 0.32 | 0.85 | 0.96 |

Four cathepsin K concentrations were evaluated to optimize the activation time. Seven different activation times were tested for each concentration. Maximum activity was reached when the enzyme was activated under acidic conditions between 2.5 and 3 hours. Furthermore, the fragility of the enzyme was quite evident at longer incubation times (activation times: 3.5 and 4 hours). Activation time was set up to be between 2 and 2.5 hours. Nevertheless, activation time was determined empirically for each procathepsin K.

*Cathepsin K Kinetic and Inhibition Studies*

Cathepsin K studies were carried out using a 96-well microplate fluorometric based assay. Conditions, reagents, and experiments can be seen in chapter three under the subsection "Cathepsin K kinetic and inhibition studies".

*Determination of $K_{M,}$ $V_{MAX}$ and $k_{CAT}$*

Kinetic parameters for cathepsin K were determined using GraphPad 5.0 (see Chapter Two, for a complete explanation). Z-FR-AMC, a broad specificity substrate for cysteine proteases was used as a fluorogenic substrate for the recombinant protein (Figure

23, equation 1.1, page 35). The release of AMC observed a linear trend when 0.5 μM Z-FR-AMC was used during these experiments. The determination of $K_M$, $V_{MAX}$ and $k_{CAT}$ was possible by using substrate concentrations that varied between 1 and 75 μM. Substrate inhibition was also observed when [Z-FR-AMC] ≥ 150 μM. To date, this is the first report of the effect of substrate inhibition on the activity of cathepsin K.

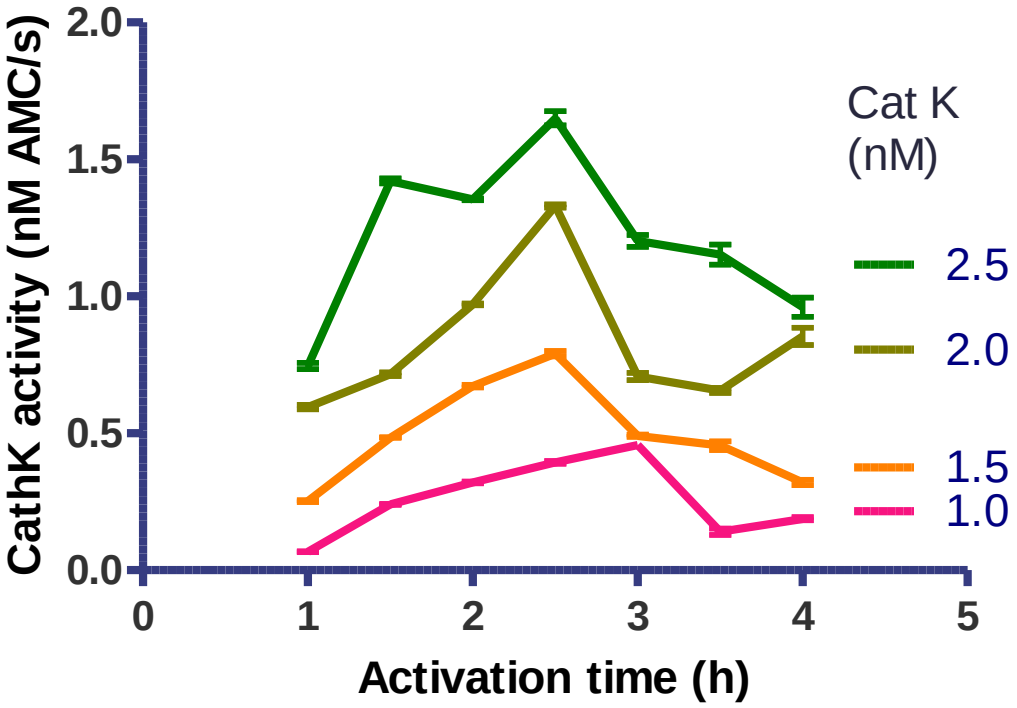



Figure 91. Cathepsin K Activity vs Enzyme Concentration

Experiments were carried out in triplicate. The parameters $K_M$, $V_{MAX}$ and $k_{CAT}$ were determined by using equations 1.1 and 3.1. Their values were found to be 13.5 ± 2.1 μM, 0.52 ± 0.05 nM AMC/sec and 0.52 $s^{-1}$ respectively [362] A comparison between cathepsins K and L kinetics values reveals a considerable difference between affinities. Z-FR-AMC represents an excellent cathepsin L substrate ($K_M$: 1 μM); however, that is not the case for cathepsin K.

Its $K_M$ is 10 times higher than cathepsin L and this discrepancy might be explained with substrate affinity. Cathepsin K has a specific preference for proline, and other substrates (Z-GPR-AMC) have been developed to test cathepsin K activity in other studies.[359]

$$v_o = \frac{v_{MAX}[S]}{K_M + [S]} \quad (1.1) \qquad k_{CAT} = \frac{v_{MAX}}{[cathepsin\ K]} \quad (3.1)$$

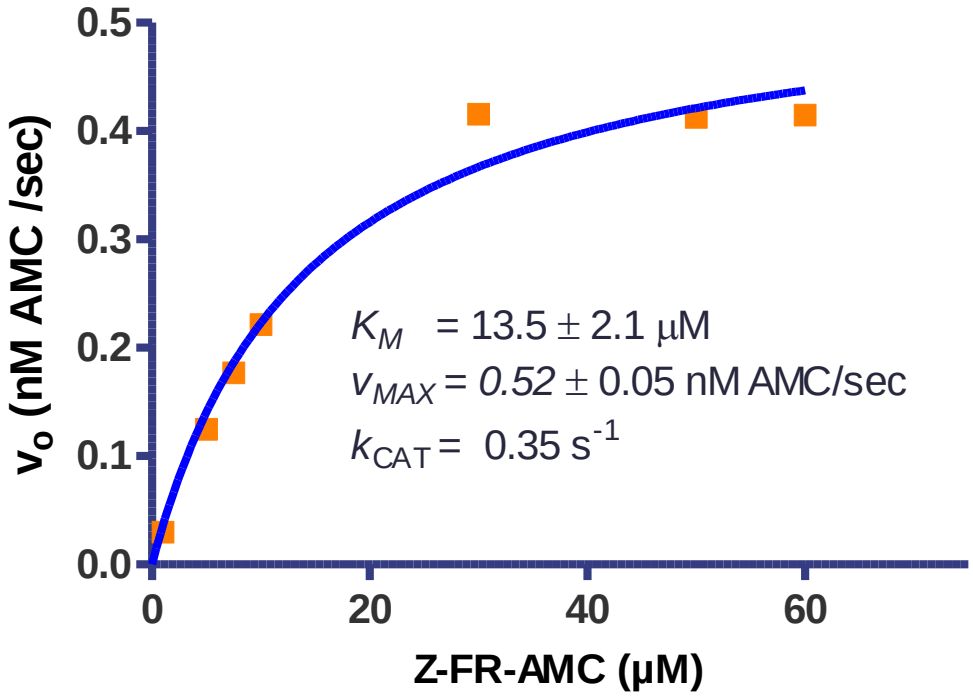



Figure 92. Determination of $K_M$ and $V_{MAX}$ for Human Cathepsin K Using Z-FR-AMC as a Fluorogenic Substrate

*Substrate Inhibition*

As previously described, substrate inhibition was observed at higher concentrations ([Z-FR-AMC) ≥ 75 μM). Figure 79 shows cathepsin K activity versuss substrate concentration. Inhibition at 75 μM is slightly evident, and a dramatic inhibition is seen when substrate concentration is doubled.

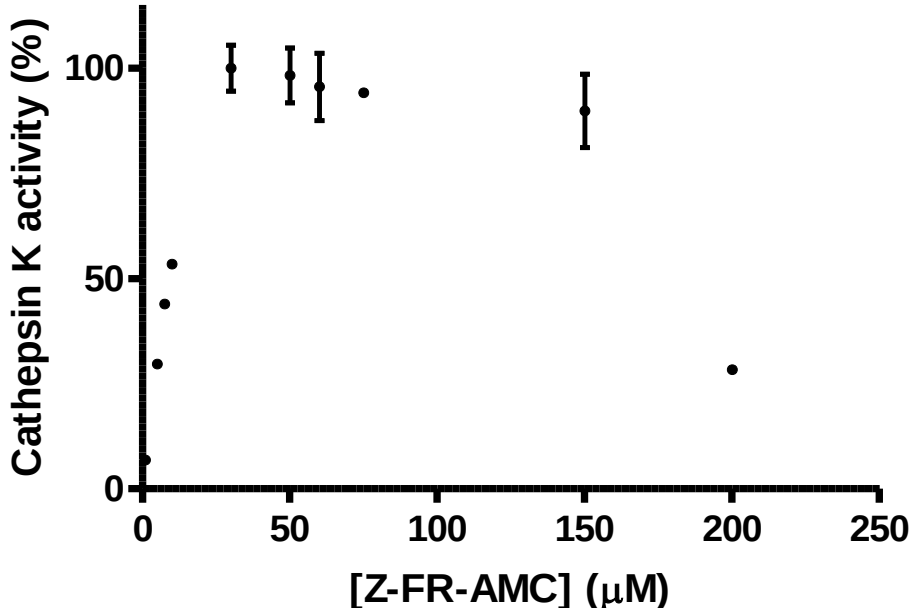



Figure 78. Effect of Substrate Inhibition on Human Cathepsin K

*Determination of Inhibitory Efficacy of Thiosemicarbazone Analogs at 10 μM*

The first set of experiments that were performed was to determine if individual compounds in the series of thiosemicarbazone derivatives were potential cathepsin K inhibitors, that is if the catalytic activity of the cysteine protease was inhibited by 50% or more by a fixed concentration (final concentration: 10 μM) of the potential inhibitor. Experimental details and conditions are similar to those described previously in the Material and Methods Section of Chapter Two. Three independent sets of experiments of untreated ([I]: 0 μM) and treated ([I]:10 μM) samples were preincubated with 1.5 nM activated cathepsin K for 5 min at 25 °C. Inhibitory activities were monitored when reactions were started by adding 50 μM Z-FR-AMC as a fluorogenic substrate. Reactions observed a linear behavior for the first minutes. Catalytic rates of uninhibited and inhibited samples were calculated by linear regression of the data. ([AMC]: dependent variable and time (sec): the independent variable). Compounds were

considered as potential inhibitors if more than 50% cathepsin K activity was inhibited by 10 μM in solution. Thus, they were further analyzed to determine an exact $IC_{50}$ value. If the ratio $v_i/vo \leq 0.5$, then the compounds were not considered potential inhibitors and an approximated $IC_{50}$ value of 10000 nM was assigned to them.

*Determination of $IC_{50}$ Values*

Several sub-libraries of thiosemicarbazones including of *meta*-bromosubstituted benzophenone thiosemicarbazones (Compounds **1-44)** among others were analyzed to investigate their ability to inhibit cathepsin K. The synthetic compounds were synthesized under the guidance of Dr. Kevin G. Pinney by several members of his research group.[3,6,7,10–12] A 96-well microplate fluorometric based assay was utilized to determine the inhibitory activity of each of these inhibitors. A comprehensive explanation of the experimental procedure can be found in the Material and Methods section of this chapter. Uninhibited cathepsin K catalytic activity showed a linear behavior when 50 μM Z-FR-AMC was used for reactions times that was up to 5 minutes long. The determination of the $IC_{50}$ is possible with experiments that observe the inhibitory capacity of the synthetic compounds when a fixed concentration of cathepsin K (1.5 nM) is preincubated for 5 minutes at 25 °C. Final concentration of each compound varied between 10 pM and 10 μM. Experiments were carried out in triplicate. Equation 1.2 was used to determined $IC_{50}$ values for these compounds. A complete structure-activity relationship (SAR) is shown in tables that group these inhibitors by functional groups or families (See Table 94-Table 113).

$$Y = \frac{v_{MIN} + (v_{MAX} - v_{MIN})}{1 + 10^{(\log(IC_{50} - X) * Hillslope)}} \tag{1.2}$$

Table 94. Inhibition of Human Cathepsin K by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12]

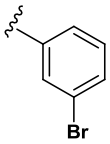
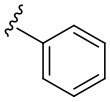
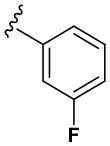
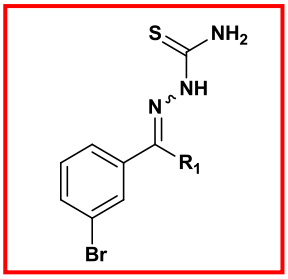
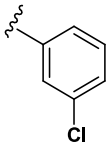
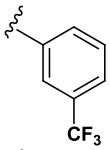
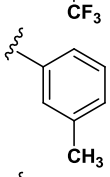
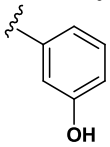
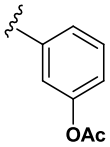
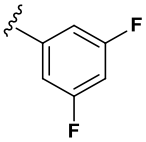
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **1** | |  | 35.2 ± 2.5 |
| **2** | |  | 660.4 ± 64.1 |
| **4** | |  | 104.2 ± 6.2 |
| **5** |  |  | 259.0 ± 19.4 |
| **6** | |  | 65.3 ± 5.4 |
| **7** | |  | 133.3 ± 4.3 |
| **8** | |  | 53.3 ± 3.4 |
| **9** | |  | 98.9 ± 8.4 |
| **10** | |  | 26.0 ± 1.0 |

Table 94. Inhibition of Human Cathepsin K by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

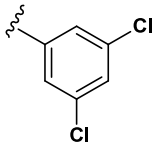
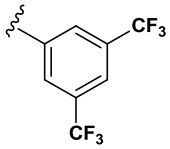
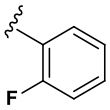
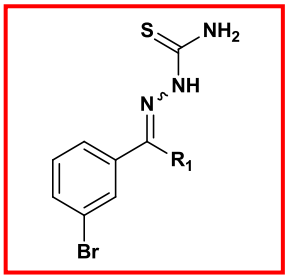
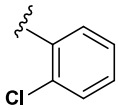
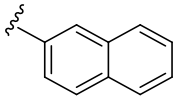
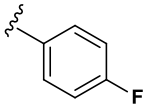
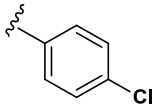
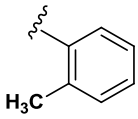
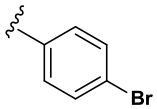
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **12** | |  | 335.6 ± 13.7 |
| **13** | |  | 11.3 ± 0.9 |
| **14** | |  | 61.8 ±5.5 |
| **15** |  |  | 312.9 ± 6.9 |
| **16** | |  | 92.5 ±2.8 |
| **17** | |  | 65.6 ±5.0 |
| **18** | |  | 626.4 ± 19.3 |
| **20** | |  | 838.9 ± 80.3 |
| **21** | |  | 49.1 ± 2.5 |

Table 94. Inhibition of Human Cathepsin K by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

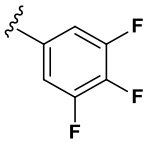
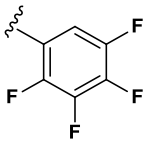
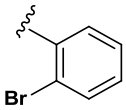
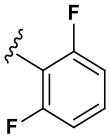
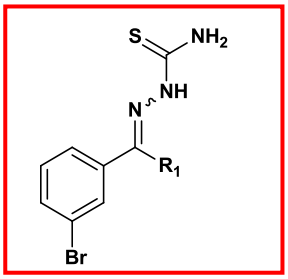
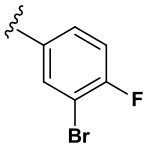
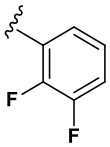
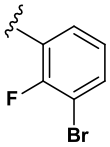
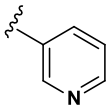
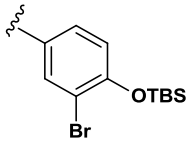
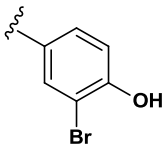
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **22** | |  | 32.6 ± 3.5 |
| **23** | |  | 22.6 ± 0.7 |
| **24** | |  | 1300 ± 99.6 |
| **25** | |  | 107.3 ± 4.0 |
| **26** |  |  | 109.0 ± 5.1 |
| **27** | |  | 40.3 ± 1.4 |
| **28** | |  | 23.0 ± 0.4 |
| **29** | |  | 66.7 ± 6.2 |
| **30** | |  | 118.8 ± 8.9 |
| **31** | |  | 10.2 ± 1.0 |

Table 94. Inhibition of Human Cathepsin K by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

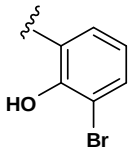
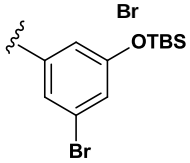
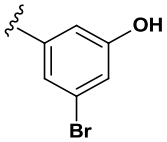
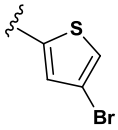
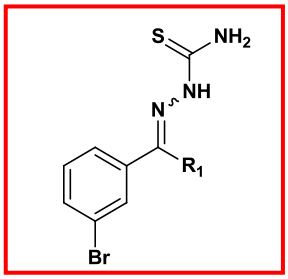
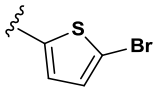
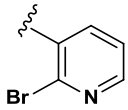
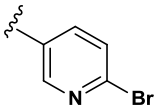
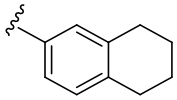
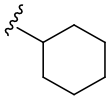
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **32** | |  | 29.0 ± 2.1 |
| **33** | |  | 65.8 ± 3.7 |
| **34** | |  | 15.6 ± 1.2 |
| **35** | |  | 524.8 ± 31.2 |
| **36** |  |  | 168.5 ± 13.6 |
| **37** | |  | 150.1 ± 4.3 |
| **38** | |  | ≥ 10000 |
| **39** | |  | ≥ 10000 |
| **40** | |  | 30.3 ± 1.1 |
| **41** | |  | 6154 ± 36.7 |

Table 94. Inhibition of Human Cathepsin K by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

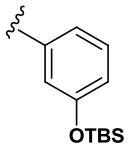
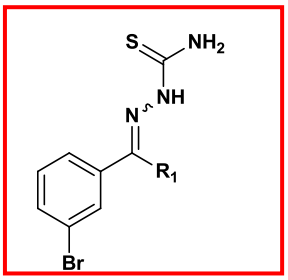
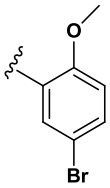
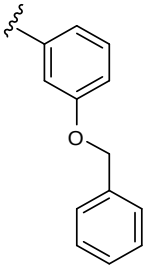
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **42** | |  | 4190 ± 286.5 |
| **43** |  |  | 4118 ± 408.4 |
| **44** | |  | 140.9 ± 3.2 |

Table 95. Inhibition of Human Cathepsin K by *para*-Bromo Functionalized Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

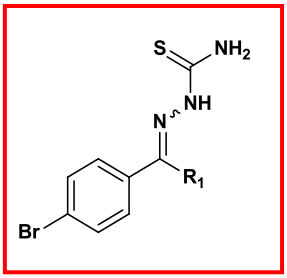
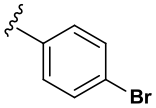
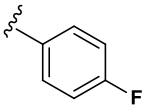
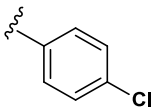
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **47** |  |  | ≥ 10000 |
| **48** | |  | 2086 ± 178.3 |
| **49** | |  | 4413. ± 333.2 |

Table 96. Inhibition of Human Cathepsin K by *para*-Bromo Functionalized Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

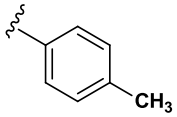
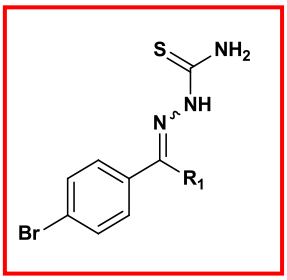
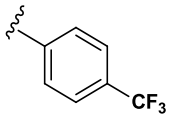
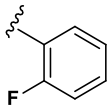
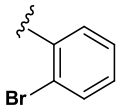
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **50** | |  | 1183 ± 81.4 |
| **51** |  |  | 146.5 ± 2.4 |
| **52** | |  | 368.6 ± 34.6 |
| **53** | |  | ≥ 10000 |

Table 97. Inhibition of Human Cathepsin K by Dihalogen-substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

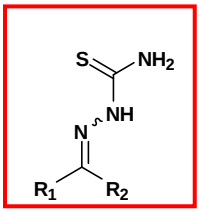
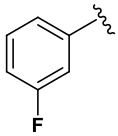
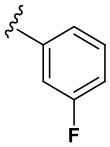
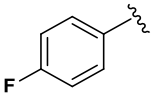
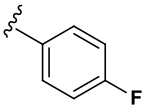
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **54** |  |  | 754.8 ± 22.2 |
| **55** |  |  | 2946 ± 261.1 |

Table 98. Inhibition of Human Cathepsin K by Dihalogen-substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

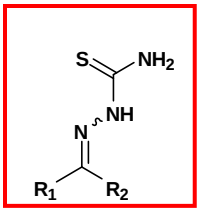
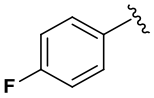
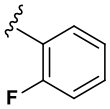
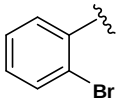
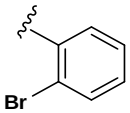
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **56** |  |  | 1004 ± 71.6 |
| **57** |  |  | 524.8 ± 31 |

Table 99. Inhibition of Human Cathepsin K by 3,3'-Dibromo-*N*-Substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

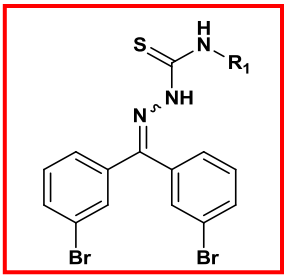
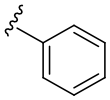
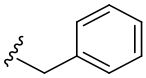
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **58** |  |  | ≥ 10000 |
| **59** | |  | ≥ 10000 |
| **60** | |  | ≥ 10000 |

Table 100. Inhibition of Human Cathepsin K by Thiosemicarbazones Containing a Phenyl Group. For Synthesis of Compounds: [3,6,7,10–12]

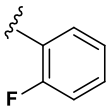
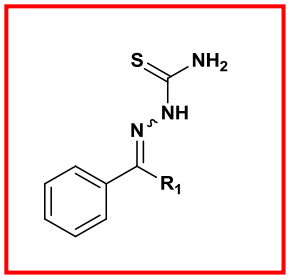
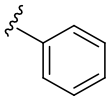
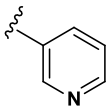
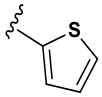
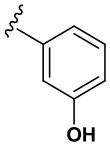
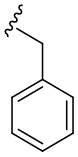
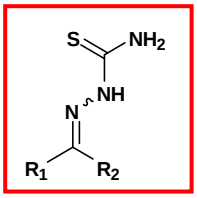
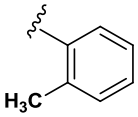
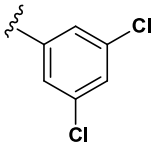
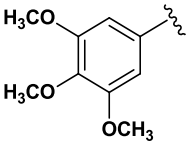
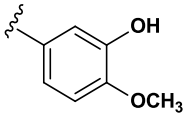
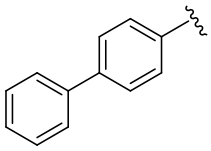
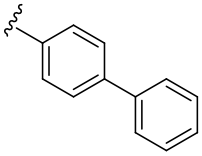
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **61** | |  | 3651 ± 44.6 |
| **62** |  |  | ≥ 10000 |
| **63** | |  | ≥ 10000 |
| **64** | |  | ≥ 10000 |
| **65** | |  | ≥ 10000 |
| **66** | |  | 5410 ± |

Table 101. Inhibition of Human Cathepsin K by Substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **73** |  |  | 57.3 ± 5.2 |
| **75** |  |  | 156.7 ± 14.7 |
| **76** |  |  | ND (Insoluble) |

*ND: Not Determined

Table 102. Inhibition of Human Cathepsin K by Functionalized Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

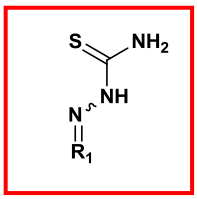
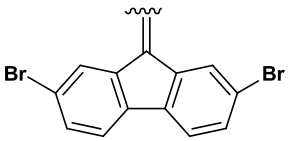
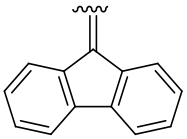
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **80** |  |  | ≥ 10000 |
| **81** | |  | ≥ 10000 |

Table 103. Inhibition of Human Cathepsin K by Functionalized Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

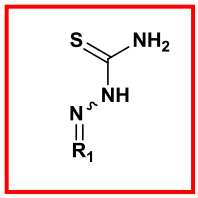
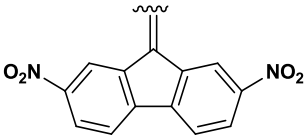
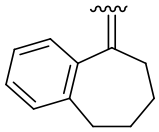
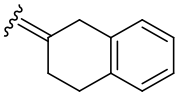
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **82** |  |  | ≥ 10000 |
| **85** | |  | ≥ 10000 |
| **90** | |  | ≥ 10000 |

Table 104. Inhibition of Human Cathepsin K by Substituted Quinolone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

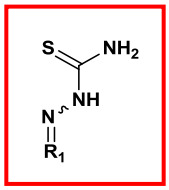
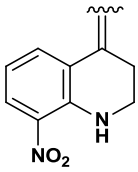
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **98** |  |  | ≥ 10000 |

Table 105. Inhibition of Human Cathepsin K by Non-Thiosemicarbazone Based Analogs. For Synthesis of Compounds: [3,6,7,10–12]

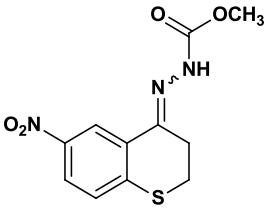
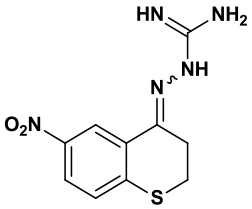
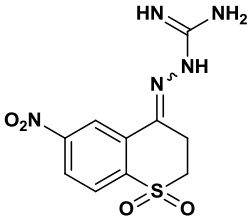
| *Compound* | *Structure* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|
| **101** |  | ≥ 10000 |
| **102** |  | ≥ 10000 |
| **103** |  | ≥ 10000 |

Table 106. Inhibition of Human Cathepsin K by Substituted Tetralone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

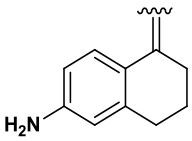
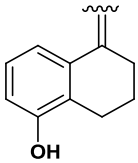
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **106** |  |  | ≥ 10000 |
| **109** | |  | ≥ 10000 |

Table 107. Inhibition of Human Cathepsin K by Functionalized Chromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

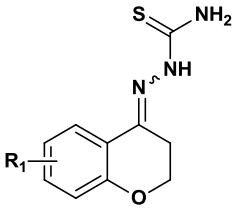
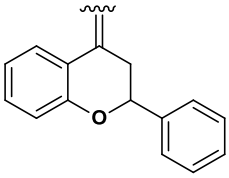
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **112** |  |  | ≥ 10000 |

Table 108. Inhibition of Human Cathepsin K by Substituted Thiochromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

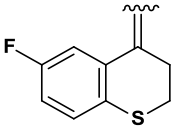
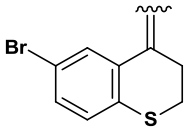
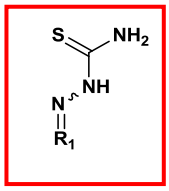
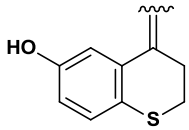
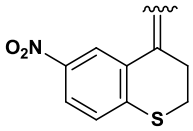
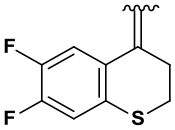
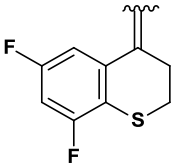
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **114** | |  | 1069 ± 104 |
| **115** | |  | 499.5 ± 34.0 |
| **117** |  |  | ≥ 10000 |
| **118** | |  | 164.7 ± 14.9 |
| **119** | |  | 176.6 ± 14.3 |
| **120** | |  | 1558 ± 156 |

Table 108. Inhibition of Human Cathepsin K by Substituted Thiochromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

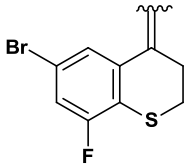
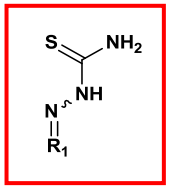
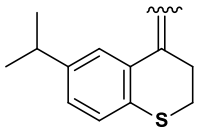
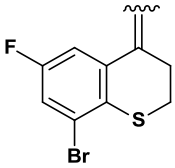
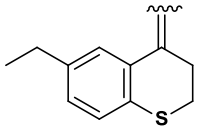
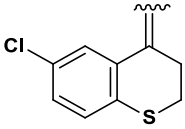
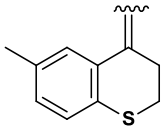
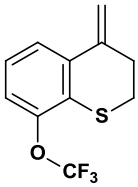
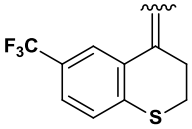
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **122** | |  | 44.2 ± 2.9 |
| **123** |  |  | 21.2 ± 1.9 |
| **124** | |  | 419.4 ± 21.7 |
| **125** | |  | 29.0 ± 1.8 |
| **126** | |  | 62.9 ± 2.3 |
| **127** | |  | 80.8 ± 2.6 |
| **128** | |  | ≥ 10000 |
| **129** | |  | 21.5 ±2.1 |

Table 108. Inhibition of Human Cathepsin K by Substituted Thiochromanone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

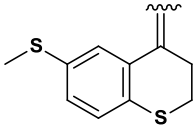
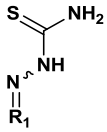
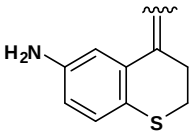
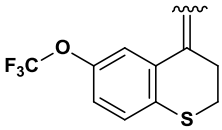
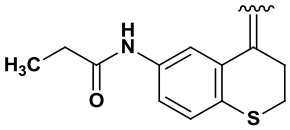
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **130** | |  | 56.7 ± 4.2 |
| **132** |  |  | ≥ 10000 |
| **133** | |  | 179.4 ± 10.7 |
| **134** | |  | ≥ 10000 |

Table 109. Inhibition of Human Cathepsin K by Substituted Sulfone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

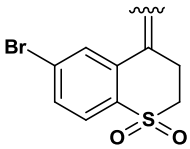
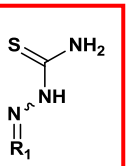
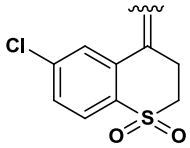
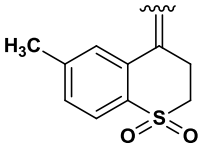
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **136** | |  | 149.9 ± 7.9 |
| **138** |  |  | 1441 ± 48.8 |
| **139** | |  | 2981 ± 41 |

Table 110. Inhibition of Human Cathepsin K by Substituted Sulfone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

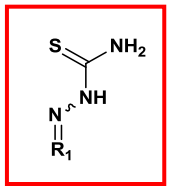
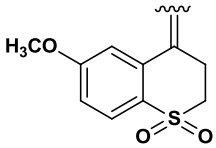
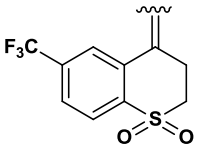
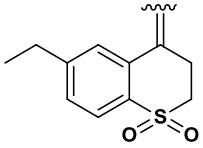
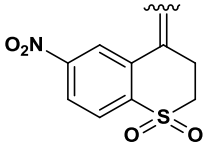
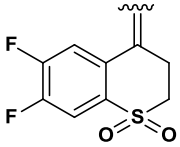
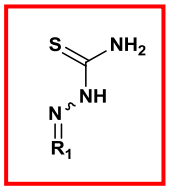
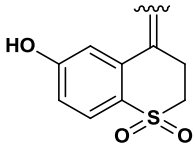
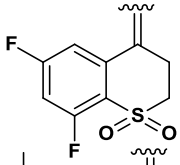
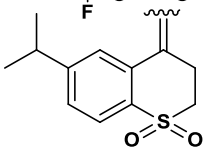
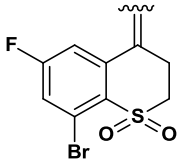
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **140** |  |  | ≥ 10000 |
| **141** | |  | 204.2 ± 2.6 |
| **142** | |  | 941.7 ± 70.0 |
| **143** | |  | ≥ 10000 |
| **144** | |  | 1273 ± 13.3 |
| **145** |  |  | ≥ 10000 |
| **146** | |  | ≥ 10000 |
| **147** | |  | 110.9 ± 11.6 |
| **148** | |  | ≥ 10000 |

Table 109. Inhibition of Human Cathepsin K by Substituted Sulfone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

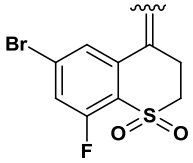
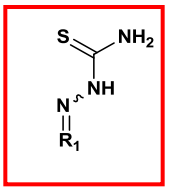
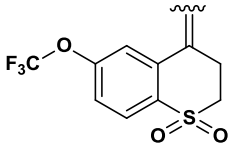
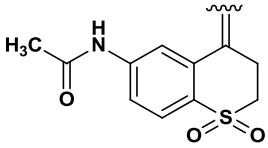
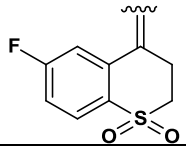
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **149** | |  | 649.6 ± 37.3 |
| **150** |  |  | 471.4 ± 14.6 |
| **151** | |  | ≥ 10000 |
| **152** | |  | 5918 ± 415.7 |

Table 111. Inhibition of Human Cathepsin K by Functionalized Benzoyl-Benzophenone Thiosemicarbazones*

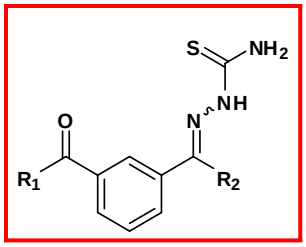
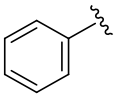
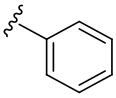
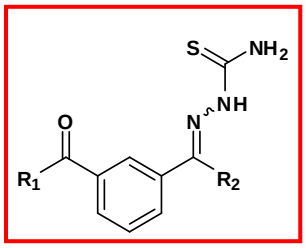
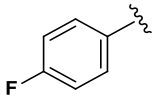
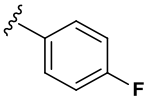
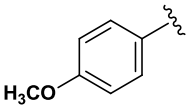
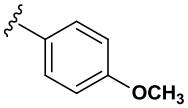
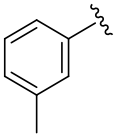
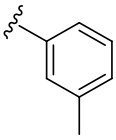
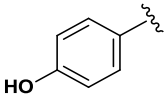
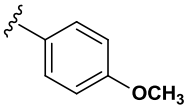
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **156** |  |  | 17.4 ± 1.4 |

Table 112. Inhibition of Human Cathepsin K by Functionalized Benzoyl-Benzophenone Thiosemicarbazones (Continued)

| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **157** |  |  | 8500 ± 663.7 |
| **158** |  |  | 2092 ± 17.5 |
| **160** |  |  | 1034 ± 68.2 |
| **161** |  |  | 6162 ± 471 |

*Compounds **156-162** have been synthesized by members of the Kevin G. Pinney laboratory (Baylor University)

Table 113. Inhibition of Human Cathepsin K by Substituted-Benzoyl-Benzophenone Thiosemicarbazones.* For Synthesis of Compounds: [3,6,7,10–12]

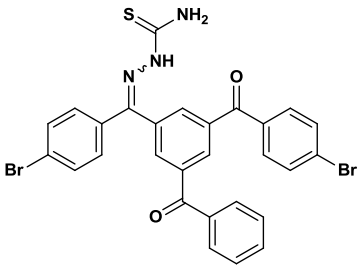
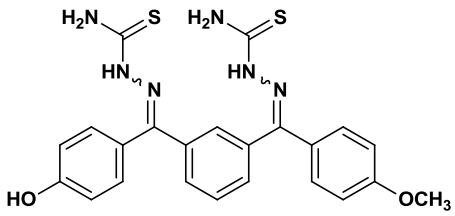
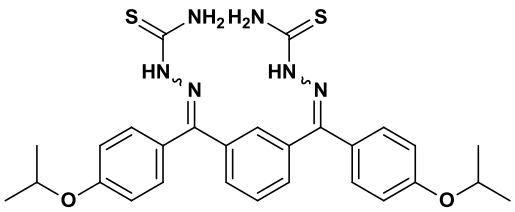
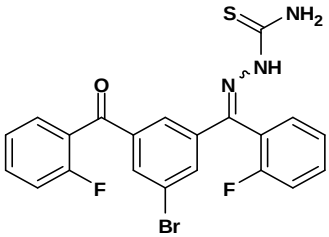
| *Compound* | *Structure* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|
| **165** |  | 2796 ± 19.8 |

Table 114. Inhibition of Human Cathepsin K by Substituted-Benzoyl-Benzophenone Thiosemicarbazones.* For Synthesis of Compounds: [3,6,7,10–12]

| *Compound* | *Structure* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|
| **166** |  | 3204 ± 356 |
| **167** |  | ≥ 10000 |
| **168** |  | 48.9 ± 3.2 |

*Compounds **163**-**168** have been synthesized by members of the Kevin G. Pinney laboratory (Baylor University)

*Structure-Activity Relationship (SAR) of Thiosemicarbazones as Cathepsin K Inhibitors*

Almost seventy-five percent of the library that was screened for cathepsin L was evaluated using cathepsin K as a target. A total of 120 thiosemicarbazone analogs were screened using activated cathepsin K during the preliminary screening in order to determine the potency of these compounds. A selected group of four different

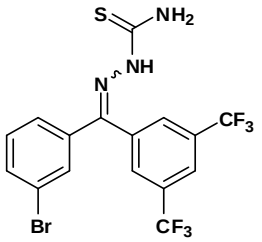
compounds showed to be excellent cathepsin K inhibitors with $IC_{50}$ between 10 and 20 nM. Eleven compounds demonstrated to be good inhibitors ($20 \leq IC_{50} \leq 100$ nM).

Thirty percent of the samples did not show a significant inhibitory potency and a general $IC_{50}$ value greater than 10,000 nM was assigned. On the other hand, compounds that showed inhibitory activity greater than 50% were considered to calculate an $IC_{50}$ value that was expected to be less than 10,000 nM. One fourth of the library (31 analogs) had a modest activity. $IC_{50}$ values of these synthetic compounds were ranged between 100 and 1000 nM. Interestengly, the entire sub-library of functionalized *meta*-bromobenzaldeyde thiosemicarbazones (Table 94, page 258) showed inhibitory activity toward cathepsin K ($IC_{50} < 6.5$ µM).

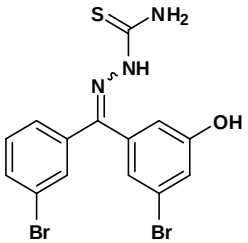
*Potent cathepsin K inhibitors*. Finally, four thiosemicarbazones showed very potent inhibitory activities against this cathepsin K (CK). Figure 93 summarizes the structures of the compounds with $IC_{50}$ values less than 20 nM. Three analogs belong to the subfamily of *meta*brominated benzophenone TSC analogs. Ditrifluoromethyl, and two brominated phenols, all halogenated moieties are the substituents in this subfamily. The fourth compound consists of the unsubstituted benzoyl benzophenone thiosemicarbazone. The $IC_{50}$ values obtained using cathepsin L (CL) are also included in this section for comparison purposes between these two cysteine proteases. This selected group also showed good inhibitory activity towards cathepsin L. Compounds **13**, **34** and **31** were better cathepsin K inhibitors when compared to their respective cathepsin L activities (i.e. up to 15-fold more potent in the case of **34**). Interestingly, **156,** the second most potent inhibitor of cathepsin L in the library, is also one of the most potent inhibitors of cathepsin K with an $IC_{50}$ value less than 20 nM.

*Compounds with $IC_{50}$ less than 100 nM*. Eleven compounds proved to have good inhibitory activity (less than 100 nM). This list included: 3-3´-dibromobenzophenone TSC (**1**), 3-bromo-3´-trifluoromethylbenzophenone TSC (**6**), 3-bromo-3´-hydroxybenzophenone TSC (**8**), 3-bromo-3´-acetobenzophenone TSC (**9**), 2-methyl-3´,5´-dichlorobenzophenone TSC (**73**), 6-isopropylsulfide TSC (**123**), 6-ethylsulfide TSC (**125**), 6-trifluoromethylsulfide TSC (**129**).

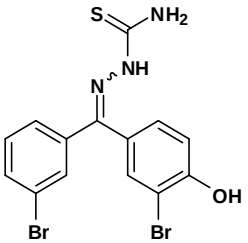
*Compounds with modest activity*. Thirty-one TSC analogs showed $IC_{50}$ values between 100 and 1000 nM. The list included, but not limited: *meta*brominated benzophenone (**2**, **4-5)** TSCs, sulfide (**115**, **118**, **124**), and sulfone (**149**, **150**) TSCs



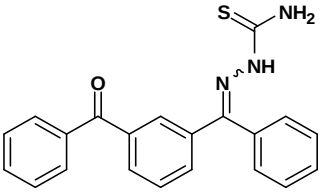
(**13**) CK: 11.3 nM
CL: 96.0 nM



(**34**) CK: 15.6 nM
CL: 232.4 nM



(**31**) CK: 10.2 nM
CL: 126.1 nM



(**156**) CK: 17.4 nM
CL: 9.85 nM

Figure 93. Thiosemicarbazone Analogs with Potent Inhibitory Activity against Human Cathepsin K[9–12]

*General Remarks of the Structure-Activity Relationship*

Further analysis of the structure-activity relationship revealed that thiosemicarbazone inhibitory activity can be enhance or reduced dramatically by the substituting numerous moieties. The series of substituted *meta*-brominated benzophenone thiosemicarbazones is the largest subgroup among the compounds that were tested. Halogenated substituents greatly enhanced the activity of the thiosemicarbazones. Compounds **1**, **2**, **4**, **5** (Table 115) were compounds sharing the main moiety but varying the *meta* halogenated substituent. Compound **2** was used as a comparison to complete the series. The unsubstituted analog **2** showed modest activity compared to **1** the brominated analog with one of the best activities in the entire library. Once again, **1** was the best halogenated analog of the series with these types of substituents. The brominated analog was 3-fold more active than its fluorinated analog. The activity of these compounds between cathepsins K and L showed that **1** is the most potent inhibitor in both cases. Analog **2** was 15-fold more potent against cathepsin K when compared to cathepsin L inhibitory activity.

Five fluorosubstituted benzophenone TSC analogs (**4**, **10**, **22**, **23**, **25**, Table 116) were also compared. Fluorination of the second phenyl rings indicated that cathepsin K-inhibitory activity of the compounds can be increased with the addition of electronegative substituents, similar toin the cathepsin L inhibitory activity. However, a deeper comparison between the two difluoro benzophenone TSC also revealed the position of the substitution plays a key role in the activity of these analogs. The di*ortho*-fluoro analog is 4 times less active than the di*meta*-counterpart. The activity of these inhibitors is greatly enhaced by the presence of bromine in the *meta* position of the second ring

(**28**). 3-bromo-2',3'-difluorobenzophenone thiosemicarbazone (**28**) is 2-fold more active than also 3-3'-difluorobenzophenone thiosemicarbazone (**27**).

Table 115. Inhibition of Human Cathepsin K by 3-Bromo-3'-Halogen Benzophenone Thiosemicarbazones

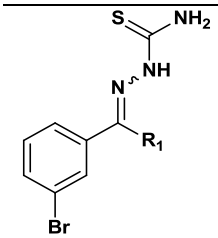
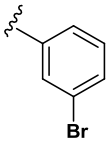
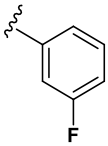
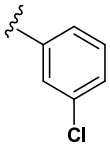
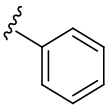
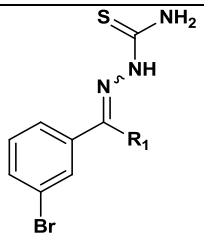
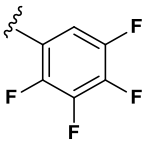
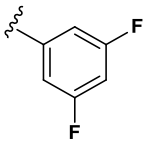
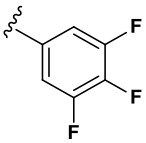
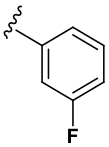
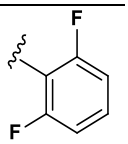
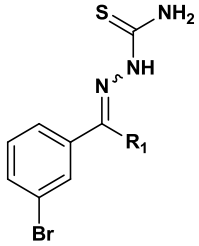
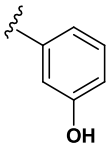
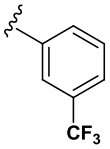
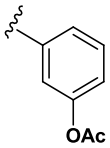
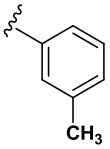
| | *R1* | | | |
|---|---|---|---|---|
|  |  < |  < |  < |  |
| $IC_{50}$ (nM) | **1** | **4** | **5** | **2** |
| CK | 35.2 | 104.2 | 259 | 660.4 |
| CL | 16.7 | 250.3 | 131.4 | ≥ 10000 |

Table 116. Inhibition of Human Cathepsin K by 3-Bromo-Fluorinated-Benzophenone Thiosemicarbazones

| |  | | | | |
|---|---|---|---|---|---|
| *R1* |  > |  > |  > |  > |  |
| $IC_{50}$ (nM) | **23** | **10** | **22** | **4** | **25** |
| CK | 22.6 | 26.0 | 32.6 | 104.2 | 107.0 |
| CL | 63.2 | 59.4 | 118.1 | 250.3 | 609.7 |

Various groups were substituted at the *meta* position of the second phenyl rings (Table 117). The groups include trifluoromethyl, methyl, hydroxyl, and acetate moieties. The best compound is the phenolic analog (**8**).

Table 117. Inhibition of Human Cathepsin K by 3-Bromo-3'-Heteroatomic Groups Benzophenone Thiosemicarbazones

| | *R1* | | | |
|---|---|---|---|---|
|  |  | <  | <  | <  |
| $IC_{50}$ (nM) | **8** | **6** | **9** | **7** |
| CK | 53.3 | 65.3 | 98.9 | 133.3 |
| CL | 131.4 | 46.5 | 150.8 | 224.4 |

A small trifluoromethyl series of three compounds (**6**, **11**, and **13**, Table 118) also revealed that substituted compounds often demonstrate better activity against cathepsin K when compared to monosubstituted 3-bromobenzophenone thiosemicarbazones. In fact, analog **13** is the second most potent inhibitor of cathepsin K with an outstanding $IC_{50}$ value of 11.3 nM.

In the *ortho*-substituted series, the inhibitory activity of the compounds increases with more electronegative substituents. Additionally, the exceptional case of **21**, 3-4'-dibromobenzophenone thiosemicarbazone is remarkable when comparing its activity toward cathepsins L and K. This analogs shows a remarkable selectivity toward cathepsin K (Table 119).

Finally, the subfamily of substituted *meta*-brominated thiosemicarbazones was completed with unbrominated and brominated pyridines and thiophenes. Pyridine TSCs demonstrated to have good inhibitory activity while unbrominated pyridines showed no activity against cathepsin K. These small series also demonstrated that benzophenone group plays a key role in the activity of TSCs (Table 120).

Table 118. Inhibition of Human Cathepsin K by 3-Bromo-Trifluoromethyl Benzophenone Thiosemicarbazones

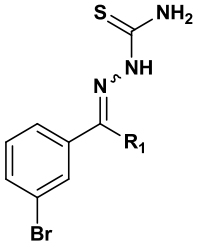
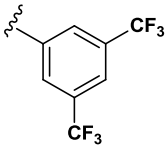
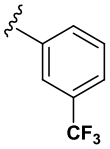
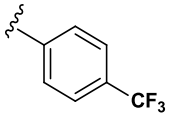
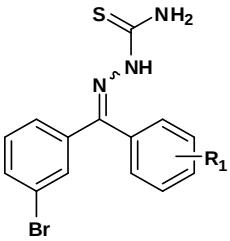
| | *R1* | | |
|---|---|---|---|
|  |  < |  < |  |
| $IC_{50}$ (nM) | **13** | **6** | **11** |
| CK | 11.3 | 65.3 | ND |
| CL | 96.0 | 46.5 | 520.9 |

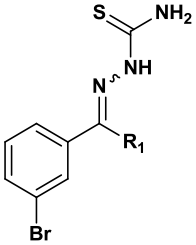
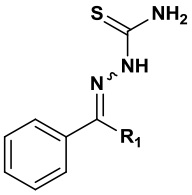
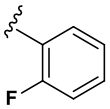
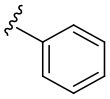
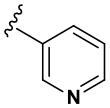
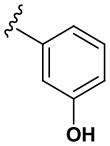
Table 119. Effect of the position of substituents in the inhibitory activity of 3-Bromo-Benzophenone Thiosemicarbazones



| $R_1$ | $IC_{50}$ (nM) *ortho* | | | *meta* | | | *para* | | |
|---|---|---|---|---|---|---|---|---|---|
| -F | **14** | CK | 61.8 | **4** | CK | 104.2 | **17** | CK | 65.6 |
| | | CL | 23.9 | | CL | 250.3 | | CL | 79.6 |
| -Cl | **15** | CK | 312.9 | **5** | CK | 259.0 | **18** | CK | 626.4 |
| | | CL | 1610 | | CL | 131.4 | | CL | 327.1 |
| -Br | **24** | CK | 1300 | **1** | CK | 35.2 | **21** | CK | 49.1 |
| | | CL | 2600 | | CL | 16.7 | | CL | ≥ 10000 |
| $CH_3$ | **20** | CK | 838.9 | **7** | CK | 133.3 | **19** | CK | ND |
| | | CL | ≥ 10000 | | CL | 224.4 | | CL | 2156 |

Another interesting comparison is between *meta* and *para* brominated benzophenones (Table 121) with different halogenated and aliphatic substituents. It is evident that *meta* bromination offers the best activity and a change in this position reduces drastically the potency of any substitutent. For example, 3,4´-dibromobenzophenone thiosemicarbazone is more potent than its *para* brominated counterpart, 4,4´-dibromobenzophenone that showed no activity towards cathepsin K.

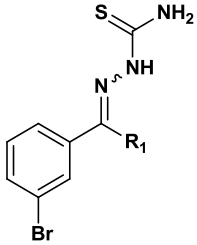
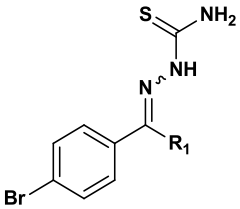
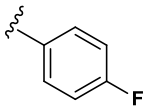
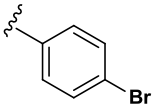
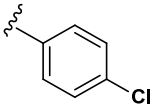
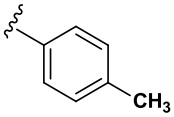
Table 120. Comparison between Brominated and Unbrominated Functionalized Benzaldehyde Thiosemicarbazones for Cathepsins K and L

| $R_1$ | |  | | |  | |
|---|---|---|---|---|---|---|
| | | | $IC_{50}$ (nM) | | | |
|  | **14** | CK | 61.8 | **61** | CK | 3651 |
| | | CL | 23.8 | | CL | ≥ 10000 |
|  | **2** | CK | 660.4 | **62** | CK | ≥ 10000 |
| | | CL | ≥ 10000 | | CL | ≥ 10000 |
|  | **29** | CK | 66.7 | **63** | CK | ≥ 10000 |
| | | CL | 1000 | | CL | ≥ 10000 |
|  | **8** | CK | 53.3 | **65** | CK | ≥ 10000 |
| | | CL | 188.7 | | CL | ≥ 10000 |

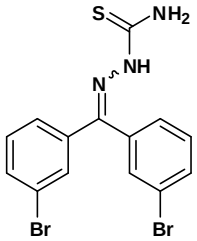
Bromination of the benzophenone group is critical for the activity of TSC derivatives as potential cathepsin K inhibitors. An examination of halogenated compounds with similar chemical structures reveals that a different halogen substituent, such as fluorine, greatly reduces the activity of the compounds. The 3,3´-difluorobenzophenone TSC is almost 22 times less active than 3,3´-dibromobenzophenone TSC. Interestingly, 4,4´-difluorobenzophenone TSC was a better inhibitor than 4,4´-dibromobenzophenone TSC that was inactive.

Compounds **61-65** also demonstrated the importance of the presence of at least one bromine in one of the phenyl rings in order to enhance the potency of the thiosemicarbazone derivatives. In this series, only the fluorinated analog of the benzophenone thiosemicarbazone **61** showed inhibitory activity.

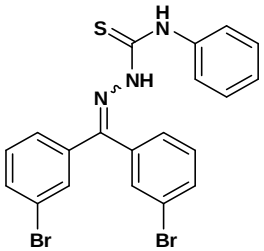
Table 121. Comparison of the Inhibitory Activity between *meta* and *para*Bromination of Benzophenone Thiosemicarbazones towards Cathepsins K and L

| $R_1$ |  | | |  | | |
|---|---|---|---|---|---|---|
| | | | IC$_{50}$ (nM) | | | |
|  | **17** | CK | 65.6 | **48** | CK | 2086 |
| | | CL | 79.6 | | CL | 3320 |
|  | **21** | CK | 49.1 | **47** | CK | ≥ 10000 |
| | | CL | 327.1 | | CL | ≥ 10000 |
|  | **18** | CK | 626.4 | **49** | CK | 4413 |
| | | CL | ≥ 10000 | | CL | ≥ 10000 |
|  | **19** | CK | ND | **50** | CK | 1183 |
| | | CL | 2156 | | CL | 4570 |

The TSC group is a key feature for a successful inhibition of cathepsin K. Compound **1**, the reference compound, is a potent inhibitor with an IC$_{50}$ value of 35.2 nM, while analogs of the thiosemicarbazone derivative were inactive (**58**-**60**, Table 99)**.**



(**1**) IC$_{50}$: 35.2 nM



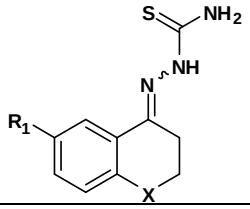
(**58**) IC$_{50}$ ≥ 10000

Figure 94. Comparison between Inhibitory Activities of **1** and **58**

The compounds with different scaffolds were examined (Table 102). The series included fluorenes, naphthalenes, indenes, annulenes, chromanones, aminoquinolines, among others. These compounds were inactive.

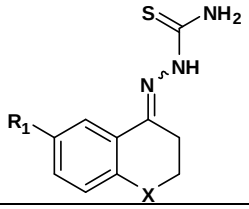
Thiochromanones and sulfones are important subclasses of thiosemicarbazones that were screened in this library. Their $IC_{50}$ values ranged between 21.2 and 10000 nM. In general, thiochromanones showed better inhibitory activities when compared with sulfones. The majority of thiochromanone were monosubstituted in the C6 position. The list of substituents includes halogen, hydroxyl, aliphatic, nitro, and acetate groups. Perhaps the most interesting observation is that aliphatic substituted analogs presented the best inhibitory activity (**123**, **125**, and **127**) with $IC_{50}$ values less than 100 nM. The potency is enhanced when the substituent is branched in both cases. However, thiochromanones are up to 37 times more potent than sulfones (Table 122). Another selective inhibitor of cathepsin K was found in this series. Compound **123** showed a good $IC_{50}$ values of 21 nM, while it was inactive as a cathepsin L inhibitor (See Table 122).

Table 122. Comparison of the Activity of Aliphatic Thiochromanone Versus Aliphatic Sulfone TSCs against Cathepsins K and L

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | $R_1$ | | | | | |
| $IC_{50}$ (nM) | *X* | | i-propyl | | | ethyl | | | methyl | | |
|  | S | **123** | CK | 21.2 | **125** | CK | 29 | **127** | CK | 81 |
| | | | CL | ≥ 10000 | | CL | 2720 | | CL | 214 |
| | $SO_2$ | **147** | CK | 111 | **142** | CK | 942 | **139** | CK | 2981 |
| | | | CL | ≥ 10000 | | CL | 6521 | | CL | ≥ 10000 |

Halogenated compounds do not follow a defined trend (Table 123). The activity in the halogenated sulfones decreases with more electronegative substituents at the C6 position; but this does not apply to the thiochromanone analogs. The best compound of this small series is the 6-chlorothiochromanone thiosemicarbazone with an $IC_{50}$ value of 63 nM.

Table 123. Comparison of the Activity of Halogenated Thiochromanone vs Halogenated Sulfone TSCs against Cathepsins K and L

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | *$R_1$* | | | | |
| $IC_{50}$ (nM) | *X* | | -F | | | -Cl | | | -Br | |
|  | S | **114** | CK | 1069 | **126** | CK | 62.9 | **115** | CK | 499.5 |
| | | | CL | 742 | | CL | 228 | | CL | 152 |
| | $SO_2$ | **152** | CK | 5918 | **138** | CK | 1441 | **136** | CK | 150 |
| | | | CL | ≥10000 | | CL | ND | | CL | 574 |

Substituted analogs with nitro and hydroxyl groups also were tested. Three out of four compounds showed no inhibition towards cathepsin K. The 6-nitrothiochromanone thiosemicarbazone was the only compound with an $IC_{50}$ value (164.7 nM, Table 124). Compound **118** also showed activity toward cathepsin L when compared to cathepsin K inhibition (2.5-fold).

Interestengly, the electrophilic trifluoromethyl thiochromanone and sulfone thiosemicarbazones (**129**, Table 125) showed remarkable activity against cathepsin K with activities as low as 21.5 nM.

Table 124. Comparison of the Activity of Nitro and Hydroxyl-Susbtituted Thiochromanone Versus Substituted Sulfone TSCs against Cathepsins K and L

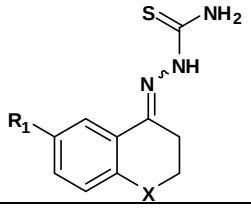
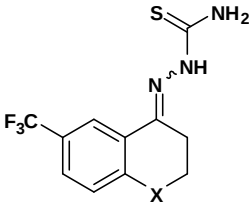
| | | | | $R_1$ | | | |
|---|---|---|---|---|---|---|---|
| $IC_{50}$ (nM) | *X* | | $-NO_2$ | | | -OH | |
|  | S | **118** | CK | 164.7 | **117** | CK | ≥10000 |
| | | | CL | 67.85 | | CL | ≥10000 |
| | $SO_2$ | **143** | CK | ≥10000 | **145** | CK | ≥10000 |
| | | | CL | 112 | | CL | ≥10000 |

Table 125. Comparison of the Activity of Trifluoromethyl Thiochromanone Versus Trifluoromethyl Sulfone TSCs against Cathepsins K and L

| | X | | |
|---|---|---|---|
|  | S | < | $SO_2$ |
| | **129** | | **141** |
| $IC_{50}$ (nM) | | | |
| CK | 21.5 | | 204.2 |
| CL | 284.1 | | 259.8 |

The position of the fluoro substituents in both thiochromanones and sulfones were also examined (Table 126). In general, dihalogenation did not improve the activity of thiochromanone. However, the 6,7-difluoro analogs (**119, 144**) were more potent than their 6,8-difluoro analogs (**120**, **146**).

The effect of substitution using different halogen substituents was also assessed by examining halogenated isomers (Table 127). Interestingly, inhibitor activity increases when bromine is at C6, confirming what was observed with monohalogenated thiochromanones and sulfones, where brominated compounds were more active than their

fluorinated counterparts. Compound **149** is 15-fold more active than its isomer analog **148**.

Table 126. Comparison of the Activity of the Effect of the Position of Fluorosubstituents in Thiochromanone and Sulfone TSCs against cathepsins K and L

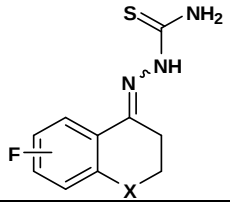
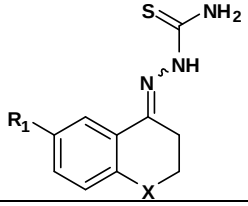
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | *F* | | | | |
| $IC_{50}$ (nM) | *X* | | 6,7-$F_2$ | | | 6,8-$F_2$ | |
|  | S | **119** | CK | 177 | **120** | CK | 1558 |
| | | | CL | 54 | | CL | 1500 |
| | $SO_2$ | **144** | CK | 1273 | **146** | CK | ≥10000 |
| | | | CL | 3650 | | CL | ≥10000 |

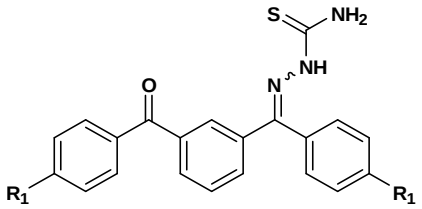
Table 127. Comparison of Halogenated Thiochromanone Isomers Versus Halogenated Sulfone Isomers against Cathepsins K and L

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | *$R_1$* | | | | |
| $IC_{50}$ (nM) | *X* | | 6-fluoro-8-bromo | | | 6-bromo-8-fluoro | |
|  | S | **124** | CK | 419.2 | **122** | CK | 44.2 |
| | | | CL | ≥ 10000 | | CL | 434.2 |
| | $SO_2$ | **148** | CK | ≥ 10000 | **149** | CK | 649.6 |
| | | | CL | ≥ 10000 | | CL | 1117 |

Lastly, the recent discovery of a new potent thiosemicarbazone derivative scaffold with an $IC_{50}$ value of 9.85 nM (Compound **156**) for cathepsin L, led to the synthesis of a new generation of thiosemicarbazones that were also tested with cathepsin K (unpublished results). Chemically, compound **156** is a benzoyl benzophenone thiosemicarbazone. The small subclass of TSC that was screened is composed of eight analogs with similar structures.

Unsubstituted and *para*-substituted analogs were also examined (Table 128). The unsubstituted analog showed an extraordinary potency with an $IC_{50}$ less than 20 nM ($IC_{50}$: 17.4 nM). Substituted analogs did not prove to have comparable activities relative to the unsubstituted benzoyl benzophenone thiosemicarbazones. $IC_{50}$ values varied between 1000 and 10000 nM. The most important feature of these compounds is their improved selectivity toward cathepsin L compared to their cathepsin K activity. Compound **157** is 340-fold more active toward cathepsin L. This compound could be used a potentially good cathepsin L inhibitor with improved selectivity.

Table 128. Inhibition of Human Cathepsins K and L by Difunctionalized Benzoyl Benzophenone Thiosemicarbazones




| | $R_1$ | -H **(156)** | -F **(157)** | $-OCH_3$ **(158)** |
|---|---|---|---|---|
| $IC_{50}$ (nM) | CK | 17.4 | 8500 | 2092 |
| | CL | 9.9 | 24.3 | 587 |

*Advanced Kinetic Studies*

Previous results shown in Chapter Two indicated the 3-3'-dibromobenzophenone thiosemicarbazone (**1)** was a slow-binding, time-dependent, reversible, competitive inhibitor of cathepsin L. It also was able to arrest cell migration and cell invasion of MDA-MB-231 cells. These promising results led to the additional research using cathepsin K as a molecular target in order to characterize its mode of inhibition. Fluorometric-based assays were used to investigate and characterize compound **1.** The

possibility of the inhibition of the activation of procathepsin K was also explored by western blotting techniques.

### *Kinetic Analysis of 3-Bromo-3'-Bromobenzophenone Thiosemicarbazones (**1**) as Cathepsin K Inhibitor*

#### *Effect of Inhibitor Concentration on Cathepsin K Progress Curves*

Cathepsin K (1.5 nM) was added to four reactions with different concentration of the inhibitor, ranging from 0 to 10 μM of **1**. Reactions were initiated by the rapid addition of the substrate (Z-FR-AMC, final concentration: 50 μM). The release of AMC from the nonfluorogenic substrate was monitored every 3 seconds for fifty minutes. Figure 95 shows the effect of inhibitor concentration on cathepsin K progress curve reactions. The time dependence inhibition of cathepsin K by **1** was more easily observed at lower inhibitor concentrations (100 nM), because enzymatic activity was completely inhibited at higher concentrations of **1** and the AMC release was not observed. Data sets for 100 nM were fitted to equation 1.7, by nonlinear regression analysis using GraphPad 5.0. P is the concentration of product (μM), $v_i$ and $v_s$ are the initial and steady-state velocities (μM/s), t is the time in seconds and $k_{obs}$ is the rate constant for conversion of the initial velocity $v_o$ to the steady state velocity vs. $k_{obs}$ units are given in $s^{-1}$. The equation was entered into the computational software (GraphPad 5.0), knowing that P and t are the dependent and independent variables, while keeping the velocities and the rate constant as unknown. For each case, the constraints for their calculation were to give a possible positive value (i.e. $k_{obs} \geq 0$). It is also worth noting that equation 1.7 is only valid when substrate depletion is insignificant. Therefore, some points were excluded in

every case for data fitting. Velocities, rates, $r^2$, and points analyzed for 50 μM Z-FR-AMC are shown in Table 129.

$$P = v_s t \frac{(v_o - v_{s)}}{k_{obs}} \left(1 - e^{-k_{obs} t}\right) \tag{1.7}$$

Table 129. Calculated Kinetic Parameters from Eq.1.7 for Cathepsin K Progress Curves with 50 μM Z-FR-AMC

| [I] (μM) | 0.1 |
|---|---|
| $v_s$ (μM/s) | 0.0005618 |
| $v_i$ (μM/s) | 0.5440 |
| $k_{obs}$ ($s^{-1}$) | 0.6773 |
| $r^2$ | 0.9102 |
| Points analyzed | 336 |

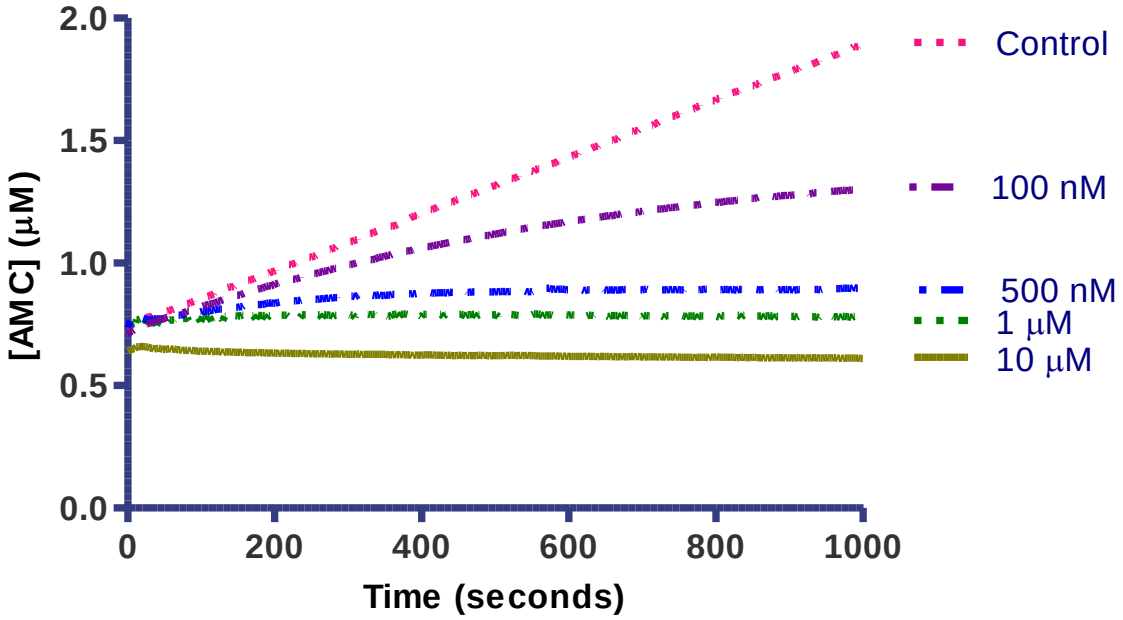



Figure 95. Cathepsin K Progress Curves with **1** Using 50 μM Z-FR-AMC

Data sets were fitted into equation 1.7. Only one concentration (0.1 μM) was able to be resolved. The results indicate **1** is a time-dependent inhibitor of cathepsin K. Additional studies will be required to confirm these results.

*Effect of Preincubation Studies on Cathepsin K Inhibition Assays using **1***

Experimental results using cathepsin K progress curves using **1** provided strong evidence that **1** is a time-dependent inhibitor. To confirm the time-dependende of **1**, the influence of preincubation time on the inhibitory potency of thiosemicarbazones and its $IC_{50}$ values was studied. Compound **1** $IC_{50}$ values were determined at 6 different preincubation times ranging between 0 and 180 minutes. Substrate concentration was 50 μM. Inhibitor final concentration varied between 0 and 10 μM. Table 130 and Figure 96 show the results of the studies.

Table 130. Effect of Pre-Incubation Times on $IC_{50}$ Values of **1** against Cathepsin K

| *Pre-incubation times (min)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 0 | 529.3 ± 44.2 |
| 5 | 35.2 ± 2.4 |
| 15 | 13.9 ± .5 |
| 45 | 5.1 ± 0.3 |
| 90 | 3.3 ± 0.5 |
| 180 | 1.6 ± 0.2 |

The effect of preincubation on the potency of **1** as seen in Table 130 is quite evident. The dibromobenzophenone thiosemicarbazone **1** had an $IC_{50}$ of 529 nM when cathepsin K was not preincubated with the compound. Compound **1** improved its activity 15-fold with only five minutes preincubation time. The potency of **1** increased with longer preincubation times. Finally, the lowest $IC_{50}$ value was determined when **1** was preincubated for forty-five minutes with cathepsin K. These results confirmed the strong

dependence of $IC_{50}$ value determination with respect to the preincubation time parameter. Controls (i.e. uninhibited cathepsin K) were monitored at every preincubation time. There was no significant loss of catalytic activity between 0 and 45 minutes. On the other hand, the activity of the enzyme decreases significantly at longer preincubation times (90 and 180 minutes), confirming the fragility of this enzyme. Maximum activity was found at 45 minutes. However, its activity was reduced by 50% and 75% for 90 and 180 minutes, respectively. Thus, the $IC_{50}$ values do not represent an accurate measurement of the inhibitory activity of **1**.

A general comparison between the results shown in Figure 90 and Figure 96 showed a discrepancy in results. It was previously stated activated cathepsin K is extremely unstable and lost 30% of its activity within the first ten minutes of activation. However, the pH of the activation was under acidic conditions ([NaOAc]: 5.5 mM; pH: 3.5) while preincubation studies where activated cathepsin K is kept for longer preincubation times are at pH 5.5 with inhibitor and detergent ([NaOAc]: 5.5 mM; pH: 3.5). Similar findings have also been reported. Lecaille and cowokers found that the activity of active cathepsin K is greatly enhanced under slightly acidic conditions ($5 \leq pH \leq 7$).[359] Furthermore, the effect of DMSO in the stability of the enzyme could not be ruled out. Cathepsin K was preincubated for longer periods in the presence of the aprotic solvent ([DMSO]: 2.8%) which could have also helped in the stability of the protease.

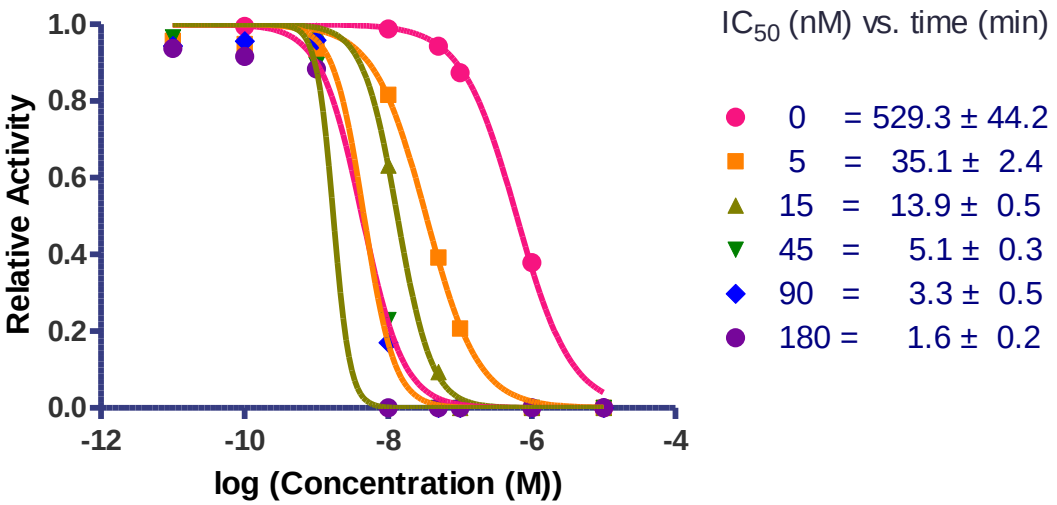



Figure 96. Effect of Preincubation Time on $IC_{50}$ Values of **1** against Cathepsin K

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation. Effect of Preincubation Time using **1***

The data obtained from the preincubation were further analyzed. The possibility that **1** was a tight-binding inhibitor was analyzed with the Morrison's quadratic equation which describes this behavior. (See equation 1.11).

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \quad (1.11)$$

The rates $v_i$ and $v_o$ are the inhibited and uninhibited cathepsin K velocities (RFU/s); $[\mathrm{E}]_T$ (nM), the total concentration of enzyme found in solution (free enzyme and inhibitor-enzyme complex); $[\mathrm{I}]_T$ (nM) is the total concentration of inhibitor present in solution (free inhibitor and inhibitor-enzyme complex) and $K_I$ (nM) is the inhibition constant, often referred as the dissociation constant. The equation may be solved to give two possible answers. However, the equation is written so that only there is an only

possible answer that fits physiological conditions (i.e. $K_I^{app} > 0$). GraphPad 5.0 was used to fit the data by non-linear regression after data manipulation for every preincubation time point. Inhibitor, substrate, and enzyme concentrations, as well as $K_M$, and, $K_I^{app}$, were all in micromolar units (μM). The residual activity (or $v_i/v_o$) was normalized to 1 (i.e. $v_o$: 1 and $0 \leq v_i \leq 1$). Normalized residual activity and [I] were defined as the dependent and independent variables respectively. Nonlinear regression was applied using the following conditions: [S]: 50 μM, $K_M$: 11.8 μM, $[E]_T$: 0.0015 μM and $v_o$: 1. Normalization and nonlinear conditions were used for three or four independent experiments per preincubation time. Average and standard errors can be seen in Table 131 and Figure 97.

Table 131. Effect of Preincubation Time in $K_I^{app}$ Values of **1** against Cathepsin K

| *Pre-incubation times (min)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 0 | 103.7 ± 12.8 |
| 5 | 6.5 ± 0.2 |
| 15 | 2.2 ± 0.1 |
| 45 | 0.8 ± 0.02 |
| 90 | 0.5 ± 0.05 |
| 180 | 0.3 ± 0.01 |

The data also show the apparent inhibition dissociation constants ($K_I^{app}$) are also time-dependent. No preincubation allows having a good inhibition constant with a value of 104 nM. Furthermore, the potency is extremely enhanced (more than 15-fold) for the calculated value at standard preincubation time (5 minutes). The apparent $K_I$ for maximum enzymatic activity (preincubation time: 45 minutes; $K_I^{app}$ : 0.8 nM) was 130-fold better when compared with the corresponding value at no preincubation time.

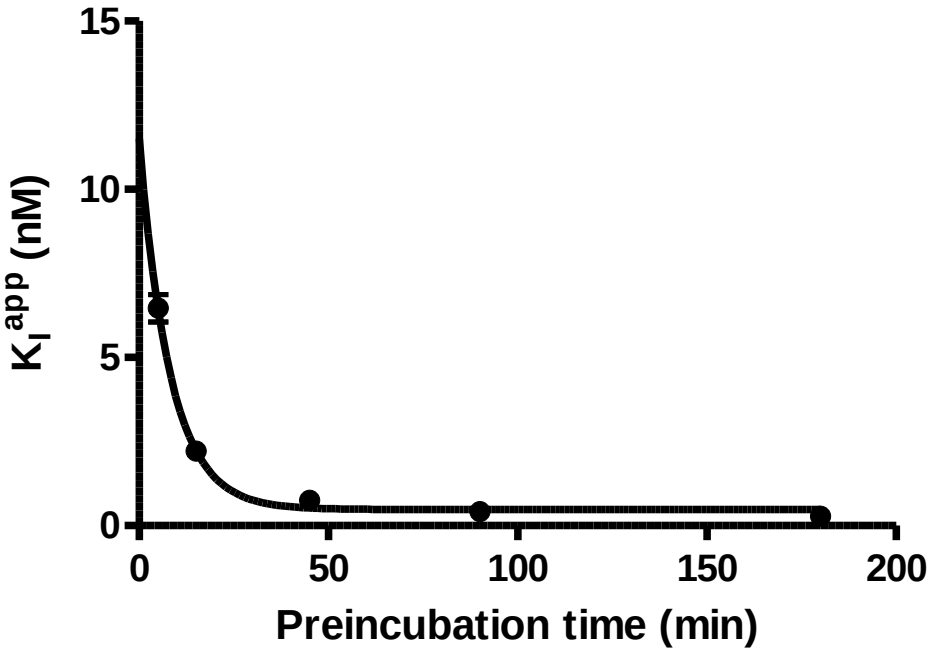



Figure 97. Effect of Preincubation Time in $K_I^{app}$ Values of **1** against Cathepsin K

*Cathepsin K Reversibility Studies*

Compound **1** was found to be a time-dependent inhibitor in the preincubation studies. The next stage in the characterization of **1** as a cathepsin K inhibitor was to demonstrate if the benzophenone thiosemicarbazone inhibits cathepsin K in a reversible fashion. A mixture containing 100 X cathepsin K and 10 X $IC_{50(\text{preincubation time: 5 min})}$, which are 150 and 350 nM respectively, were incubated at 25 °C for 60 minutes. The inhibition of cathepsin K by **1** was able to be monitored by the rapid dilution of the mixture (100-fold) with assay buffer containing Z-FR-AMC. Final conditions were 1.5 nM cathepsin K, 1.7 nM of dibromobenzophenone thiosemicarbazone (**1**) and 50 μM Z-FR-AMC. Additionally, a control experiment (cathepsin K with 4% DMSO as a control vehicle) was also set up. Complete experimental details can be found in the Material and Methods sections. Figure 98 shows the release of AMC for the first 2000 seconds. However,

uninhibited and inhibited reactions were followed for a total time of four hours. A concentrated solution of cathepsin K was able to recover its catalytic activity very slowly after the rapid dilution with assay buffer containing Z-FR-AMC. Thus, a linear regression was applied to the first 2000 seconds of the reactions to determine cathepsin K activity. Cathepsin K activity for untreated reaction was found to be 837 pM AMC/s. Similarly, the inhibited reaction when the system was preincubated for one hour was 87 nM AMC/s. The results showed cathepsin K recovered its activity less than 10% when **1** was present in solution**.** Enzymatic activity can be seen within the first 300 seconds.

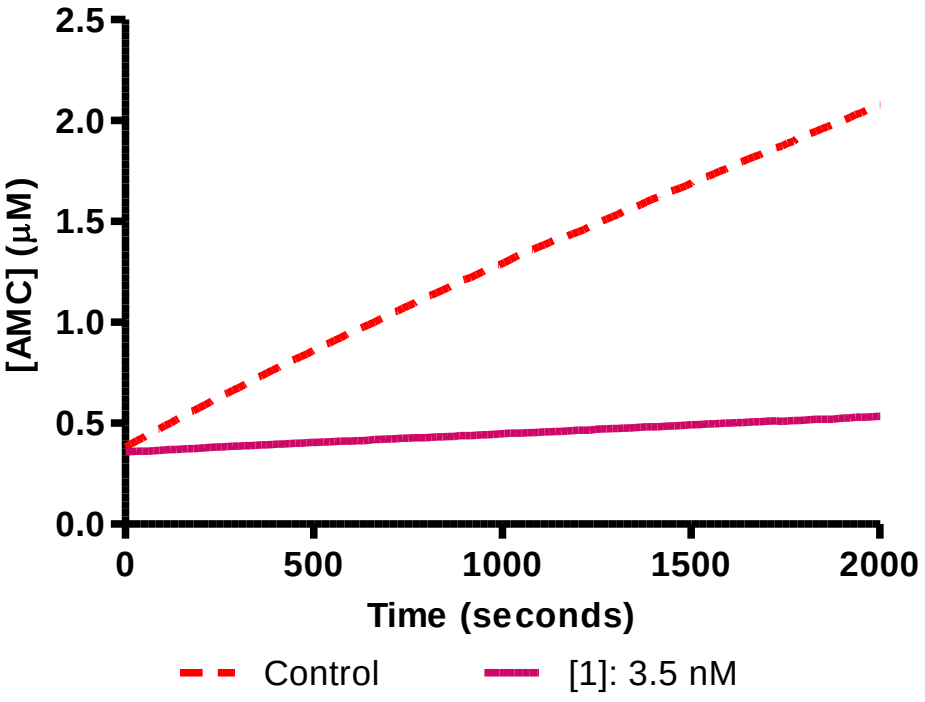



Figure 98. Cathepsin K Reversibility Studies with **1** Using 50 μM Z-FR-AMC

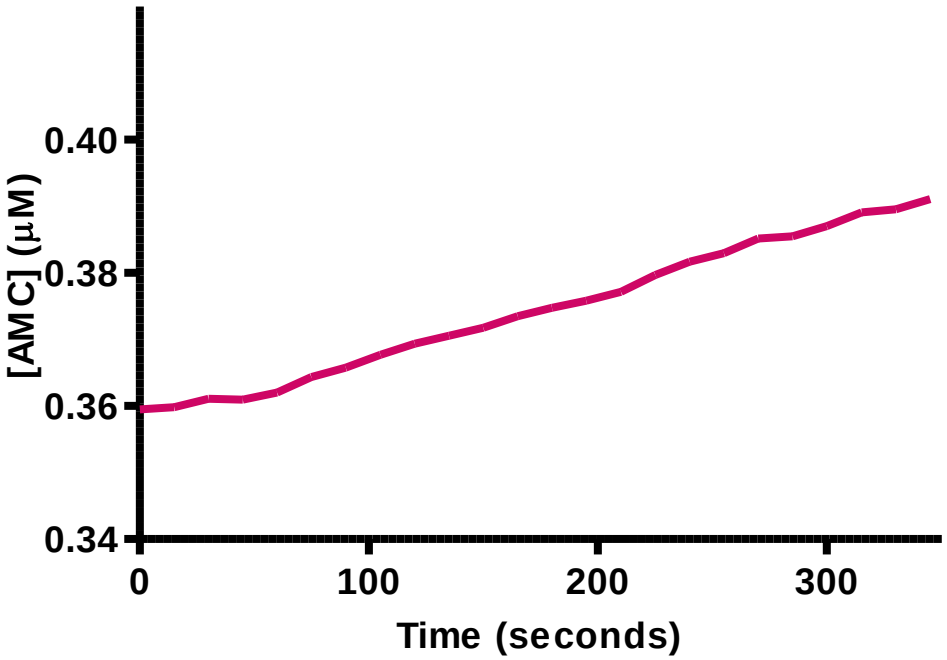



Figure 99. Cathepsin K Reversibility Studies with **1** using 50 μM Z-FR-AMC

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

We finally tried to determine the mode of inhibition of **1** as a cathepsin K inhibitor. Therefore, we investigated the effect of substrate concentration to verify if these synthetic compounds were competing for the same site on the enzyme. Three substrate concentrations (final concentrations: 50, 25, and 10 μM) were used to determine $IC_{50}$ values for compound **1.** Cathepsin K and the series of inhibitors were incubated for five minutes. Results can be found in Table 132 and Figure 100.

Table 132. Effect of Substrate Concentration on $IC_{50}$ values of **1** against cathepsin K

| *[Z-FR-AMC] (μM)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 50 | 35.1 ± 2.5 |
| 25 | 25.8 ± 1.7 |
| 10 | 17.2 ± 1.1 |

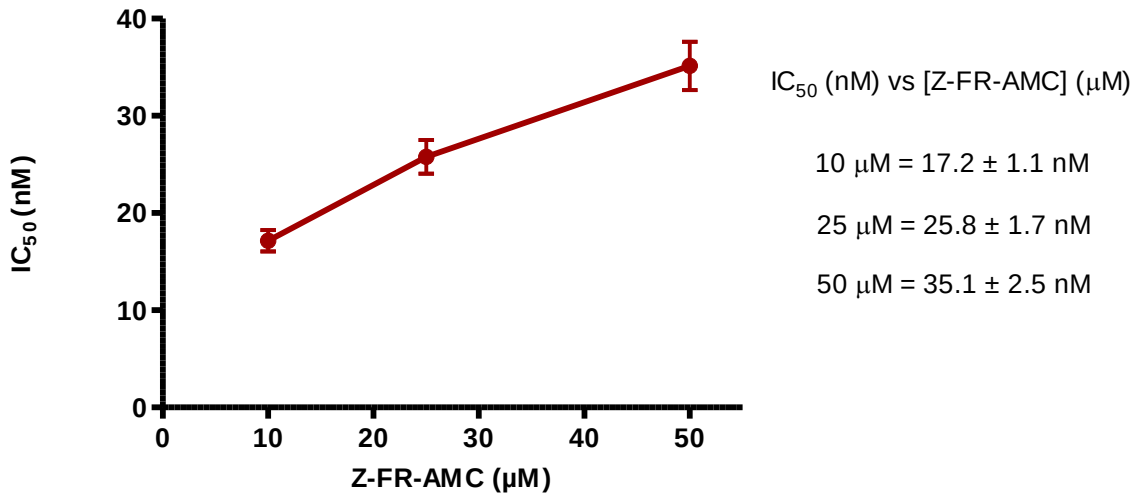



Figure 100. Effect of Substrate Concentration on $IC_{50}$ Values of **1** against Cathepsin K

The effect of substrate concentration was investigated using **1** as one of the lead compound of the series of thiosemicarbazones. Lowering the amount of substrate in solution gives a subtle, yet evident increase in the potency of the dibrominated analog (**1**). According to literature, a positive linear behavior (i.e. $IC_{50}$ α [S]) is an indication the compound acts as competitive inhibitor, that is, both substrate and the compound compete for cathepsin K active site.

*Detection Limits of Procathepsin K by Fluorescent Western Blotting*

Finally, we decided to follow the conversion of procathepsin K to its mature form, using molecular biological techniques. However, it was first necessary to determine the detection limits of both species, procathepsin K and mature cathepsin K when using Western Blotting. The objective of this experiment was the detection limits for the activation of procathepsin K under acidic conditions at room temperature. For this experiments, Western Blotting by using fluorescent-labeled antibodies were used. Inactive human procathepsin K was detected in the nanomolar range using the protocol

previously described in this chapter. Fourteen samples ranging from 5 ng to 200 ng were loaded in a 4-12 % SDS-PAGE gel, blotted and detected. Figure 101 shows the detection limits of procathepsin K using fluorescent Western Blotting. All the samples, with the exception of 5 ng (lane 2) were detected. The intensity of the bands was increased with the protein sample concentration. The western blotting image also revealed that small amounts of procathepsin K had been cleaved and mature cathepsin K was present (lanes 9: 45 ng -14: 200 ng).

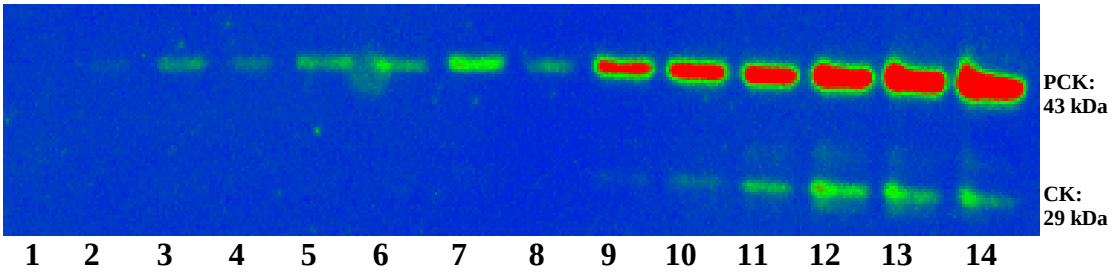



Figure 101. Western Blotting of the Detection of Procathepsin K. Legend: **1**: 5 ng (not detected); **2**: 10 ng; **3**: 15 ng; **4**: 20 ng; **5**: 25 ng; **6**: 30 ng; **7**: 35 ng; **8**: 40 ng; **9**: 45 ng; **10**: 55 ng; **11**: 65 ng; **12**: 100 ng; **13**: 150 ng; **14**: 200 ng; PCK: procathepsin K (mW~ 43 kDa); CK: active cathepsin K (MW ~29 kDa)

*Inhibition of the Activation of Human Procathepsin K*

Cysteine proteases are expressed as proenzymes in mammalian tissues. Cathepsins can be activated under acidic conditions (i.e. lower pH) or by using a procathepsin as a substrate in a hydrolytic reaction (See Chapter 1). The objective of this experiment was to investigate the possibility of the inhibition of procathepsin K under acidic conditions in the presence of **1**, a slow-binding, reversible, competitive inhibitor of mature cathepsin K.

The inhibition of the activation of human procathepsin K was followed at low temperatures. Previous reports have found the activation time increases when activation under acidic conditions is performed at low temperatures. The reaction was followed by using two different techniques: western blotting and using enzyme activity assays with Z-FR-AMC. Briefly, two samples of procathepsin K were activated at 4 °C at pH 3.5. Activation was followed for 6.75 hours and two aliquots were taken every 45 minutes to be monitored using enzymatic assays and western blotting. Final conditions during the activation process are: 1 µM active cathepsin K, 0.7 mM EDTA, 333.3 mM NaCl, 5.4 mM NaOAc pH 3.5, 4% DMSO, and 20 µM of **1** (for the treated samples). Samples for western blotting were inactived with 1.0 % SDS and heat, while samples for the fluorometric enzyme activity assay were used as a stock solution for standard kinetic assays. Final conditions for this assay are: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 DTT, 0.01 % Brij, 1.5 nM inhibited (or control) cathepsin K, 4% DMSO, and 50 µM Z-FR-AMC. Standard procedures using semi-dry protein transfer and a fluorescent western blotting kit were utilized to detect 200 ng of protein that could have been activated or inhibited. Results of western blotting experiments can be seen in Table 108.

Figure 104-A shows the first 270 minutes of activation at the specified conditions without any inhibitor. Three bands can be observed at time 0 minutes. Calculated molecular weights were 43, 35 and 28 kDa. Two of these bands correspond to molecular weights of procathepsin K (MW ~ 43 kDa) and mature cathepsin K (MW ~ 29 kDa) respectively according to suggested values found in literature. The third band (MW: 35 KDa) has also been found in other experimental procedures and suggests that cathepsin K activation is a multistep process where multiple cleavage points produce the final mature

macromolecule. The presence of a band at 28 kDa suggests that procathepsin K had already started the activation process. As previously stated, activation of cathepsin K can be a spontaneous process under acidic conditions whereby an active molecule of cathepsin K can facilitate the activation of other proenzyme molecules. Results confirmed activation of the proenzyme is a lengthy process and can take several minutes. The presence of the 43 kDa band can be observed between 2.5 and 3 hours. The activation of the enzyme produces two clear bands at 35 and 28 kDa. The formation of a transient mature cathepsin K is evident and stable for the first three hours.

The presence of this band was not observed during the rest of reaction as seen in Figure 104-A. The third band (28 kDa) was observed during the entire experiment and the intensity of the bands was increasing at longer activation time points suggesting that the proenzyme was completely activated between 5 and 6 hours.

Table 108-B shows the experimental results of the inhibition of the activation of proenzyme by analog **1**. Once again, three well defined bands with molecular bands were observed (43, 35, and 28 kDa) corresponding to proenzyme, transient mature cathepsin K and fully activated cathepsin K, respectively. However, the biggest difference between both experiments is that the presence of the proenzyme was detected for periods of times longer than 3 hours. The intensity second band (35 kDa) was not as evident as in the case of the untreated samples. Finally, the bands corresponding to the fully activated cathepsin K were barely seen at time points up to 180 minutes. Activated cathepsin K started to appear after two hours. These findings suggest that analog **1** was able to inhibit the activation of procathepsin K. It is unclear whether the synthetic compound is able to inhibit the activation of the proenzyme by direct inhibition of the macromolecule, or if

the compound inhibits only activated cathepsin K, which is partially responsible of the activation of other molecules that are in the proform.

The first set of samples was used to test the activity of mature cathepsin K under acidic conditions at low temperatures using standard fluorometric enzymatic assays. Figure 103 depicts the untreated and treated samples respectively. Briefly, untreated and treated samples were not preincubated in order to measure the activity of activated enzyme under low temperatures. Reactions started by the addition of 500 µM Z-FR-AMC. Final conditions for cathepsin K and Z-FR-AMC are 1.5 nM and 50 µM, respectively. Measurements were taken every 6 seconds for 1200 seconds.

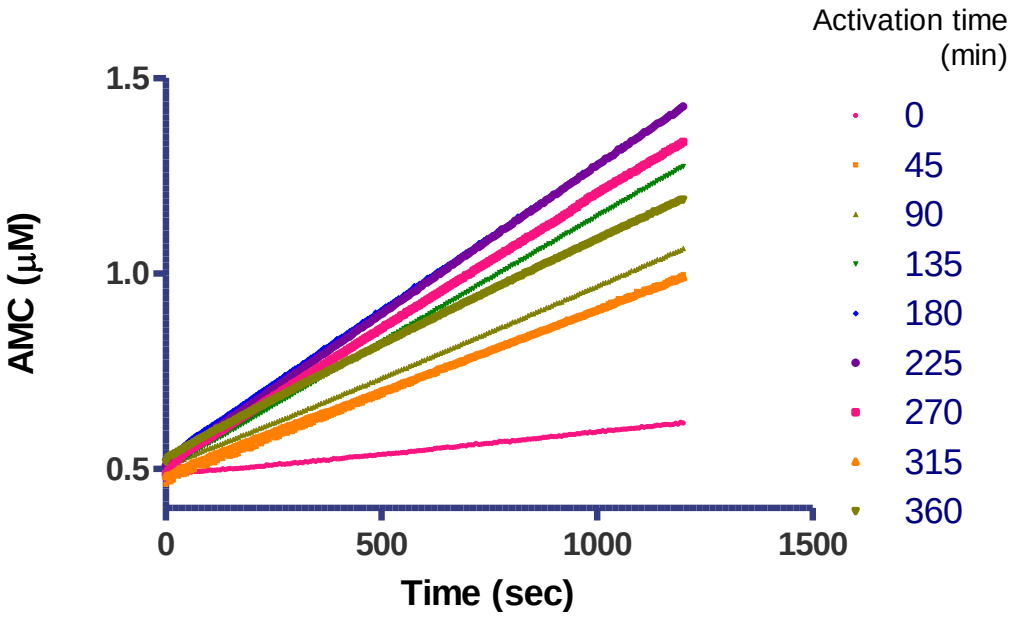



Figure 102. Activation of Procathepsin K Under Acidic Conditions at 3 °C

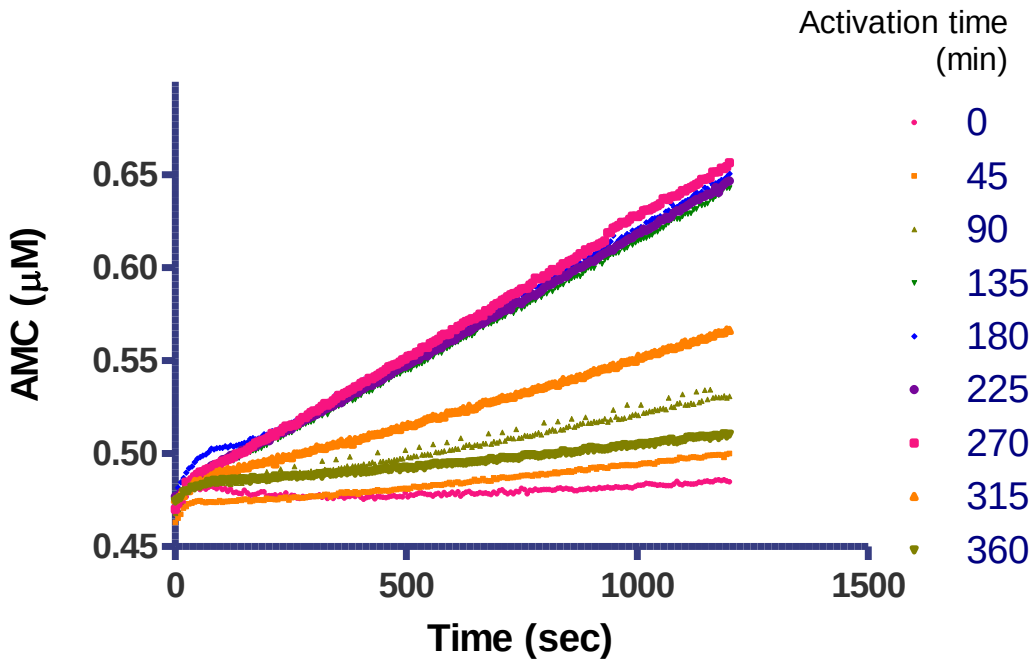



Figure 103. Inhibition of the Activation of procathepsin K by **1** under acidic conditions at 3 °C

Cathepsin K activity was also estimated in nM AMC/s rates. Maximum activities for untreated and treated samples were 0.77 and 0.13 nM AMC/s respectively. Table 116 and figure 90 show the complete evaluation of cathepsin K activity. This represents an 80% inhibition compared to the untreated sample.

Activation of untreated samples occurs rapidly within the first 45 minutes, possibly due to the abundance of natural substrate (procathepsin K) present in solution. The activity increases up to 4 times in that period of time (0.11 vs 0.44 nM AMC/s). Maximum activity was stable between 180 and 225 minutes of activation time, which corresponds with visual observations for the western blotting experimental results. Activity of the active enzyme slowly decreases after 270 minutes. The activity after 6 hours of activation time (3 hours after maximum activity) is 80% with respect to the activated enzyme at 180 minutes.

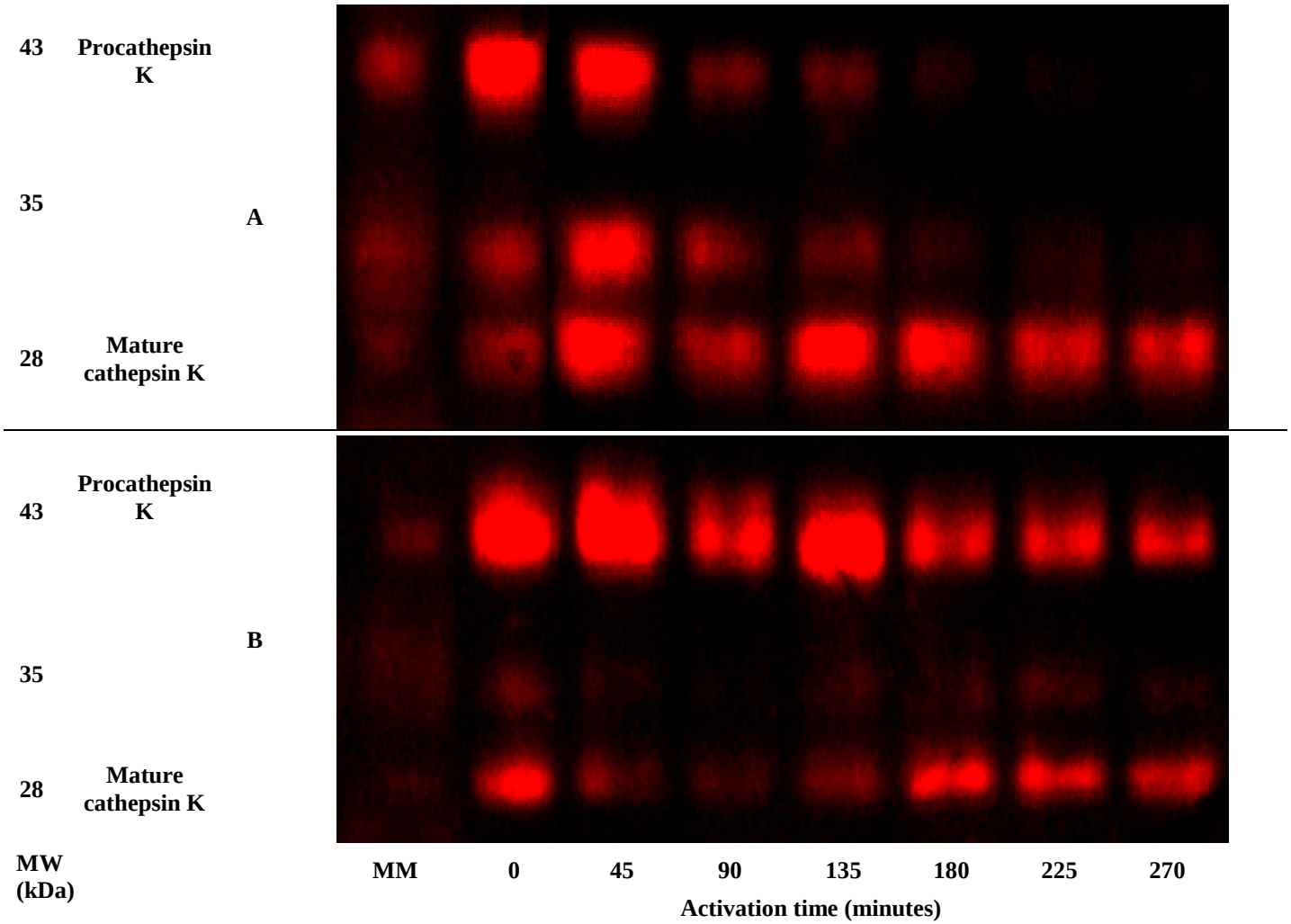



Figure 104. Western Blotting of the Inhibition of Procathepsin K Activation by **1**. A. Untreated Sample. B. Treated Sample with **1**

On the other hand, activation of treated samples occurs very slowly. Overall activities of the reactions for the first 90 minutes showed no catalytic activity in any sample. The inhibitory potency of **1** with regard to the activation of procathepsin K is clearly seen in the activity of the rest of the samples. Activity of cathepsin K was stable for more almost two hours and maximum activity was observed at 270 minutes, which is 1.5 hours after maximum activity for untreated samples was completed. Samples with inhibitor showed little or no activity after 5 hours of activation time.

Closer examination of the samples reveals that samples showed catalytic activity, probably due to the presence of activated cathepsin K that was not inhibited. This activity is quickly inhibited by analog one within the first 35 seconds and inhibition is observed for the rest of the reaction See Figure 105.

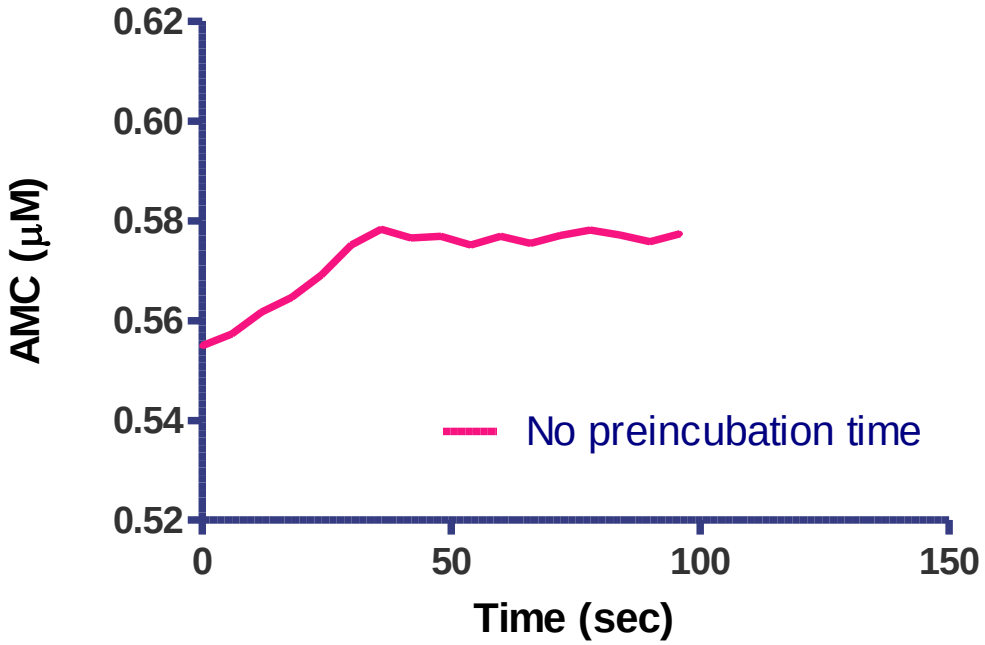



Figure 105. Inactivation of the Activation of Procathepsin K by **1** Under Acidic Conditions at 3 °C (No Preincubation Time).

Table 133. Catalytic Activity of Activated Cathepsin K Measure in nM AMC/s

| Activation time (min) | Untreated samples (nM AMC/s) | Treated samples (nM AMC/s) |
|---|---|---|
| 0 | 0.11 | 0.00 |
| 45 | 0.44 | 0.00 |
| 90 | 0.46 | 0.00 |
| 135 | 0.65 | 0.13 |
| 180 | 0.77 | 0.13 |
| 225 | 0.78 | 0.13 |
| 270 | 0.73 | 0.14 |
| 315 | 0.44 | 0.07 |
| 360 | 0.60 | 0.00 |

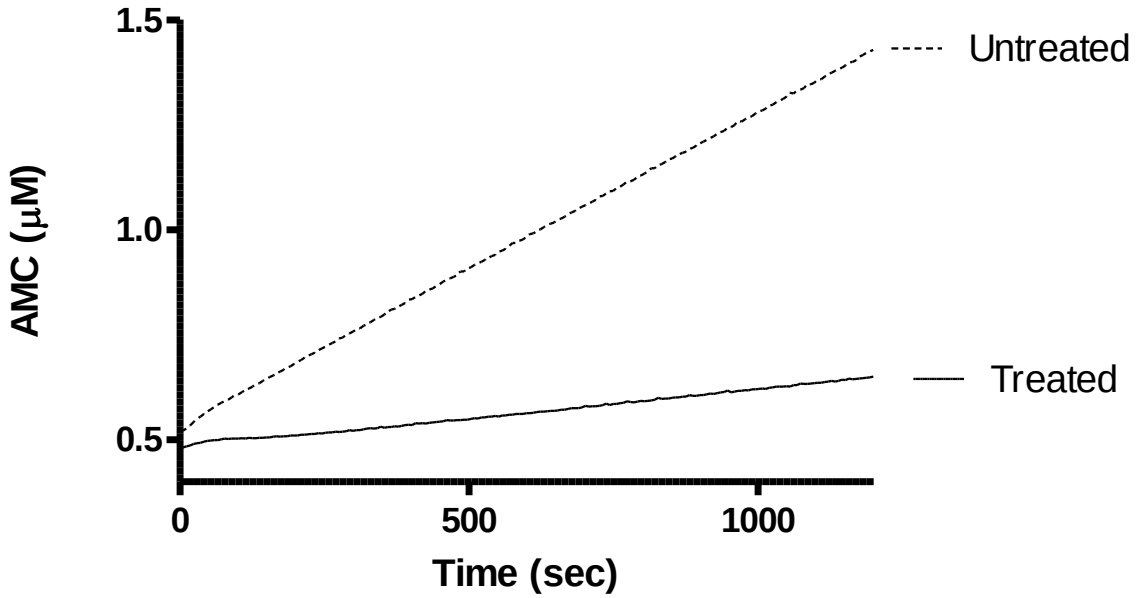



Figure 106. Untreated and Treated Cathepsin K (1.5 nM) with **1** (30 nM) at 3 Hours After Reactions of Activation Process Started

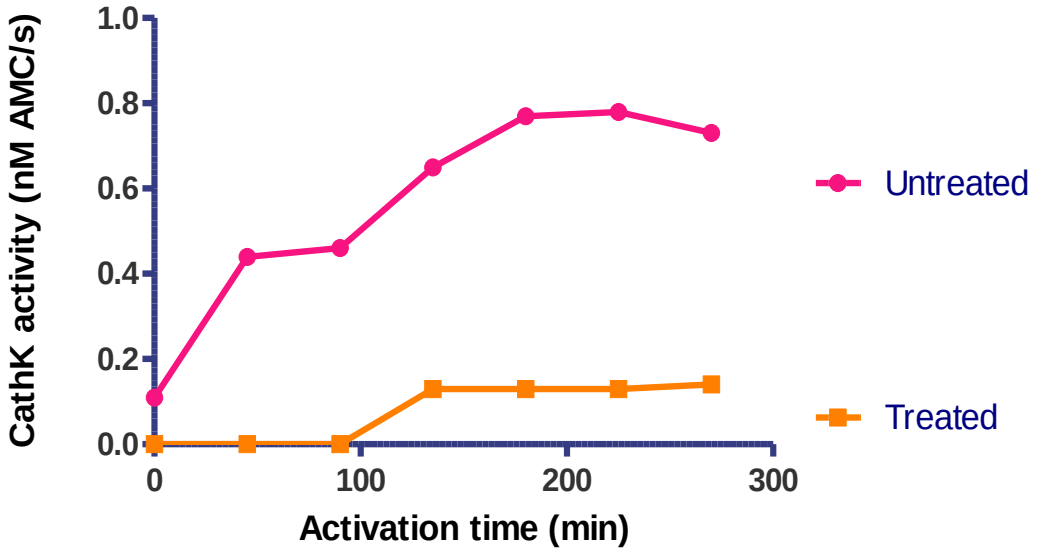



Figure 107. Catalytic activity of Activated Cathepsin K Measure in nM AMC/s

*Inhibition of Cathepsin K Proteolytic Activity by a Thiosemicarbazone Derivative*

Cathepsin K is the main protease found in osteoclasts, bone cells responsible for bone resorption processes. Research has shown that cathepsin K is capable of degrading high molecular weight proteins found in the extracellular matrix such as type I and type IV collagens.[372] Type I collagen is a major component of the ECM, while type IV is the main macromolecule found in the basal membrane. Thus, we explored the catalytic activity of cathepsin K with type IVcollagen from human placenta under physiological-like conditions. Additionally, calf brain tubulin (MW: 55 kDa) was used as a potential cathepsin K substrate for comparison. To date, there is no report of explaining the proteolytic ability of cathepsin K to degrade tubulin. However, cathepsin D, an aspartyl protease, was able to degrade microtubulin-associated protein 2 (MAP-2) and tubulin.[422] We also tested the ability of one of the lead inhibitors (**1,** $IC_{50}$: 35.2 nM), to inhibit the type IV collagenase catalytic activity of cathepsin K. Untreated and treated samples were incubated at 37 °C under acidic conditions (pH 5.5). Six sets of samples were prepared

for this experiment. Three sets (one per substrate) were not treated with **1** (DMSO was used as control vehicle). The fourth (type I collagen), fifth (type IV collagen), and sixth (calf brain tubulin) sets were treated with 20 μM of **1** in 4% DMSO. Activated cathepsin K was added to the samples containing 0.15% chondroitin 4-sulfate with one substrate. The inhibitor **1** was not preincubated with activated cathepsin K in any of these experiments. The role of C4-S was previously established by Li in recent reports. The collagenase activity of cathepsin K is enhanced by the presence of C4-S while cathepsin L can be inhibited by this glycosaminoglycan. Additionally, C4-S increases the stability of the enzyme when C4-S is added to samples containing cathepsin K.[371,423]

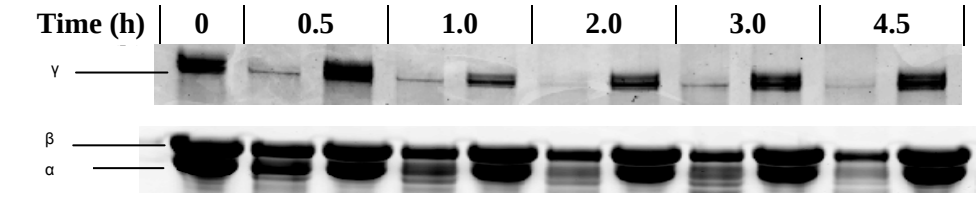
*Inhibition of Cathepsin K Type IV Collagenase Activity by Thiosemicarbazone Derivative **1**. (Preincubation Time: 0 Hours)*

Figure 108 shows a gel image of inhibition of the collagenase activity of cathepsin K with **1 (**no preincubation time). The figure shows the results of degradation of type IV collagen from human placenta using activated recombinant human cathepsin K soluble in acidic solutions.[330] The first sample is untreated type IV collagen that was not treated with cathepsin K, inactivated at the beginning of the experiment in the vehicle control (4% DMSO). Reported literature establishes that collagen is large protein that is divided into three defined chains: α, β, and γ heavy chains and is crosslinked and polymerizes into fibrils. The molecular weights of the heavy chains are approximately 92, 169 and 207 kDa α, β, and chains.[424] Results show the presence of three major bands with high molecular weights that are consistent with the reported values for α, β, and γ heavy chains in the literature. Approximate molecular weights are: α ≈ 100 kDa, β ≈ 120-170 kDa and γ ≈ 270 kDa (lane 1). The second and third bands of the Figure 108 represent untreated and treated samples that were stopped after 30 minutes of the natural

substrate. Untreated samples contained type IV collagen, C4-S and cathepsin K in DMSO, while treated samples consisted of type IV collagen, cathepsin K, C4-S and compound **1** (20 μM**)** in DMSO. In the untreated sample ([**1**]: 0 μM) the γ chain was degraded in as little as 30 minutes. The α, and β, bands look significantly less intense than their respective control at 30 minutes. The γ chain is completely degraded, while the α, and β bands are very significantly degraded. There is evidence of degradatation products when compared to untreated collagen at time 0 (bands with smaller molecular weight, data not shown). Similar results can be seen after three and 4.5 hours of reaction time. However, that is not the case for the treated samples with 20 μM of cathepsin inhibitor **1**. These results indicate that the dibromobenzophenone thiosemicarbazone (**1**) is capable of inhibiting cathepsin K type IV collagenase activity more than fifty percent. A comparison when using human cathepsin L as the target protease also demonstrates the differences of their collagenase activities. Cathepsin K (70 nM) was able to degrade 3 micrograms of type IV collagen, while 270 nM of cathepsin L (4-fold) were needed to degrade the same amount of protein in the same time.

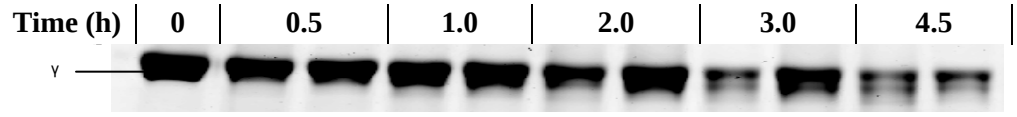
Additionally, Figure 109 shows the results of samples containing cathepsin K with **1** (20 μM) and tubulin from calf brain that was immediately added without preincubation. Cathepsin K was able to degrade the protein in a lesser extent. Untreated tubulin shows one defined band with a molecular weight of 55 kDa. Significant degradation of the untreated samples was observed after two hours of reactions. Dibromobenzophenone thiosemicarbazone was able to inhibit the proteolytic activity of cathepsin K up to three hours when using tubulin as a substrate. To date, there is no report showing similar experiments.

Additionally, parallel experiments were conducted to verify the stability of cathepsin K when using C4-S. The stability of the enzyme was dramatically increased by the addition of the glycosaminoglycan. Similar reports were found and the activity of cathepsin K is not affected by C4-S.[371] Eighty percent of the activity remained after 4.5 hours of reaction when compared to the activity at 0 hours. Furthermore, the activity of the enzyme was constant for the first three hours. Previous experiments showed that cathepsin K is stable for approximately 45 minutes when C4-S is not present in solution (See *Effect of Preincubation Studies on Cathepsin K Inhibition Assays using **1*** for a complete list of experimental conditions).




| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Collagen IV** | + | + | + | + | + | + | + | + | + | + | + |
| **Cathepsin K** | - | + | + | + | + | + | + | + | + | + | + |
| **1** | - | - | +- | - | + | - | + | - | + | - | + |
| **C4-S** | + | + | + | + | + | + | + | + | + | + | + |

Figure 108. Inhibition of Collagenase Activity of Cathepsin K by **1**, Preincubation Time: 0 hours




| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Tubulin** | + | + | + | + | + | + | + | + | + | + | + |
| **Cathepsin K** | - | + | + | + | + | + | + | + | + | + | + |
| **1** | - | - | +- | - | + | - | + | - | + | - | + |
| **C4-S** | + | + | + | + | + | + | + | + | + | + | + |

Figure 109. Inhibition of the Proteolytic Activity of Cathepsin K by **1**, No Preincubation Time

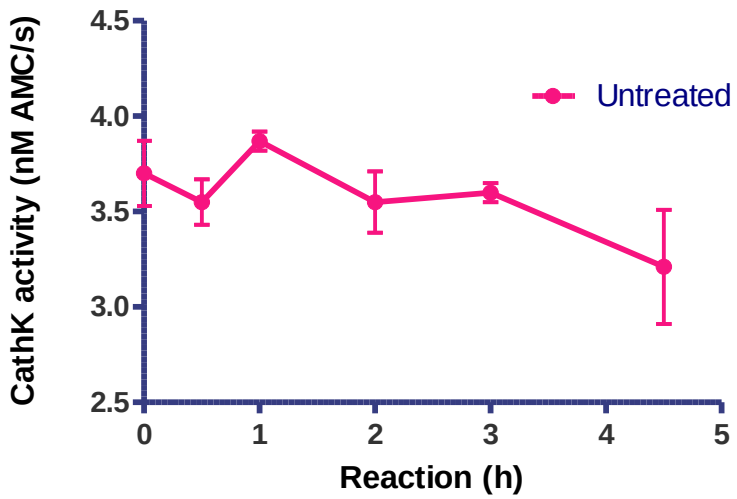
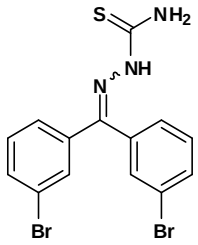



Figure 110. Effect of Chondroitin 4-Sulfate on Cathepsin K Activity

*Effect of Chondroitin 4-Sulfate on Cathepsin K Inhibition Assay*

Finally we decided to explore the effect of chondroitin 4-sulfate in the inhibitory activity of **1.** The standard $IC_{50}$ assay were conducted using two sets of the thiosemicarbazone analog with the presence or absence of C4-S. Preincubation time was set to 5 minutes. Final conditions were: 150 mM NaOAc pH 5.5, 2.5 mM EDTA, 2.5 mM DTT, 4% DMSO, 50 μM Z-FR-AMC, 1.5 nM cathepsin K and 0.15% C4-S. A control $IC_{50}$ (i.e. no C4-S) was also tested for comparison purposes. Results can be seen in Figure 111.

Results clearly showed that C4-S did not interfere in the inhibitory activity of the dibromobenzophenone thiosemicarbazone (**1**). Previous reports done by Li and coworkers established the catalytic activity of cathepsin K is not affected when the protease is bound to C4-S. [373] They also found C4-S interacts with the R-domain of the cysteine protease, quite far from the active site of cathepsin K. Our studies also confirmed these statements due to the independence of $IC_{50}$ with respect to C4-S.[373]




| Without Chondroitin 4-Sulfate | With Chondroitin 4-Sulfate (0.15%) |
|---|---|
| $IC_{50}$: 31.9 ± 2.9 nM | $IC_{50}$: 29.9 ± 3.6 nM |

Figure 111. Effect of Chondroitin 4-Sulfate on Cathepsin K Inhibition Assay using **1**

*Conclusions*

More than one hundred TSC analogs were evaluated as inhibitors of recombinant, activated cathepsin K from human liver. The library contained a select group of potent, cathepsin K inhibitors. Figure 112 shows the Structure-Activity Relationship of thiosemicarbazones based on their chemical structures.

Advanced kinetics were used to investigate the mechanism of **1**. The compound was determined to be a time-dependent, slow, reversible, competitive inhibitor of the fluorogenic inhibitor Z-FR-AMC. The increase in $IC_{50}$ values as a function of Z-FR-AMC (substrate) concentration indicates that **1** competes with the substrate for binding to the enzyme active site. Fluorescent western blotting demonstrated that **1** was able to inhibit the cleavage (substrate) and activation of human procathepsin K under acidic conditions and low temperatures.

Finally, the proteolytic activity of mature cathepsin K was examined by using a series of natural substrates type IV collagen from human placenta. The collagenase activity of cathepsin K was inhibited by 3,3'-dibromobenzophenone thiosemicarbazone. Additionally, tubulin from calf brain was used as a potential substrate for the enzyme.

Results indicate tubulin underwent proteolytic degradation by cathepsin K. Analog **1** was also able to delay this activity.

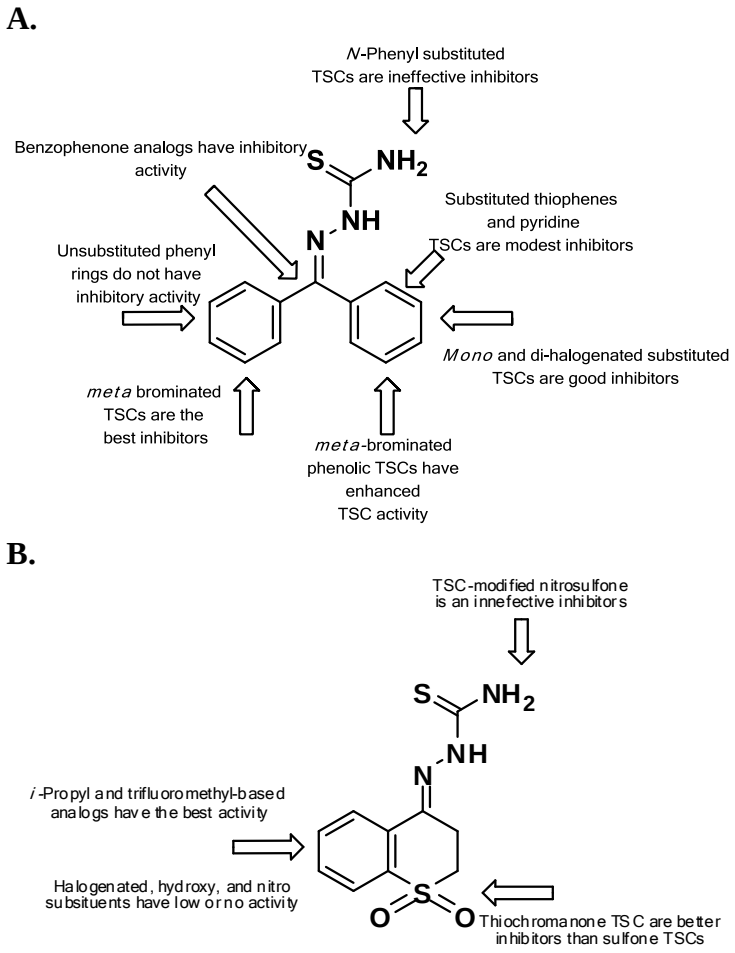



Figure 112. Structure-Activity Relationship for TSCs as Cathepsin K Inhibitors. **A**. Benzophenone TSCs. **B**. Sulfone TSCs

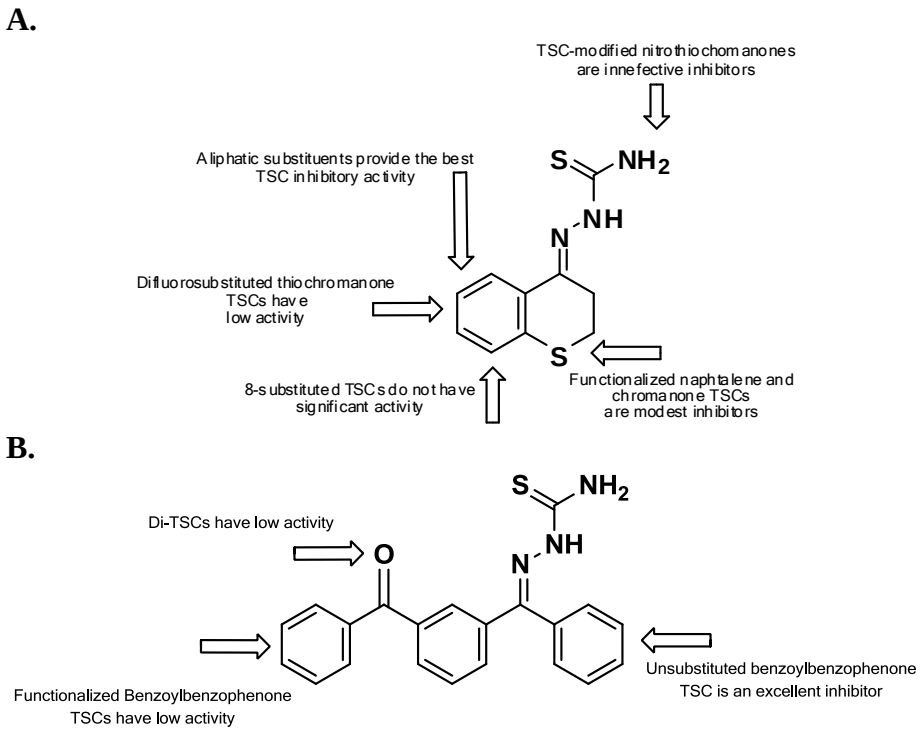



Figure 113. Structure-Activity Relationship for TSCs as Cathepsin K Inhibitors. **A**. Thiochromanone TSCs. **B**. Benzoylbenzophenone TSCs

# CHAPTER FOUR

## Evaluation of Thiosemicarbazones as Cruzain Inhibitors

### *Role of Proteases in Parasitic Diseases*

Pathological diseases in mammalians are often caused by parasites, organisms that get their nutrients at the expense of their hosts. Parasites can be classified as endoparasites or ectoparasites, depending if they live inside or outside the body of the hosts respectively. There are numerous transmission vectors; however, contaminated water, food, insect bites and direct contact between healthy and infected organisms are the most common transmission methods. The list of parasitic diseases includes amoebiasis, enterobiasis, giardiasis, scabies, sleeping sickness, toxoplasmosis among others.[425]

Parasitic diseases are extremely difficult to understand due to the complexity in the pathology and biology of individual diseases. Some parasites can inhabit their host organisms for years, even decades, being asymptomatic. The lack of effective treatments makes the eradication of these diseases a challenging task. Malaria, as an example, is considered an endemic disease that affects more than 500 million people worldwide.[426]

One of the main functions of proteases in parasitic diseases is parasite invasion. Parasitic migration and invasion are intrinsically related. Parasites, usually in their larval stage, migrate through organs or tissues. This phenomenon is due to the action of powerful proteases, usually metalloproteases, cysteine, and serine proteases. Research has proven that inhibition of some of these targets might delay or arrest parasitic invasion

in human skin. Recently, Lopez-Quezada was able to arrest skin invasion by *Schistosoma mansoni* by using a serine protease inhibitor.[427]

## *General Considerations of Chagas' Disease*

### *History, Statistics, and Geographical Distribution of Chagas' Disease*

Carlos Chagas', a Brazilian physician discovered this tropical disease more than one hundred years ago.[428] Chagas' disease is also known as mal de Chagas' (in Spanish and Portuguese) and *American trypanosomiasis*.[429] Recent demographics showed than more than 8 million people are infected with the parasite and one quarter of the total population in Latin America are at potential risk to be affected.[430]

Chagas' disease is widely spread in several regions of Latin America, with endemic proportions in more than 18 countries. However, an increasing number of cases have been detected in non-endemic areas, especially in those countries with higher number of immigrants carrying the disease. Patients with Chagas' disease have been reported in Netherlands, Spain, France, Portugal, Switzerland, Germany and United Kingdom.[431] More than 3 million immigrants live in these countries and the expected number of infected patients could be as high as one hundred thousand cases. This problem has also been extended in the United States and other parts of the world.[432] Eleven possible vectors of *Trypanosoma cruzi* have been discovered in characterized in southern states of the nation. Higher prevalence of the vectors (insects infected with the protozoa) can be found in Texas, California and Arizona. A lower number of vectors is also found in Florida, Georgia, Tennessee, Alabama and Louisiana.[433]


Figure 114. Estimated Number of Immigrants with *Trypanosoma cruzi* Infection Living in Non-Endemic Countries. (Reproduced from Rassi, page 1391 [432])

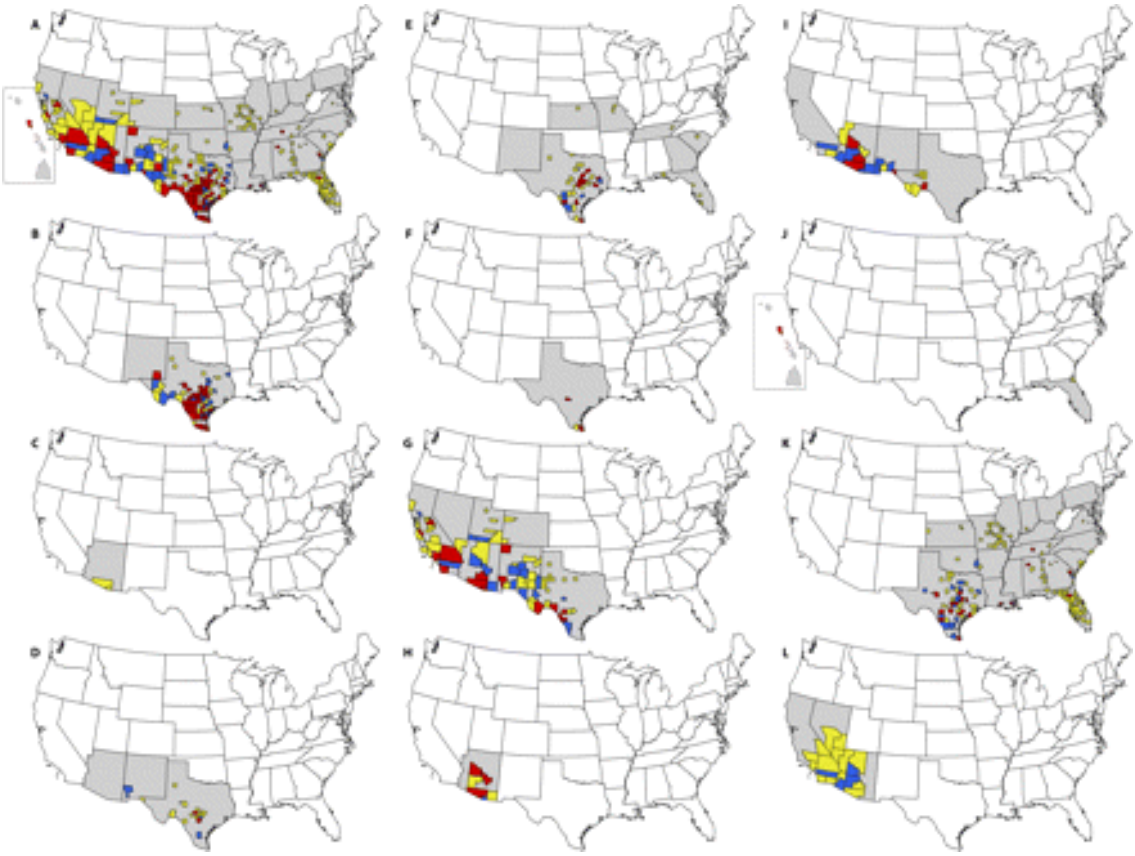


Figure 115. Triatomine Species Geographic Distribution by State (gray areas) and County and *Trypanosoma cruzi* Infection Status by County in the Continental United States and Hawaii. (A) All species; (B) *Triatoma gerstaeckeri*; (C) *T. incrassata*; (D) *T. indictiva*; (E) *T. lecticularia*; (F) *T. neotomae*; (G) *T. protracta*; (H) *T. recurva;* (I) *T. rubida*; (J) *T. rubrofasciata*; (K) *T. sanguisuga*; (L) *Paratriatoma hirsuta*. Red, *T. cruzi*-positive specimens; blue, negative specimens; yellow, no testing reported. (Reproduced from Bern, page 660[433])

*Etiology of Trypanosoma cruzi and its Vectors*

The protozoan *T. cruzi* belongs to the *Trypanosomatidae* family, or trypanosomes, and can be divided into two major groups: *Stercoraria* and the *T. cruzi* clade.[434] Some of the transmission vectors of the trypanosomes found in South America are: *Triatoma infestans, Rhodnius prolixus,Triatoma dimidiata, Mepraia spinolai and Mepraia gajardoi*.[435] Other names given to the insects include: “kissing bugs” and vinchuca. The former name is given to these insects due to their preference to feed on people’s faces.

*Mode of Transmission and Life Cycle of T. cruzi*

Several mammals, *Homo sapiens* included, are considered common hosts for Chagas' disease. Mammalian blood is part of the triatomid's dietary habits. Usually, insects that are infected with *T. cruzi* transmit the infection when they bite healthy mammals. The spread of the protozoan occurs when the transmission vectors defecate contaminated feces over the skin of the mammal. The trypomastigotes (first morphological stage of *T. cruzi*) enters in the blood stream by physical and biochemical processes (self scratching and protealytic degradation by enzymes found in vector's saliva respectively). Then, trypomastigotes differentiate into epimastigotes, the second biological stage of the protozoan via a phagocytosis process by macrophages and leucocytes. The last biological stage of *T. cruzi*, amastigote, multiplies inside of their host cells and differentiates into new trypomastigotes. The new generation of trypomastigotes is transported by the blood system and invades other target organs such as heart, muscles, placenta, glial cells among others.[436–438] Figure 95 depicts the major stages of the life cycle of *Trypanosoma cruzi*.

Non-traditional modes of transmission have also been detected. Literature reports cases of infected people with Chagas' disease by blood transfusion, organ transplants, contaminated food and placental infection to newborns.[439–442]

*Clinical Stages and Symptoms*

American trypanosomiasis is divided into three clinical stages: acute, asymptomatic and chronic stages.[429,443] However, the incubation stage is remarkably important because it happens prior to the acute stage. The mode of transmission determines the duration of that stage. The incubation period varies between five to

fourteen days, if the patient was infected via triatomid feces, and can last up to forty days if the transmission happened due to blood transfusion.

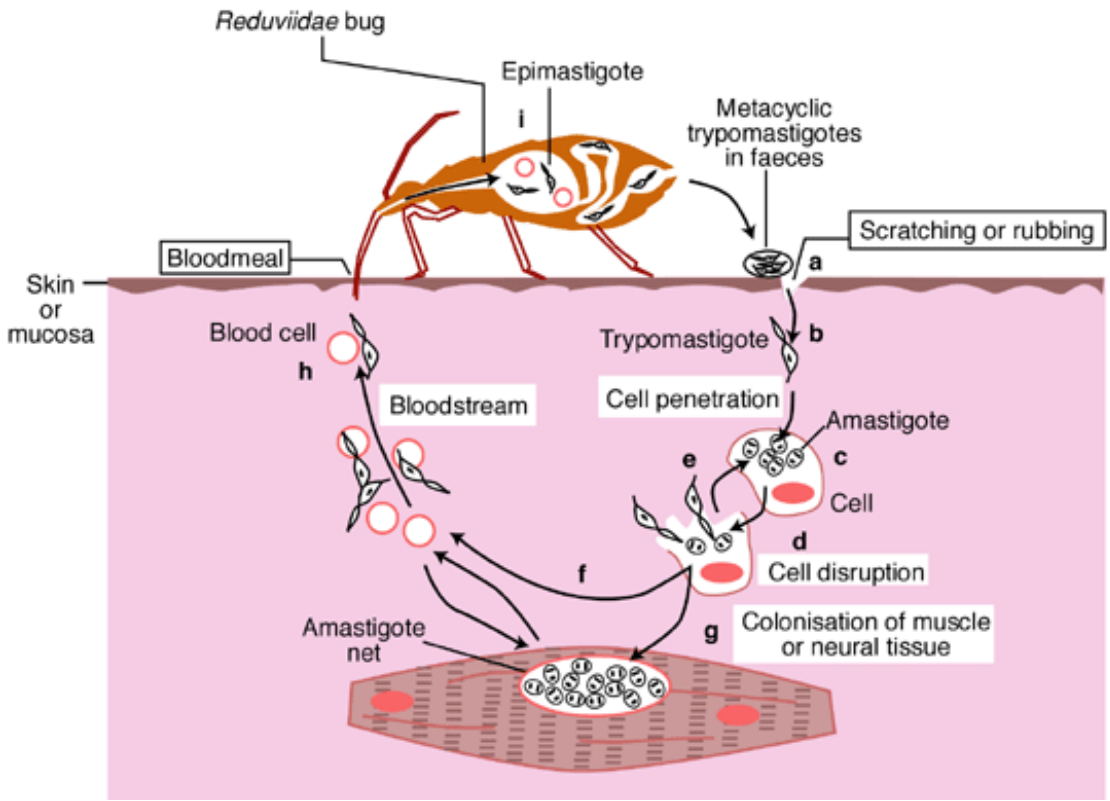



Figure 116. Schematic Representation of the Life Cycle of the Flagellate Protozooan *Trypanosoma cruzi* (reproduced from Macedo, page 3[438])

The main characteristic that defines the acute phase is the ability to find the parasites within the blood stream. Patients can develop several symptoms that include: fever, anorexia, myalgia, joint pain, nausea, vomiting, and an unilateral edema, also known as Romana's sign. The latter symptom is particularly helpful for the detection of the infection during this stage.

The asymptomatic stage is the second stage of the disease.[444] It is also called indeterminate stage because patients do not show any symptoms. In this stage, parasites

practically disappear from the blood stream and they are more difficult to detect. The duration of the stage can vary from a few months to decades.

The chronic stage occurs when some organs start failing. Cardiac and digestive systems are usually are the most vulnerable organs to this infection. Patients in a chronic stage develop arrhythmias, cardiac failure, embolic disease, and death in many cases. Digestive anomalies include problems when swallowing, chest pain, and excessive production of saliva among others. Severe constipation can last for months.

Pregnant women and patients with AIDS or weak immune systems are very vulnerable to Chagas' disease.[445] Newborn infants can get the infection from their mother and develop meningoencepalitis, retina changes and cardiac insufficiency.[446]

*Methods of Detection*

There are several methods to detect Chagas' disease. The most important are: microscopic techniques, parasite isolation, serology, molecular biology techniques and clinical examination.

A. Clinical diagnosis at a late stage is usually by electrocardiograms. These examinations look for abnormal arrhythmias, tachycardia, heart block and ventricular aneurysms.[447]

B. Microscopic techniques are able to detect *T. cruzi* using staining protocols such as Giemsa or Wright tests. The protozoan can be detected in several human fluids and tissues (blood, and cerebrospinal fluid).[448]

C. Parasite isolation can be lengthy, and expensive. Standard experiments require the inoculation of healthy mammals with blood from infected patients. Other protocols include cell culture, and xenodiagnosis.[449]

D. Serology is one of the most used and precise techniques that are available for the diagnosis and detection of Chagas' disease. Indirect Immunofluorescent Assay (IFA), Enzyme-linked immunosorbent assay (ELISA), hemagglutination, radioimmunoprecipitation are commonly used with high sensitivity and specificity rates.[450] However, false positives may happen with other parasites from the Leishmania family.[451]

E. Molecular biology techniques are quite useful due to the detection of the parasite using polymerase chain reaction (PCR) and immunoblotting protocols.[452]

*Prevention*

Preventive techniques are usually suggested in countries when the disease is considered endemic. Some people use permethrin, an insecticide, in their beds to repel the transmission vectors.[453]

*Treatment*

Currently, there is no effective treatment to either prevent or cure Chagas ' disease. Thus, mortality rates are between 5% and 8% in some regions. Unfortunately, more than twenty percent of infected patients develop the chronic stage. Currently, nifurtimox (Nx, a nitrofuran derivative) and benznidazole (Bz, a nitroimidazole derivative) are currently used for the treatment of patients diagnosed with Chagas' disease in its acute stage as effective chemotherapeutic agents.[454,455] However, treatments with nifurtimox are usually painful, toxic, lengthy and show poor efficacy.[456] Side effects of these treatments include stomach pain and rash in the case of Nx, while Bz may cause digestive problems in some patients.

A

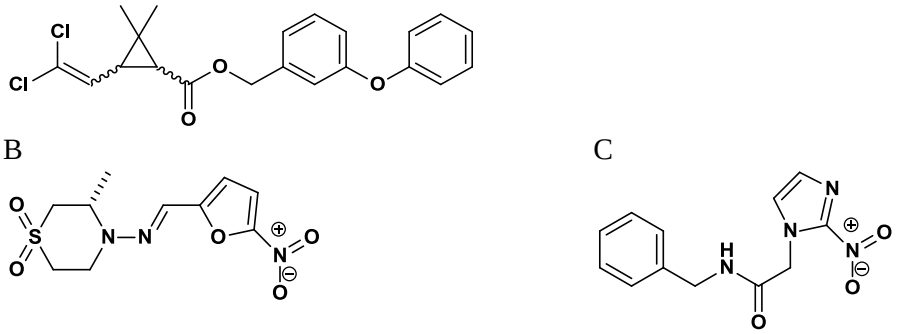



Figure 117. Chemical Structures of A. Permethrin. B. Nifurtimox. C Benznidazole

*Mechanisms of Action of Nifurtimox and Benznidazole*

Maya and coworkers showed that Nx and Bz act against the formation of electrophilic metabolites within the parasite. They have done an extensive review about the mechanism of action of these anti-parasitic agents. The presence of a nitro group allows nifurtimox and benznidazole to be potent agents. NADPH dependent oxidoreductases, form very reactive radicals and molecules(R-NO). Molecular oxygen facilitates the recovery of the compound via reduction of the former intermediates. Superoxide dismutase utilizes these radicals, in order to form molecular oxygen and hydrogen peroxide. Hydroxyl free radicals bind to critical cell components such as lipids and DNA producing irreversible changes to these cellular components.[457]

*Introduction to Cruzain and its Importance in Chagas' disease*

*Nomenclature, Classification and Historical Background*

Cruzain, the recombinant form of cruzipain, is a parasitic hydrolase found in *Trypanosoma cruzi* with an EC number of 3.4.22.51.[458] Both names are used in literature. Cruzain is also considered a cathepsin-L like enzyme due to their structural similarities and is also considered as an endopeptidase.

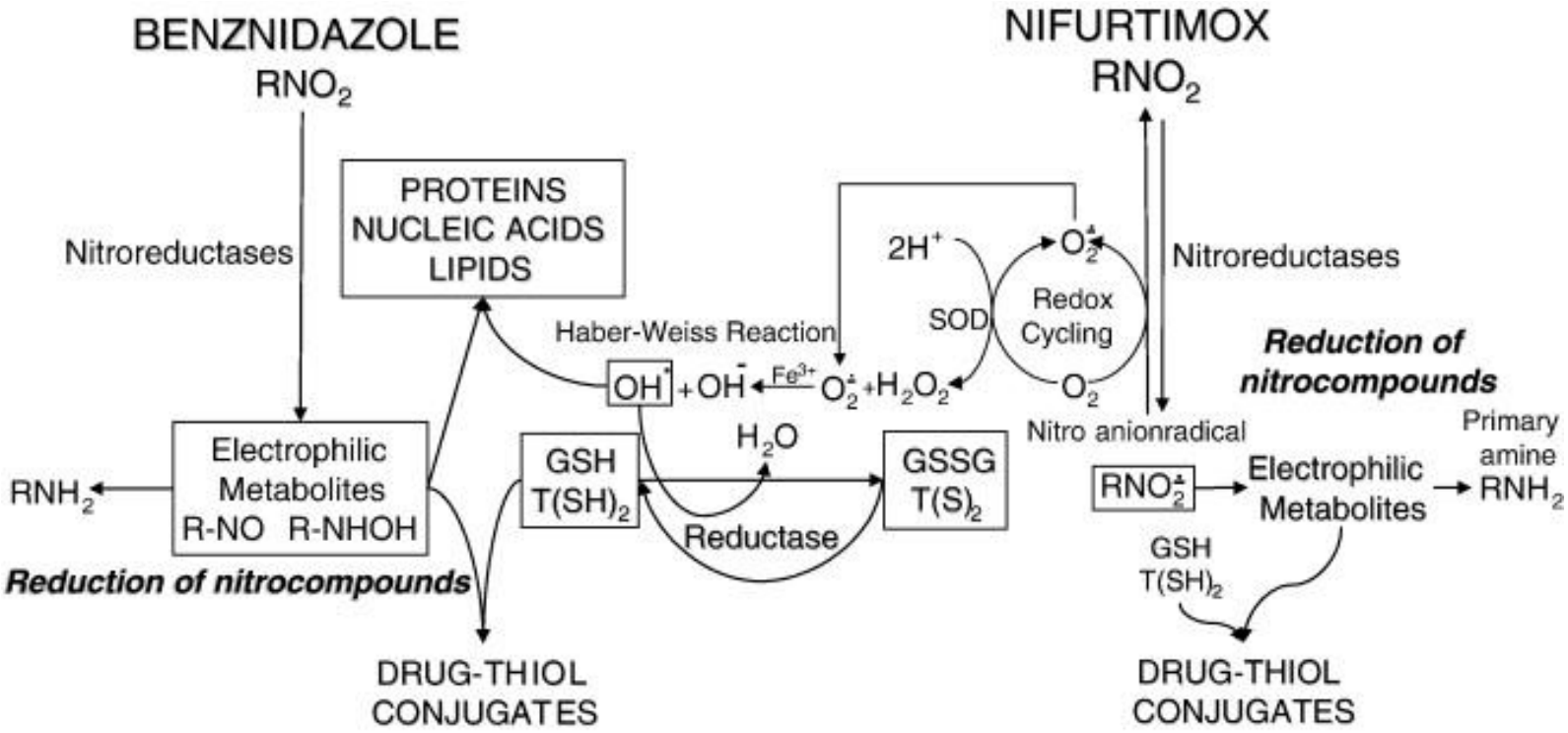



Figure 118. Role of Glutathione and Trypanothione in the Action and Metabolism of the Antichagasic Drugs Nifurtimox and Benznidazole (Reproduced from Maya, page 603 [457])

The official name of the protein is cruzipain. Other names for this protease are cruzain, and cruzaine among others. First reports of cruzain activity were by Itow and coworkers in 1977.[459] Furthermore, Rangel reported a purified form of cruzain few years later. Cruzain is considered the most important and powerful protease found in *Trypanosoma cruzi*.[460]

*Localization of Cruzain*

Cruzain is mainly found in parasitic lysosomes, similar to cathepsin L, which is expressed in mammalian organisms. An in-depth investigation confirmed cruzain trafficking is extremely complex. Cruzain is transported to the epimastigote lysosome via endoplasmic reticulum and Golgi apparatus.[461–463] Immunoblotting data indicate cruzain is also expressed in every stage of the protozoan life cycle. Specifically, cruzain was found at the surfaces of epimastigotes and amastigotes, as well in transitional forms of the trypomastigote-amastigote pseudo stage.[464–466]

*Biological Roles of Cruzain*

Cruzain is considered the most important protease due to its high catalytic activity found in *T. cruzi*. Reports showed cruzipain is mainly involved in cell invasion, immune evasion and disease pathogenesis.

Cruzain plays a key role in the parasite's ability to invade host cells. Studies showed cruzain mediates cell invasion by two independent pathways. In the first one, cruzain stimulates the activity of B2 kinin receptor, while the second one is completely kinin independent but helps in cell invasion processes.[467]

Cruzain is also involved in immunogenic processes. Investigations reported that mice immunized with cruzain showed strong responses against muscle and/or cardiac

myosin. The studies also demonstrated that mice immunized with cruzain showed a clear enlargement of their spleens.[468,469]

Parasite life cycles are well regulated by cytokine activated macrophages.[470] This is one of the key roles in the pathogenesis of Chagas' disease. However, recent theory showed cruzain is also important in the regulation of macrophages. For example, macrophages were activated when they were cultured with purified cruzipain.[471] Evident signs of this activation include L-arginase activity, and expression of IL-20 and TGFβ.[472] Nevertheless, these results have not been proven in parasite infection in *in vivo* studies.

Cruzain's kininogenase activity releases Lys-bradykinin, an inflammatory peptide and prekallikrein which also generates Lys-bradykinin.[473] These two proteins are responsible for increasing parasitic uptake in host cells.[426]

Other roles of cruzain include: cell remodeling in the epimastigote, activation of bradykinin receptors, and mediation of apoptotic processes.[474–476] Researchers have also worked with genetically-modified knockout models of *Trypanosoma cruzi*. Unsuccessful living models indicate that cruzain plays important roles in the life cycle of the parasite.[477] Engel and coworkers were able to eradicate *in vitro* infections with cysteine proteases, supporting the previous statement.[462]

### *Catalytic Activity of Cruzain and Substrate Specificity*

Cruzain is a cathepsin L-like cysteine protease. Its amino acid sequences is 45% identical compared to mouse cathepsin L and its substrate specificity is very similar when their catalytic activities are compared.[478] Several studies have been carried out to understand cruzain's substrate specificity compared to other cysteine proteases members such as cathepsin B, cathepsin L and cathepsin F, which are identical as much as 50%.

Gillmor explored the ability of this parasitic protease to cleave synthetic fluorogenic substrates.[479] Essentially, they explored five different fluorogenic substrates with the form Z-XR-AMC where X where hydrophilic or hydrophobic residues at the P2 position. Their studies were extended to test cruzain's enzymatic activity over a broad pH range (3 to 8). In summary, cruzain prefers substrates with hydrophobic residues at the P2 position (X= Phe, Ile and Tyr) over hydrophilic ones (Arg, Gln). Nevertheless, the protease is able to cleave hydrophilic residues to a lesser extent. Cruzain's activity was well defined over a broad pH range when hydrophobic residues were tested reaching 50% activity under acidic conditions (pH 5.5). Studies showed the phenolic ring of phenylalanine can be easily accommodated in a hydrophobic pocket formed by Ala133, Leu157, Gly160 and Glu205 and Met68.

*Natural Substrates of Cruzain*

Cruzain's enzymatic activity has been well studied due to its major roles in cell invasion and immunogenic procedures. The enzyme is able to cleave proteins found in the extracellular matrix such as collagen, and gelatin, among others.[480] Overall, *Trypanosoma cruzi* is capable of degrading selective members of the extracellular matrix, including collagen, laminin, and heparin sulfate.[481] Interestingly, *in vitro* experiments have shown cruzain can degrade fibronectin, which is one of the few macromolecules that is not affected during parasitic cell invasion. Cruzain also possesses caseinase properties and can also cleave hemoglobin, bovine serum albumin and immunoglobulin.

*Structure of Cruzain*

Cruzipain is expressed in *Trypanosoma cruzi* as a preproenzyme of 467 amino acids residues containing four very well defined portions: a signal peptide (19 residues),

the propeptide (104 residues), the mature portion (215 residues) and a carboxy-terminal domain (129 residues).[460] The signal and prodomain, which account for 123 amino acids, are cleaved under acidic conditions. The function domain is to protect unnecessary hydrolytic activity of the enzyme inside of the eukaryotic cells of the parasite, similarly to other cysteine proteases found in mammals. Mature cruzain is a monomer containing 215 residues with a molecular weight of 23 kDa; however, glycosylated cruzain molecular weight is 51 kDa. Finally, the 129-residue carboxyl terminal domain is an unusual feature that is a characteristic of cruzain. To date, there are no other known enzymes sharing this portion. Researchers have unsuccessfully tried to crystallize cruzain with its C-terminal. The challenges in the elucidation of the biological functions of the C-terminal have increased due to the lack of validated crystal structures and the existence of other modifications (i.e. isotypes) are also known, yet difficult to understand their implications in *in vivo* studies. Nevertheless, it has been concluded cruzain's catalytic activity is C-terminal independent.[482]

The amino acid sequence of mature cruzain is shown in Figure 120.[131] The amino acid sequence of both cysteine cathepsins (cathepsin L and cruzain) have been identified and confirmed. Literature reports identities between 45% and 50%. The figure shows the amino acid sequence of mature cruzain containing 215 residues.

The three most abundant residues are alanine, glycine, asparagine and valine (25, 23 and 21 residues respectively). On the other hand, the least abundant residues are histidine (4 residues), arginine and phenylalanine (three residues each). Hydrophobic residues (A, F, I, L, M, P, V and W) account for forty percent of cruzain's composition. Also, 38 percent of the composition is entirely made of polar, uncharged residues.

The secondary structure also reveals that seven alpha helices account for 25% of the structure (55 residues). However, 24% of the sequence forms beta sheets distributed in seventeen strands (52 residues).

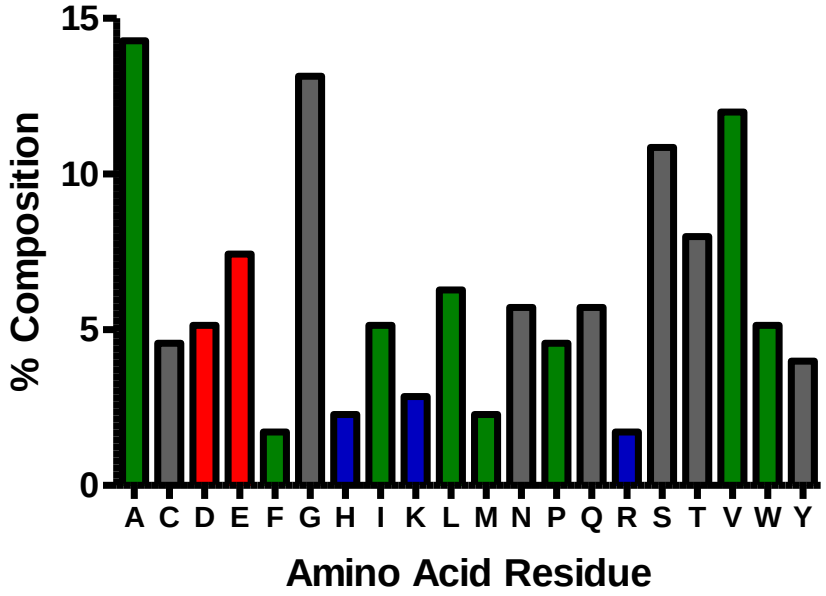



Figure 119. Composition of Mature Cruzain. Legend= blue: basic residues; red: acidic residues; green: hydrophobic (non polar) residues; gray: neutral, polar residues

Table 134. Amino Acid Compositition of Mature Cruzain

| AA | # | % | AA | # | % | AA | # | % |
|---|---|---|---|---|---|---|---|---|
| Ala | 25 | 11.6 | Ile | 9 | 4.2 | Arg | 3 | 1.4 |
| Cys | 8 | 3.7 | Lys | 5 | 2.3 | Ser | 19 | 8.8 |
| Asp | 9 | 4.2 | Leu | 11 | 5.1 | Thr | 14 | 6.5 |
| Glu | 13 | 6.1 | Met | 4 | 1.9 | Val | 21 | 9.8 |
| Phe | 3 | 1.4 | Asn | 10 | 4.7 | Trp | 9 | 4.2 |
| Gly | 23 | 10.7 | Pro | 8 | 3.7 | Tyr | 7 | 3.3 |
| His | 4 | 1.9 | Gln | 10 | 4.7 | Total | 215 | 100 |

The complete amino acid sequence of preprocruzain is shown in Figure 121. The macromolecule is made of 467 residues and divided in four different regions: The preregion, the propeptide, the mature enzyme and the carboxy-terminal. Eakin and

coworkers offers a complete description of the molecule. Nevertheless, its crystal structure has not been elucidated.[460]

```
1     APAAVDWRARGAVTAVKDQGQCGSCWAFSAIGNVECQWFLAGHPLTNLSE
51    QMLVSCDKTDSGCSGGLMNNAFEWIVQENNGAVYTEDSYPYASGEGISPP
101   CTTSGHTVGATITGHVELPQDEAQIAAWLAVNGPVAVAVDASSWMTYTGG
151   VMTSCVSEQLDHGVLLVGYNDSAAVPYWIIKNSWTTQWGEEGYIRIAKGS
201   NQCLVKEEASSAVVG
```

Figure 120. Amino Acid Sequence of Cruzain (PDB: 1ME3[131])

```
1     MSGWARALLLAAVLVVMACLVPAATASLHAEETLTSQFAEFKQKHGRVYESAAEEAFRLS
61    VFRENLFLARLHAAANPHATFGVTPFSDLTREEFRSRYHNGAAHFAAAQERARVPVKVEV
121   VGAPAAVDWRARGAVTAVKDQGQCGSCWAFSAIGNVECQWFLAGHPLTNLSEQMLVSCDK
241   PQDEAQIAAWLAVNGPVAVAVDASSWMTYTGGVMTSCVSEQLDHGVLLVGYNDSAAVPYW
301   IIKNSWTTQWGEEGYIRIAKGSNQCLVKEEASSAVVGGPGPTPEPTTTTTTSAPGPSPSY
361   FVQMSCTDAACIVGCENVTLPTGQCLLTTSGVSAIVTCGAETLTEEVFLTSTHCSGPSVR
421   SSVPLNKCNRLLRGSVEFFCGSSSSGRLADVDRQRRHQPYHSRHRRL
```

Figure 121. Amino Acid Sequence of Preprocruzain. Legend: Green: Signal Peptide; Blue: Propeptide; Red: Mature Cruzain; Purple: C-Terminal.[460]

*Crystal Structure of Cruzain*

The first crystal structure of cruzain was reported McGrath in 1995, when they solved the structure at a 2.35 Å.[483] The monomer is folded into two well defined domains: the L-domain which is mainly α-helical and the R-domain with extended antiparallel β-sheet interactions. Similar to cathepsin L, the catalytic triad is composed of Cys25, His159 and Asn175 (papain numbering). The active site is located at the inner face of the two domains. Cruzain's structure and domains are very similar when compared to papain, the major cysteine protease. Nevertheless, cruzain and papain differ in their loops and turns.

The amino acid sequence shows cruzain possesses eight cysteine residues. For example, Cys153 and Cys200 form a disulfide bond and are located in close proximity to

the active site (less than 2.0 Å). Interestingly, four residues (148, 159, 190 and 202) are isostructural in both cysteine proteases (papain and cruzain).

One of the unique features of cruzain is the presence of one cysteine residue at position 36. (Cys36). To date, there are no other cysteine proteases with this residue located at that position. The thiol group was found to have weak interactions with two carbonyl oxygen atoms (Ala12, Gly32) and one water molecule.

A comparison between papain and cruzain molecules inhibited by synthetic inhibitors reveals similarities between both proteases. For example, McGrath proposed that inactivation occurs by alkylation of the active residue Cysteine 25.

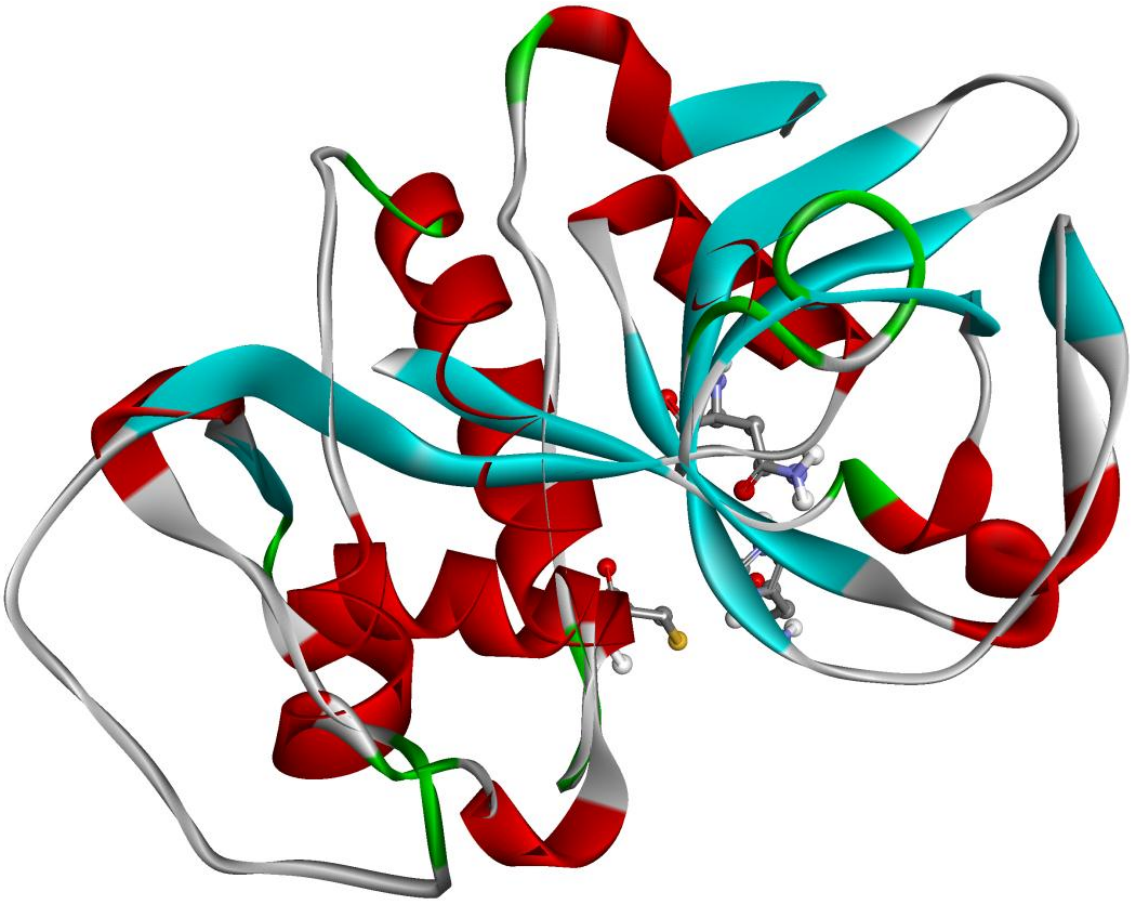


Figure 122. Crystal Structure of Cruzain PDB: 1AIM[479]

## *Inhibitors of Cruzain*

Extensive research related to Chagas' disease is found in recent literature due to the significant number of cases. Thus, a number of synthetic inhibitors have been developed to elucidate possible chemotherapeutic agents against the disease. In this case, a review of cruzain inhibitors will be discussed due to the proteolytic importance of the enzyme in the parasitic disease. Nevertheless, natural inhibitors are also known to be good and potent inhibitors of cruzipain.

### *Natural Inhibitors of Cruzain*

Cruzain is expressed in parasitic organisms as procruzain, similarly to cathepsins L and K, which also belong to the subclan of cysteine proteases. One of the major functions of the propeptide is the prevention of enzymatic activity of cruzain within the parasite. Thus, activation occurs within the parasitic lysosome where cruzain activity is mainly found. Cleavage and activation of the enzyme occur under acidic conditions. Studies showed that this proregion, (103 residues) with a molecular size of 14 kDa, is a potent inhibitor of mature cruzain. Dissociation constant was calculated as low as 18 pM in *in vitro* studies.[484] Similar studies also confirmed that regions of the propeptides can also inhibit mature cruzain as well.[485]

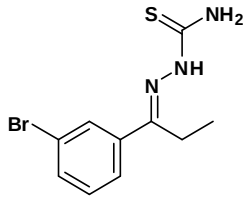
Other natural inhibitors of cruzain can be found within the protozoan. Cystatins, are small-size macromolecules with regulatory functions. For example chagasin is a potent and reversible inhibitor of cruzain.[486] These proteins are excellent cruzain inhibitors with dissociation constants in the picomolar range.

The p41 fragment of the MHC class-II associated molecule is a tight-binding, competitive, reversible inhibitor of cruzain. The affinity for cruzain is clearly seen in its dissociation constant ranging in the picomolar values as well ($K_I$: 58 pM).[487]

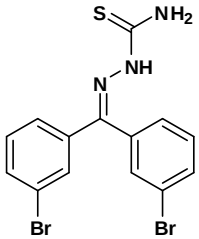
*Synthetic Inhibitors of Cruzain*

A selected group of synthetic inhibitors that have been designed for the inhibition of cruzain can be found in Figure 124.

The discovery of nonpeptidic thiosemicarbazones as cysteine protease inhibitors was reported by Du and coworkers in 2002. They found 3'-bromopropiophenone thiosemicarbazone was able to inhibit cruzain.[8] Similar findings have been also reported by Chiyanzu et al. and Siles and coworkers in independent studies.[9,488] The first collaborative work between Dr. Mary Lynn Trawick and Dr. Kevin G. Pinney reported the synthesis of novel functionalized thiosemicarbazones. The list included naphthalenes, benzophenone, propiophenones, among others. They were the first to report the potency of 3,3'-dibromobenzophenone (**1**) as a cysteine protease inhibitor. They tested it against recombinant cruzain obtaining an $IC_{50}$ value of 24 nM.[3–5,9]

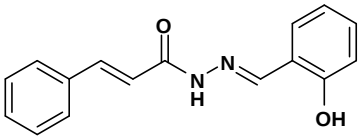



3'-Bromopropiophenone Thiosemicarbazone ($IC_{50}$: 100 nM)[8]

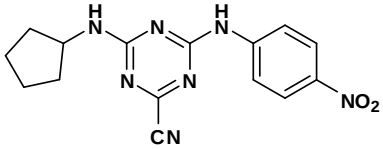



3,3'-Dibromobenzophenone Thiosemicarbazone ($IC_{50}$: 24 nM)[9]

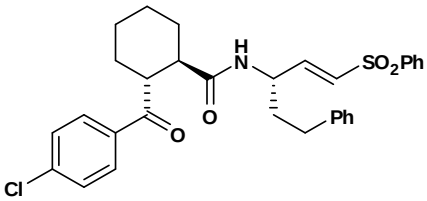
Figure 123. Nonpeptidic Thiosemicarbazone Inhibitors of Cruzain



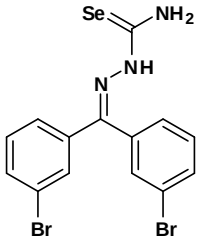
Cinnamic *N*-acylhydrazone ($IC_{50}$: 52 μM)[489]



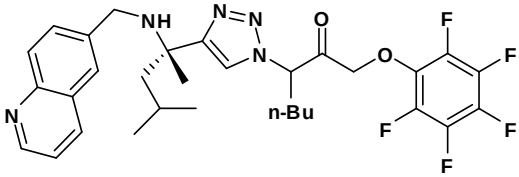
Nitro-substituted triazine ($IC_{50}$: 1 nM)[490]



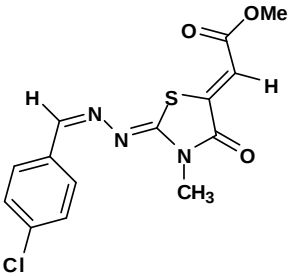
Vinylsulfone Derivative ($IC_{50}$: 50 nM)[491]



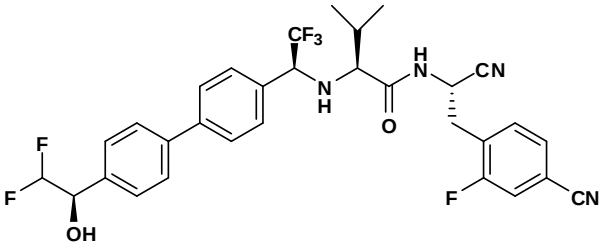
Selenosemicarbazone
(99% inhibition, [I]: 10 μM)[492]



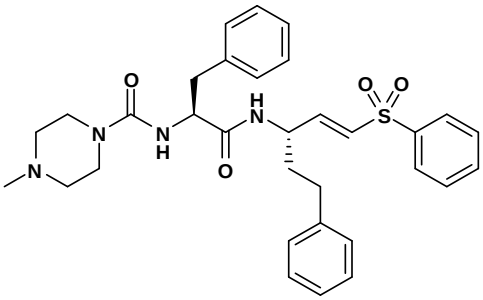
Tetrafluorophenoxymethyl ($K_I$: 100 nM)[493]



Thiazolidinones (75% inhibition, [I]: 100 μM)[494]



Amino nitrile derivative ($IC_{50}$: 0.4 nM)[391]



K777 (vinyl sulfone)[495,496]

Figure 124. Selected Cruzain Inhibitors Reported Between 2009 and 2012

*Material and Methods for the Biological Evaluation of Thiosemicarbazones Derivatives as Inhibitors of Cruzain*

*Materials*

The comprehensive list of materials and equipment is listed in chapter two under the Materials and equipment section. Recombinant cruzain (1.1 μM) was expressed and purified by Dr. Wara M. Arispe and Lauren Adamson under the guidance and supervision of Dr. MaryLynn Trawick.[5] Cruzain DNA plasmid was kindly donated by Elizabeth Hansell and Dr. James McKerrow from the University of California at San Francisco, CA. Thiosemicarbazone derivatives were synthesized by members of Dr. Kevin G. Pinney's laboratory and have been previously reported.[3,6,7,9–12]

*Experimental Section*

*Kinetic Cruzain Assay*

Kinetic analysis of cruzain was carried out by using a Thermo Fluoroskan microplate reader, 3691 96-well black microplates. Volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 10 µl of 35% DMSO solution, 70 µl cruzain stock solution, and 20 µl of Z-FR-AMC stock solution. A mixture containing assay buffer, 35% DMSO, and was preincubated at 25 °C during 5 minutes using 96-black microplates. Triplicate sets of the catalytic activity of cruzain were monitored by adding a concentration of Z-FR-AMC to every reaction. The production of AMC was monitored for 5 minutes at 25 °C using excitation and emission references of 355 and 460 nm respectively. Readings were taken every 10 seconds for five minutes. Reactions were carried out in triplicate. The final concentrations of the kinetic assay are: 100 mM

NaOAc pH 5.5, 1 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 0.1 nM cruzain. Final concentrations of Z-FR-AMC varied between 0.2 and 15 µM.

*Preliminary Inhibition Studies*

Thiosemicarbazone analogs (provided by Dr. Kevin G. Pinney's laboratory) were prescreened to determine if they have inhibitory activity against cruzain.[3,6,7,9–12]

Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 10 µl of 35% DMSO or 10 µl of dilution "1" (final concentration: 10 µM, See Material and Methods of Chapter Two for complete details), 70 µl cruzain stock solutions, and 20 µl of Z-FR-AMC stock solution. The enzyme-inhibitor mixtures (180 µ) assay buffer, 35% DMSO or inhibitor, and cruzain was preincubated at 25 °C during 5 minutes using 96-black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and uninhibited samples were monitored for five min. Final reactions were started by adding 20 µl of Z-FR-AMC. The final concentrations of the preliminary inhibitory reactions were: 100 mM NaOAc pH 5.5, 1 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 0.1 nM cruzain, 10 µM of the screened inhibitor and 15 µM of Z-FR-AMC. Readings were taken every 25 seconds for five minutes and reactions were carried out in triplicate. Compounds that did not have cruzain inhibitory activity more than 50% (i.e. $v_i/v_o \leq 0.5$) were considered as 'inactive' compounds and a general $IC_{50}$ value greater than 10000 nM was assumed. Compounds that inhibited cruzain inhibitory activity more than 50% were further considered for cruzain inhibitory studies and an exact $IC_{50}$was determined.

*Cruzain Inhibition Assay*

Inhibitory analysis of cruzain and its inhibitors was carried out by using a modified protocol of the kinetic cruzain assay, previously described. Total volume of the reaction was 200 µl. Every well contained 100 µl of assay buffer, 10 µl of 35% DMSO or 10 µl of inhibitor dilutions, 70 µl cruzain stock solutions, and 20 µl of Z-FR-AMC stock solution. An 180 µl mixture containing assay buffer, 35% DMSO or inhibitor, and cruzain was preincubated at 25 °C during 5 minutes using 96-black microplates. Reactions were started by adding 20 µl of Z-FR-AMC. The release of AMC by inhibited and uninhibited samples were monitored for five minutes. The final concentrations of the inhibitory cruzain assay are: 100 mM NaOAc pH 5.5, 1 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 0.1 nM cruzain and 15 µM of Z-FR-AMC. Final concentrations of the inhibitors varied between 10 µM and 10 pM. Readings were taken every 25 seconds for five minutes and reactions were carried out in triplicate. Table 135 summarizes final volumes and final conditions for the kinetic and cruzain inhibition assays.

*Construction of AMC Calibration Curve*

The calibration curve obtained for cathepsin L experiments (Chapter 2) was used for similar cruzain experiments. Complete details and results of these experiments can be found in chapter 2 under the subsection "Construction of AMC calibration curve).

*Effect of Inhibitor Concentration on Cruzain Progress Curves*

Final concentrations, conditions, and volumes are similar to the cruzain inhibition assay. Assay buffer, inhibitors (final concentrations varied between 100 nM and 10 µM) and Z-FR-AMC (final concentration: 5 µM) were added to the 96-well black plates

(volume of the substrate-inhibitor mixture: 180 µl). Then, 20 µl of cruzain stock solution were added immediately without preincubation time. Readings were taken every 3 seconds for fifty minutes.

Table 135. Preparation Table for Kinetic, Cruzain Inhibition Assays and Construction of AMC Calibration Curves

| Reagent | Kinetic assay (µl) | Inhibitory assay (µl) | AMC Curve (µl) |
|---|---|---|---|
| Assay Buffer | 100 | 100 | 100 |
| Water | 0 | 0 | 60 |
| Control | 10 | 0 | 0 |
| Inhibitor | 0 | 10 | 0 |
| CZ | 70 | 70 | 20 |
| Z-FR-AMC | 20 | 20 | 20 |
| Total (µl) | 200 | 200 | 200 |

*Effect on Preincubation Studies on Cruzain Inhibition Assays*

Preincubation studies with several compounds were carried out by modifying cruzain inhibition assay. Different sets of mixtures containing assay buffer, inhibitor and cruzain were preincubated at various periods of times between 0 and 240 minutes. Reactions were taken every 25 seconds for five minutes and carried out in triplicate.

*Determination of $K_i^{app}$ Using Morrison's Quadratic Equation*

Data that were obtained on the effect of preincubation studies for numerous compounds were further analyzed. The possibility of these compounds to be tight binding inhibitors was examined by fitting the data by a nonlinear regression using Morrison's quadratic equation.

*Cruzain Reversibility Studies*

Three milliliters of assay buffer for reversibility studies were prepared. Twenty five microliters of cruzain (100X: 100 nM) assay buffer for reversibility studies were pre-incubated with an equal amount of a concentrated solution of the inhibitor (100X: 0.1 $IC_{50}$) at 25 °C between one and four hours. Then, two microliters of the enzyme-inhibitor mixture were rapidly mixed with 198 µl of ASR in order to start the reaction. Total reaction volume was 200 µl. Readings were taken every twenty five seconds for four hours. Final concentrations are similar as described previously. Final conditions were: 100 m*M* NaOAc, pH 5.5, 1 m*M* EDTA, 2.5 m*M* EDTA, 0.01% Brij 35, 0.1 n*M* cruzain and 15 µ*M* of Z-FR-AMC. Table 136 describes required volumes for this experiment.

Table 136. Preparation Table for Cruzain Reversibility Assay

| Reagent | Reversibility studies (µl) |
|---|---|
| Assay Buffer | 198 |
| Inhibitor/CZ | 2 |
| Total | 200 |

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

The effect of [Z-FR-AMC] was studied with compounds **1** and **17**. Minor modifications of the cruzain inhibition assay were carried out. Different sets of mixtures containing assay, inhibitor and cruzain were preincubated at standard preincubation times (5 minutes). Reactions were initiated by the addition of different concentrations of Z-FR-AMC . Reading were taken every 25 seconds for five minutes and carried out in triplicate.

*Effect of Substrate Concentration (Z-FR-AMC) on Cruzain Progress Curves*

Final concentrations, conditions, and volumes are similar to those previously described for the cruzain inhibition assay. Assay buffer, one inhibitor concentration (Final concentration for **10**: 100 nM) and Z-FR-AMC (final concentrations varied between 0.5 and 15 μM) were added to the 96-well black plates (volume of the substrate-inhibitor mixture: 180 μl). Then, 20 μl of cruzain stock solution was added immediately without preincubating. Readings were taken every 3 seconds for fifty minutes.

*Inhibition of Cruzain Collagenase Activity by Thiosemicarbazone Derivatives*

A solution of human type I collagen ([ACI]: 0.4 mg/ml) and cruzain stock solution (CLI) were prepared. The inhibitor stock solution was prepared by mixing 1.04 μl of 20 m*M* stock solution of the inhibitor in DMSO, 19.8 μl of DMSO and 179.2 μl of water. Conditions of this stock solution are: 10.4% DMSO and 104 μM of the inhibitor. The experiment was started by preincubating 3 μl of CZI and 2.5 μl of the inhibitor stock solution in a microcentrifuge tube at 37 °C for 0.5 hours. Similarly, another group of control samples (without the inhibitor) was also set up. A 1X staining solution was prepared by diluting 10 μl of 5000X SYPRO® red protein gel staining dye with 49.9 ml of 7.5% acetic acid. Then, reactions were initiated by adding 7.5 μl of type I collagen and monitored between 0 and 420 minutes at 37 °C. Final conditions of the reactions were: 100 m*M* NaOAc pH 5.5, 1 m*M* EDTA, 2.5 m*M* DTT, 20 n*M* cruzain, 2% DMSO, 20 μ*M* of the tested inhibitor, and 0.01 mg/ml type I collagen. Reactions were stopped by adding 2 μl of LDS NuPAGE® sample buffer, heated at 90 degrees for ten minutes and immediately stored at -80°C. Inactivated samples (15 μl) were loaded onto 4-12 % NuPAGE® Bis-Tris gels. Electrophoresis was carried out by using MOPS running buffer

at 200 V for 50 minutes. Gels were stained in 50 ml of 1X SYPRO® staining solution at room temperature between one and two hours followed by a destaining process using water (1X) and 7.5% acetic acid (1X). Finally, gel imaging was performed with a GE Typhoon 9400 FL with excitation and emission wavelengths of 550 and 630 nm, respectively. Table 137 describes conditions for inhibition of collagenase activity studies.

Table 137. Preparation Table for Cruzain Collagenase Activity Assay

| Reagent | Control (μl) | Inhibited (μl) |
|---|---|---|
| CZI | 3.0 | 3.0 |
| Control (Water) | 2.5 | 0 |
| Inhibitor | 0 | 2.5 |
| TIC | 7.5 | 7.5 |
| Preincubation time (h) | 0.5 | |
| Total | 13 | 13 |

*Molecular Modeling Studies*

Computational studies were performed using Discovery Studio 3.0. First, a crystal structure was obtained from the PDB database. The chosen molecule was 1ME3.[131]

Similar experiments were performed with this molecule by Chen.[4] His results were used as a comparison point to verify the accuracy of the results.

*Preparation of the Protein*

The selected crystal structure is related to the analysis of recombinant cruzain bound to a synthetic ketone inhibitor. Therefore, the protein needed to be prepared and validated using a high affinity substrate as reported.[131] Water molecules and other ligands (synthetic ketones) were removed. The structure was then examined in order to

correct possible structural disorders, protein residue connectivity, bond-orders, and missing side-chain or backbone atoms and to correct the predicted pKa values (This option was not performed due to limited options in licensing). The pH of the system was also changed in order to modify the protonation state of cruzain's termini and side chain residues. The pH value was set to 6.8. Finally, the binding site was defined according to default specifications from PDB records. This was done in order to create a spherical binding site object. The binding site appears as a red transparent red sphere.

*Preparation of the Ligand (Thiosemicarbazones)*

Seven molecules (**1**, **8**, **9**, **10**, **17**, **36** and **58**) were selected to be modeled with cruzain in order to verify their mode of inhibition and major interactions. Analog **1** was used as a comparison because it was previously modeled with cruzain.[4] Each molecule was drawn in ChemDraw 6.0 and copied into Discovery Studio Client 3.0. Each molecule was prepared as 'ligands'. For each ligand, changes in ionization states, and canonical tautomers were allowed to be carried out. At the end of the run, a list of tautomers was generated for each compound.

*Docking Simulations*

CDOCKER, a dock ligand protocol was used to docking the selected ligands into the macromolecule. The algorithm allows the simulation of several ligands with a single receptor protein simultaneously. The input receptor was the cruzain molecule (1ME3) and the ligands were the entire list of tautomers for each compound. The total number of top hits was set up to 200. The site sphere coordinates were set: 5.1024, 9.25127, 6.43005, 8.4. These coordinates were set according to the PDB files default conditions

and they established the permitted area where both ligand and receptor (thiosemicarbazones and cruzain respectively) could interact. The number of 'orientations to refine' was set to 20. Similarly, the 'maximum bad orientation' and 'orientation vdW Energy Threshold numbers were 800 and 800 respectively. The selected forcefield was CHARMm and the ligand partial charge method was changed to MMFF. CHARMm was used due to its extensive use when modeling organic molecules with proven accuracy and works quite well with a variety of solvents used in *in vitro* studies in several receptor-ligand studies where ligands are synthetic molecules. Similarly, MMFF is type of forcefield derived from experimental calculations. MMFF is extensively used in pharmaceutical industries where for studies involving changes in conformation energies and nonbonded interactions. MMFF also gives accurate results for a wide range of organic molecules but fails when parallel processing is required. Docking simulations took several minutes to be performed. For each compound, the number of conformations varied from a couple of dozen to hundreds to possible conformations. These conformations were arranged based on the interaction energies to localize the conformations with highest interaction energies for every compound. A visible conformation was selected in order to observe interactions between cruzain and its inhibitors. Finally, a selection filtering residues having interactions with every ligand was set. For each modeling, only residues that were located less than 5.0 Å were shown. Distances for selected atoms were selected and calculated. Also, hydrogen bonds in that region were chosen and selected. Discovery Studio 3.0 Client set hydrogen bonds less than 2.5 Å, but this distance was also modified depending on each case.

*Results and Discussion*

Previous research has indicated that some thiosemicarbazones are lead compounds as cruzain inhibitors.[9] Thus, new generations of thiosemicarbazones required a complete screening to verify their potency. The Structure-Relationship Activity (SAR) was characterized after completing *in vitro* testing against cruzain ($IC_{50}$ values). The project was completed in collaboration between Dr Kevin G Pinney and Dr. Mary Lynn Trawick groups. Synthetic nonpeptidic thiosemicarbazones were provided by members of the laboratory of Dr. Kevin G. Pinney at Baylor University.[3,6,7,9–12]

Fluorometric based assays were utilized to study various assay parameters in inhibitory activities, determination of $K_I$, and reversibility of thiosemicarbazone inhibitors. To understand the interactions between these inhibitors and cruzain, molecular modeling of thiosemicarbazones was used. These studies will also aid in the understanding of developing inhibitory moieties that could be explored in combination with existing inhibitors. Cruzain studies were carried out using a 96-well microplate fluorometric based assay. Z-FR-AMC is a fluorogenic substrate that has been used with serine and cysteine proteases. However, Salvati and Ascenzi observed substrate inhibition when [Z-FR-AMC] > 15 μM.[497,498] Preliminary experiments also confirmed their findings and substrate concentration was set to 15 μM. Final conditions for the cruzain fluorometric assay are: 100 mM NaOAc pH 5.5, 1 mM EDTA, 2.5 mM DTT, 0.01% Brij 35, 2% DMSO, 0.1 nM cruzain, and 0.5 – 15 μM Z-FRM-AMC.

*Assay Optimization. Effect of DMSO on Cruzain Inhibition Assays*

Siles and coworkers reported an $IC_{50}$ value of 24 nM for analog **1**.[9] The effect of DMSO during enzymatic catalysis was also evaluated using cruzain as our molecular

target. Experiments were conducted using analog **1**, by increasing the concentration of DMSO in solution. The $IC_{50}$ of compound **1** was revisited and it was found to be between 10 and 12 nM (See Table 138).

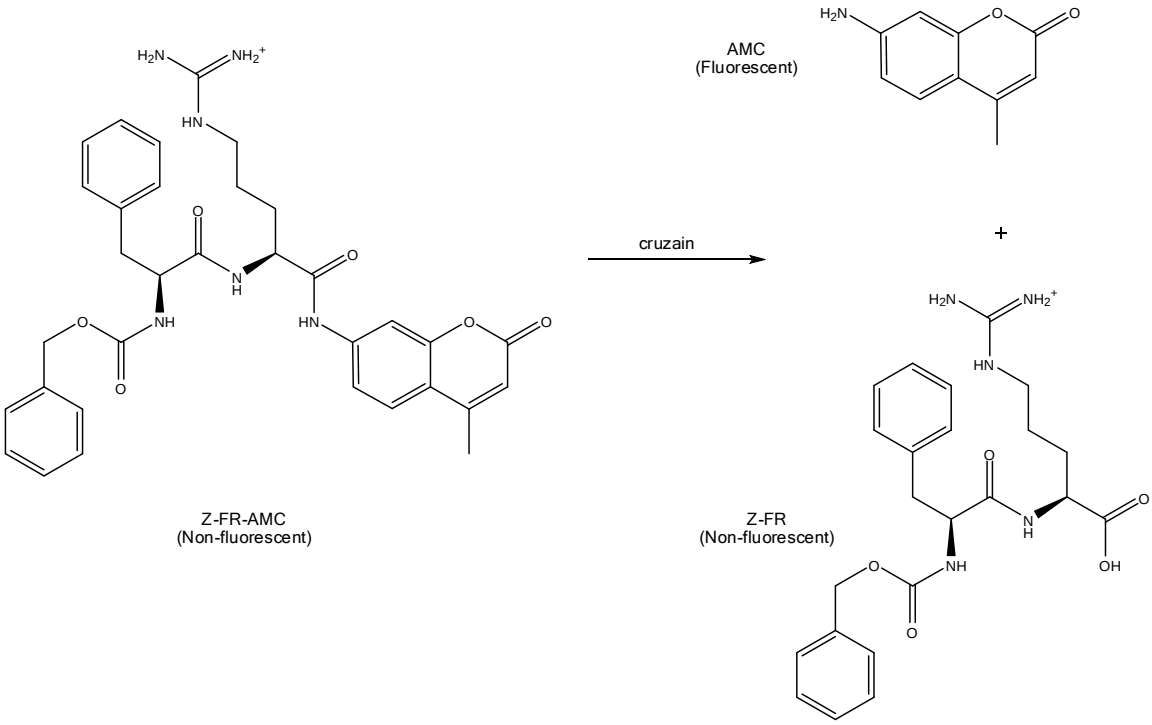
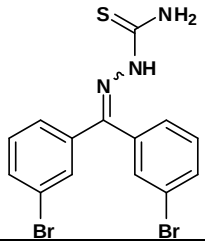



Figure 125. Hydrolysis of the Z-FR-AMC Using cruzain

Table 138. Effect of DMSO Concentration on Cruzain Inhibition Studies

| DMSO (%) |  |
|---|---|
| | **1** ($IC_{50}$, nM) |
| 0.7 | 24[9] |
| 2 | 12 |

The obtained results showed the effect of DMSO in the inhibition of cruzain. Solubility of the lead compound (**1**) was increased upon addition of DMSO and their inhibitory activity was enhanced accordingly. An increase in cruzain activity and stability was also observed when DMSO concentration was increased. The efficiency of **1** is 2 times better when comparing its $IC_{50}$ at 0.7 %.

*Determination of $K_M$, $V_{MAX}$ and $k_{CAT}$*

Determination of kinetic constants was based on the analysis assumption that enzymatic activity of cruzain is substrate-concentration dependent. Thus, a high concentration of the substrate (up to 15-fold $K_M$ value) was used for studies related to cruzain. Cruzain catalytic activity showed a linear behavior when 0.5 μM Z-FR-AMC was used for reaction times that were at least 5 minutes long. The determination of $K_M$, $V_{MAX}$ and $k_{CAT}$ was possible with experiments that observe the catalytic activity of a fixed concentration of cruzain (0.1 nM) but vary the concentration of Z-FR-AMC (0.2 -15 μM). Experiments were carried out in triplicate. Catalytic rates were calculated by applying linear regression of the data. ([AMC] is the dependent variable and time (seconds) is the independent variable). Then, a nonlinear regression analysis of the Michaelis-Menten equation (Eq. 1.1) was performed to calculate $K_M$ and $V_{MAX}$ values with the aid of commercially available software (GraphPad 5.0). The $k_{CAT}$ constant value was determined using equation 4.1. The parameter $v_o$ is the initial rate velocity at a specific substrate concentration. The $V_{MAX}$ is the maximum velocity, $K_M$ is the Michaelis-Menten constant, [S] is the substrate concentration, and $k_{CAT}$ is the catalytic rate constant of the reaction.

$$v_o = \frac{V_{MAX}[S]}{K_M + [S]} \quad (1.1) \qquad k_{CAT} = \frac{V_{MAX}}{[cruzain]} \quad (4.1)$$

$K_M$, $v_{MAX}$ and $k_{CAT}$ values were found to be 1.01 ± 0.1 μM, 1.73 ± 0.08 AMC nM/seconds and 17.3 $s^{-1}$, respectively. These values are similarly to those reported in the literature ($K_M$: 0.96 μM).[499]

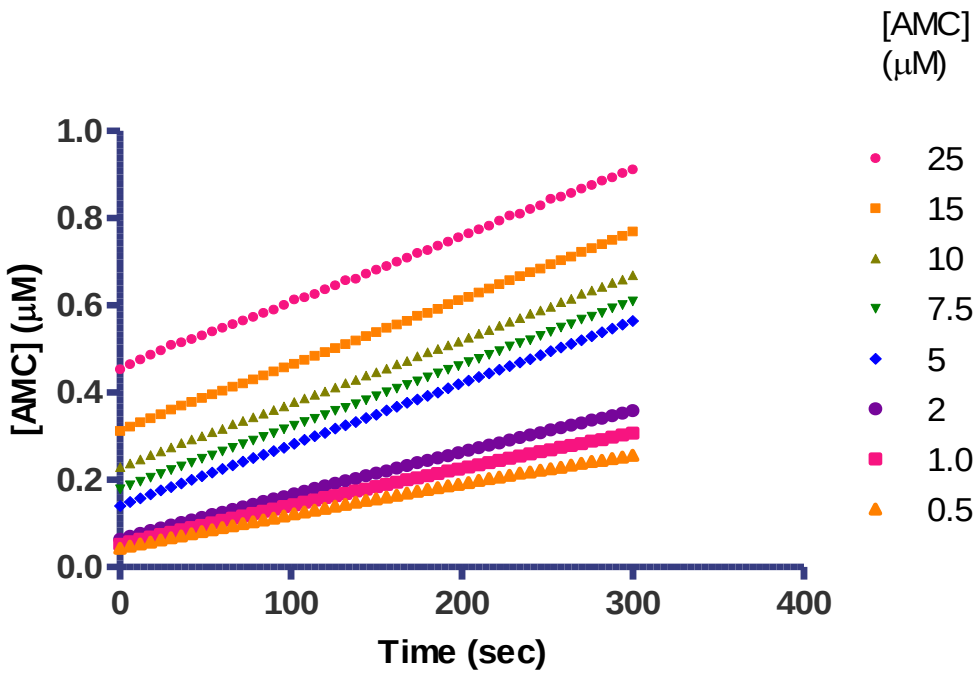



Figure 126. Catalytic Activity of Cruzain Using Z-FR-AMC as the Fluorogenic Substrate

*Determination of Inhibitory Efficacy of Thiosemicarbazone Analogs at 10 μM*

The first set of experiments that were performed to verify if the individual compound of the library of thiosemicarbazones could be potential cruzain inhibitors. If the catalytic activity of cruzain was inhibited by 50% or more by a fixed concentration (Final concentration: 10 μM) of the potential inhibitor. Three independent sets of experiments of untreated ([I]: 0 μM) and treated samples ([I]:10 μM) were preincubated with 0.1 nM cruzain for 5 minutes at 25 °C. Inhibitory activities were monitored when

reactions were started by adding 15 μM Z-FR-AMC as a fluorogenic substrate. Reactions demonstrated a linear behavior for at least the first five minutes. Active compounds were further analyzed to determine an exact $IC_{50}$ value. If the ratio $v_i/v_o \leq 0.5$, then the compounds were not considered potential inhibitors and an approximated $IC_{50}$ value ≤ 10000 nM was assigned to them.

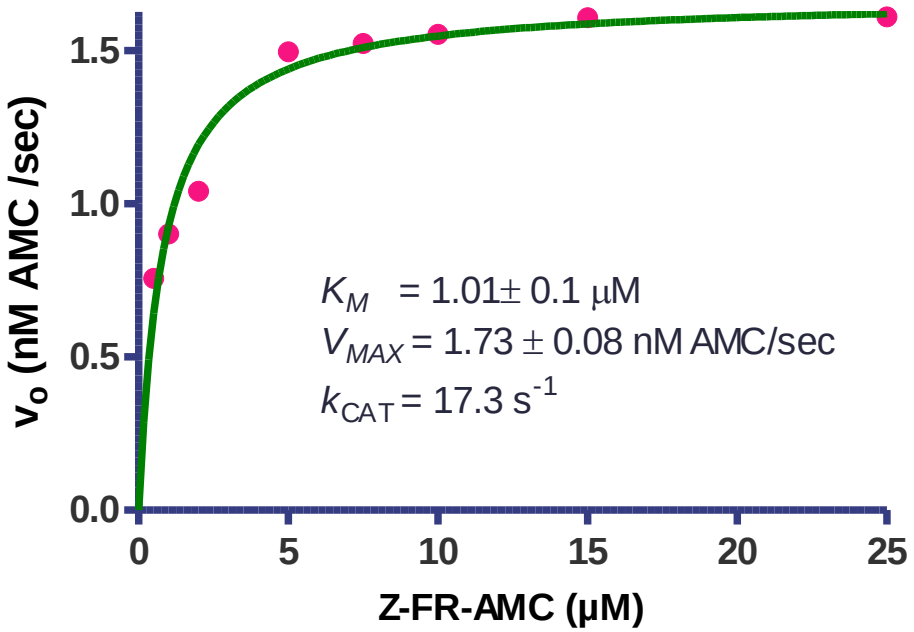



Figure 127. Determination of $K_M$ and $V_{MAX}$ for Cruzain Using Z-FR-AMC as a Fluorogenic Substrate

*Determination of $IC_{50}$ values*

A section of the library comprising more than sixty compounds was analyzed to verify their efficacy to inhibit the catalytic activity of cruzain. The synthetic compounds were synthesized through a collaborative project under the guidance of Dr. Kevin G. Pinney by several members of his research group.[9–12] A 96-well microplate fluorometric based assay was utilized to determine the inhibitory activity of each of these inhibitors. Uninhibited cruzain catalytic activity showed linear behavior when 15 μM Z-FR-AMC

was used for reactions times that were 5 minutes long. The determination of the $IC_{50}$ values was carried out with experiments that observe the inhibitory capacity of the synthetic compounds when a fixed concentration of cruzain (0.1 nM) was preincubated for 5 minutes at 25 °C. The final concentration of each compound varied between 10 pM and 10 μM. Three experiments were carried out for each tested inhibitor. Catalytic rates of uninhibited and inhibited samples were calculated by linear regression of the data. ([AMC]: dependent variable and time (seconds): the independent variable). The data followed a typical sigmoidal dose response, and therefore, a nonlinear regression of the equation 1.2 was performed to calculate $IC_{50}$ values with the aid of commercially available software (GraphPad 5.0). The value Y represents the inhibited activity (normalized relative to control, X is log([inhibitor]) in *M*. The velocities $v_{min}$ and $v_{max}$ represent when cruzain was preincubated with the highest and lowest inhibitor concentrations respectively (10 μM and 0 pM or control). The Hillslope value is the slope of the sigmoidal curve. $IC_{50}$ ± S.E. values represent the average and standard errors of at least three independent experiments. A complete structure-activity relationship (SAR) is shown in Table 139 to Table 147 that group these inhibitors by functional groups or families.

$$Y = \frac{v_{MIN} + (v_{MAX} - v_{MIN})}{1 + 10^{(\log(IC_{50} - X) * Hillslope)}} \quad (1.2)$$

Table 139. Inhibition of Cruzain by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12]

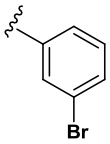
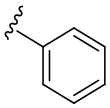
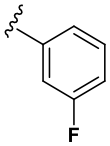
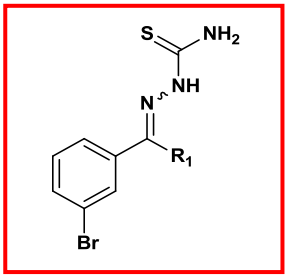
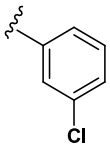
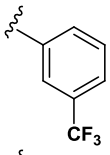
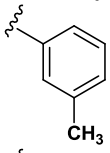
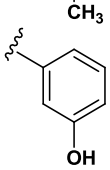
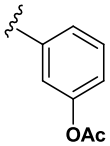
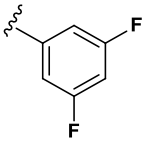
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E, (nM)* |
|---|---|---|---|
| **1** | |  | 10.61 ± 0.4 |
| **2** | |  | 73.0 ± 4.7 |
| **3**[8] | |  | 221.7 ± 8.98 |
| **4** | |  | 110.8 ± 7.4 |
| **5** |  |  | 30.2 ± 2.8 |
| **6** | |  | 64.4 ± 2.9 |
| **7** | |  | 94.4 ± 6.5 |
| **8** | |  | 211.9 ± 16.4 |
| **9** | |  | 12.1 ± 0.2 |
| **10** | |  | 17.6 ± 1.5 |

Table 139. Inhibition of Cruzain by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

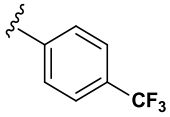
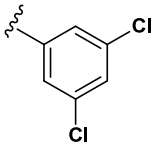
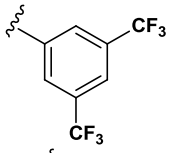
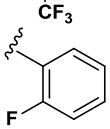
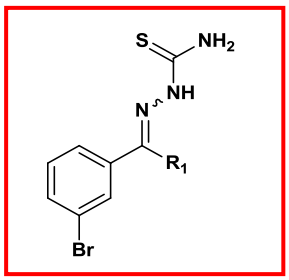
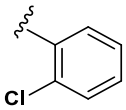
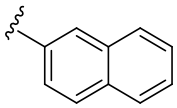
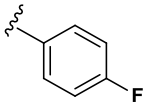
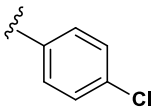
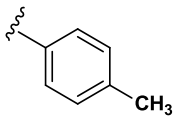
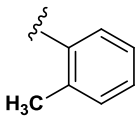
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **11** | |  | 637.5 ± 47.3 |
| **12** | |  | 97.2 ± 8.9 |
| **13** | |  | 678.2 ± 69.8 |
| **14** | |  | 70.37 ± 6.4 |
| **15** |  |  | 182.4 ± 11.7 |
| **16** | |  | ND |
| **17** | |  | 14.3 ± 1.5 |
| **18** | |  | 275.3 ± 27.3 |
| **19** | |  | 526.9 ± 33.1 |
| **20** | |  | 770.5 ± 176.8 |

Table 139. Inhibition of Cruzain by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

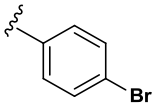
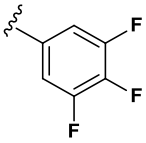
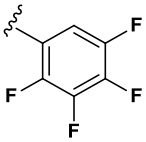
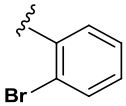
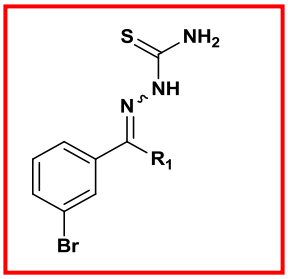
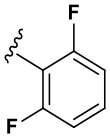
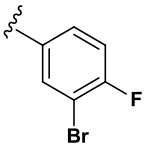
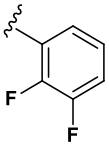
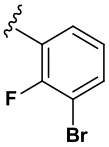
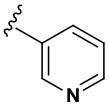
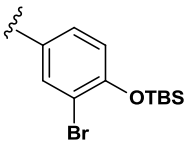
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.), ( nM)* |
|---|---|---|---|
| **21** | |  | 127.0 ± 4.5 |
| **22** | |  | 134.9 ± 5.2 |
| **23** | |  | 83.7 ± 5.6 |
| **24** | |  | ND |
| **25** |  |  | 266.8 ± 21.7 |
| **26** | |  | 47.7 ± 4.6 |
| **27** | |  | 34.2 ± 2.8 |
| **28** | |  | 113.6 ± 7.2 |
| **29** | |  | ND |
| **30** | |  | 64.4 ± 4.1 |

Table 139. Inhibition of Cruzain by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

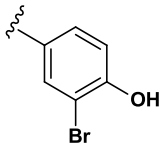
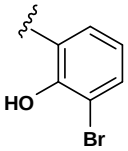
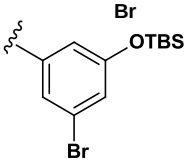
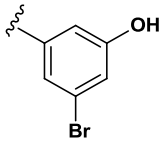
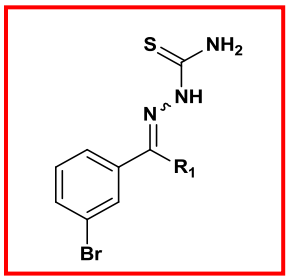
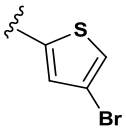
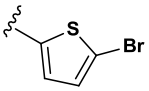
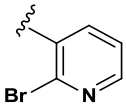
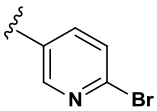
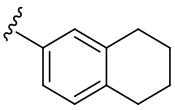
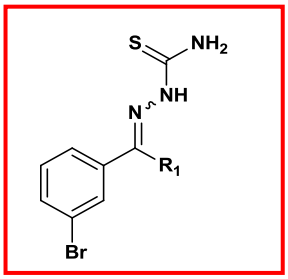
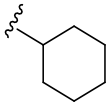
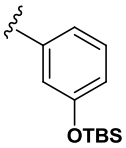
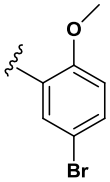
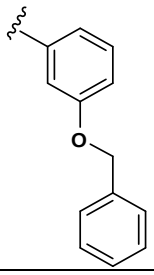
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **31** | |  | ND |
| **32** | |  | ND |
| **33** | |  | ND |
| **34** | |  | 97.8 ± 8.3 |
| **35** |  |  | 100.1 ± 2.4 |
| **36** | |  | 20.2 ± 0.14 |
| **37** | |  | 31.6 ± 1.6 |
| **38** | |  | ND |
| **39** | |  | ≥ 10000 |
| **40** | |  | ND |

Table 139. Inhibition of Cruzain by Thiosemicarbazones Containing a *meta*-Bromophenyl Substituent Group. For Synthesis of Compounds: [3,6,7,10–12](Continued)

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E., (nM)* |
|---|---|---|---|
| **41** |  |  | ≥ 10000 |
| **42** | |  | ND |
| **43** | |  | ND |
| **44** | |  | 46.0 ± 2.1 |

*ND: Not determined

Table 140. Inhibition of Cruzain by *para*-Bromo Functionalized Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

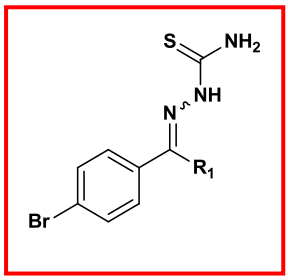
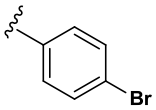
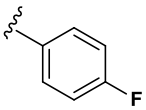
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **47** |  |  | ≥ 10000 |
| **48** | |  | 1453 ± 130.5 |

Table 141. Inhibition of Cruzain by *para*-Bromo Functionalized Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

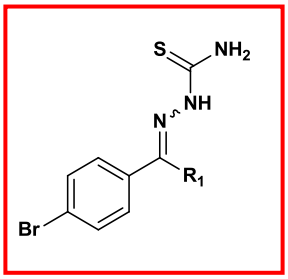
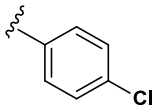
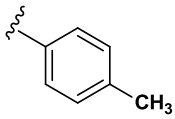
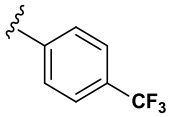
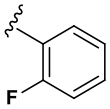
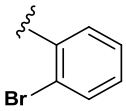
| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **49** |  |  | ≥ 10000 |
| **50** | |  | 1970 ± 151.1 |
| **51** | |  | 2262 ± 247.9 |
| **52** | |  | 466.4 ± 35.8 |
| **53** | |  | ≥ 10000 |

Table 142. Inhibition of Cruzain by Dihalogen-substitutedBenzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

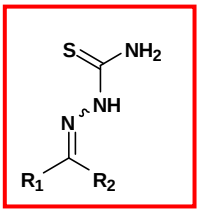
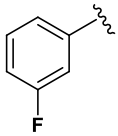
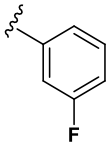
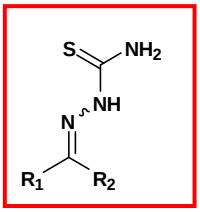
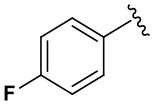
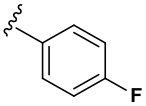
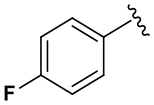
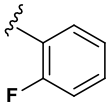
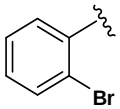
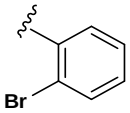
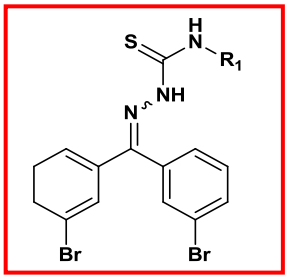
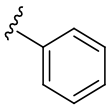
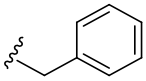
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **54** |  |  | 484.3 ± 46.5 |

Table 143. Inhibition of Cruzain by Dihalogen-substitutedBenzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12] (Continued)

| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **55** |  |  | ≥ 10000 |
| **56** |  |  | 494.4 ± 40.9 |
| **57** |  |  | ND |

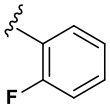
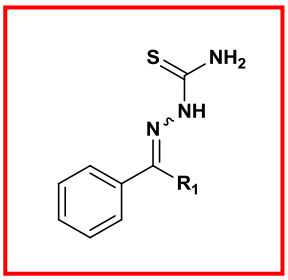
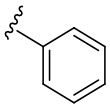
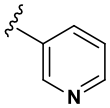
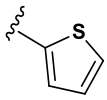
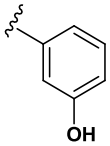
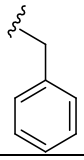
*ND: Not Determined

Table 144. Inhibition of Cruzain by Dibromo-*N*-Substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **58** |  |  | ≥ 10000 |
| **59** | |  | ≥ 10000 |
| **60** | |  | ND |

*ND: Not determined

Table 145. Inhibition of Cruzain by Thiosemicarbazones Containing a Phenyl Group. For Synthesis of Compounds: [3,6,7,10–12]

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **61** | |  | ND |
| **62** |  |  | ≥ 10000 |
| **63** | |  | ≥ 10000 |
| **64** | |  | ≥ 10000 |
| **65** | |  | ≥ 10000 |
| **66** | |  | ≥ 10000 |

*ND: Not Determined

Table 146. Inhibition of Cruzain by Substituted Benzophenone Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

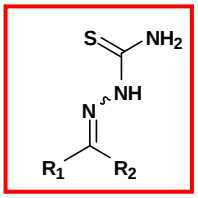
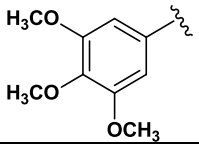
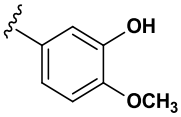
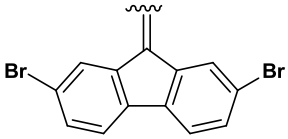
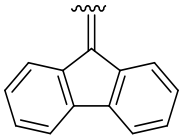
| *Compound* | *$R_1$* | *$R_2$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| |  | | |
| **75** |  |  | 96.0 ± 1.9 |

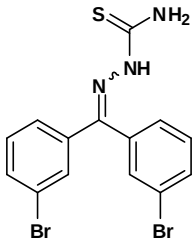
Table 147. Inhibition of Cruzain by Functionalized Fluorene Thiosemicarbazones. For Synthesis of Compounds: [3,6,7,10–12]

| *Compound* | *Structure* | *$R_1$* | *$IC_{50}$ ± S.E.,(nM)* |
|---|---|---|---|
| **80** |  |  | ≥ 10000 |
| **81** | |  | ≥ 10000 |

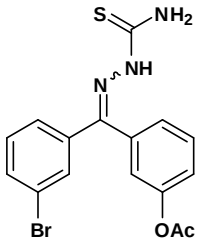
*Structure-Activity Relationship (SAR) of Thiosemicarbazones as Cruzain Inhibitors*

More than 60 thiosemicarbazone (TSC) synthetic compounds were analyzed during the preliminary screening in order to determinate the potency of these compounds. A considerable number of compounds showed a significant inhibitory potency, and $IC_{50}$ values were calculated. A selected group of five different compounds were shown to be outstanding cruzain inhibitors with $IC_{50}$ between 10 and 20 nM. More than half of the screened inhibitors show good inhibitory activities toward cruzain ($21 \leq IC_{50} \leq 1000$ nM). One fifth of the library (36 analogs) had a moderate activity. Finally, eighteen compounds show low or no cruzain inhibitory activity ($IC_{50} \geq 1000$ nM).

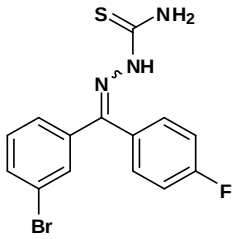
*Potent cruzain inhibitors*. Five thiosemicarbazones showed very potent inhibitory activities against the cysteine protease. Figure 128 summarizes the structures of the compounds with $IC_{50}$ values less or equal to 20 nM. Four analogs belong to the subfamily of *meta*brominated benzophenone TSC analogs (**1**, **9**, **17**, **10**). Bromine, acetate and fluorine moieties are the substituents in this subfamily. The fifth analog is a brominated thiophene thiosemicarbazone TSC (**36**).



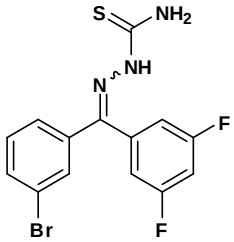
(**1**) CZ: 10.6 nM
CL: 16.7 nM



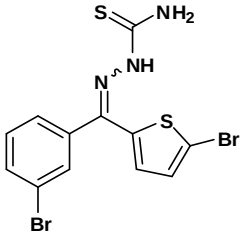
(**9**) CZ: 12.1 nM
CL: 150.8 nM



(**17**) CZ: 14.3 nM
CL: 79.6 nM

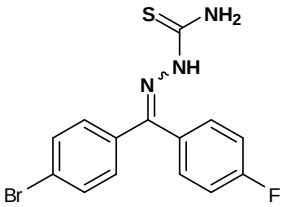


(**10**) CZ: 17.7 nM
CL: 59.4 nM

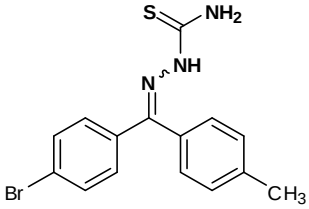


(**36**) CZ: 20.2 nM
CL: 931.3 nM

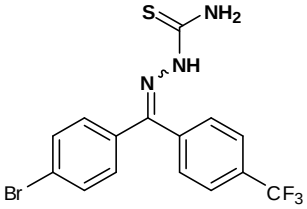
Figure 128. Thiosemicarbazone analogs with Potent Inhibitory Activity against Cruzain[9–12]



(**48**) CZ: 1453 nM
CL: 3320 nM



(**50**) CZ: 1970 nM
CL: 2828 nM



(**49**) $IC_{50}$: 2262 nM
CL ≥ 10000 nM

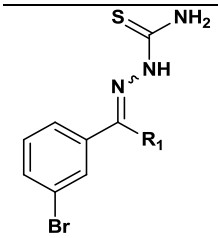
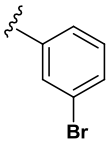
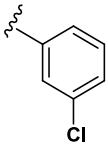
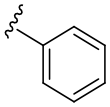
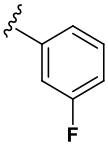
Figure 129. Thiosemicarbazones with Low Activity toward Cruzain

*General remarks of the Structure-Activity Relationship*

Further analysis of the structure-activity relationship revealed that thiosemicarbazone inhibitory activity can be enhanced or reduced dramatically by moiety substitution. The series of substituted *meta*brominated thiosemicarbazones is the largest subgroup among the compounds that were tested. It consists of forty-four different

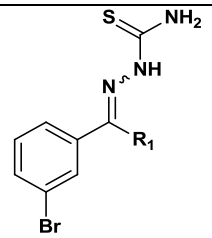
analogs with inhibitory activities that varied between 16.7 and ≥ 10000 nM. Halogenated substituents greatly enhanced the activity of the thiosemicarbazones. Compounds **1**, **2**, **4**, **5** (Table 148) are thiosemicarbazones sharing the main moiety (3-bromobenzophenone thiosemicarbazone) but varying the 3'-halogenated substituent. Compound **2** is used as a comparison to complete the series. All of these compounds possessed good inhibitory activity with $IC_{50}$ values less than 150 nM. The brominated analog is 10-fold more active than its fluorinated analog. Interestingly, the monobromobenzophenone thiosemicarbazone (**2**) showed good activity against the parasitic enzyme when compared to cathepsin L (See Table 148). Increased solubility and and lower enzyme concentration may contribute to its activity. The dependence of $IC_{50}$ as a function of enzyme concentration is also known in literature. Shoited reported that a 10-fold increase in enzyme concentration during *in vitro* assays enhances the potency of tetraiodophenolphtalein as an AmpC inhibitors by a factor of five.[500]

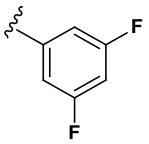
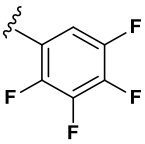
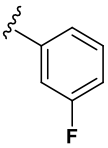
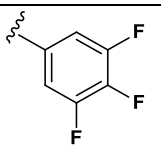
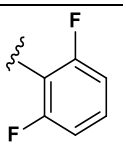
Table 148. Inhibition of Cruzain by 3-Bromo-3'-Halogen Benzophenone Thiosemicarbazones

| | *$R_1$* | | | |
|---|---|---|---|---|
|  |  |  | <  | <  |
| $IC_{50}$ (nM) | **1** | **5** | **2** | **4** |
| CZ | 10.2 | 30.2 | 73.0 | 110.8 |
| CL | 16.7 | 131.4 | ≥ 10000 | 250.3 |

Five fluorinated-substituted benzophenone TSC analogs (**4**, **10**, **22**, **23**, and **25**; Table 149(Table 149) were also compared. Polyfluorination of the benzyl rings proves that cruzain-inhibitory activity of the compounds can be increased with the addition of electronegative substituents. However, a deeper comparison between the two difluoro benzophenone TSCs also reveals that position of the substitutions plays a key role in the activity of these analogs. The *ortho*-difluoro analog is 14 times less active than the di*meta* counterpart. The effect of other substituents was also evaluated (Table 150). Various groups were substituted at the *meta* position of one of the phenyl rings. The group includes trifluoromethyl, methyl, hydroxyl, and acetate moieties. Two moieties showed good activities compared to the acetate analog, which is a potent inhibitor of cruzain with an $IC_{50}$ value less than 20 nM (Compounds **6** and **7**).

Table 149. Inhibition of Cruzain by 3-Bromo-poly'-fluoro-Benzophenone Thiosemicarbazones




| $R_1$ |  | > |  | > |  | > |  | > |  |
|---|---|---|---|---|---|---|---|---|---|
| $IC_{50}$ (nM) | **10** | | **23** | | **4** | | **22** | | **25** |
| CZ | 17.6 | | 22.6 | | 104.2 | | 134.9 | | 266.8 |
| CL | 59.4 | | 63.2 | | 250.3 | | 118.1 | | 609.7 |

A small trifluoromethyl series of three compounds (**6**, **11**, and **13**; Table 151) differs in their inhibition activity when cruzain and cathepsin L $IC_{50}$ values are compared.

Only **6,** the *meta*-substituted analog, showed good inhibition values towards cruzain. Its inhibitory activity is 10-fold higher when compared with **13.**

Table 150. Inhibition of Cruzain by 3-Bromo-3'-Heteroatomic Groups Benzophenone Thiosemicarbazones

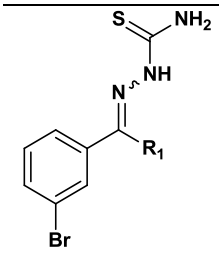
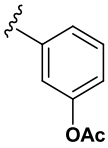
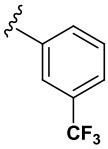
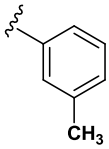
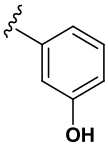
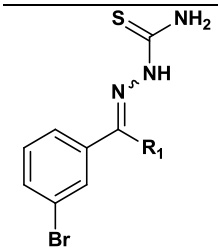
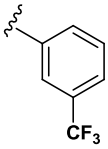
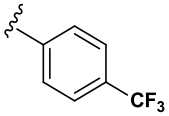
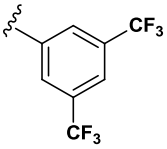
| | *R1* | | | |
|---|---|---|---|---|
|  |  | <  | <  | <  |
| $IC_{50}$ (nM) | **9** | **6** | **7** | **8** |
| CZ | 12.1 | 64.4 | 94.4 | 211.9 |
| CL | 150.8 | 46.5 | 224.4 | 131.4 |

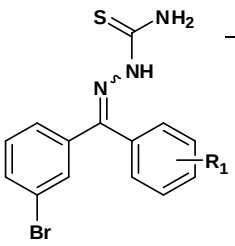
Table 151. Inhibition of Cruzain by 3-Bromo-Trifluoromethyl Benzophenone Thiosemicarbazones

| | *R1* | | |
|---|---|---|---|
|  |  | <  | <  |
| $IC_{50}$ (nM) | **6** | **11** | **13** |
| CZ | 64.4 | 637.5 | 678.2 |
| CL | 46.5 | 520.9 | 96.0 |

However, the effect of position of the substituent can be analyzed in depth with the series of twelve 3-bromobenzophenone thiosemicarbazones that are also monosubstituted with halogen and aliphatic groups. (F, Cl, Br, and $CH_3$). Table 152 compares $IC_{50}$ values for these compounds. The inhibitory activity of the compounds increases with more electronegative substituents if they are positioned in the *ortho*-position. This trend is not observed for the *meta* and *para* positions on the second phenyl

ring. Furthermore, the *meta* brominated analog is the most potent inhibitor found in this library. In general, *meta*substituted TSCs analogs are better inhibitors than *ortho* and *para*-substituted TSCs. Interestingly, **17** is the only *para*-halogenated analog with excellent inhibition properties.

Table 152. Effect of the Position of Substituents in the Inhibitory Activity of 3-Bromo-Benzophenone Thiosemicarbazones

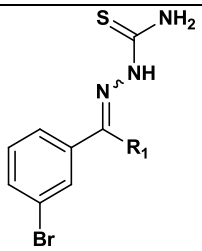
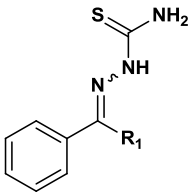
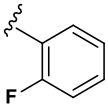
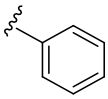
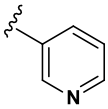
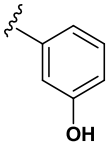



| $R_1$ | | | $IC_{50}$ (nM) *ortho* | | | *meta* | | | *para* |
|---|---|---|---|---|---|---|---|---|---|
| -F | **14** | CZ<br>CL | 70.4<br>23.9 | **4** | CZ<br>CL | 110.8<br>250.3 | **17** | CZ<br>CL | 14.3<br>79.6 |
| -Cl | **15** | CZ<br>CL | 182.4<br>1610 | **5** | CZ<br>CL | 30.2<br>131.4 | **18** | CZ<br>CL | 275.3<br>327.1 |
| -Br | **24** | CZ<br>CL | ND<br>2600 | **1** | CZ<br>CL | 10.6<br>16.7 | **21** | CZ<br>CL | 127.0<br>≥ 10000 |
| $CH_3$ | **20** | CZ<br>CL | 770.5<br>≥ 10000 | **7** | CZ<br>CL | 94.4<br>224.4 | **19** | CZ<br>CL | 526.9<br>2156 |

Compounds **61-65 (**Table 153**)** also demonstrated the importance of the presence of bromine in one of the phenyl rings in order to enhance the potency of thiosemicarbazones. An examination of unsubstituted and bromo-substituted analogs unsubstituted and halogenated benzophenones, phenols, pyridines, and thiophenes; revealed than none of the unsubstituted compounds showed significant inhibitory activity. Bromination in the 3 position of one of the phenyl rings of the benzophenone group is also critical for the activity of TSC as potential cruzain inhibitors. A closer examination of halogenated compounds with similar chemical structures reveals that a different halogen substituent, such as fluorine, greatly reduces the activity of the compounds. The 3,3'-Difluorobenzophenone TSC (**54**) is almost 48 times less active than 3,3'-dibromobenzophenone TSC (**1**). Furthermore, both 4,4'-difluorobenzophenone

TSC (**55**) and 4,4'-dibromobenzophenone TSC (**47**) showed no activity against cruzain, showing that *para*-disubstitution with halogen substituents cannot be used as a possible route for the design of similar compounds.

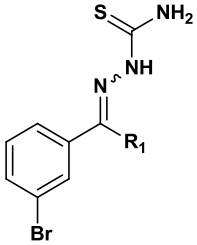
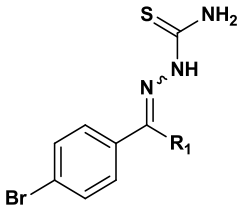
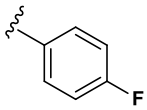
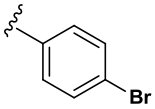
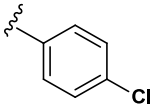
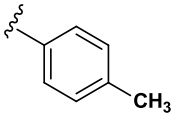
Table 153. Comparison between Brominated and Unbrominated Benzophenone Thiosemicarbazones for Cruzain

| $R_1$ | |  | | |  | |
|---|---|---|---|---|---|---|
| | | | $IC_{50}$ (nM) | | | |
|  | **14** | CZ | 70.4 | **61** | CZ | ND |
| | | CL | 23.8 | | CL | ≥ 10000 |
|  | **2** | CZ | 73.0 | **62** | CZ | ≥ 10000 |
| | | CL | ≥ 10000 | | CL | ≥ 10000 |
|  | **29** | CZ | ND | **63** | CZ | ≥ 10000 |
| | | CL | 1000 | | CL | ≥ 10000 |
|  | **8** | CZ | 211.9 | **65** | CZ | ≥ 10000 |
| | | CL | 188.7 | | CL | ≥ 10000 |

Compounds **58**, **59**, and **60** (Table 144) demonstrated that modification of the thiosemicarbazone moiety (i.e. substitution on the $NH_2$ group of the thiosemicarbazone moiety) was detrimental in the activity of the synthetic compounds. None of these compounds showed significant activity towards cruzain.

Finally, Table 147 shows a small series of fluorene thiosemicarbazones. Both of them showed any activity towards cruzain (Analogs **80** and **81**).

Table 154. Comparison of the Inhibitory Activity between *meta* and *para*-Bromination of Benzophenone Thiosemicarbazones

| $R_1$ |  | | |  | | |
|---|---|---|---|---|---|---|
| | | | $IC_{50}$ (nM) | | | |
|  | **17** | CZ | 14.3 | **48** | CZ | 1453 |
| | | CL | 79.6 | | CL | 3320 |
|  | **21** | CZ | 127.0 | **47** | CZ | ≥ 10000 |
| | | CL | 327.1 | | CL | ≥ 10000 |
|  | **18** | CZ | 275.3 | **49** | CZ | ≥ 10000 |
| | | CL | ≥ 10000 | | CL | ≥ 10000 |
|  | **19** | CZ | 770.5 | **50** | CZ | 2262 |
| | | CL | 2156 | | CL | 4570 |

*Advanced Kinetic Studies*

Four compounds, **1**, **9**, **10** and **17 (**Figure 128**)** were screened against recombinant cruzain and demonstrated to be potent inhibitors of this protease. Their $IC_{50}$ values varied between 10.6 and 25 nM. Several experiments were carried out to characterize their mode of inhibition against recombinant cruzain. Compounds **1**, **10** and **17** are halogen-substituted bromobenzophenone thiosemicarbazone analogs. Compounds **1**, **10** and **17** are 3,3'-dibromobenzophenone TSC; 3-bromo-3',5'-difluorobenzophenone TSC and 3-bromo-4'-fluorobenzophenone TSC respectively. Finally, analog **9** is 3-bromo-4'-acetobenzophenone TSC.

*Kinetic Analysis of 3-Bromo-3'-Bromobenzophenone Thiosemicarbazone (**1**) as a Cruzain Inhibitor*

*Effect of Preincubation Studies on Cruzain Inhibition Assays using **1***

The effect of preincubation time in the determination of $IC_{50}$ values was examined using analog **1**. Experimental details are described in Chapter Two. $IC_{50}$ values were determined at eight different preincubation times ranging between 0 and 240 minutes. Final concentration of the inhibitor varied between 0 and 10 μM. Table 155 and Figure 130 show the results of the studies. Final conditions are: 1 mM EDTA, 2.5 DTT, 0.01% Brij 30, 2% DMSO, 15 μM Z-FR-AMC and 0.1 nM cruzain.

Table 155. Effect of Preincubation Time on $IC_{50}$ Values of **1** against Cruzain

| *Pre-incubation times (min)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 240 | 0.51 ± 0.050 |
| 120 | 0.93 ± 0.08 |
| 90 | 1.22 ± 0.04 |
| 60 | 1.19 ± 0.08 |
| 30 | 2.12 ± 0.20 |
| 5 | 10.6 ± 0.4 |
| 1 | 21.0 ± 1.9 |
| 0 | 207.5 ± 17.6 |

Preincubation time plays a vital role in the determination for any *in vitro* assay. The effect of preincubation on the potency of **1** as seen in Table 155 is quite evident. The $IC_{50}$ value is enhanced 10 fold by incubation as little as one minute. ($IC_{50(0min)}$: 208 and $IC_{50(1min)}$: 21.0 nM). A similar trend can be observed when preincubation times were increased. The $IC_{50}$ value when cruzain was preincubated four hours with compound **1** was 0.5 nM. These results confirmed the strong dependence of $IC_{50}$ value on the preincubation time parameter. Controls (i.e. uninhibited cruzain) were monitored at

every preincubation time. There was no significant loss of catalytic activity at longer preincubation times.

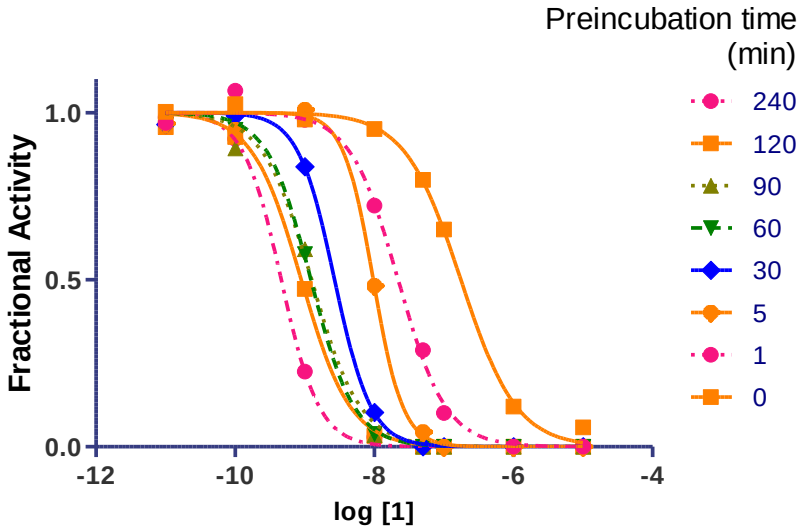



Figure 130. Effect of Preincubation Time on $IC_{50}$ Values of **1** against Cruzain

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation. Effect of Preincubation Time using **1**.*

The data obtained from the preincubation were further analyzed. The possibility that **1** was a tight-binding inhibitor was analyzed with Morrison's quadratic equation. (See equation 1.11).

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \quad (1.11)$$

The rates $v_i$ and $v_o$ are the inhibited and uninhibited cruzain velocities (RFU/s); $[\mathrm{E}]_T$ (nM) is, the total concentration of enzyme found in solution (free enzyme and inhibitor-enzyme complex); $[\mathrm{I}]_T$ (nM) is the total concentration of inhibitor present in

solution (free inhibitor and inhibitor-enzyme complex); and $K_I$ (nM) is the inhibition constant, often referred as the dissociation constant. The equation may be solved to give two possible answers. However, the equation is written so that there is only one possible answer that fits physiological conditions (i.e. $K_I^{app} > 0$). GraphPad 5.0 was used to fit the data after data manipulation for every preincubation time point. Inhibitor, substrate, and enzyme concentrations, as well as $K_M$, $K_I^{app}$, were all in micromolar units (μM). The residual activity (or $v_i$/vo) was normalized to 1 (i.e. $v_o$: 1 and $0 \leq v_i \leq 1$). Normalized residual activity and [I] were defined as the dependent and independent variables respectively. Nonlinear regression was applied using the following conditions: [S]: 15 μM, $K_M$: 1.01 μM, $[E]_T$: 0.0001 μM and $v_o$: 1. Normalization and nonlinear conditions were used to three or four independent experiments per preincubation time. Results and standard errors can be seen in Table 156 and Figure 131.

Table 156. Effect of Preincubation Time in $K_I^{app}$ Values of **1** against Cruzain

| *Pre-incubation times (min)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 0 | 12.82 ± 1.04 |
| 1 | 1.14 ± 0.16 |
| 5 | 0.59 ± 0.03 |
| 30 | 0.15 ± 0.01 |
| 60 | 0.08 ± 0.01 |
| 90 | 0.08 ± 0.01 |
| 120 | 0.06 ± 0.01 |

The $K_I^{app}$ values are shown in Table 156. It is apparent that the inhibition constants are also time-dependent. One minute of preincubation shows an excellent inhibition constant with a value of 1.14 nM. Standard preincubation times (5 minutes) enhanced more than 20-fold) when compared the calculated parameter for no preincubation time. The best inhibition constant value was obtained at 2 hours.

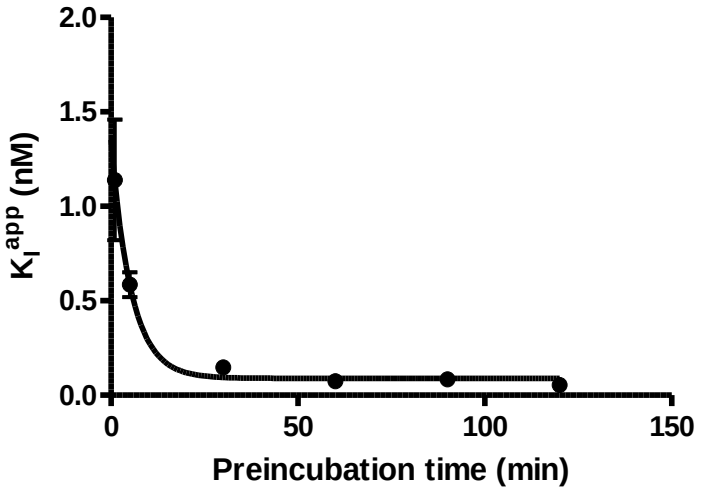



Figure 131. Effect of Preincubation Time in $K_I^{app}$ values of **1** against Cruzain

The apparent $K_I$ was 200-fold better when compared with the corresponding value at no preincubation time ($K_I^{app}$ = 12.8 nM). The 240 minutes data did not fit into the equation and $K_I^{app}$ could not be estimated. Inhibition constants were calculated based on the assumption that compound **1** might be a tight binding inhibitor. The obtained results confirmed that dibromobenzophenone thiosemicarbazone is a very tight inhibitor with $K_I^{app}$ as low as 60 pM.

*Cruzain Reversibility Studies*

Compound **1** was found to be a time-dependent inhibitor in the preincubation studies. Therefore, we decided to explore if this specific compound was a reversible inhibitor of cruzain. A mixture containing 100 X cruzain and 10 X $IC_{50(\text{preincubation time: 5 minutes})}$, which are 10 and 106 nM respectively, were incubated for different preincubation times ranging between 0 and 240 minutes at 25 °C. The inhibition of cruzain by **1** was able to be monitored by the rapid dilution of the mixture (100-fold) with assay buffer containing Z-FR-AMC. Experiment was set up in order to monitor these reactions almost

immediately after adding cruzain to the assay. Final conditions are 0.1 nM cruzain, 1.1 nM of the dibromobenzophenone thiosemicarbazone (**1**), and 15 μM Z-FR-AMC. Additionally, a control experiment (cruzain with DMSO as control vehicle) was also carried out for every preincubation time. Figure 132 shows the release of AMC for the first 3500 seconds. Readings were taken every 50 seconds. Uninhibited and inhibited reactions were followed for a total time of four hours. Cruzain was able to recover its catalytic activity after the rapid dilution with assay buffer containing Z-FR-AMC. Apparent substrate depletion can be observed in the uninhibited reaction after 3500 seconds. Thus, linear regression was applied to the first 3500 seconds of the reactions to determine cruzain activity. Table 157 shows cruzain activities for unhibited and inhibited reactions. Rates for unhibited and inhibited (preincubation time: 240 minutes) are 1.89 and 0.62 nM AMC/s respectively (remaining activity: 33%). However, similar experiments reported reversibility studies using a preincubation time of 60 minutes. According to Table 157, cruzain's remainin activity was 40.7% when the mixture cruzain-inhibitor was preincubated for one hour at the specified conditions. Interestingly, the inhibitory activity of **1** was not improved by longer preincubation times suggesting the slow release of the compound (See Figure 132, Figure 133, and Figure 134).

Further observations helped to identify if cruzain was able to recove after 240 minutes of preincubation time with dibromobenzophenone thiosemicarbazone (SeeFigure 133). Cruzain was able to recover its inhibited activity after 400 seconds of reactions.

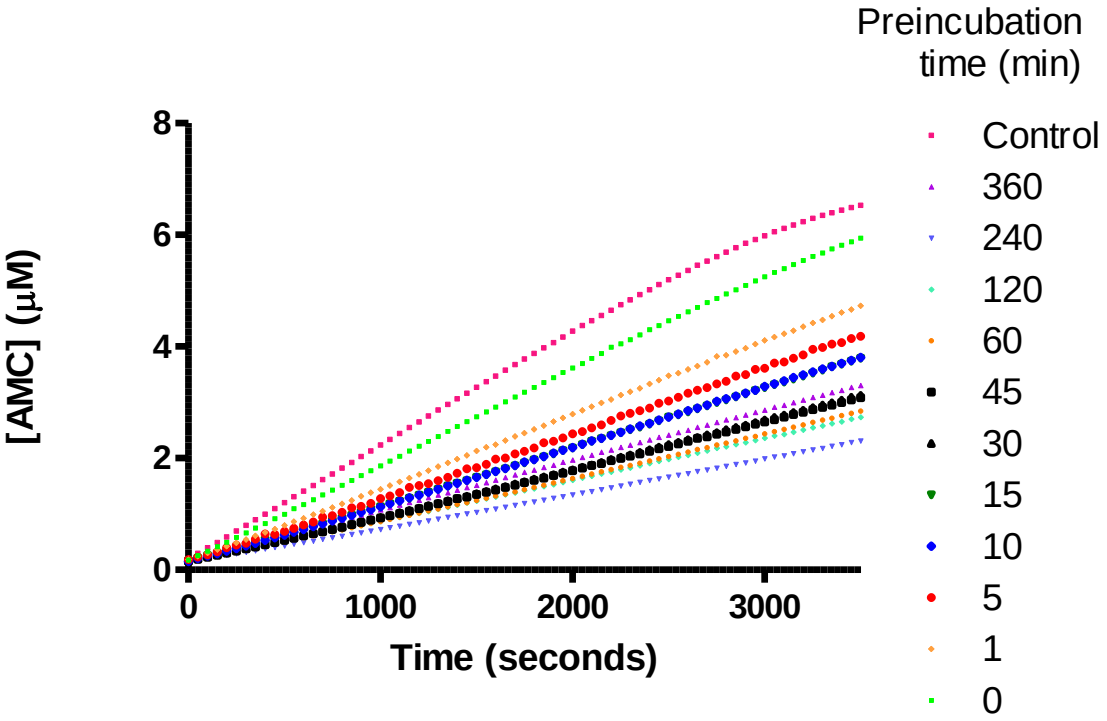



Figure 132. Effect of Preincubation Studies in Cruzain Reversibility Studies with **1** using 15 μM Z-FR-AMC

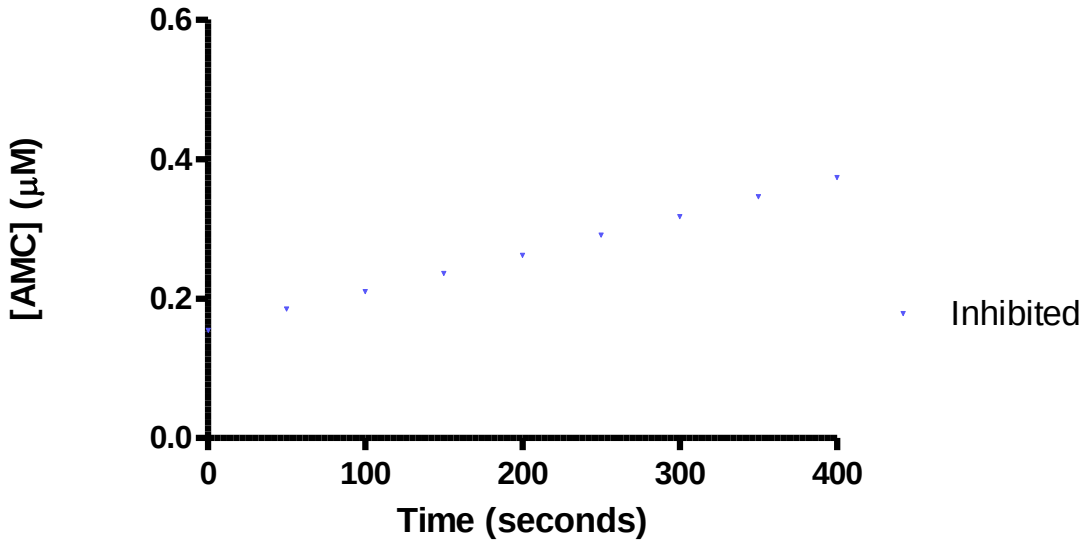



Figure 133. Cruzain Reversibility Studies with **1** Using 15 μM Z-FR-AMC

Table 157. Effect of Preincubation Studies in Cruzain Reversibility Studies with **1**

| *Pre-incubation times (min)* | *Inhibited activity (nM AMC/s)* | *Remaining activity (%)* |
|---|---|---|
| 0 (unhibited) | 1.89 | 100 |
| 0 (inhibited) | 1.69 | 89.4 |
| 1 | 1.32 | 69.8 |
| 5 | 1.16 | 61.4 |
| 15 | 1.05 | 55.6 |
| 30 | 0.85 | 45 |
| 45 | 0.84 | 44.4 |
| 60 | 0.77 | 40.7 |
| 120 | 0.73 | 38.6 |
| 240 | 0.62 | 32.8 |

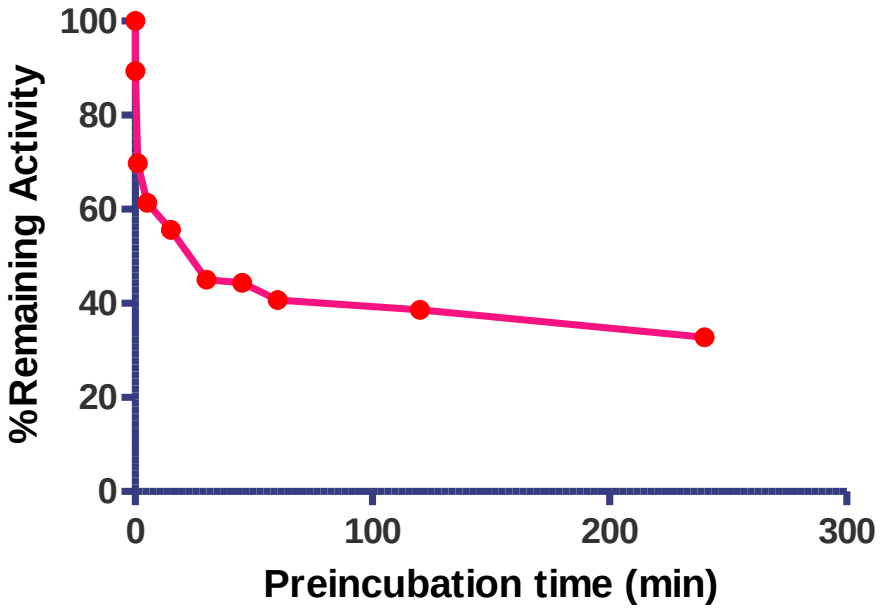



Figure 134. Effect of Preincubation Times on the Activity of Cruzain with **1** Using 15 μM Z-FR-AMC

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

We finally tried to determine the mode of inhibition of thiosemicarbazones as slow-binding inhibitors of cruzain. Therefore, we investigated the effect of substrate concentration on $IC_{50}$ values.[177] Classical methods for the determination of mode of inhibition (steady-state kinetics) fail when studying a slow, tight-binding inhibitors. In

these cases, Copeland offers a suitable alternative for the investigation. $IC_{50}$ values decrease hyperbolically if the compound is an uncompetitive inhibitor. In the case of competitive inhibitors, $IC_{50}$ values increase in a linear trend with higher substrate concentrations. Six substrate concentrations (Final concentrations: 15, 10, 7.5, 5.0 2.5 and 1.0 μM) were used to determine $IC_{50}$ values for compound **1.** Cruzain and the series of inhibitors were incubated for five minutes. Results can be found in Table 158 and Figure 135.

Table 158. Effect of Substrate Concentration on $IC_{50}$ Values of **1** against Cruzain

| *[Z-FR-AMC] (μM)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 15 | 11.3 ± 1.2 |
| 10 | 10.9 ± 0.2 |
| 7.5 | 9.5 ± 0.7 |
| 5.0 | 7.6 ± 0.7 |
| 2.5 | 6.5 ± 0.1 |
| 1.0 | 4.1 ± 0.2 |

The effect of substrate concentration was investigated using **1** as a lead compound of the series of thiosemicarbazones. A linear behavior can be observed with the values. Similarly, to cathepsin L case, the inhibitor acts as a competitive inhibitor with respect to the substrate. According to literature, a positive linear behavior (i.e. $IC_{50}$ α [S]) is an indication the compound acts as competitive inhibitor, that is, both substrate and the compound compete for cruzain active site. Although the compound is a slow-binding inhibitor, the reduction of substrate concentration (15-fold) improves the $IC_{50}$ value by almost 3-fold ($IC_{50([Z\text{-}FR\text{-}AMC]:\ 15\ \mu M)}$: 11.3 nM vs $IC_{50([Z\text{-}FR\text{-}AMC]:\ 1\ \mu M)}$: 4.1 nM).

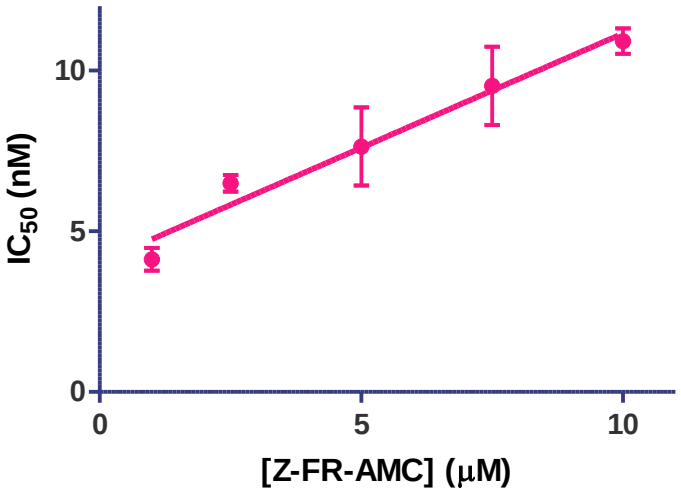



Figure 135. Effect of Substrate Concentration on $IC_{50}$ Values of **1** against Cruzain

*Kinetic Analysis of 3-Bromo-3',5'-Difluorobenzophenone Thiosemicarbazone (**10**) as a Cruzain Inhibitor*

*Effect of Preincubation Times on Cruzain Inhibition Assays using **10***

The effect of preincubation time in the determination of $IC_{50}$ values was examined using analog **10**. Experimental details are described in chapter 4 (See Effect of Preincubation Times on Cruzain Inhibition Assays using **1**). Results can be seen in Table 159 and Figure 136.

Table 159. Effect of Preincubation Time on $IC_{50}$ Values of **10** towards Cruzain

| *Pre-incubation times (min)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 240 | 0.8 ± 0.03 |
| 120 | 1.6 ± 0.1 |
| 90 | 2.2 ± 0.2 |
| 60 | 2.3 ± 0.2 |
| 30 | 4.4 ± 0.4 |
| 5 | 17.7 ± 1.5 |
| 1 | 37.3 ± 3.6 |
| 0 | 222.5 ± 6.7 |

The effect of preincubation times on the potency of **10** can be seen in Table 159 is quite evident. The $IC_{50}$ value is enhanced 6-fold by incubation as little as one minute. ($IC_{50(0min)}$: 222.5 and $IC_{50(1min)}$: 37.3 nM). A similar trend can be observed when preincubation time was increased and the $IC_{50}$ value when cruzain was preincubated four hours with compound **10** was 0.8 nM.

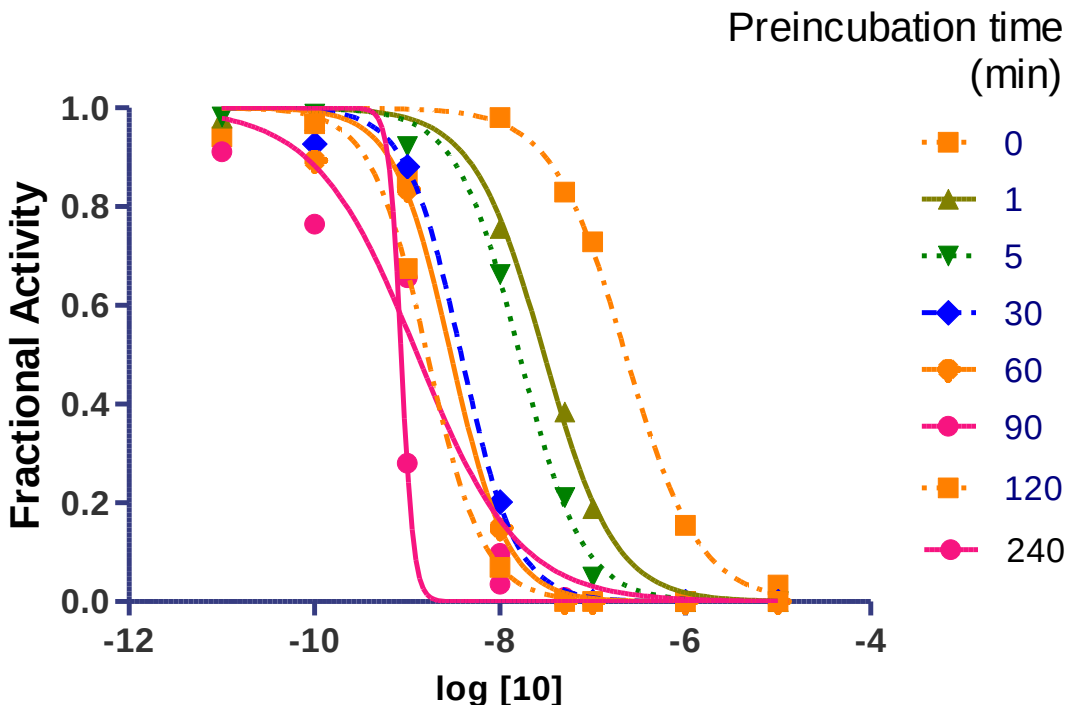



Figure 136. Effect of Preincubation Time on $IC_{50}$ Values of **10** against Cruzain

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation. Effect of Preincubation Time using **10**.*

The data obtained from the preincubation studies were further analyzed. The possibility that **10** was a tight-binding inhibitor, like analog **1**, was analyzed with Morrison's quadratic equation. (See equation 1.11). Results can be found in Table 160 and Figure 137.

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \quad (1.11)$$

Table 160. Effect of Preincubation Time on $K_I^{app}$ values of **10** against Cruzain

| *Pre-incubation times (min)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 0 | 12.23 ± 1.92 |
| 1 | 2.11 ± 0.25 |
| 5 | 1.03 ± 0.09 |
| 30 | 0.23 ± 0.01 |
| 60 | 0.14 ± 0.01 |
| 90 | 0.12 ± 0.01 |
| 120 | 0.10 ± 0.01 |

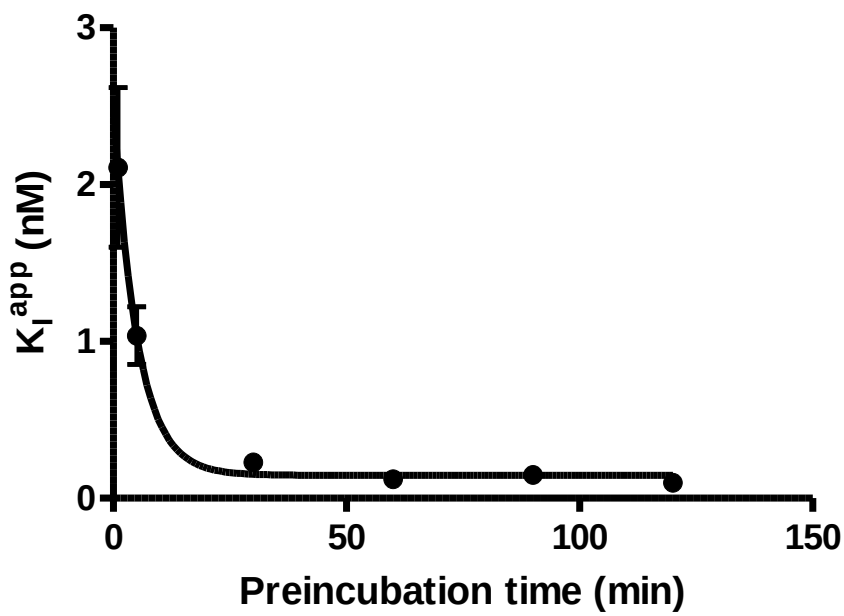



Figure 137. Effect of Preincubation Time in $K_I^{app}$ values of **10** against Cruzain

The $K_I^{app}$ values of **10** are also time-dependent. One minute of preincubation shows an excelletn inhibition constant with a value of 2.1 nM. The standard preincubation time (5 minutes) enhanced $K_I^{app}$ more than 12-fold) when compared the calculated parameter for no preincubation time ($K_I^{app:}$ 0.10 nM). The best inhibition

constant value was obtained at 2 hours. The apparent $K_I$ was 120-fold better when compared with the corresponding value at no preincubation time ($K_I^{app}$ = 12.2 nM). The 240 minutes data did not fit into the equation and $K_I^{app}$ could not be estimated. The obtained results confirmed that **10**, a polyhalogenated benzophenone thiosemicarbazone is a very tight inhibitor with $K_I^{app}$ as low as 100 pM.

*Cruzain Reversibility Studies*

Compound **10** was found to be a time-dependent inhibitor in the preincubation time studies. Therefore, we decided to explore if this specific compound was a reversible inhibitor of cruzain. A mixture containing 100 X cruzain and 10 X $IC_{50(\text{preincubation time: 5 minutes})}$, which are 10 and 180 nM respectively, were incubated for different preincubation times ranging between 0 and 240 minutes at 25 °C. The inhibition of cruzain by **10** was able to be monitored by the rapid dilution of the mixture (100-fold) with assay buffer containing Z-FR-AMC. The experiment was set up in order to monitor these reactions almost immediately after adding cruzain to the assay. Final conditions are 0.1 nM cruzain, 1.8 nM of **10**, and 15 μM. Additionally, a control experiment (cruzain with DMSO as control vehicle) was also carried out for every preincubation time. Figure 138 shows the release of AMC for the first 3500 seconds. Readings were taken every 50 seconds. Uninhibited and inhibited reactions were followed for a total time of four hours. Cruzain was able to recover its catalytic activity after the rapid dilution with assay buffer containing Z-FR-AMC. Apparent substrate depletion can be observed in the uninhibited reaction after 3500 seconds. Thus, linear regression was applied to the first 3500 seconds of the reactions to determine cruzain activity. Figure 138 shows cruzain activities for uninhibited and inhibited reactions. Rates for uninhibited and inhibited (preincubation

time: 240 minutes) are 1.89 and 0.52 nM AMC/s respectively (Remaining activity: 73%). According to Table 161, the polyhalogenated thiosemicarbazone was able to inhibit cruzain activity by 54.0% when the mixture cruzain-inhibitor was preincubated for two hours at the specified conditions. Interestingly, the inhibitory activity of **10** was not improved by longer preincubation times suggesting the slow release of the compound (See Figure 138, Figure 139 and Figure 140). Further observations helped to identify if cruzain was able to recovery after 240 minutes of preincubation time with dibromobenzophenone thiosemicarbazone (See Figure 139, and Figure 140). Cruzain was able to recover its inhibited activity after 400 seconds of reactions.

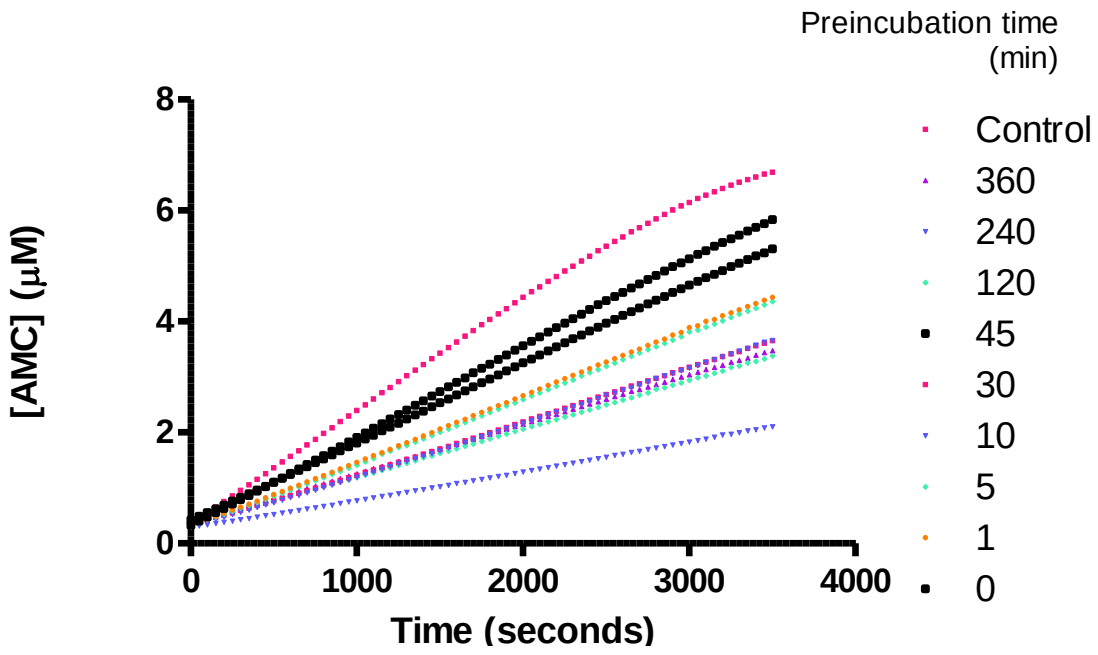



Figure 138. Effect of Preincubation Studies in Cruzain Reversibility Studies with **10** Using 15 μM Z-FR-AMC

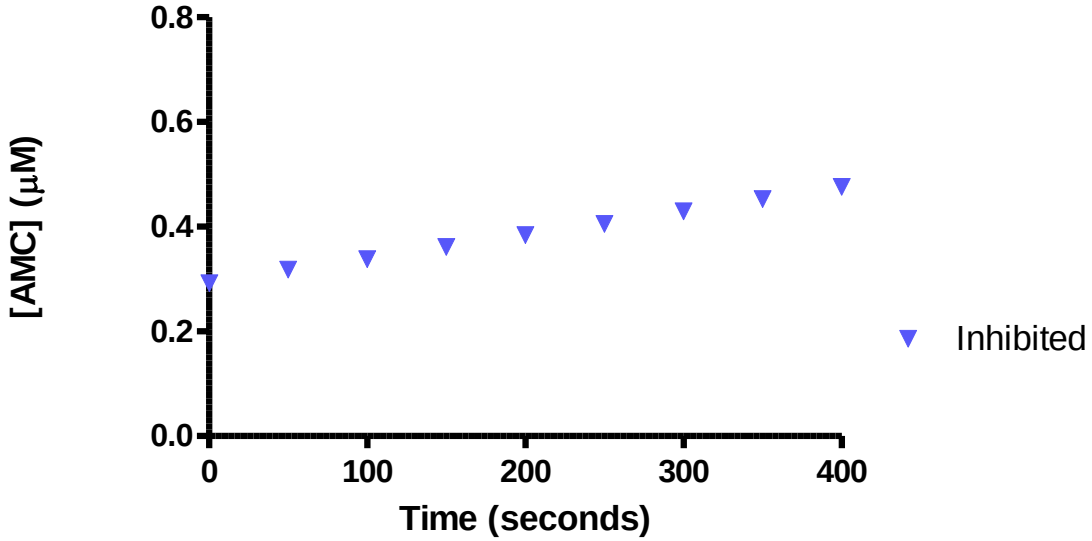



Figure 139. Cruzain Reversibility Studies with **10** Using 15 μM Z-FR-AMC

Table 161. Effect of Preincubation Studies on Cruzain Reversibility Studies with **10**

| *Pre-incubation times (min)* | *Inhibited activity (nM AMC/s)* | *Inhibited activity (%)* |
|---|---|---|
| 0 (unhibited) | 1.89 | 0 |
| 0 (inhibited) | 1.60 | 15.3 |
| 1 | 1.19 | 37.0 |
| 5 | 1.17 | 38.1 |
| 10 | 0.97 | 48.7 |
| 30 | 0.96 | 49.2 |
| 120 | 0.87 | 54.0 |
| 240 | 0.52 | 72.5 |

*Effect of Substrate Concentration (Z-FR-AMC) on $IC_{50}$ Values*

We finally tried to determine the mode of inhibition of **10** as a slow-binding inhibitor of cruzain. Six substrate concentrations (Final concentrations: 15, 10, 7.5, 5.0 2.5 and 1.0 μM) were used to determine $IC_{50}$ values for compound **10.** Cruzain and the series of inhibitors were incubated for five minutes. Results can be found in Table 162

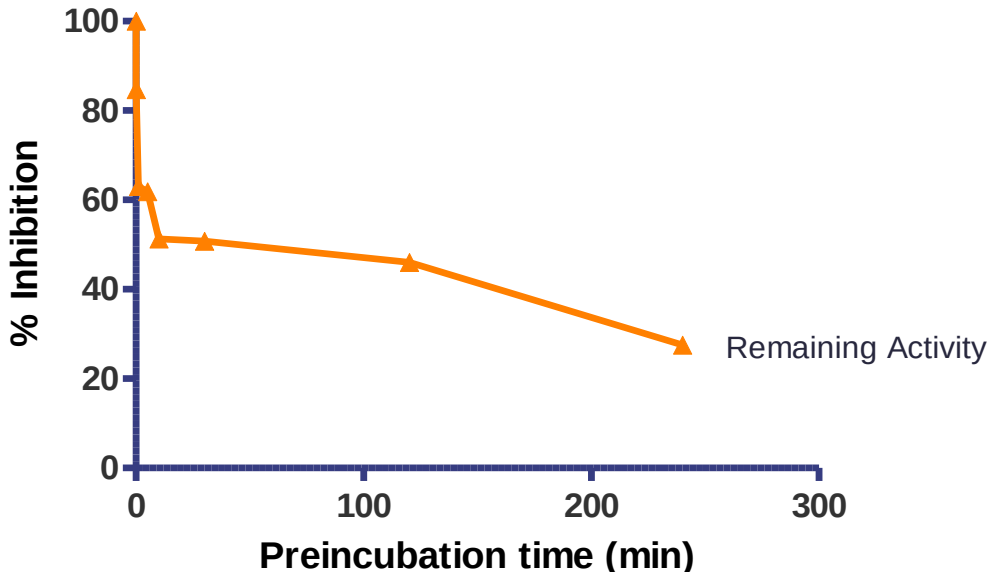



Figure 140. Effect of Preincubation Times on the Activity of Cruzain with **10** using 15 μM Z-FR-AMC

Table 162. Effect of Substrate Concentration on $IC_{50}$ Values of **10** against Cruzain

| *[Z-FR-AMC] (μM)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 15 | 17.6 ± 0.6 |
| 10 | 17.5± 1.15 |
| 7.5 | 14.9 ± 0.8 |
| 5.0 | 13.1 ± 0.4 |
| 2.5 | 13.4 ± 0.6 |
| 1.0 | 4.75 ± 0.3 |

The effect of substrate concentration was investigated using **10** as a lead compound of the series of thiosemicarbazones. Similar to **1**, **10** acts as a competitive slow-binding inhibitor. The trend shows an increase in $IC_{50}$ value with increased substrate concentration. The reduction of substrate concentration (15-fold) improves the $IC_{50}$ value by almost 4 times. ($IC_{50([Z\text{-}FR\text{-}AMC]:\ 15\ \mu M)}$: 17.6 nM vs $IC_{50([Z\text{-}FR\text{-}AMC]:\ 1\ \mu M)}$: 4.8 nM).

*Kinetic Analysis of 3-Bromo-3'-Acetobenzophenone Thiosemicarbazone (**9**) as a Cruzain Inhibitor*

*Effect of Preincubation Times on Cruzain Inhibition Assays Using **9***

The effect of preincubation time in the determination of $IC_{50}$ values was examined using analog **9**. Experimental details are described in chapter 4 (See Effect of Preincubation Times on Cruzain Inhibition Assays using **1**). Preincubation times ranged between 0 and 120 minutes. Results can be seen in Figure 141 and Table 163.

Table 163. Effect of Preincubation Time on $IC_{50}$ Values of **9** against Cruzain

| *Pre-incubation times (min)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 120 | 3.8 ± 0.1 |
| 60 | 4.2 ± 0.4 |
| 30 | 8.8 ± 0.5 |
| 5 | 15.4 ± 1.2 |
| 1 | 132.5 ± 6.2 |
| 0 | 4935 ± 122 |

The effect of preincubation times on the potency of **9** can be seen in Table 159. The $IC_{50}$ value is enhanced 37-fold by incubation as little as one minute ($IC_{50(0min)}$: 4935 and $IC_{50(1min)}$: 132.5 nM). A similar trend can be observed when preincubation time was increased and the $IC_{50}$ value when cruzain was preincubated two hours with compound **9** was 3.8 nM.

*Determination of $K_i^{app}$ using Morrison's Quadratic Equation. Effect of Preincubation Time using **9**.*

The data obtained from the preincubation studies were further analyzed. Analogs **1** and **10**, both halogenated thiosemicarbazones inhibitors were tight binding inhibitors of cruzain. The addition of a nonhalogenated group to the main moiety might provide a

different behavior. Thus, the data were analyzed with Morrison's quadratic equation. (See equation 1.11). Results can be found in Table 164 and Figure 142.

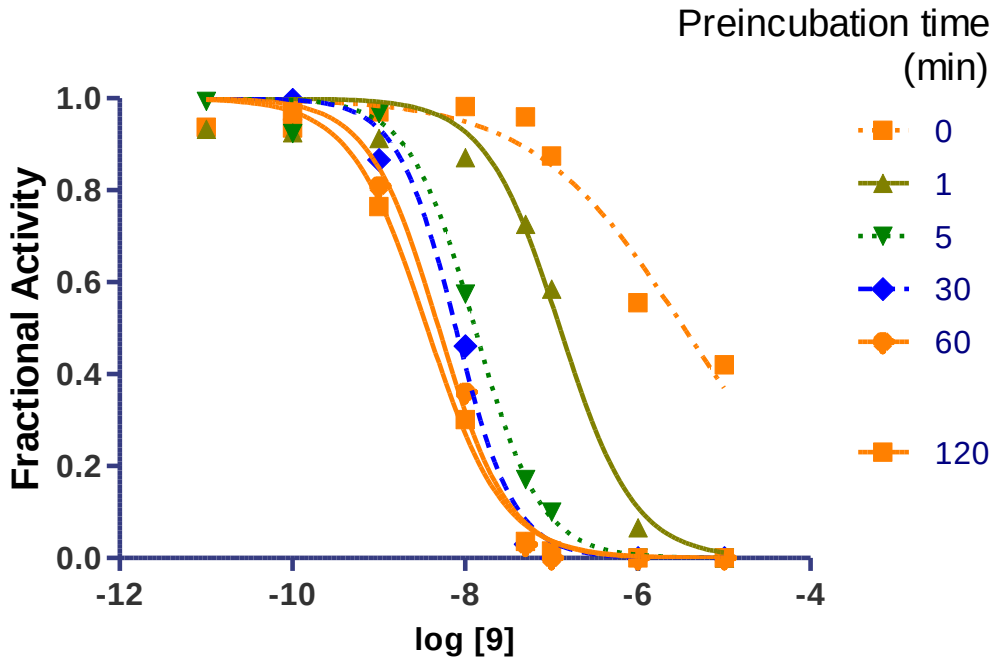



Figure 141. Effect of Preincubation Time on $IC_{50}$ Values of **9** against Cruzain

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \quad (1.11)$$

Table 164. Effect of Preincubation Time on $K_I^{app}$ Values of **9** against Cruzain

| *Pre-incubation times (min)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 1 | 11.19 ± 2.21 |
| 5 | 0.92 ± 0.06 |
| 30 | 0.42 ± 0.05 |
| 60 | 0.28 ± 0.03 |
| 120 | 0.27 ± 0.02 |

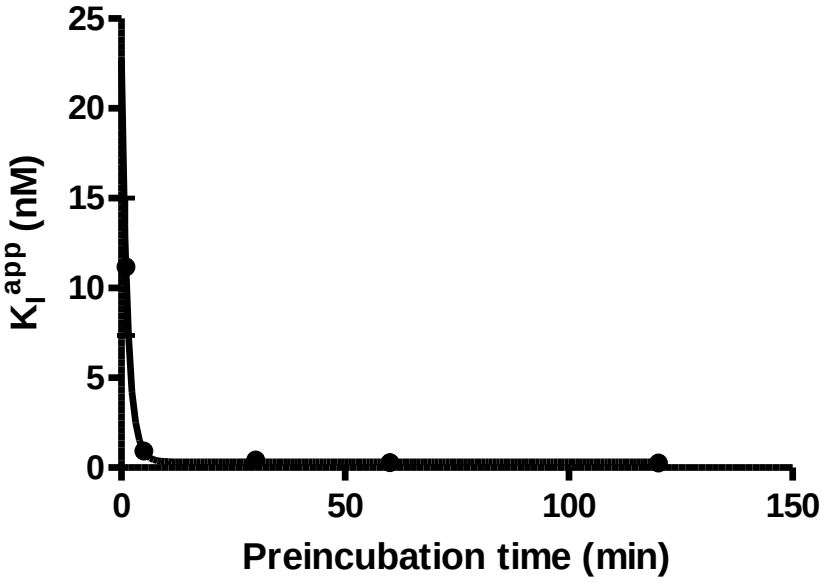



Figure 142. Effect of Preincubation Time in $K_I^{app}$ values of **9** against Cruzain

The $K_I^{app}$ values of **9** are also time-dependent. One minute of preincubation shows a good inhibition constant with a value of 11.2 nM. Standard preincubation times (5 minutes) enhanced the $K_I^{app}$ value more than 12-fold) when compared the calculated parameter for one minute of preincubation time. The dissociation constant for the data set corresponding to 0 minutes preincubation time could not be determined. The best inhibition constant value was obtained at 2 hours. The apparent $K_I$ was 37-fold better when compared with the corresponding value at one minute preincubation time ($K_I^{app}$ = 11.2 nM). The obtained results confirmed that **9**, an 3-bromo-3'-acetobenzophenone thiosemicarbazone is an excellent inhibitor with $K_I^{app}$ as low as 270 pM.

*Kinetic Analysis of 3-Bromo-4'-Fluorobenzophenone Thiosemicarbazone (**17**) as a Cruzain Inhibitor*

*Effect of Inhibitor Concentration on Cruzain Progress Curves*

Compound **17**, is a dihalogenated benzophenone thiosemicarbazone. Its value was determined to be 12.1 nM and is the second most potent inhibitor found in this library. Cruzain (0.1 nM) was added to six different concentration curves ranging from 0 to 10 μM of **17**. Reactions were initiated by the rapid addition of Z-FR-AMC (final concentration: 15 μM). The release of AMC from the nonfluorescent substrate was monitored every 3 seconds for fifty minutes. Figure 143 shows the effect of inhibitor concentration on the cruzain progress curves at a fixed substrate concentration. Uninhibited and inhibited reactions were monitored for 3000 seconds. Visual observations showed analog **17** is a time-dependence inhibitor. Then, the data were fitted to equation 1.7, by nonlinear regression analysis using GraphPad 5.0. P is the concentration of product (μM), $v_o$ and $v_s$ are the initial and steady-state velocities (μM/s), t is the time in seconds and $k_{obs}$ the rate constant for conversion of the initial velocity $v_o$ to the steady state velocity $v_s$. The rate constant ($k_{obs}$) units are given in $s^{-1}$. The equation was entered into the computational software, knowing that P and t are the dependent and independent variables, while keeping the velocities and the rate constant as unknowns. For each case, the constraints for their calculation were to give positive values (i.e. $k_{obs} \geq 0$). It is also worth noting that equation 1.7 is only valid when substrate depletion is insignificant.[188] Therefore, some points were excluded in every case for data fitting. Velocities, rates, $r^2$ and points analyzed for each substrate concentration are shown in Table 165. Experiments were carried in triplicate.

$$P = v_s t \frac{(v_o - v_{s)}}{k_{obs}}\left(1 - e^{-k_{obs}t}\right) \tag{1.7}$$

Table 165. Calculated Kinetic Parameters from Eq 1.7 for Cruzain Progress Curves with 15 μM Z-FR-AMC

| [I] (μM) | 1 | 0.5 | 0.1 | 0.05 |
|---|---|---|---|---|
| $v_s$ (μM/s) | 3.45E-05 | 7.71E-05 | 0.000406 | 0.000731 |
| $v_i$ (μM/s) | 0.003057 | 0.002666 | 0.00211 | 0.002111 |
| $k_{obs}$ ($s^{-1}$) | 0.005461 | 0.002984 | 0.00096 | 0.000843 |
| $r^2$ | ≥ 0.94 | ≥ 0.98 | ≥ 0.99 | ≥ 0.99 |
| Points analyzed | 1000 | 1000 | 1000 | 1000 |

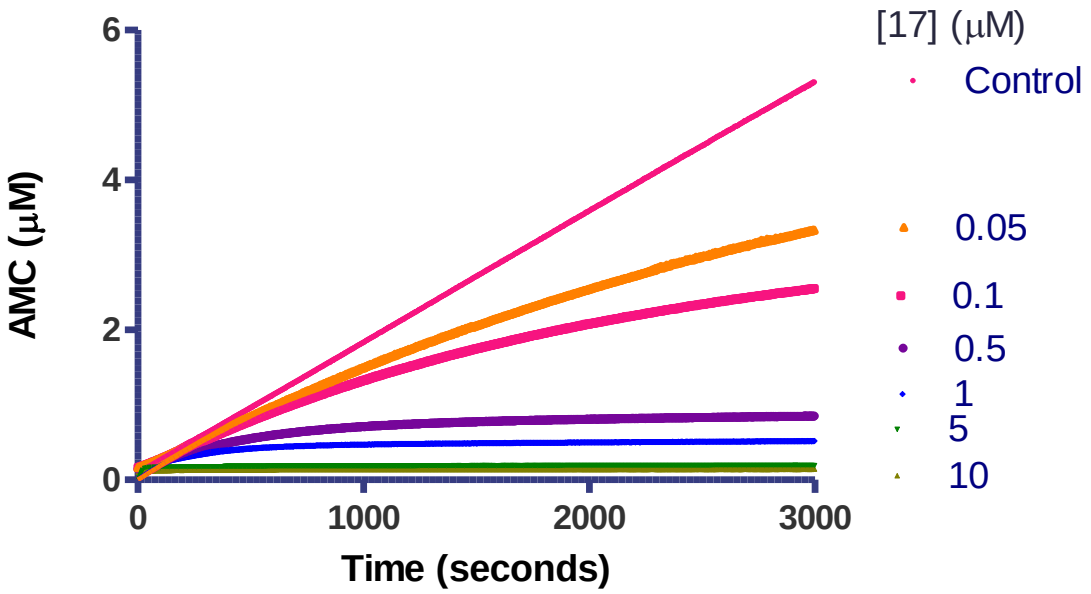



Figure 143. Cruzain Progress Curves with **17** Using 15 μM Z-FR-AMC

Approximate values for the rate constants, and velocities were obtained for every inhibitor concentration (0.05 ≤ [I] ≤ 10 μM). Data were fitted into Equation 1.7, but due to the elevated amount of inhibitor present in solution, compared to the concentration of the cruzain, the results were not reasonable. A better fit was acquired with lower concentrations of **17** (50-1000 nM). We decided to investigate the mechanism of inhibition of **17** as a slow binding inhibitor of cruzain, due to the clear time-dependence

of the progress curves. There are two possible mechanisms that are known for slow binding inhibitors.

Figure 144 shows the two proposed mechanisms that could be observed using the dihalogenated thiosemicarbazone (3-Br-4'-F-TSC) as a slow-binding inhibitor of cruzain. Mechanism A summarizes a simple reversible inhibition with $k_{on}$ and $k_{off}$ values relatively small. The parameters $k_{on}$ and $k_{off}$ are the rate constants for the formation and dissociation of the cruzain-3-Br-4'-F-TSC complex. Mechanism B offers an extra step which is more complicated to monitor or verify. This is a more general approach where it is assumed the cruzain -3-Br-4'-F-TSC complex undergoes an auto-isomerization, or a possible covalent modification of the enzyme due to the presence of the inhibitor.

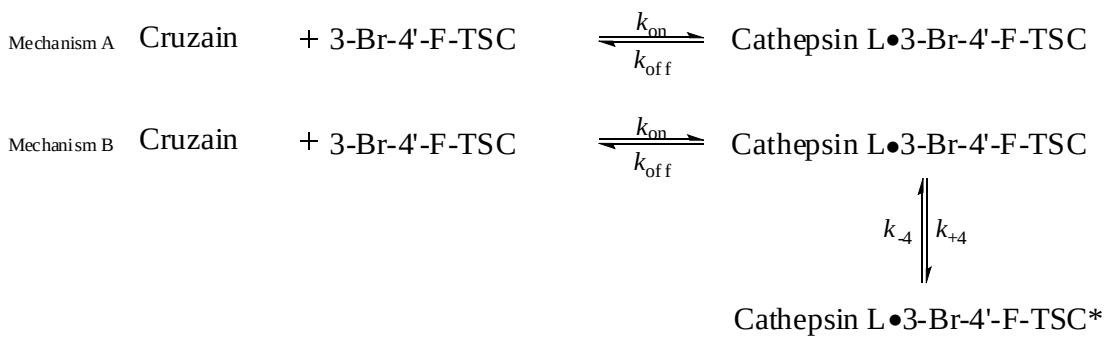



Figure 144. Possible Mechanisms of Inhibition of Cruzain by **17**

The $k_{obs}$ values were plotted versus inhibitor concentration (Figure 145). Mechanism A, (Figure 144) is carried out for slow-binding inhibitors. However, many reversible, yet covalent inhibitors conform to mechanism A, as previously discussed in Chapter Two. The linear dependence of this parameter with respect of the inhibitor concentration is a clear indication the inhibition could follow a simple mechanism. Furthermore, a visual inspection of the progress curves suggests that **17** is a time-

dependent inhibitor of cruzain, but further experiments will need to be performed to validate that mechanism. Other slow-binding inhibitors, covalent and reversible, have also been reported with similar trends but did not follow simple reversible mechanisms.[328] The figure represents the average value of three independent experiments with standards errors.

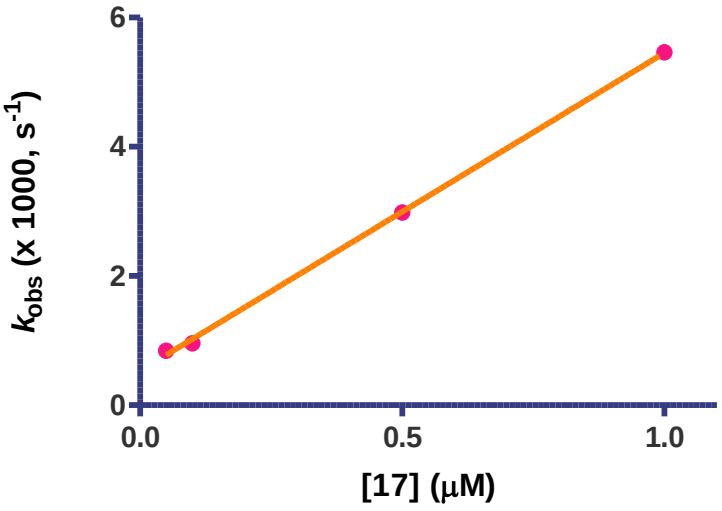



Figure 145. Calculated $k_{obs}$ from Eq. 1.7 Cruzain Progress Curves with **17**

Table 166. Calculated $k_{obs}$ from Eq. 1.7 Cruzain Progress Curves with **17**

| [I] (μM) | $k_{obs}$ ($s^{-1}$) |
|---|---|
| 1 | 0.0005852 ± 0.0004082 |
| 0.5 | 0.003082 ± 0.0001027 |
| 0.1 | 0.0009589 ± 4.016e-5 |
| 0.05 | 0.0008181 ± 5.447e-5 |

Linear regression analysis was performed using GraphPad 5.0 in order to analyze the data (15 μM Z-FR-AMC). Results can be seen in Table 55. The parameters $k_{on}$ and $k_{off}$ can be then calculated using equations 2.1 and 2.2.[328]

$$k_{obs} = k_{on}^{app}[I] + k_{off} \quad (2.1)$$

$$k_{obs} = \frac{k_{on}[I]}{(1 + \frac{[S]}{K_M})} + k_{off} \quad (2.2)$$

Table 167. Calculated $k_{on}$, and $k_{off}$ values from Eq. 2.1 Using Cruzain Progress Curves with **17**

| [Z-FR-AMC] (μM) | 15 |
|---|---|
| $k_{on}^{app}$ (M·s)$^{-1}$ | 4922 |
| $k_{on}$ (M·s)$^{-1}$ | 307.62 |
| $k_{off}$ (s$^{-1}$) | 0.0005318 |
| $r^2$ | 0.9994 |

*Effect of Preincubation Studies on Cruzain Inhibition Assays Using **17***

The determination of kinetic parameters using cruzain progress curves provided strong evidence that **17** is a slow-binding inhibitor. The effect of preincubation time in the determination of $IC_{50}$ values was examined using analog **17**. Experimental details are described in Chapter 2. Compound **17** $IC_{50}$ values were determined at six different preincubation times ranging between 0 and 120 minutes. Inhibitor final concentration varied between 0 and 10 μM. Figure 146shows the results of the studies.

Table 168. Effect of Pre-incubation Times on $IC_{50}$ Values of **17** against Cruzain

| *Pre-incubation times (*minutes*)* | *$IC_{50}$ ± Standard Error (nM)* |
|---|---|
| 120 | 1.7 ± 0.1 |
| 60 | 2.2 ± 0.3 |
| 30 | 2.7 ± 0.3 |
| 5 | 12.1 ± 1.1 |
| 1 | 91.9 ± 7.4 |
| 0 | 3144 ± 41 |

The effect of preincubation on the potency of **17** can be seen in Table 168. The 3-bromo-3'-fluorobenzophenone thiosemicarbazone (**17**) showed low inhibitory activity when cruzain was not preincubated with the compound. However, the analog showed good activity with one minute preincubation time. The trend changed when the preincubation time was increased to 5 minutes. The inhibitory activity of **17** increased 32-fold with a remarkable $IC_{50}$ less than 20 nM (12.1 nM). The potency of **17** modestly increased with longer preincubation times. Finally, the best activity was found when **17** was preincubated for two hours with cruzain, the activity of the thiosemicarbazone analog increased 7-fold to give a value of 1.7 nM. These results confirmed the strong dependence of $IC_{50}$ value determination with respect to the preincubation time. Controls (i.e. uninhibited cruzain) were monitored at every preincubation time. There was no significant loss of catalytic activity at longer preincubation times.

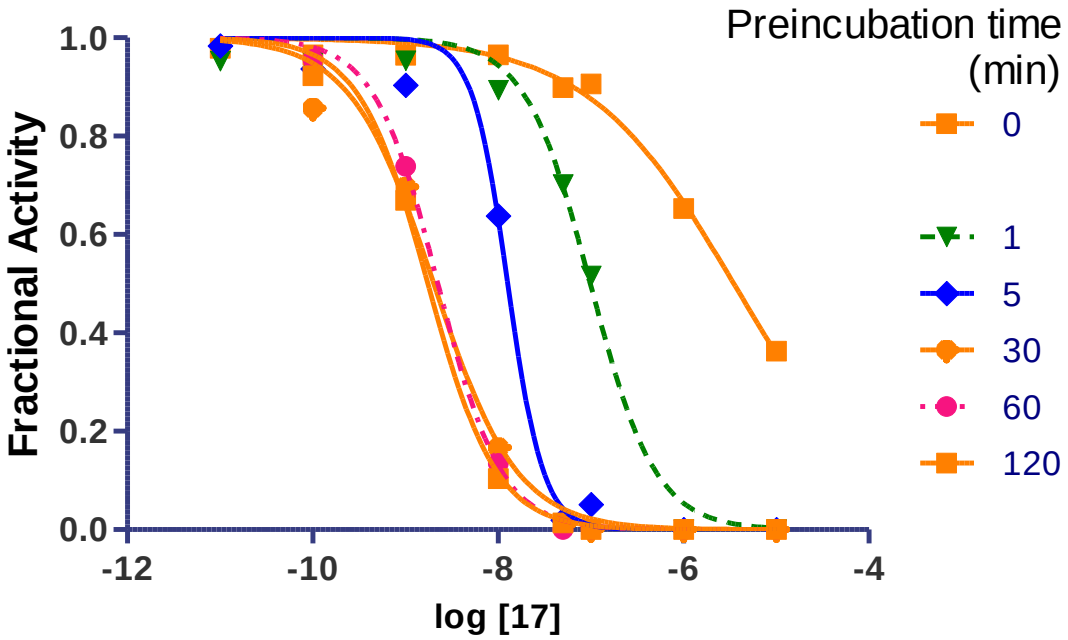



Figure 146. Effect of Preincubation Time on $IC_{50}$ Values of **17** against Cruzain

*Determination of $K_i^{app}$ Using Morrison's Quadratic Equation. Effect of Preincubation Time using **17***

The data obtained from the preincubation was further analyzed. The possibility that **17** was a tight-binding inhibitor was analyzed with Morrison's quadratic equation (See equation 1.11). Detailed analysis was described in chapter two. Results can be seen in Table 169 and Figure 147.

$$\frac{v_i}{v_o} = 1 - \frac{([\mathrm{E}]_T + [I]_T + (K_I(1 + \left(\frac{[S]}{K_M}\right)) - \sqrt{([\mathrm{E}]_T + [I]_T + (K_I\left(1 + \left(\frac{[S]}{K_M}\right)\right))^2 - 4[E]_T[\mathrm{I}]_T}}{2\,[E]_T} \quad (1.11)$$

Table 169. Effect of Preincubation Time on $K_I^{app}$ Values of **17** against Cruzain

| *Pre-incubation Times (*minutes*)* | $K_I^{app}$ ± *Standard Error (nM)* |
|---|---|
| 0 | 189.6 ± 1.04 |
| 1 | 5.3 ± 0.6 |
| 5 | 0.8 ± 0.06 |
| 30 | 0.16 ± 0.01 |
| 60 | 0.15 ± 0.02 |
| 120 | 0.11 ± 0.01 |

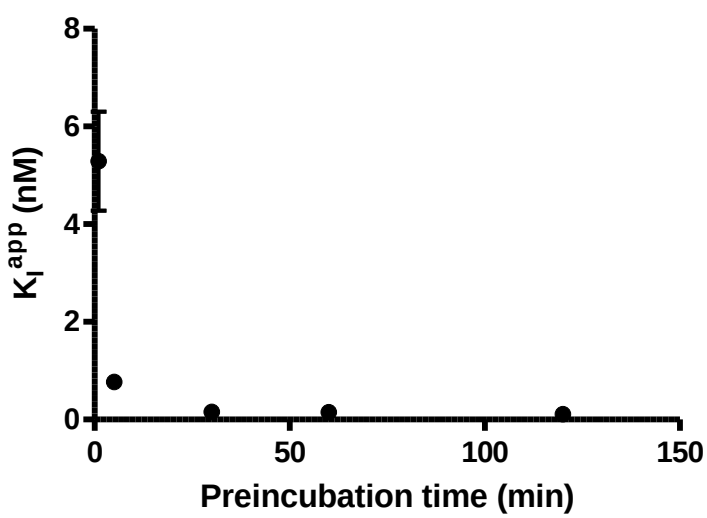



Figure 147. Effect of Preincubation Time on $K_I^{app}$ Values of **17** against Cruzain

The $K_I^{app}$ values are also time-dependent. One minute of preincubation shows an excellent inhibition constant with a value of 5.3 nM. Furthermore, the potency is extremely enhanced (more than 16-fold) for the calculated value at standard conditions (preincubation time: 5 minutes). The best inhibition constant value was obtained at the longest preincubation time (2 hours). The apparent $K_I$ was 48-fold better when compared with the corresponding value at one minute preincubation time ($K_I^{app}$ = 0.11 nM).

*Cruzain Reversibility Studies*

Compound **17** was found to be a time-dependent inhibitor in the preincubation studies. Therefore, we decided to explore if this specific compound was a reversible inhibitor of cruzain. A mixture containing 100 X cruzain and 10 X $IC_{50(\text{preincubation time: 5 minutes})}$, which are 10 and 120 nM respectively, were incubated at 25 °C for one hour. The inhibition of cruzain by **17** was able to be monitored by the rapid dilution of the mixture (100-fold) with assay buffer containing Z-FR-AMC. The experiment was set up in order to monitor these reactions almost immediately after adding cruzain to the assay. Final conditions were 0.1 nM Cruzain, 1.2 nM of **17,** and 15 μM. Additionally, a control experiment (cruzain with DMSO as control vehicle) was also carried out. Figure 148 shows the release of AMC for the first 7500 seconds. However, uninhibited and inhibited reactions were followed for a total time of four hours. Cruzain was able to recover its catalytic activity after the rapid dilution with assay buffer containing Z-FR-AMC. The remaining activity was sixty percent, when compared to control, was observed in the recovery curve when cruzain was preincubated for one hour with compound **17**. Apparent substrate depletion can be observed in the uninhibited reaction after 5000 seconds. Thus, linear regression was applied to the first 7000 seconds of the reactions to determine

cruzain activity. Rates for unhibited and inhibited (preincubation time: 60 minutes) are 1.49 and 0.91 nM AMC/s respectively. Further observations helped to identify if cruzain was able to recovery after 240 seconds of preincubation time with this synthetic compound (See Figure 149)

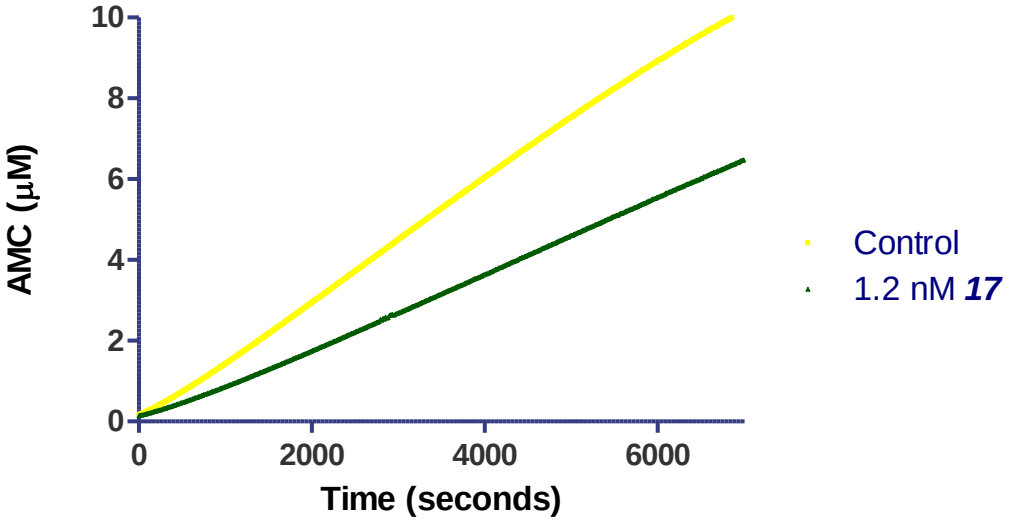



Figure 148. Cruzain Reversibility Studies with **17** Using 15 μM Z-FR-AMC

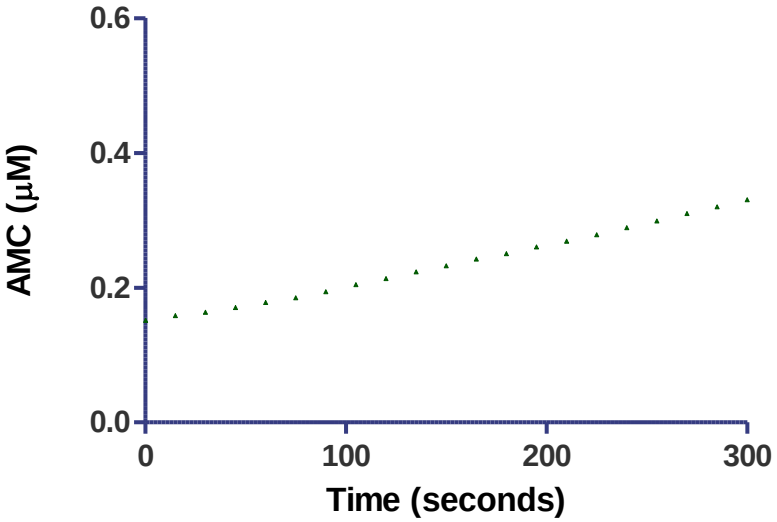



Figure 149. Cruzain Recovery Studies with **17** Ising 15 μM Z-FR-AMC

*Effect of Substrate Concentration (Z-FR-AMC) on Cruzain Progress Curves*

The mode of inhibition of **17** was explored. A fixed concentration of **17** (100 nM) was used to monitor the effect of this compound on cruzain progress curves when using different substrate concentrations. Then, by using non-linear regression analysis (equation 1.7), the parameters $k_{obs,}$ $v_o$, and $v_s$ were determined. Compound **17** (0.1 μM) and ten Z-FR-AMC concentrations (ranging from 0.5 and 15 μM) were set up in order to verify the effect of substrate on inhibited cruzain progress curves. Reactions were started by the addition of cruzain (final concentration: 0.1 nM). The release of AMC was monitored every five seconds for sixty minutes. Figure 150 shows a typical result of this experiment. The results once again showed the strong time dependence of **17** when inhibiting cruzain. Substrate depletion is negligible for every substrate concentration as seen in the figure. Data were fitted into equation 1.7, (see Effect of inhibitor concentration in Cruzain progress curves. Velocities, rates, $r^2$ and points analyzed for each substrate concentration are shown in Tables 59, 60, and 61.

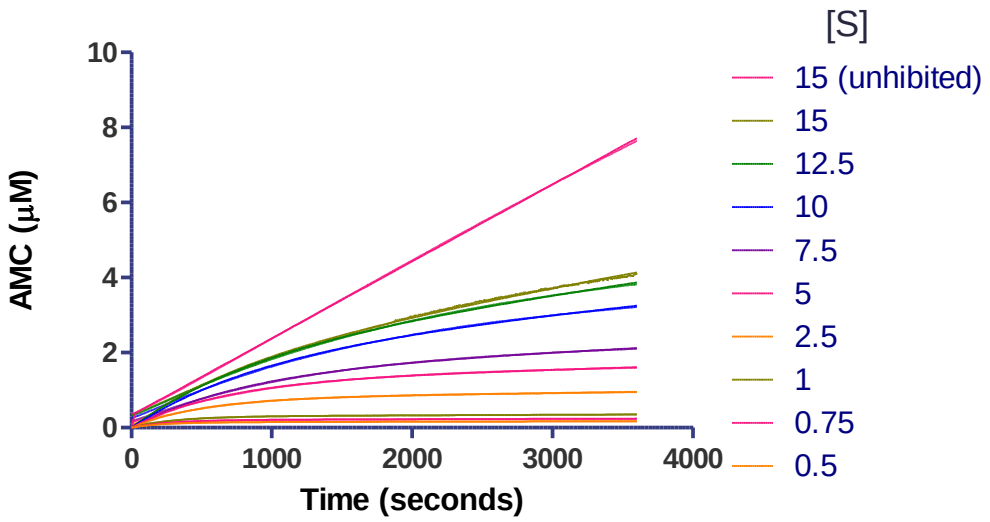



Figure 150. Cruzain Progress Curves with 0.1 μM **17** using Z-FR-AMC

Table 170 shows the average and standard error of calculated $k_{obs}$ as a function of [S]. Data were graphed as a function of the unitless substrate/Michaelis-Menten constant [S]/$K_M$, where $K_M$: 1.01 μM as previously determined. The rate constant $k_{obs}$ follows an inversely hyperbolic trend. The rate constant values increased as the concentration of Z-FR-AMC is decreased in solution. This is a clear indication that **17** is a slown binding reversible competitive inhibitor of cruzain with respect to the fluorogenic substrate.

Table 170. Effect of [Z-FR-AMC] on $k_{obs}$ Values when Using **17** against Cruzain

| [Z-FR-AMC] (μM) | $k_{obs}$ ± *Standard Error (x $10^3$, $s^{-1}$)* |
|---|---|
| 15 | 1.178 ± 0.067 |
| 12.5 | 1.039 ± 0.094 |
| 10 | 1.087 ± 0.011 |
| 7.5 | 1.187 ± 0.006 |
| 5 | 1.483 ± 0.014 |
| 2.5 | 2.101 ± 0.045 |
| 1.0 | 3.673 ± 0.148 |
| 0.75 | 4.172 ± 0.181 |
| 0.5 | 5.069 ± 0.195 |

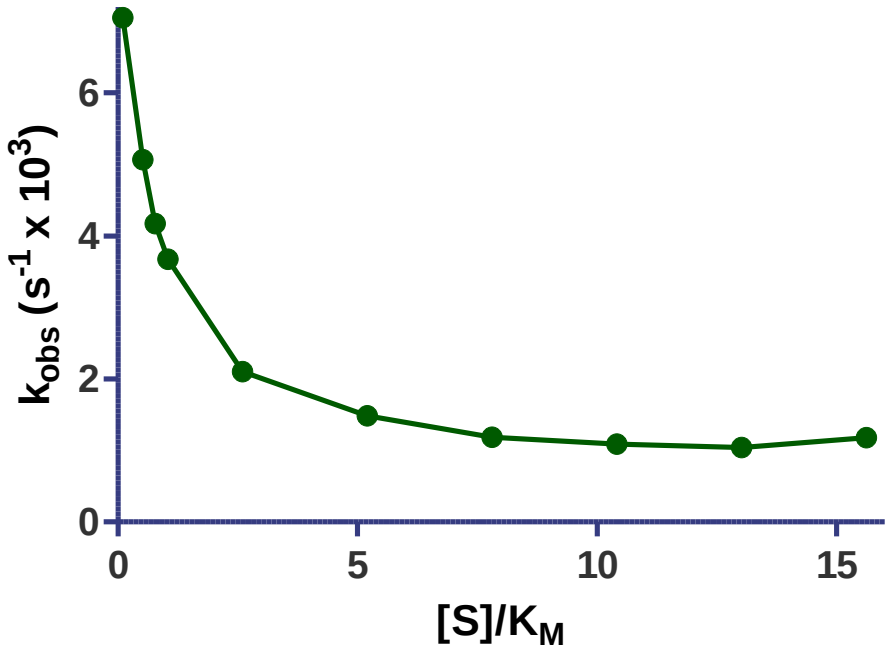



Figure 151. Effect of [Z-FR-AMC] on $k_{obs}$ values when using **17** against Cruzain

*Inhibition of Cruzain Collagenase Activity by Thiosemicarbazone Derivatives*

Cruzain is the most important protease found in *Trypanosoma cruzi*. It has also been shown that cruzain is also capable of degrading high molecular weight proteins and participates in host invasion by degrading surrounding tissue. Research has shown that cruzain is a validated target in the treatment of Chagas' disease. Thus, we explored the catalytic activity of cruzain by using a natural substrate, type I collagen from human skin. We also tested the ability of one of the lead compounds (**17,** $IC_{50}$: 12 nM), to inhibit the collagenase catalytic activity of cruzain. Untreated and treated samples were incubated at 37 °C under acidic conditions (pH 5.5). Four sets of samples were prepared for this experiment. Two sets were not treated with **17** (DMSO was used as control vehicle). The third and fourth sets were treated with 20 μM of compound **17** in 2% DMSO. The effect of preincubation activity on the catalytic activity of cruzain and the inhibitory potency of compound **17** was also tested**.** A preincubation time of 30 minutes was chosen to carry out these experiments.

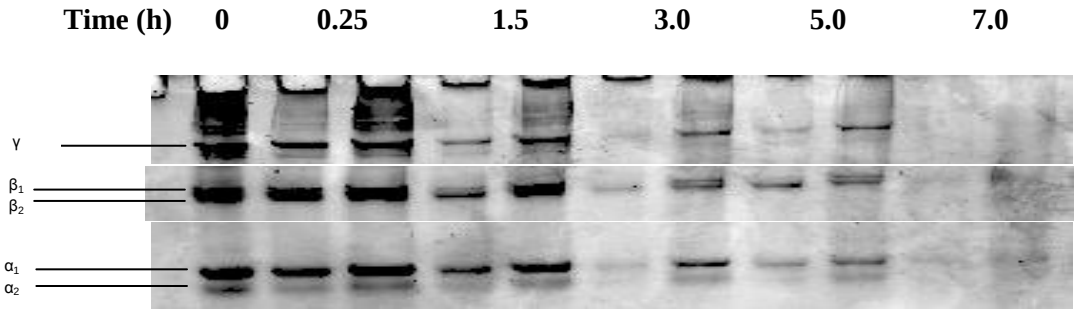
*Inhibition of Cruzain Collagenase Activity by Thiosemicarbazone Derivative **17**. (Preincubation Time: 0.5 Hours).*

A sample containing three microliters of 1.2 μM Cruzain in CLI and 2.5 μl of 104 μM compound **17** were preincubated as previously described for thirty minutes. Individual samples were prepared to complete a series of six samples. Then, 7.5μl of 0.4 mg/ml type I collagen in acetic acid were added to the cruzain-inhibitor mixture. Every sample was carefully mixed and placed in a 37°C water bath. Reactions were monitored between 0 and 7 hours, stopped with 2.5 μl LDS sample buffer and heated at 90 °C for

every time point. Inactivated samples were immediately stored at -80 °C to preserve them.

Untreated and treated samples were loaded onto 4-12% Bis-Tris SDS gel. A sample of molecular weight standards, one sample with type I collagen only, and one sample of cruzain only were also loaded as reference controls. These last two samples were inactivated after thirty minutes of preincubation time. Electrophoresis was performed at 200 V for 60 minutes. The gel was rinsed with water and placed into a 1X SYPRO® solution, a fluorescent protein staining dye. The gel was stained for a minimum of one hour at room temperature. The destaining process was made by washing the gel once with water and twice with 7.5% acetic acid. Finally, a digital image of the gel was obtained by using a GE Typhoon 9400 fluorescence scanner imager. Figure 152 shows the results of degradation of type I collagen from human skin using human cruzain soluble in acidic solutions.[330] The activity of the protease was stopped at different time points (0-7 hours). The figure shows the progress of the reaction for the first three hours after collagen I was added to the cruzain-inhibitor mixture that was previously incubated for thirty minutes. The first lane is type I collagen that was not treated with cruzain in the vehicle control (2% DMSO). Reported literature establishes that type I collagen is large protein that is divided into three defined chains: α, β, and γ heavy chains and is crosslinked and polymerizes into fibrils. The molecular weights of the heavy chains are approximately $80 \leq MW_{\alpha} \leq 125$ kDa for α chain ; $160 \leq MW_{\beta} \leq 250$ kDa for β chain; and $240 \leq MW_{\gamma} \leq 375$ kDa for γ chain.[331,332] Results show the presence of three major bands with high molecular weights that are consistent with the reported values for α, β, and γ heavy chains in the literature. molecular weights are: α ≈ 102 kDa,

β ≈ 150 kDa and γ ≈ 200 kDa (lane 1). The second and third bands of represent untreated and treated samples that were stopped after 15 minutes of the natural substrate. Untreated samples contained type I collagen, and cruzain in DMSO, while treated samples consisted of type I collagen, cruzain, and compound **17** in DMSO. The untreated sample showed a slight degradation process even in as little as 15 minutes. The α, β, and γ bands look more diminished with respect to their respective control. The degradation is more evident in the case of the α bands (lane 2). The degradation process is evident after 60 and 180 minutes. All of the bands α, β, and γ, look diminished and there is evidence that heavy bands (compared to untreated collagen at time 0, lane 1) are degraded. A comparison between treated (lanes 3, 5 and 7) and untreated (lanes 4, 6 and 8) samples reveals that compound **17** was able to inhibit the collagenase activity of cruzain. Longer reaction times (t= 7 hours) indicate that uninhibited cruzain is able to completely degrade 3 μg of type I collagen. These results indicate that the **17** is capable of inhibiting cruzain collagenase activity by approximately fifty percent up to 3 hours.




| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Collagen | + | + | + | + | + | + | + | + | + | + | + |
| Cruzain | - | + | + | + | + | + | + | + | + | + | + |
| **17** | - | - | + | - | + | - | + | - | + | - | + |

Figure 152. Inhibition of Collagenase Activity of Cruzain by **17**, Preincubation Time: 0.5 hours.

*Molecular Docking Studies with Thiosemicarbazones as Inhibitors of Cruzain*

An evaluation of the library of thiosemicarbazones showed several compounds are outstanding inhibitors of cruzain with $IC_{50}$ values less than 20 nM. Thus, the most potent five compounds in the library (**1**, **9**, **10**, **17**, and **36**) were docked with cruzain using 1ME3 as a reference structure. Also two compounds with modest (**8**) or no activity (**58**) were modeled for comparison purposes.

The entire library, with the exception of **58** suggested the formation of covalent bonds between the thiocarbonyl carbon and Cys25 (papain numbering) thiolate group of cruzain ($C—S^-$), the active reside in the catalytic triad. Calculated distances between these two groups varied between 3.144 (**36)** and 3.467 Å (**9**). The proximity between both groups promotes the formation of a transient covalent bond. Furthermore, the distance for **58** is 4.59 Å, which could have explained the inefficacy of this analog to inhibit cruzain. The modification of the thiosemicarbazone group might be the reason for this behavior. **1** was modeled in previous work ($IC_{50}$ : 12 nM) and served as a positive control for the construction of the screened library. Contacts provided electrostatic interactions with partial positive charges which also stabilize the oxyanion hole and provided evidence for the formation of the transient tetrahedral intermediate formed between the inhibitor and the active site of cruzain. The series of docked molecules with cruzain are shown as follows: cruzain is in the ribbon diagram, (with the exception of the active site residues that were ball and stick format), and thiosemicarbazones are in space-filling format. Atoms were colored by using a modified CPK color code (carbon: green; oxygen: red; sulfur: yellow; bromine: burgundy; nitrogen: blue; hydrogen: white).

*Molecular Docking of **9** with Cruzain*

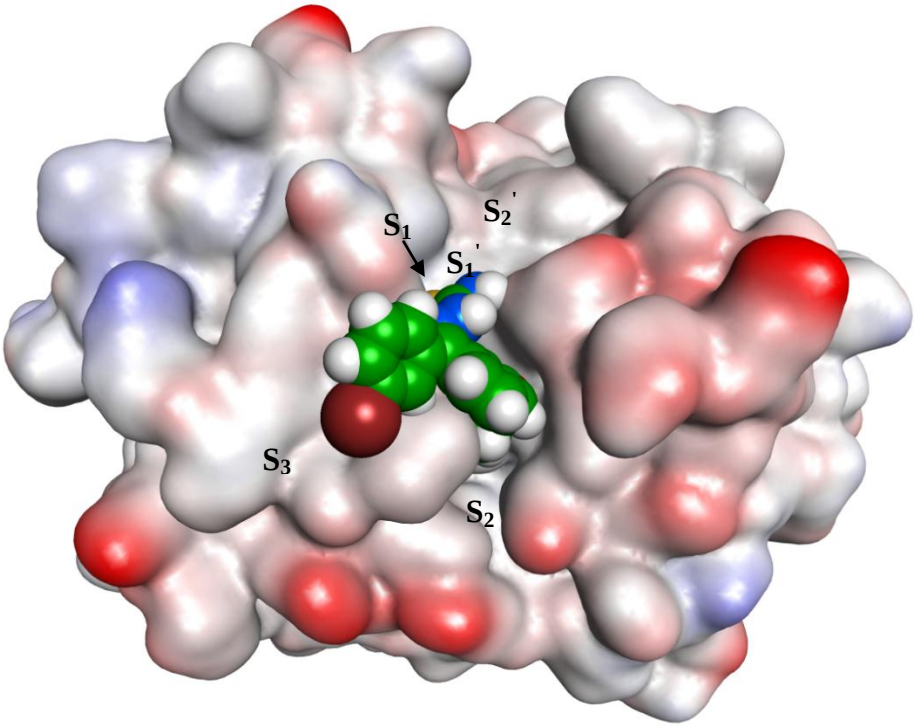
Analog **9**, 3-bromo-3´-acetobenzophenone thiosemicarbazone was the second most potent compound found in the library containing several thiosemicarbazones. Its $IC_{50}$ value was calculated to be 12.1 nM. The relative interaction energy for the top conformation was calculated to be -62.6 kcal/mol. The top conformation was examined to explore important interactions and possible hydrogen bonds. The thiocarbonyl arm is in the S1 pocket. The functionalized phenyl arm with the acetate group is in the S2 pocket, such that the acetate group is facing down and toward the S2 pocket while the bromophenyl arm is located in the right-hand side of the S3 pocket, so the bromine substituent is pointing out and facing down. Important interactions between **9** and the active site of cruzain are:

1. Hydrogen bond between the backbone carbonyl oxygen of Asp158 with one of the hydrogens of **9** $NH_2$ group (2.01 Å). This hydrogen bond helps to orient the thiosemicarbazone moiety at the active site.
2. The (C—S) distance is 3.467 Å

*Molecular Docking of **10** with Cruzain*

Analog **10**, 3-bromo-3',5'-difluorobenzophenone thiosemicarbazone was a potent inhibitor. Its $IC_{50}$ value was found to be 17.6 nM. The energy interaction for the top conformation was calculated to be -52.8 kcal/mol. The thiocarbonyl arm is in the S1 pocket. The functionalized bromophenyl arm is located deep in the S2 pocket, similar to **9**. Interestingly, the difluorophenyl arm is located along the right-edge of the S3 pocket placing the substituents (fluorines) outwards. Important interactions between **10** and the active site of cruzain are:

**A**




**B**

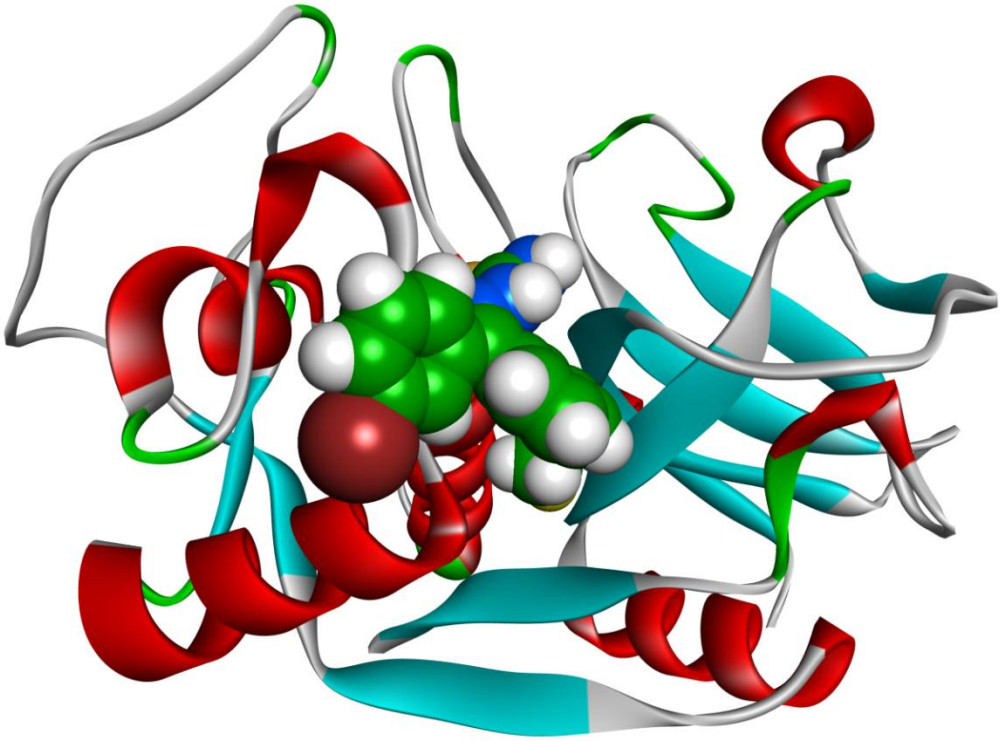


Figure 153. Molecular Docking of **9** with Cruzain. **A**. Electrostatic Surface of Cruzain Interacting with **9**. B. Ribbon Diagram.

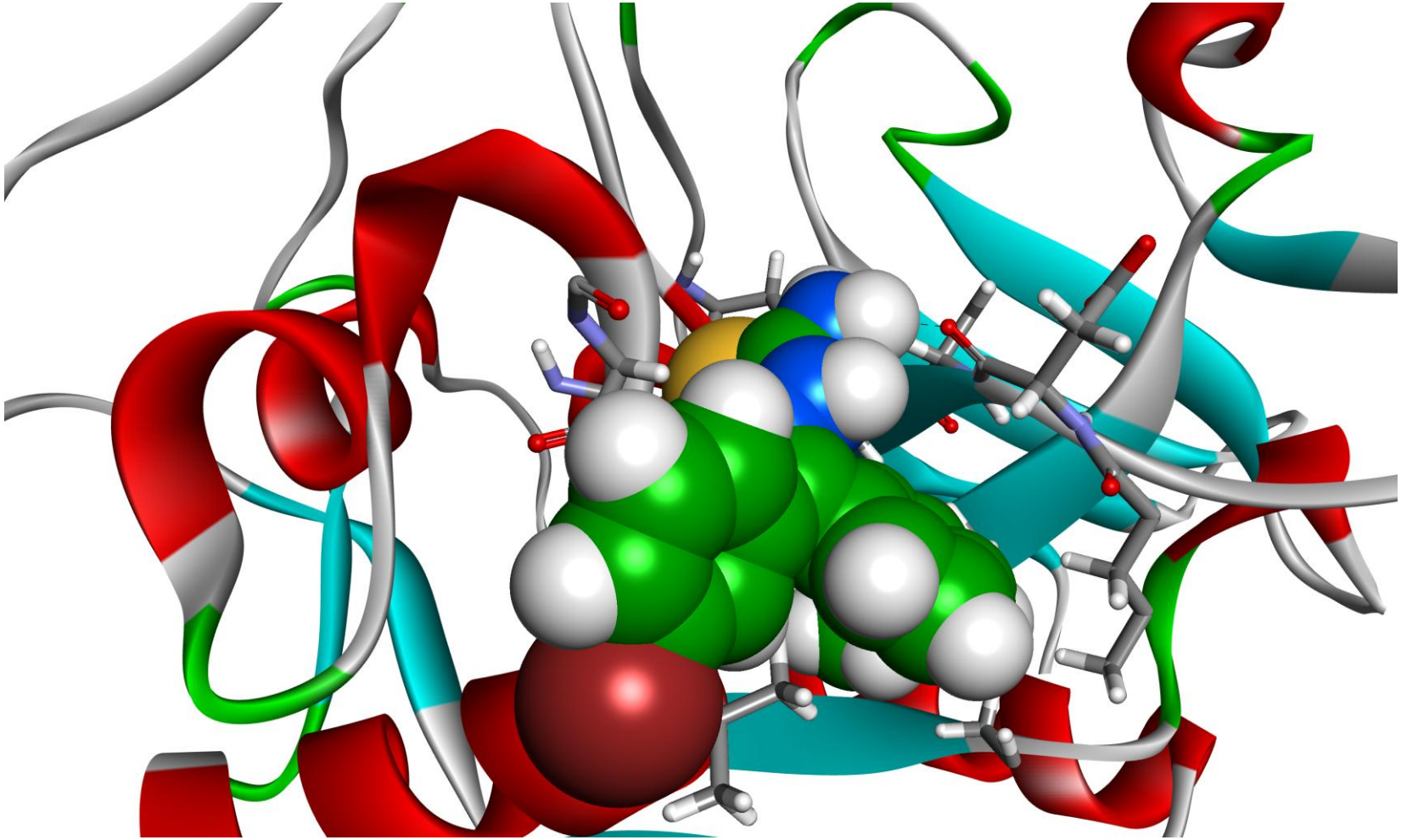


Figure 154. Molecular Docking of Cruzain with **9**

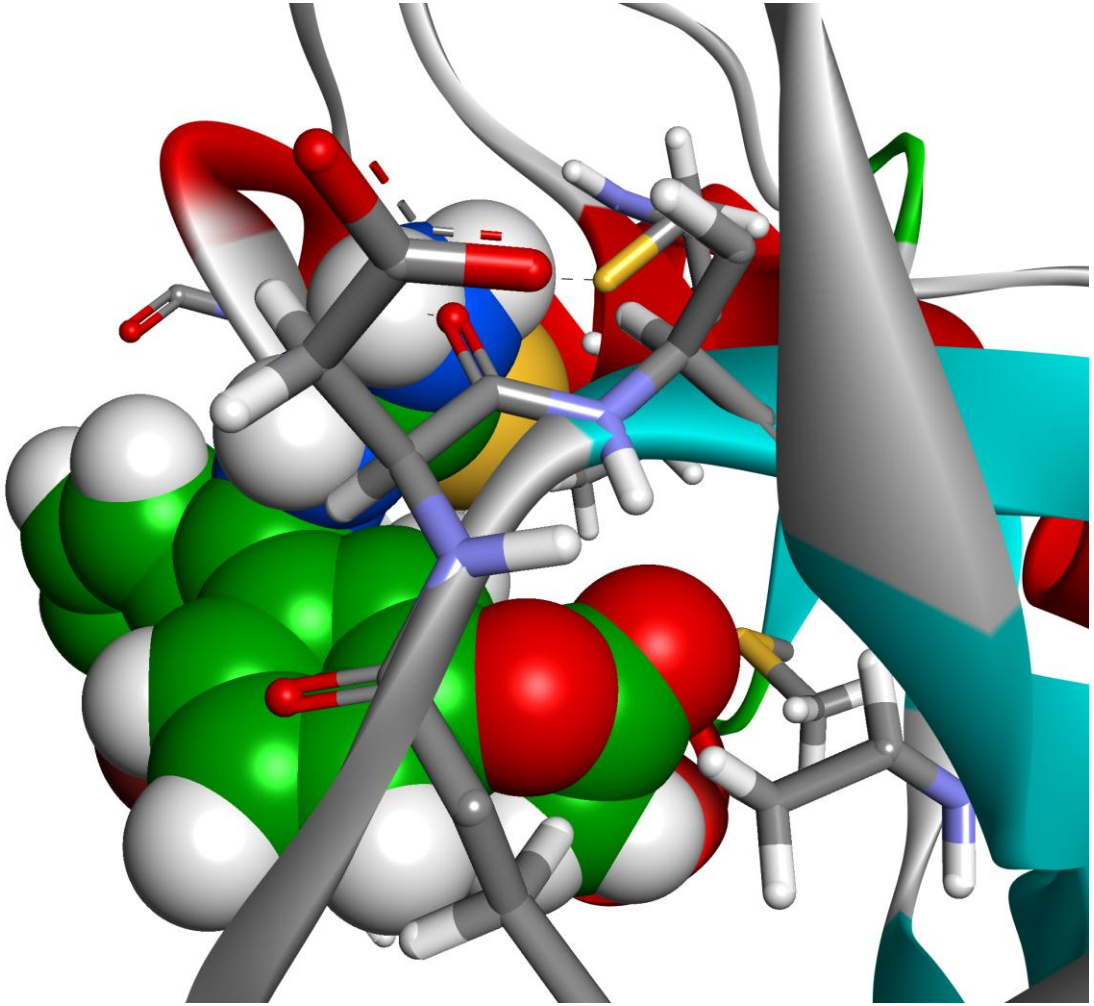


Figure 155. Molecular Docking of **9** with Cruzain. The Thiocarbonyl Carbon and Cys25 (papain numbering) Thiolate Group of Cruzain (C─S) Distance is 3.467 Å

1. The thiocarbonyl sulfur is in close proximity with the side-chain carboxamide hydrogen of Gln19 ( 2.26 Å)
2. The (C—S) distance is 3.193 Å.

*Molecular Docking of **17** with Cruzain*

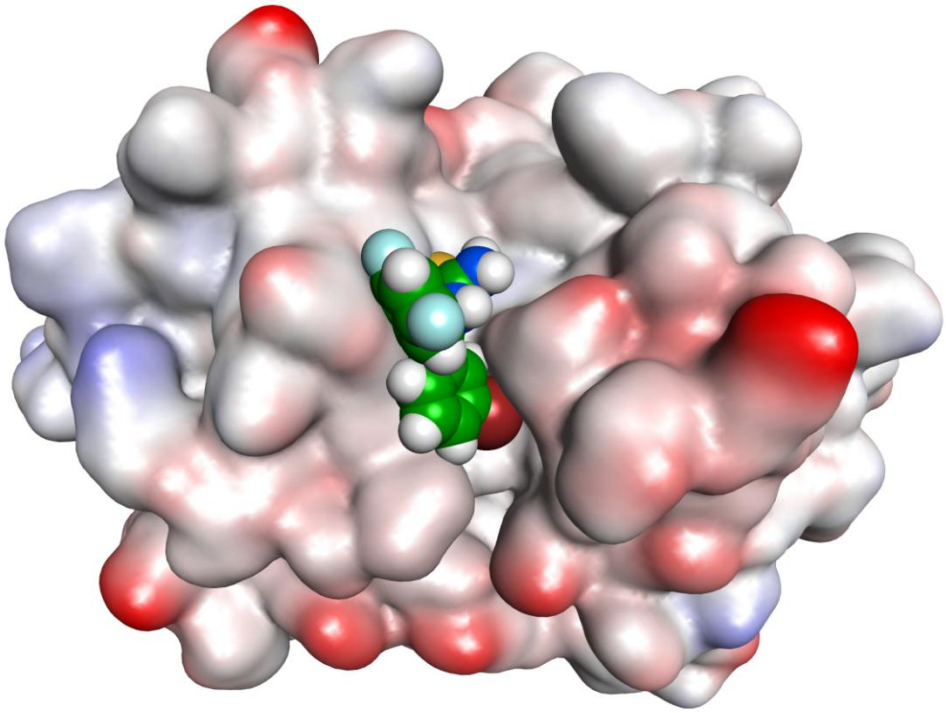
Analog **17**, 3-bromo-4'-fluorobenzophenone thiosemicarbazone was a potent inhibitor with an $IC_{50}$ value of 14.3 nM. The top energy interaction was -55.0 kcal/mol. The thiocarbonyl arm is in the S1 pocket. The functionalized bromophenyl arm is located deep in the S2 pocket, similar to **9**. Interestingly, the fluorophenyl arm is located along the right-edge of the S3 pocket placing the substituents (fluorines) outwards. Important interactions between **17** and the active site of cruzain are:

1. The thiocarbonyl sulfur is in close proximity with the side-chain carboxamide hydrogen of Gln19 ( 2.26 Å, see compound **10**)
2. The (C—S) distance is 3.212 Å.

*Molecular Docking of **36** with Cruzain*

Analog **36**, a dibrominated thiophene thiosemicarbazone had also excellent inhibitory activity towards recombinant cruzain ($IC_{50}$: 20.2 nM). The top conformation interaction energy was -52.9363kcal/mol. Overall, it was found to have its thiosemicarbazone arm in the right side of the S3 pocket (bromine facing downwards). The thiophene moiety is in the S2 pocket placing the thiophene sulfur outward and the thiophene bromine downwards. The (C—S) distance is 3.144 Å, the shortest distance found in the set of six models. Two hydrogen bonds were found between the backbone carbonyl oxygen of Asp158 and two hydrogens of the thiosemicarbazone moiety of **36**.

**A**



**B**

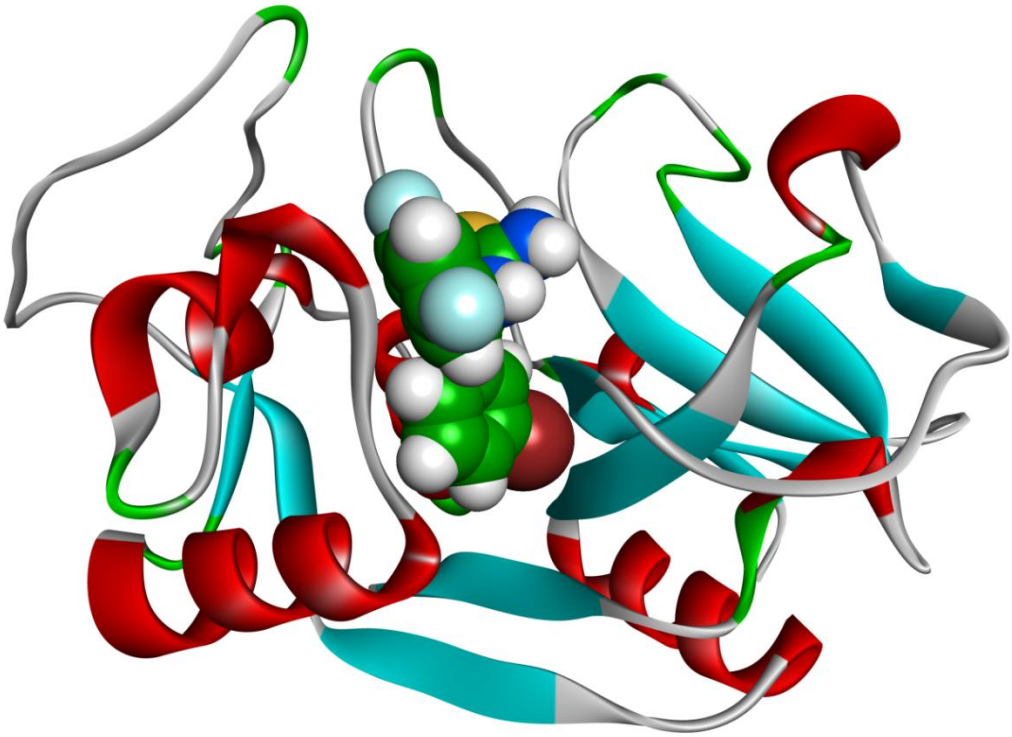


Figure 156. Molecular Docking of **10** with Cruzain. **A**. Electrostatic Surface of Cruzain Interacting with **10**. B. Ribbon Diagram

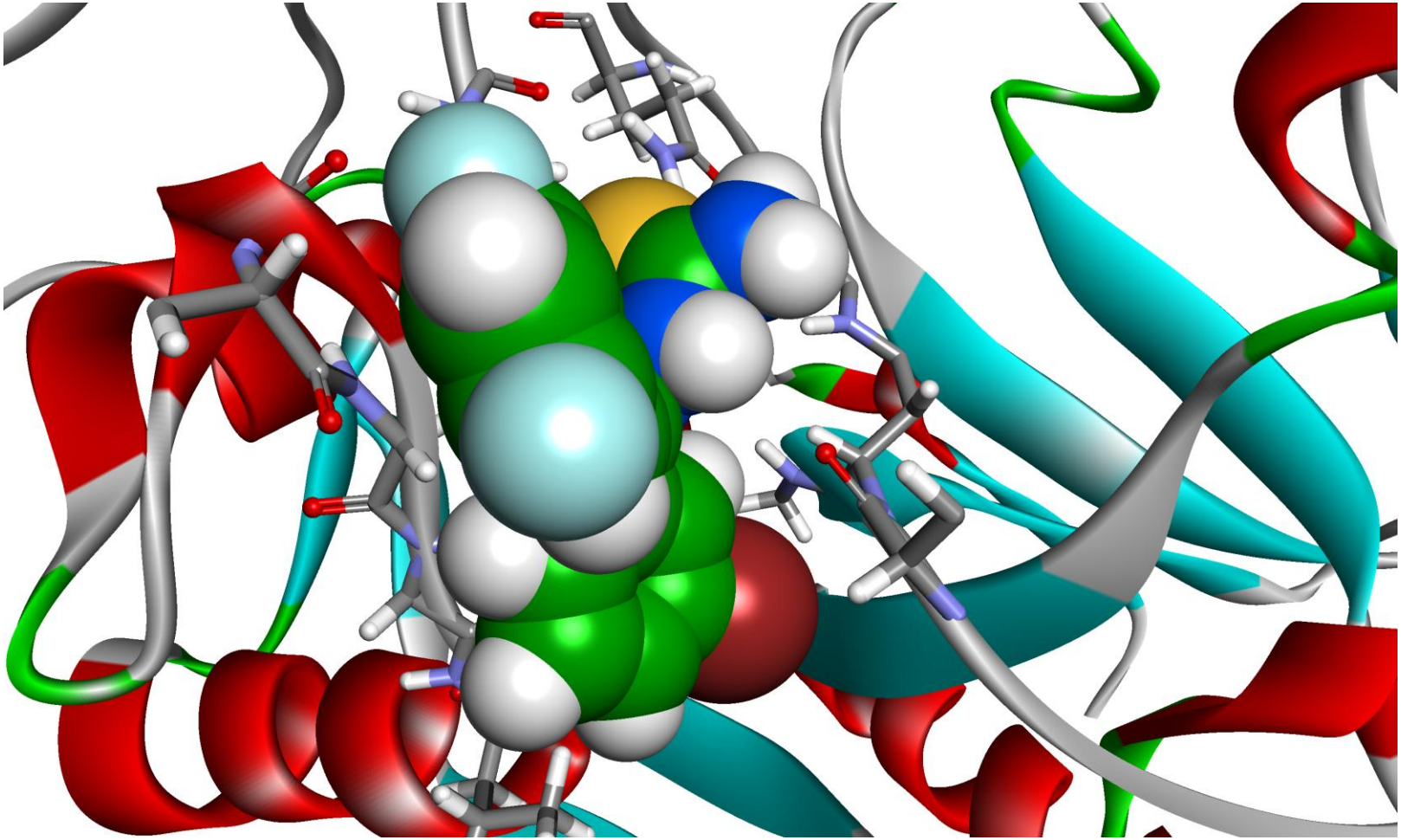


Figure 157. Molecular Docking of Cruzain with **10**

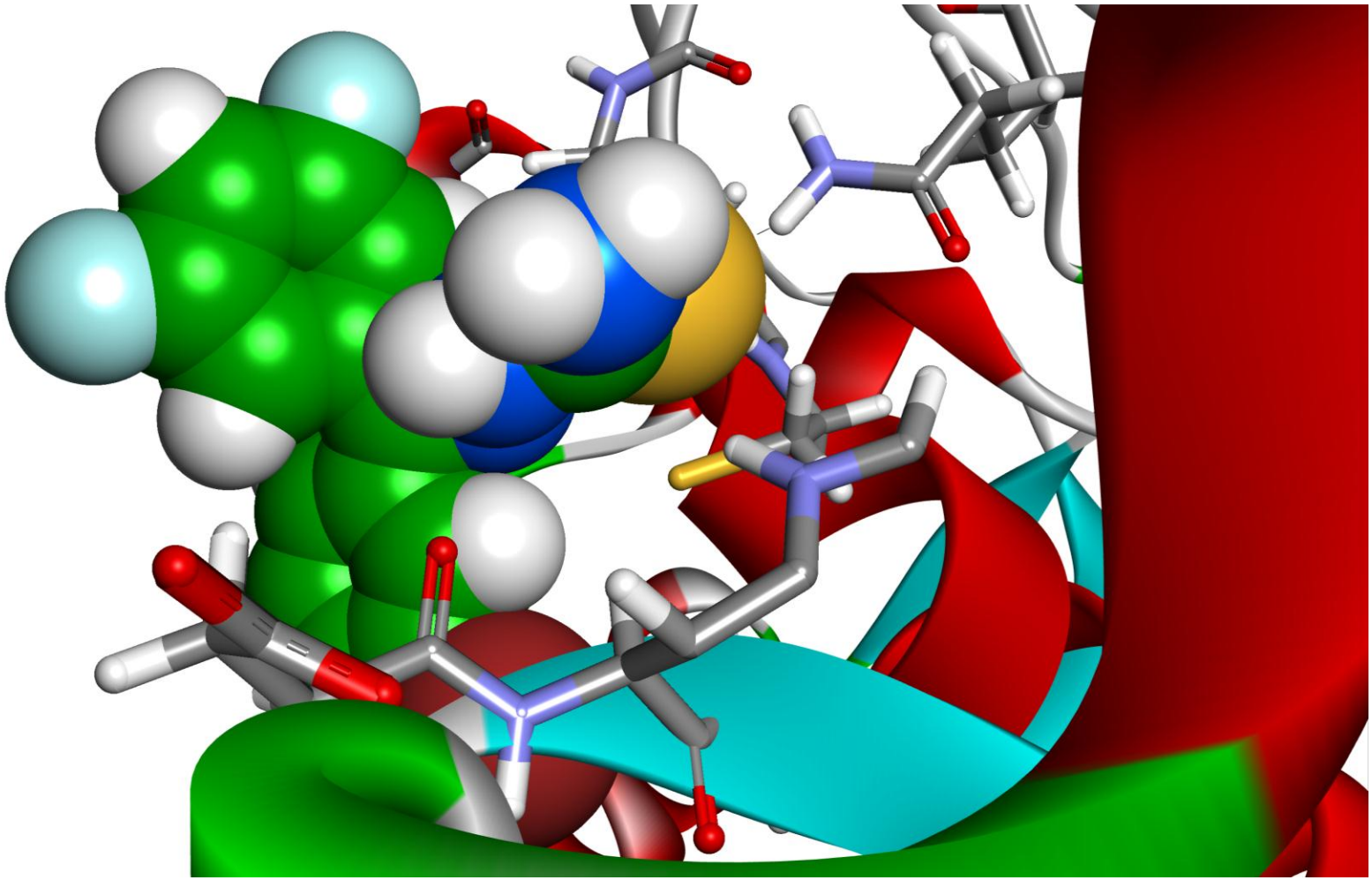
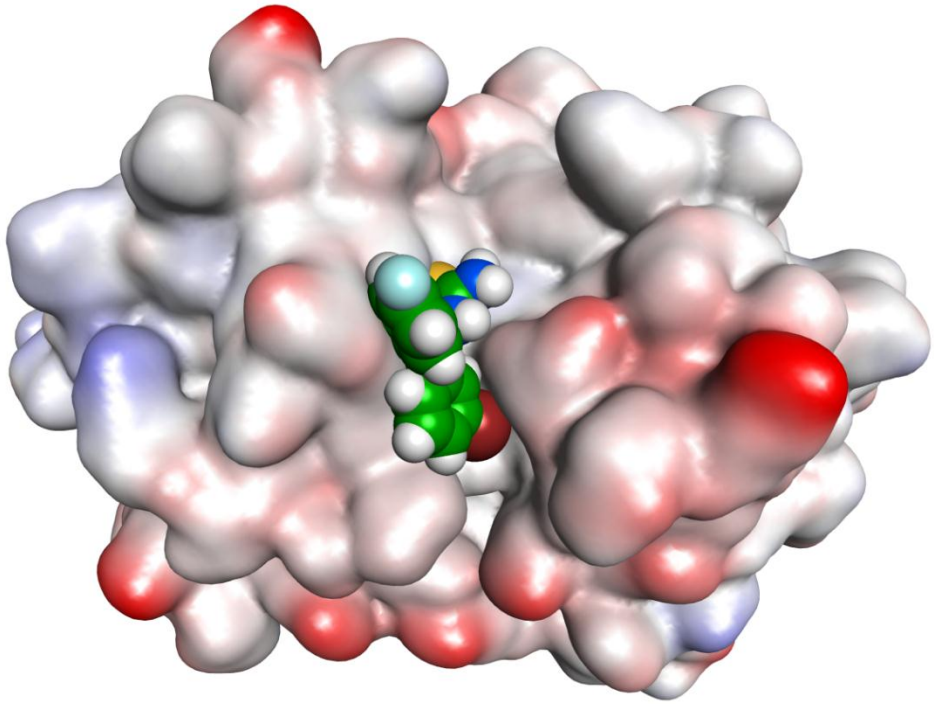


Figure 158. Molecular Docking of **10** with Cruzain. The Thiocarbonyl Carbon and Cys25 (papain numbering) Thiolate Group of Cruzain (C─S) Distance is 3.19 Å

**A**



**B**

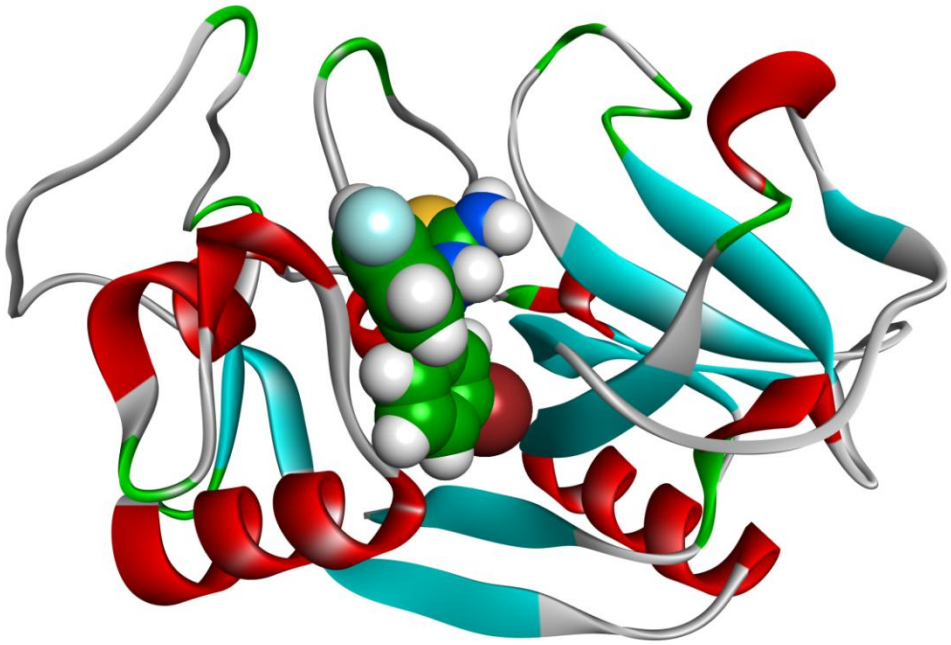


Figure 159. Molecular Docking of **17** with Cruzain. **A**. Electrostatic Surface of Cruzain Interacting with **17**. B. Ribbon Diagram

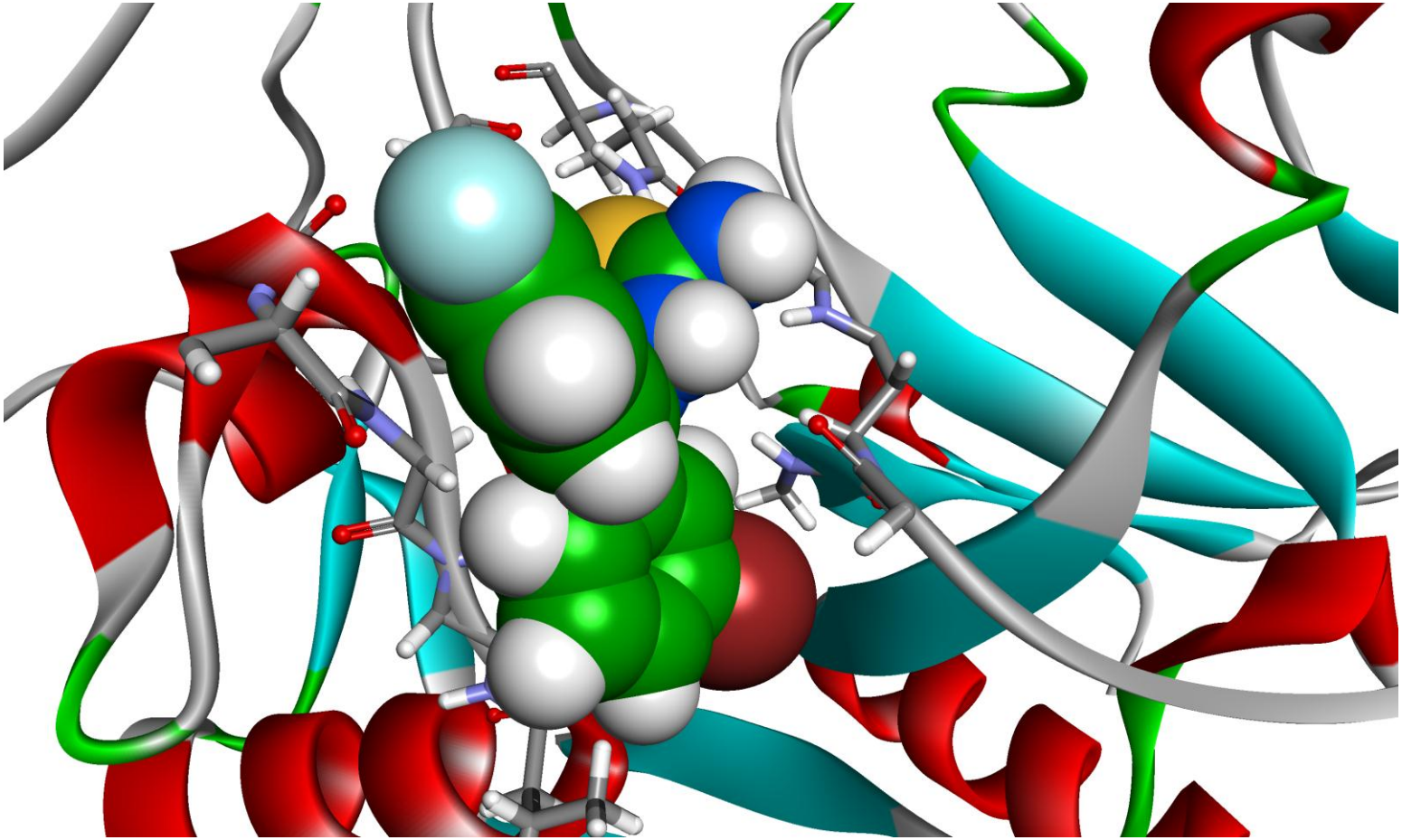


Figure 160. Molecular Docking of Cruzain with **17**

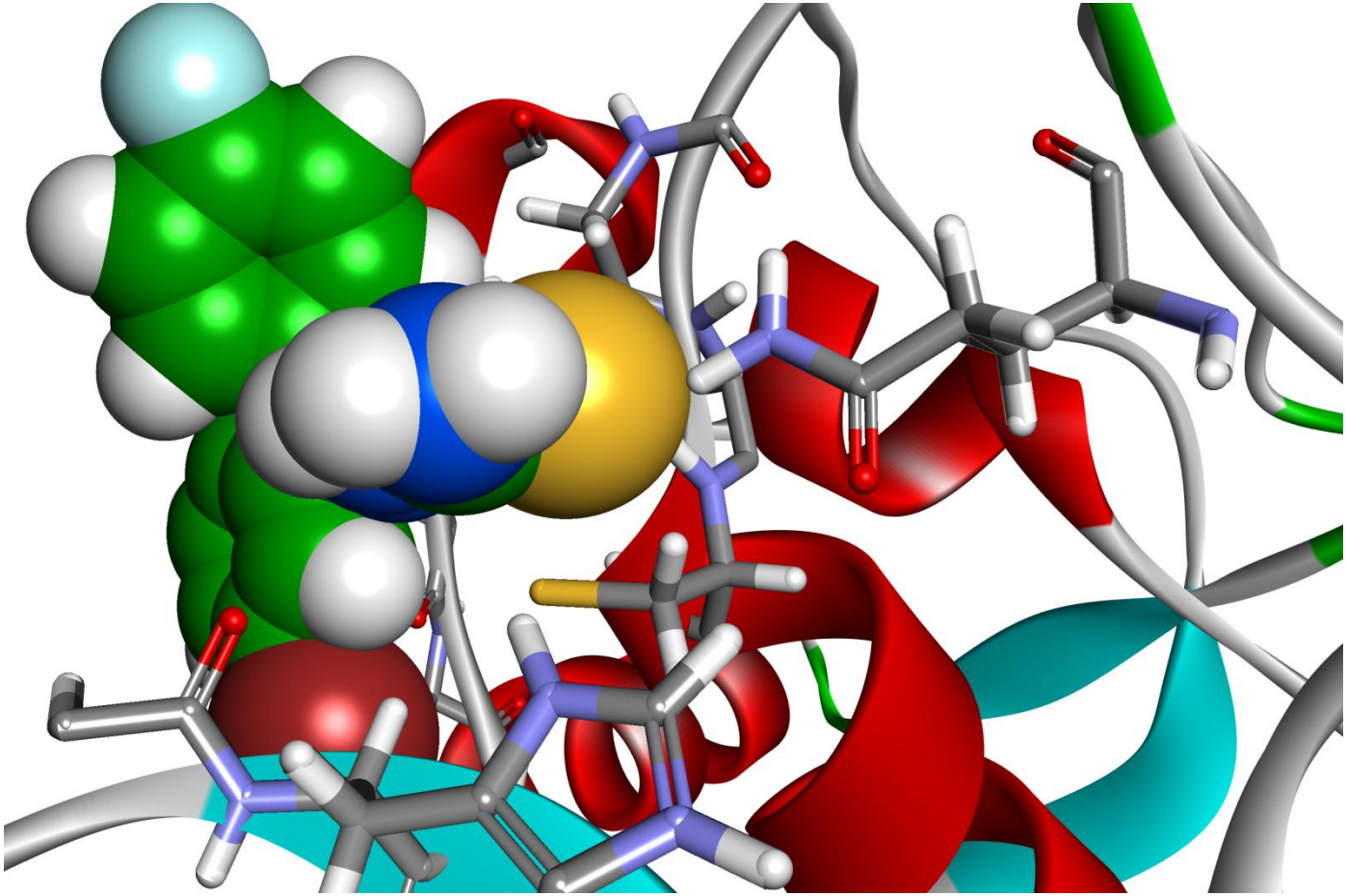


Figure 161. Molecular Docking of **17** with Cruzain. The Thiocarbonyl Carbon and Cys25 (papain numbering) Thiolate Group of Cruzain (C─S) Distance is 3.21 Å

*Molecular Docking of **58** with Cruzain*

Finally, **58**, an N-phenylated dibromobenzophenone analog with no activity was also modeled. Both bromophenyl arms are located in the S1 and S3 pockets having their substituents facing outward. The phenylated thiosemicarbazone is located in the S2 pocket and in an orientation that does not promote formation of a transient covalent bond. The (C─S) distance is: 4.597, which is considerably much larger than the distance found for the other compounds that were docked with cruzain. The lack of hydrogen bonds in the model, which are very important in orient the compound at the active site, might explain the poor activity of this compound. This model also confirmed the importance of the thiosemicarbazone group in the inhibitory activity of these nonpeptidic compounds.

## *Conclusions*

A series of benzophenone and other functionalized thiosemicarbazones were analyzed as potential inhibitors of cruzain. Five compounds, including a nonbenzophenone thiosemicarbazone showed excellent inhibitory activity toward cruzain with $IC_{50}$ values less than 20 nM.

Figure 168 shows the Structure-Activity Relationship of these functionalized benzophenone thiosemicarbazones based on their chemical structures. Advanced kinetics were used to investigate the mechanism of four of these compounds (**1, 9, 10, 17**). These compounds were determined to be time dependent, reversible inhibitors of cruzain.

Analog **1** was a competitive inhibitor by examining the effect of substrate concentration on $IC_{50}$ values. Similarly, **17** was also a competitive inhibitors of the fluorogenic substrate Z-FR-AMC. The decrease in $k_{obs}$values as a function of Z-FR-

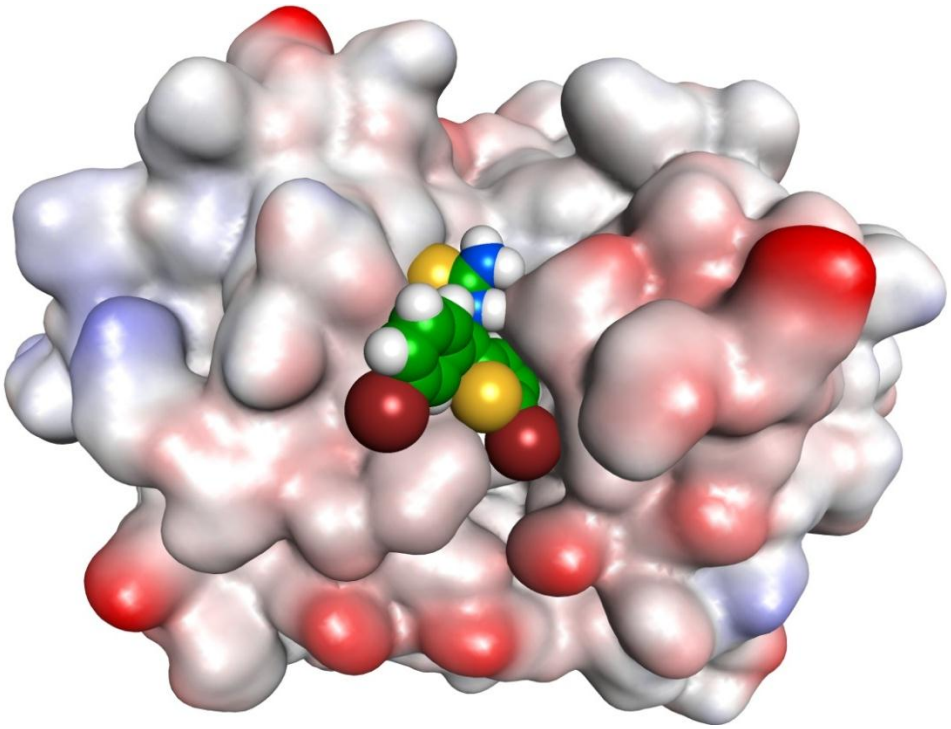
AMC (substrate) concentration in cruzain progress curves indicates that **17** competes with the substrate for binding to the enzyme active site.

The inhibition of type I collagenase activity of recombinant cruzain was examined by using a natural substrate for the enzyme, type I collagen from human skin. The proteolytic activity of cruzain was inhibited by 3-bromo-4'-fluorobenzophenone thiosemicarbazone.

Finally, six of these analogs were docked with cruzain using computational software. Modification of the thiosemicarbazone is detrimental in the activity of these compounds. Molecular docking with analog **58**, N-phenyl-3-3'-dibromobenzophenone thiosemicarbazone ($IC_{50} \geq 10000$ nM), indicated that the bulky phenyl group was located in the S2 pocket and the thiosemicarbazone was not the correct orientation for formation of a covalent bond with the enzyme. The lack of hydrogen bonds may explain the inactivity of these compounds.

For the most potent inhibitors, molecular docking with cruzain showed that thiosemicarbazones are placed at the S1 subsite in close proximity to cruzain Cys25 and in the correct orientation for a chemical reaction.

**A.**



**B.**

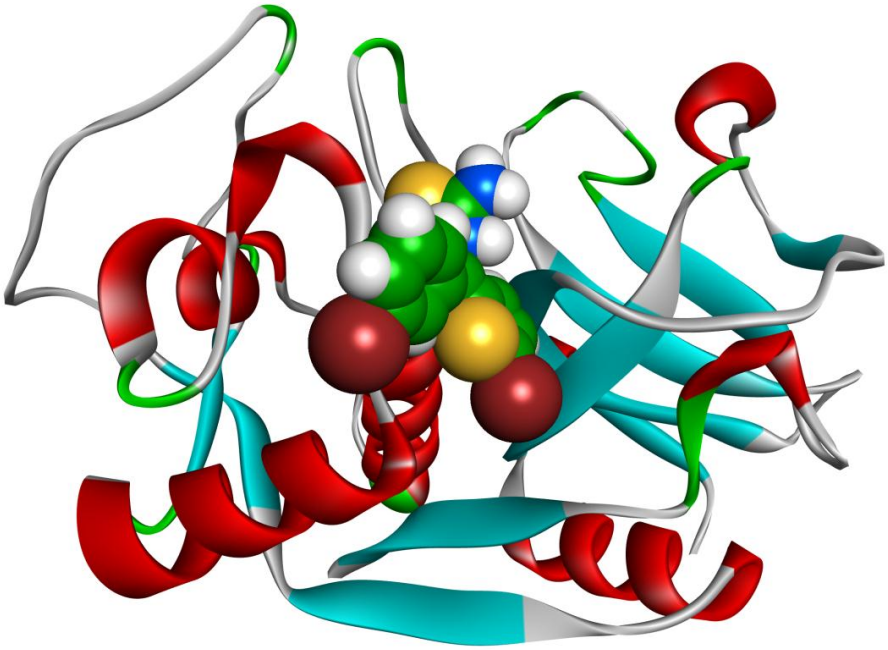


Figure 162. Molecular Docking of **36** with Cruzain. **A**. Electrostatic Surface of Cruzain Interacting with **36**. B. Ribbon Diagram

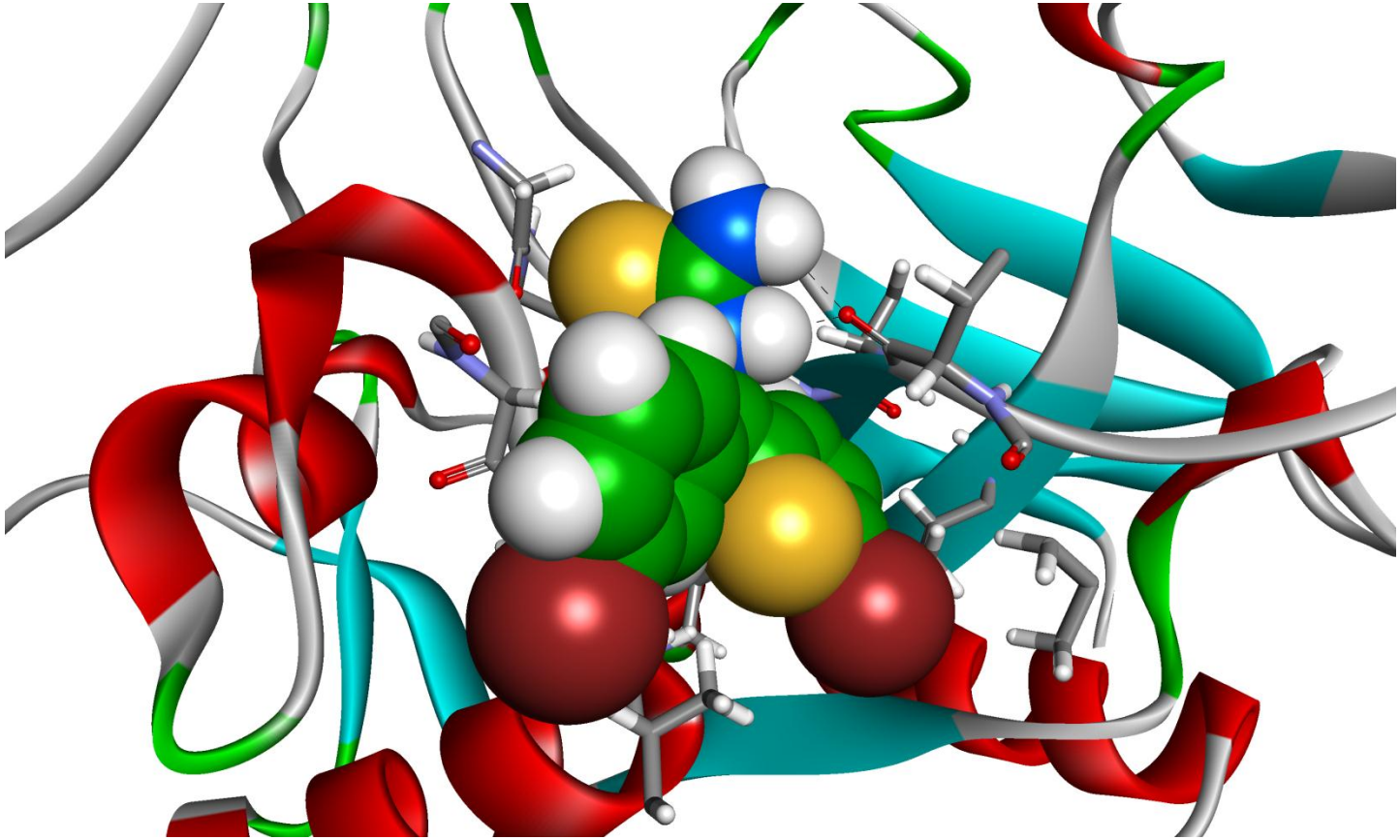


Figure 163. Molecular Docking of Cruzain with **36**

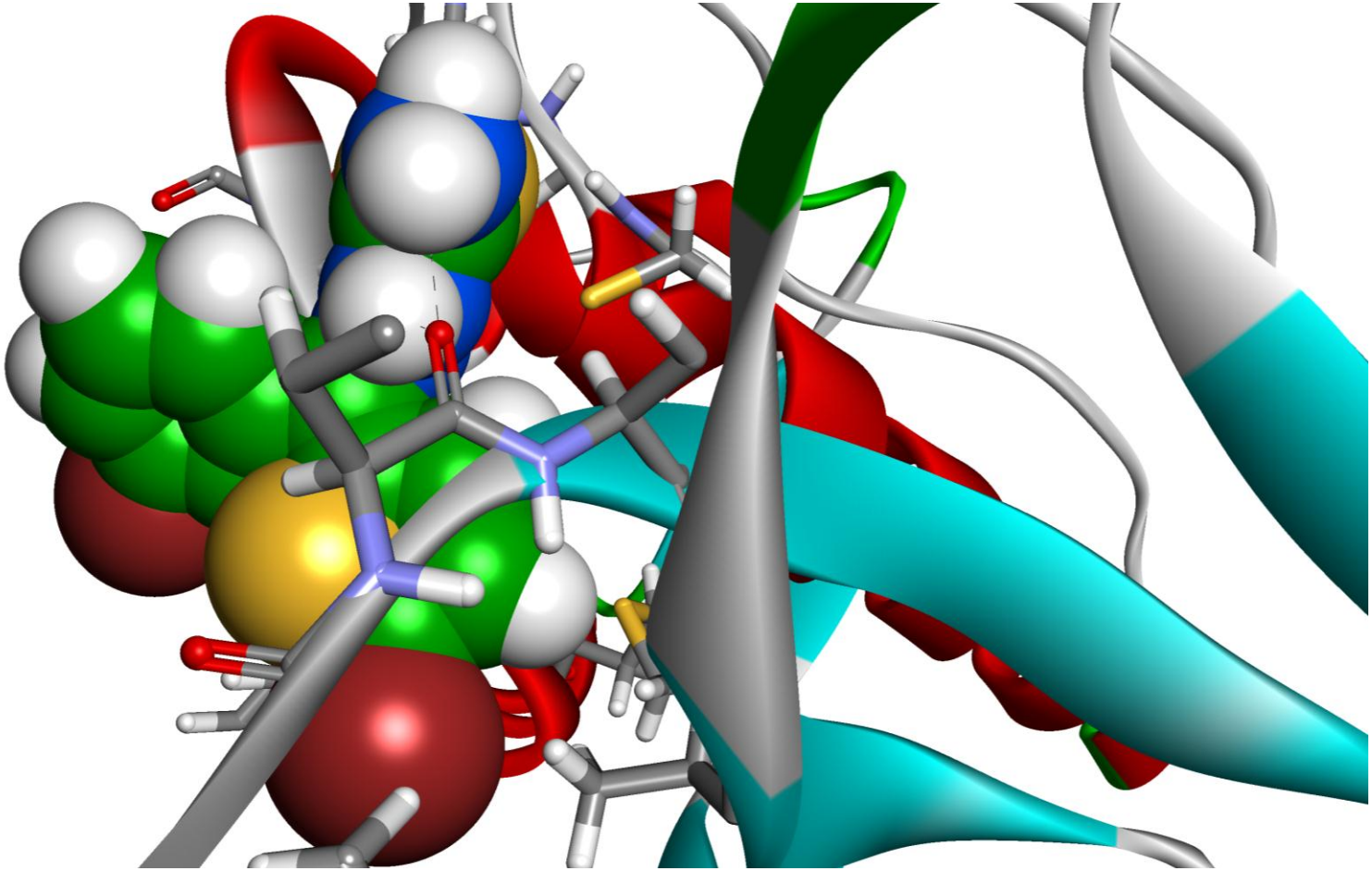
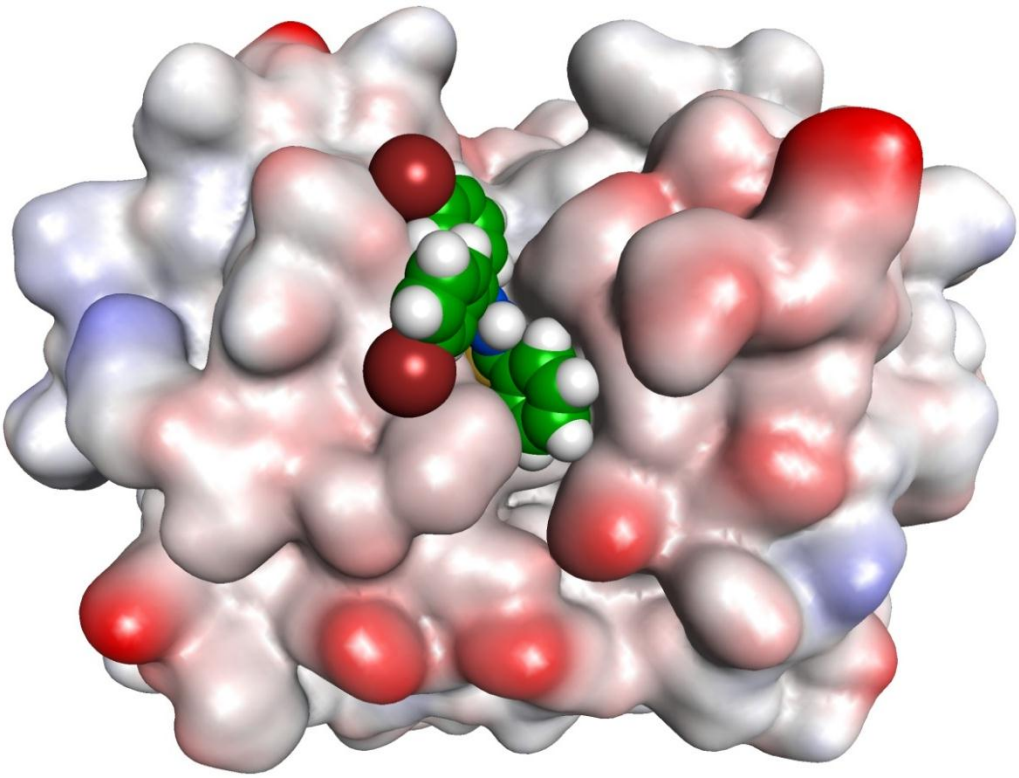


Figure 164. Molecular Docking of **36** with Cruzain. The Thiocarbonyl Carbon and Cys25 (papain numbering) Thiolate Group of Cruzain (C─S) Distance is 3.144 Å

**A**



**B**

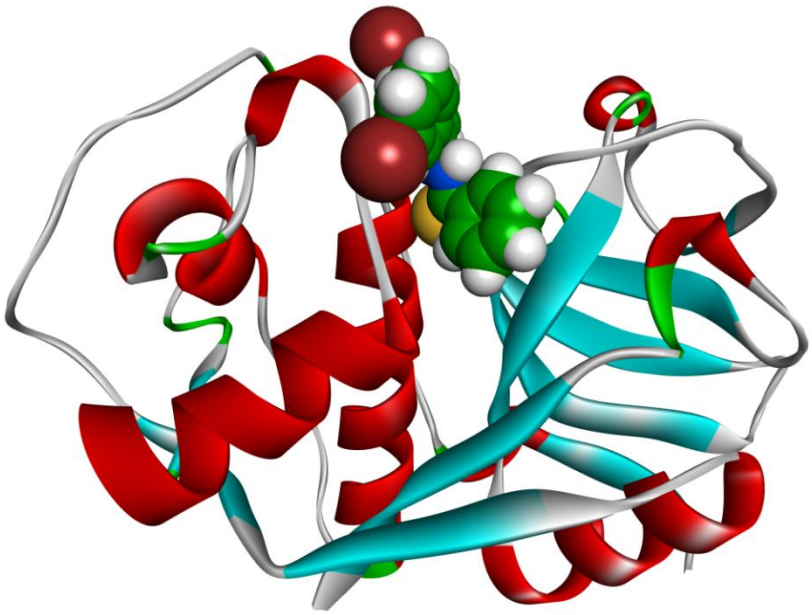


Figure 165. Molecular Docking of **58** with Cruzain. **A**. Electrostatic Surface of Cruzain Interacting with **58**. B. Ribbon Diagram

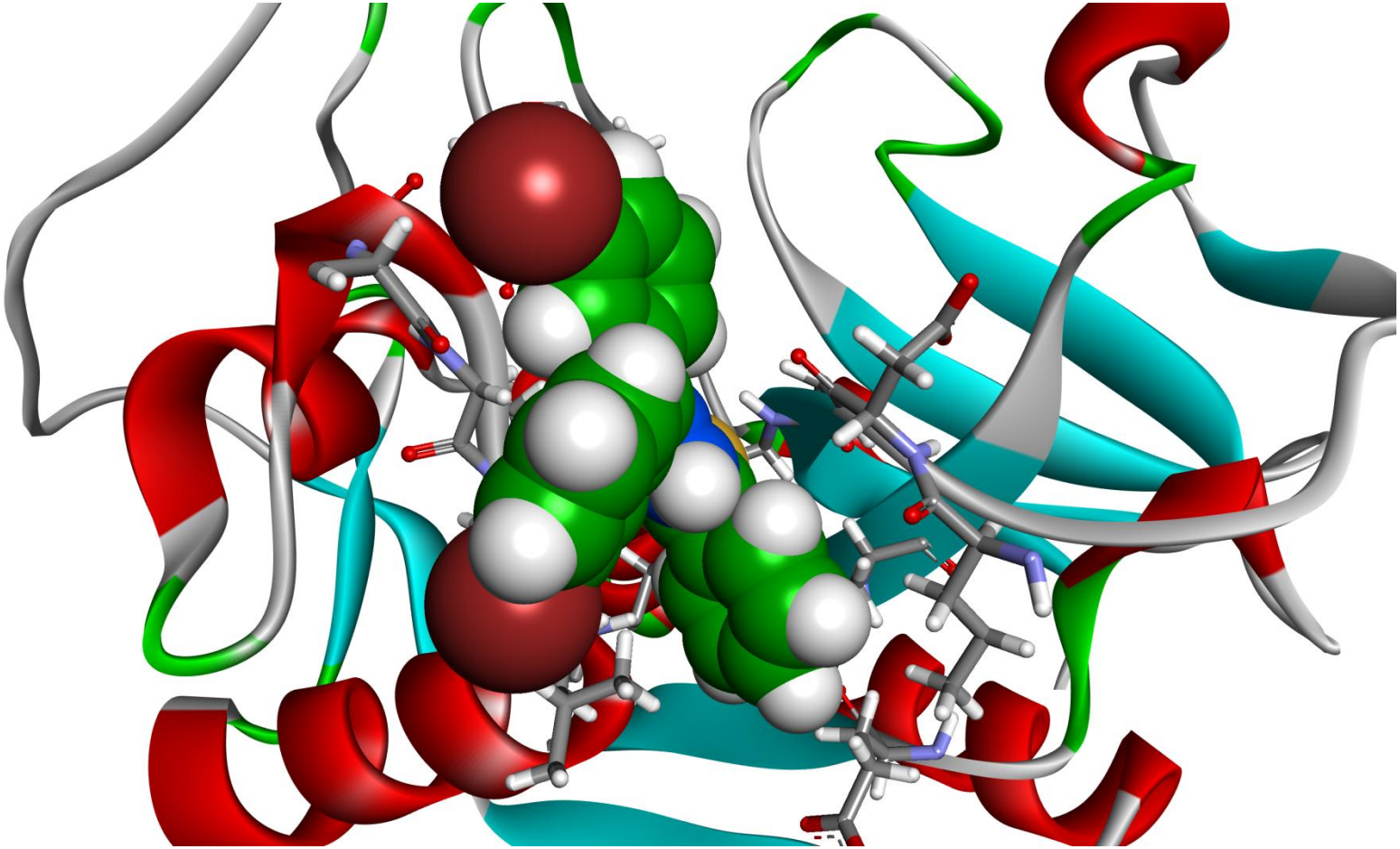


Figure 166. Molecular Docking of Cruzain with **58.** No Hydrogen Bonds were Formed Between **58** and Cruzain Active Site

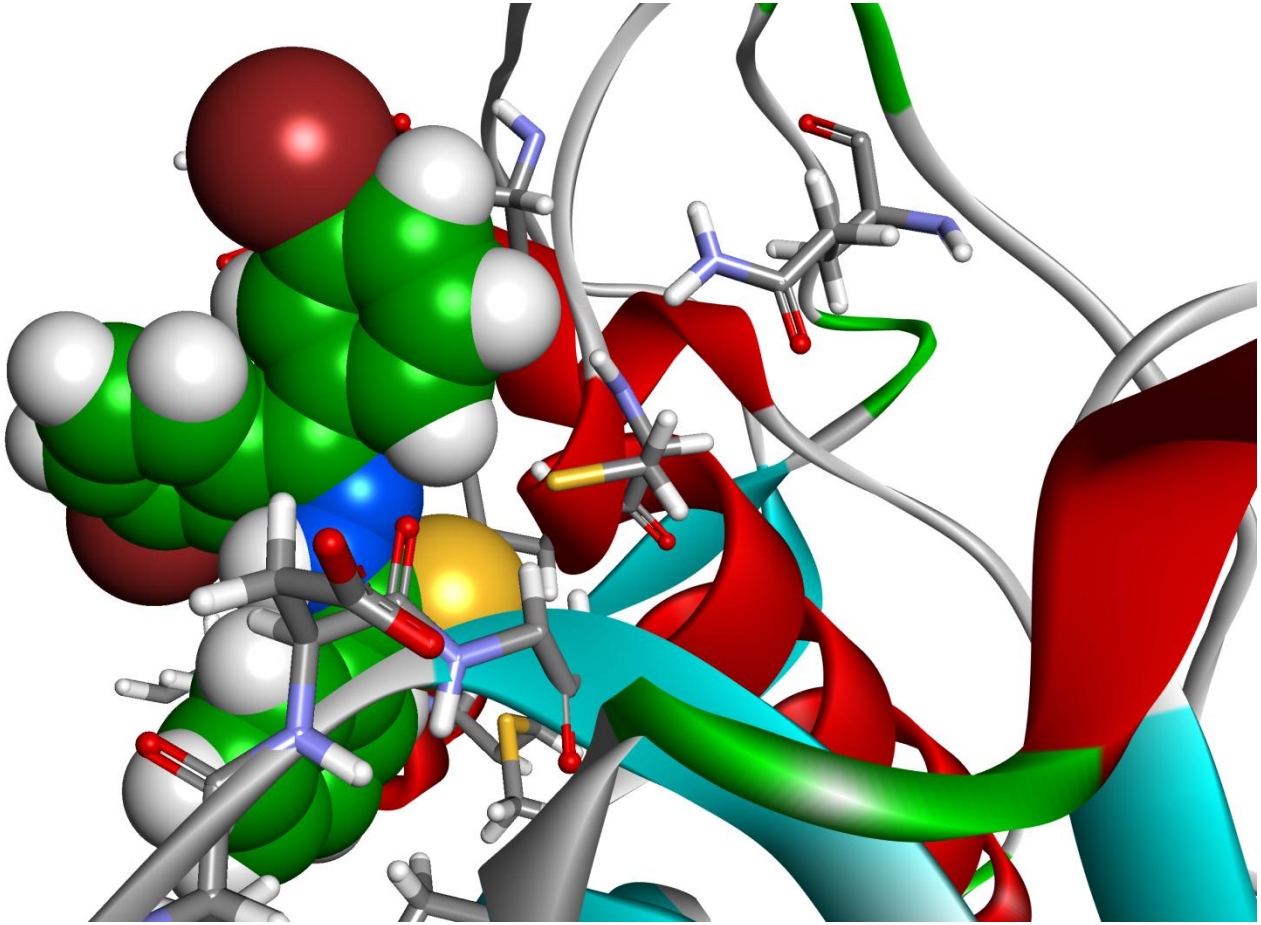


Figure 167. Molecular Docking of **58** with Cruzain. The Thiocarbonyl Carbon and Cys25 (papain numbering) Thiolate Group of Cruzain (C─S) Distance is 4.597Å

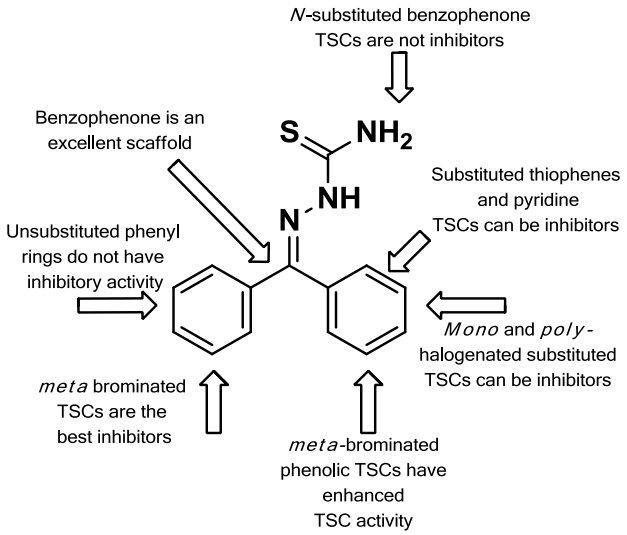



Figure 168. Structure-Activity Relationship for Functionalized Benzophenone TSCs as Cruzain Inhibitors

# CHAPTER FIVE

## Conclusions and Future Directions

Small, nonpeptidic thiosemicarbazones were tested as inhibitors of three related cysteine proteases, human cathepsins L and K, and cruzain, from *Trypanosoma cruzi*.

Seven compounds were found to be excellent inhibitors of human cahtepsin L Three analogs belong to the subfamily of *meta*brominated benzophenone TSC analogs: 3-bromo-2'-fluorobenzophenone thiosemicarbazone (**14;** $IC_{50}$: 24 nM), 3-bromo-3'-dibromobenzophenone thiosemicarbazone (**1;** $IC_{50}$: 17 nM) and 3-bromo-3'-trifluoromethyl benzophenone thiosemicarbazone (**6;** $IC_{50}$: 47 nM) have all halogenated moieties as substituents in this subfamily. The second subgroup consists of three unsubstituted and substituted benzoyl benzophenone thiosemicarbazones: unsubstituted benzoyl benzophenone thiosemicarbazone (**156;** $IC_{50}$: 9.9 nM); 5-(4-fluorobenzoyl)-4'-fluorobenzophenone thiosemicarbazone (**157**; $IC_{50}$: 24.3 nM); and 5-(2-fluorobenzoyl)-3-bromo-2'-fluorobenzophenone thiosemicarbazone (**168**; $IC_{50}$: 8.1 nM). The best two compounds found in the family of TSC are benzoyl-benzophenone TSC with $IC_{50}$ values less than 10 nM. Compound **8,** 3-bromo-3'-hydroxybenzophenone thiosemicarbazone (**14;** $IC_{50}$: 189 nM) was further examined due to the possibility of derivatization and better solubility properties.

Compounds **1** ($IC_{50}$: 17 nM) and **8** ($IC_{50}$: 189 nM), functionalized benzophenone thiosemicarbazones were lead inhibitors of cathepsin L, and were evaluated to characterize their mode of action and assess their performance and efficacy in a series of *in vitro* studies. Kinetic analysis was used to investigate the mechanism and mode of

inhibition of these compounds. Compounds **1** and **8** were determined to be slow, time-dependent, reversible, competitive inhibitor of the fluorogenic substrate, Z-FR-AMC. Additionally the mechanism of **8,** 3-bromo-3'-hydroxybenzophenone thiosemicarbazone, was determined to be a slow, simple reversible inhibition but substrate dependent. However, at lower concentrations of the substrate, the kinetics suggested a more complex two-step mechanism of inhibition. Molecular docking of **8** with cathepsin L suggests that it binds in a manner that positions the inhibitor thiocarbonyl moiety for attack by the enzyme Cys25 thiolate to form, at least, a transient tetrahedral intermediate. The kinetic and modeling studies support this proposed mechanism. Compound **8** was also able to significantly inhibit the activity of cathepsin L toward one of its natural substrates, type I collagen from human skin. A selected group of cathepsin L inhibitors (**1**, **8**, **156**, **157**, **168**) displayed potent activity in cell studies by inhibiting invasion and migration of the breast cancer MDA-MB-231 cell line.

Future studies involving these compounds include a complete kinetic characterization of two benzoylbenzophenone thiosemicarbazones (**156**, and **168)** with $IC_{50}$ values less than 10 nM as cathepsin L inhibitors. The presence of several members of the cathepsin family, including cathepsins K, S, V and F, and cathepsin L-like enzymes in human tumors needs to be considered when testing effective cathepsin L inhibitors in *in vitro* studies. The lack of selective, yet potent inhibitors for each cathepsin and limited *in vitro* cell invasion and cell migration models are the biggest obstacles when confirming effective and selective inhibition of cathepsin L by these novel nonpeptidic thiosemicarbazones. The usage of additional cell lines, like prostate, melanoma and

ovarian cancer cell lines might help to elucidate effective treatment for each one of these diseases.

3-3'-Dibromobenzophenone thiosemicarbazone (**1**) was an excellent inhibitor of human cathepsin K ($IC_{50}$: 35.2 nM). Advanced kinetics studies were used to investigate the mechanism of **1**. This compound was determined to be a slow, reversible, competitive inhibitor of the protease, similar to cathepsin L. Fluorescent Western blotting demonstrated that **1** was able to inhibit the cleavage and activation of human procathepsin K, and important consideration for use as an antimetastatic agent. The expression of cathepsin K as a proenzyme is well known, yet their mechanism of activation (cleavage of the propeptide) is poorly understood under physiological conditions. The acidic microenviroment of tumors and bone resorption cycles promotes the cleavage of overexpressed procathepsin K, which is vital in those processes. Targeting the inhibition of the activation or the expression of procathepsin K might emerge as a possible alternative for the treatment of osteoarthritis and bone metastasis due to the activity of cathepsin K.

Lastly, **1** was able to inhibit the powerful collagenase and proteolytic activity of cathepsin K when using type IV collagen, a natural substrate of this cysteine protease. Tubulin, from calf brain was susceptible to degradation by cathepsin K and compound **1** impeded the proteolytic degradation of this alternative substrate.

Interestingly, the majority of the potent compounds do not exhibit a clear selectivity toward cathepsin L or K, perhaps due to the similarities in the active sites of both cysteine proteases. Three compounds with a high degree of selectivity were found in the library of thiosemicarbazones and could be used as promising agents of cathepsins

L or K. Nevertheless, subtle differences in cathepsins L and K *in vitro* assays might explain the difference in activities toward both proteases. A more in depth screening using alternative substrates (specific substrates) might help to elucidate their activities.

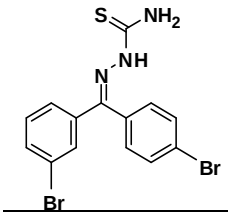
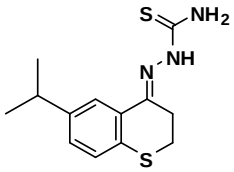
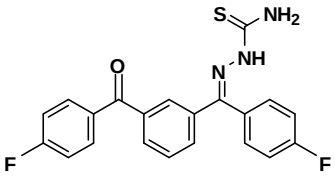
| $IC_{50}$ (nM) |  |  |  |
|---|---|---|---|
| | **49** | **123** | **157** |
| Cathepsin L | ≥ 10000 | ≥ 10000 | 50 |
| Cathepsin K | 49.1 | 21.3 | 8500 |

Figure 169. Chemical Structures of Selective Inhibitors of Human Cathepsins L and K

Four functionalized thiosemicarbazones derivatives (**1**, **9**, **10**, and **17**) were found to be time-dependent inhibitors of recombinant cruzain. Compounds **1, 9,** and **17** were slow reversible, competitive inhibitors of cruzain. Their proposed mode of inhibition is via transient covalent modification due to the interactions between the thiocarbonyl carbon of the thiosemicarbazone moiety and the thiolate group of cysteine 25, the key residue in the activity of these cysteine proteases. This mechanism is supported by advanced kinetic studies using **17** and molecular docking using the aforementioned inhibitors and cruzain. It was also demonstrated that the presence of the thiosemicarbazone moiety in these molecules is vital for their inhibitory potency. Modification of the functional group impedes both covalent and hydrogen-bond interactions between the inhibitor and cruzain's active site. Lastly, the collagenase activity of cruzain was reduced by the presence of **17,** a dihalogenated benzophenone thiosemicarbazone.

These preliminary studies have contributed to understanding the characteristics of thiosemicarbazones as potential anticancer and/or anti-Chagas' disease agents.

In summary, advanced kinetic studies and molecular modeling provide evidence for the proposal that the best thiosemicarbazone inhibitors of each enzyme form at least a transient covalent bond with the active site Cys25 (Figure 170).

The possibility that thiocarbamoylated cruzain, cathepsin L or cathepsin K is formed followed by slow hydrolysis cannot be ruled out, and future studies using mass spectrometry techniques could be addressed to elucidate this mechanism.

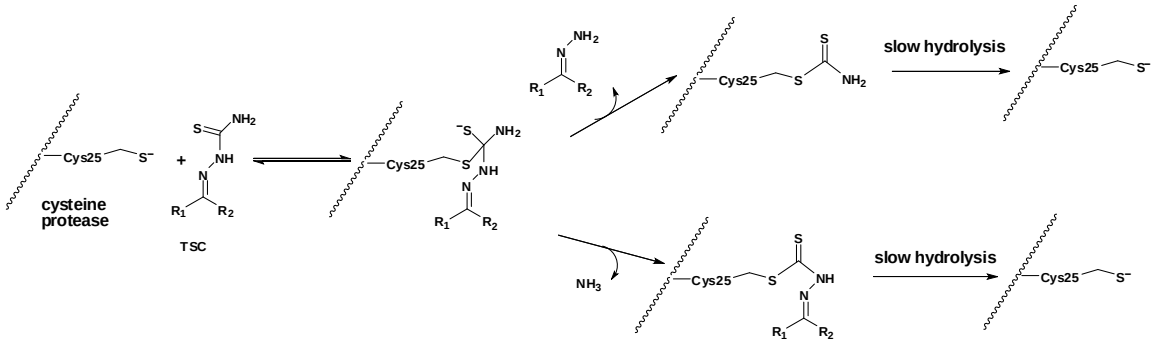
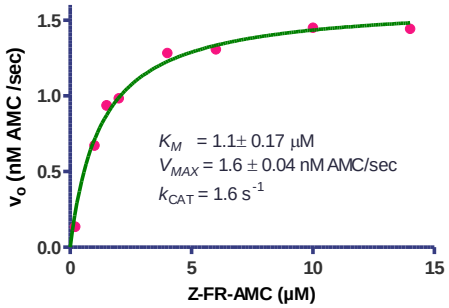



Figure 170. Proposed mechanism of Cysteine Protease Inhibition by Thiosemicarbazone Derivatives

APPENDICES

APPENDIX A

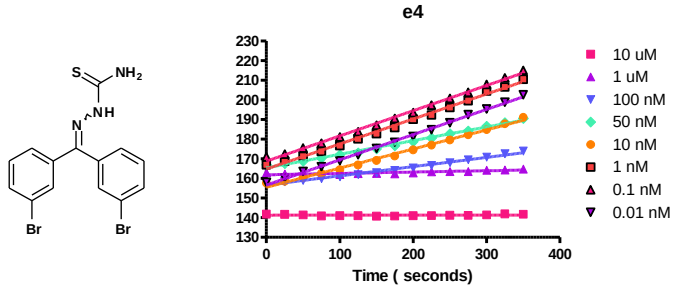
*In Vitro* Evaluations of Thiosemicarbazones as Inhibitors of Human Cathepsin L

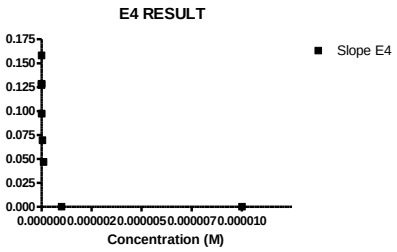


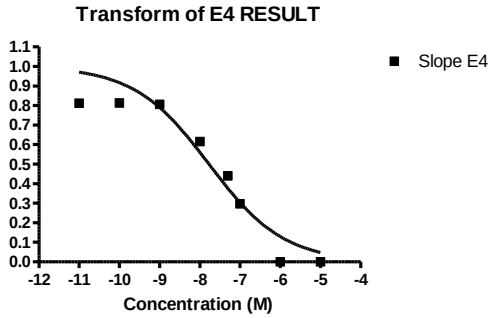

| Substrate Concentration (uM) | 2 A |
|---|---|
| 14.00000 | 1.443 |
| 10.00000 | 1.450 |
| 6.00000 | 1.307 |
| 4.00000 | 1.283 |
| 2.00000 | 0.983 |
| 1.50000 | 0.937 |
| 1.00000 | 0.671 |
| 0.20000 | 0.135 |

| | 2 A |
|---|---|
| Michaelis-Menten | |
| Best-fit values | |
| Vmax | 1.618 |
| Km | 1.278 |
| Std. Error | |
| Vmax | 0.05065 |
| Km | 0.1511 |
| 95% Confidence Intervals | |
| Vmax | 1.494 to 1.742 |
| Km | 0.9086 to 1.648 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9871 |
| Absolute Sum of Squares | 0.01838 |
| Sy.x | 0.05535 |
| Constraints | |
| Km | Km > 0.0 |
| Number of points | |
| Analyzed | 8 |

A.1 Calculation of Michaelis-Menten Constant for Cathepsin L using Z-FR-AMC as a Substrate



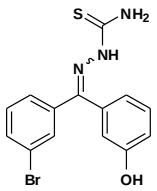

| Concentration (M) | Slope E4 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.046930 |
| 5.000000e-008 | 0.069420 |
| 1.000000e-008 | 0.097350 |
| 1.000000e-009 | 0.127300 |
| 1.000000e-010 | 0.128500 |
| 1.000000e-011 | 0.128400 |

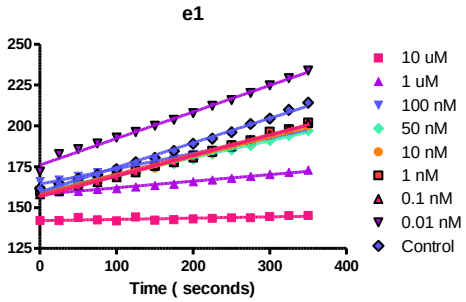
| | Concentration (M) | Slope E4 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.297 |
| | -7.301 | 0.439 |
| | -8.000 | 0.616 |
| | -9.000 | 0.805 |
| | -10.000 | 0.813 |
| | -11.000 | 0.812 |

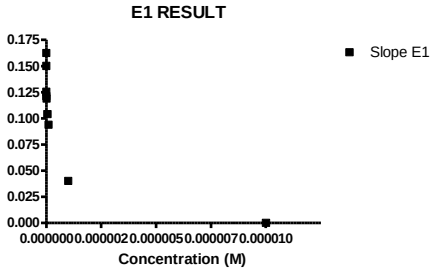
| | Slope E4 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.771 |
| HillSlope | -0.4669 |
| EC50 | 1.692e-008 |
| Std. Error | |
| LogEC50 | 0.2083 |
| HillSlope | 0.1056 |
| 95% Confidence Intervals | |
| LogEC50 | -8.281 to -7.262 |
| HillSlope | -0.7253 to -0.2086 |
| EC50 | 5.234e-009 to 5.473e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9263 |
| Absolute Sum of Squares | 0.06201 |
| Sy.x | 0.1017 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

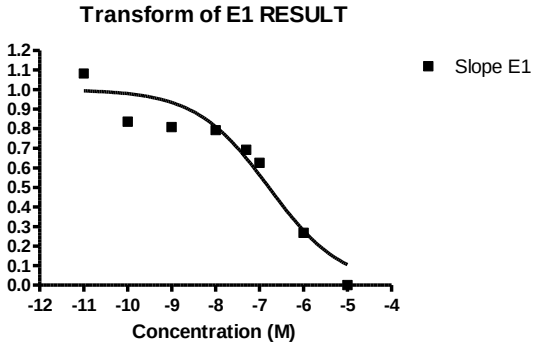
A.2 Representative $IC_{50}$ Calculation Using **1** as a Cathepsin L Inhibitor

# GDK-II-98



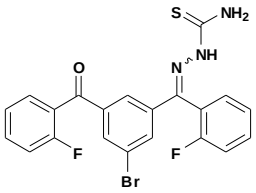

| Concentration (M) | Slope E1 |
|---|---|
| 1.00e-005 | 0.000000 |
| 1.00e-006 | 0.040220 |
| 1.00e-007 | 0.093860 |
| 5.00e-008 | 0.104000 |
| 1.00e-008 | 0.118900 |
| 1.00e-009 | 0.121300 |
| 1.00e-010 | 0.125400 |
| 1.00e-011 | 0.162500 |

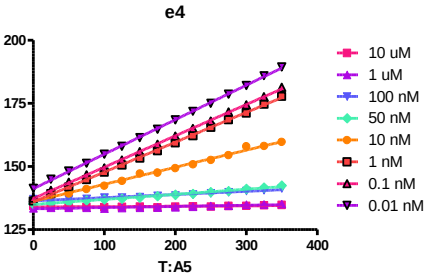
| | Concentration (M) | Slope E1 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.268 |
| | -7.000 | 0.625 |
| | -7.301 | 0.693 |
| | -8.000 | 0.792 |
| | -9.000 | 0.808 |
| | -10.000 | 0.835 |
| | -11.000 | 1.083 |

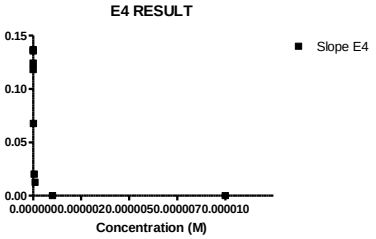
| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -6.788 |
| HillSlope | -0.5218 |
| EC50 | 1.630e-007 |
| Std. Error | |
| LogEC50 | 0.1977 |
| HillSlope | 0.1258 |
| 95% Confidence Intervals | |
| LogEC50 | -7.272 to -6.304 |
| HillSlope | -0.8297 to -0.2139 |
| EC50 | 5.350e-008 to 4.966e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9262 |
| Absolute Sum of Squares | 0.06172 |
| Sy.x | 0.1014 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

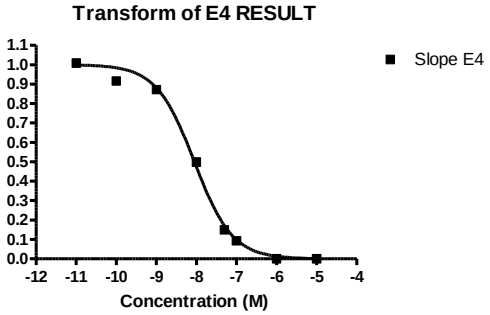
A.3 Representative $IC_{50}$ Calculation using **8** as a Cathepsin L Inhibitor

# ENP-I-151



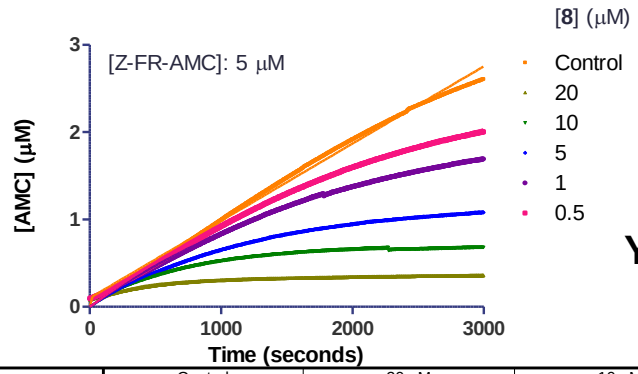



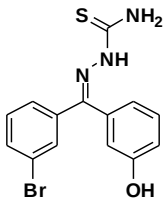

| Concentration (M) | Slope E4 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.012590 |
| 5.000000e-008 | 0.020180 |
| 1.000000e-008 | 0.067530 |
| 1.000000e-009 | 0.118100 |
| 1.000000e-010 | 0.124100 |
| 1.000000e-011 | 0.136700 |

| | Concentration (M) | Slope E4 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.093 |
| | -7.301 | 0.149 |
| | -8.000 | 0.498 |
| | -9.000 | 0.871 |
| | -10.000 | 0.915 |
| | -11.000 | 1.008 |

| | Slope E4 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -8.050 |
| HillSlope | -0.9136 |
| EC50 | 8.915e-009 |
| Std. Error | |
| LogEC50 | 0.04992 |
| HillSlope | 0.08138 |
| 95% Confidence Intervals | |
| LogEC50 | -8.172 to -7.928 |
| HillSlope | -1.113 to -0.7144 |
| EC50 | 6.729e-009 to 1.181e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9953 |
| Absolute Sum of Squares | 0.006209 |
| Sy.x | 0.03217 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

A.4 Representative $IC_{50}$ Calculation Using **168** as a Cathepsin L Inhibitor

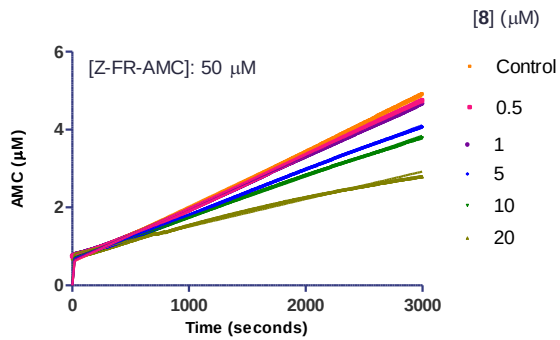


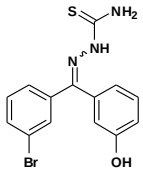




**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | Control | 20 uM | 10 uM | 5 uM | 1 uM | 0.5 uM |
|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | | | Hit constraint | |
| Best-fit values | | | | | | |
| vs | 0.0008844 | 2.732e-005 | 8.012e-006 | 4.627e-005 | ~ 1.888e-016 | 2.040e-015 |
| vi | ~ 0.09269 | 0.001111 | 0.001048 | 0.0009878 | 0.001053 | 0.001114 |
| kobs | ~ 0.9178 | 0.003824 | 0.001566 | 0.0009347 | 0.0004653 | 0.0003643 |
| Std. Error | | | | | | |
| vs | 1.789e-006 | 7.660e-007 | 2.648e-006 | 5.860e-006 | | 2.403e-006 |
| vi | ~ 0.2554 | 1.200e-005 | 6.525e-006 | 4.565e-006 | 3.956e-006 | 2.000e-006 |
| kobs | ~ 2.555 | 5.560e-005 | 2.185e-005 | 1.764e-005 | 1.864e-005 | 1.980e-006 |
| 95% Confidence Intervals | | | | | | |
| vs | 0.0008809 to 0.0008879 | 2.582e-005 to 2.882e-005 | 2.821e-006 to 1.320e-005 | 3.479e-005 to 5.776e-005 | | 0.0 to 4.710e-006 |
| vi | (Very wide) | 0.001088 to 0.001135 | 0.001035 to 0.001061 | 0.0009789 to 0.0009968 | 0.001046 to 0.001061 | 0.001110 to 0.001118 |
| kobs | (Very wide) | 0.003715 to 0.003933 | 0.001523 to 0.001609 | 0.0009001 to 0.0009693 | 0.0004288 to 0.0005019 | 0.0003605 to 0.0003682 |
| Goodness of Fit | | | | | | |
| Degrees of Freedom | 997 | 997 | 997 | 997 | 997 | 997 |
| R square | 0.9960 | 0.9682 | 0.9909 | 0.9973 | 0.9988 | 0.9989 |
| Absolute Sum of Squares | 2.378 | 0.1373 | 0.2438 | 0.2249 | 0.2790 | 0.3652 |
| Sy.x | 0.04883 | 0.01173 | 0.01564 | 0.01502 | 0.01673 | 0.01914 |
| Constraints | | | | | | |
| vs | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | |
| Analyzed | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

A.5 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 5 μM (Preincubation Time: 0 minutes)

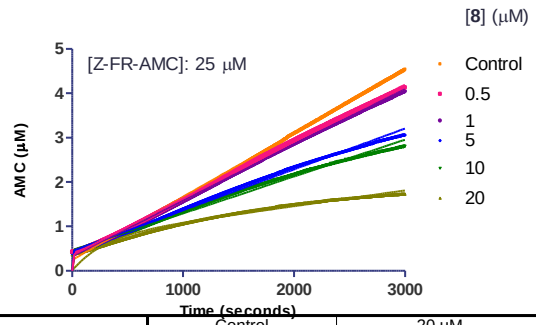


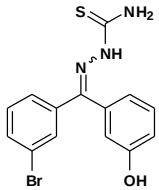




**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | Control | 20 uM | 10 uM | 5 uM | 1 uM | 0.5 uM |
|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | Ambiguous | Ambiguous | Ambiguous | Ambiguous |
| Best-fit values | | | | | | |
| vs | 0.001423 | 0.0007067 | 0.001048 | 0.001150 | 0.001340 | 0.001376 |
| vi | ~ 5.266 | 0.4482 | ~ 12.38 | ~ 6.261 | ~ 5.213 | ~ 5.182 |
| kobs | ~ 8.798 | 0.5569 | ~ 17.64 | ~ 9.394 | ~ 8.232 | ~ 8.597 |
| Std. Error | | | | | | |
| vs | 2.150e-006 | 2.212e-006 | 1.169e-006 | 1.149e-006 | 1.497e-006 | 1.668e-006 |
| vi | ~ 3.733e+008 | 0.09774 | ~ 2.060e+009 | ~ 4.754e+009 | ~ 2.460e+008 | ~ 3.505e+008 |
| kobs | ~ 6.239e+008 | 0.1218 | ~ 2.936e+009 | ~ 7.133e+009 | ~ 3.886e+008 | ~ 5.817e+008 |
| 95% Confidence Intervals | | | | | | |
| vs | 0.001419 to 0.001428 | 0.0007024 to 0.0007110 | 0.001046 to 0.001050 | 0.001148 to 0.001152 | 0.001337 to 0.001343 | 0.001373 to 0.001380 |
| vi | (Very wide) | 0.2566 to 0.6398 | (Very wide) | (Very wide) | (Very wide) | (Very wide) |
| kobs | (Very wide) | 0.3181 to 0.7956 | (Very wide) | (Very wide) | (Very wide) | (Very wide) |
| Goodness of Fit | | | | | | |
| Degrees of Freedom | 997 | 997 | 997 | 997 | 997 | 997 |
| R square | 0.9981 | 0.9904 | 0.9988 | 0.9990 | 0.9988 | 0.9987 |
| Absolute Sum of Squares | 2.871 | 3.628 | 1.018 | 0.9840 | 1.672 | 1.837 |
| Sy.x | 0.05366 | 0.06032 | 0.03196 | 0.03142 | 0.04095 | 0.04293 |
| Constraints | | | | | | |
| vs | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | |
| Analyzed | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

A.6 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 50 μM (Preincubation Time: 0 minutes)

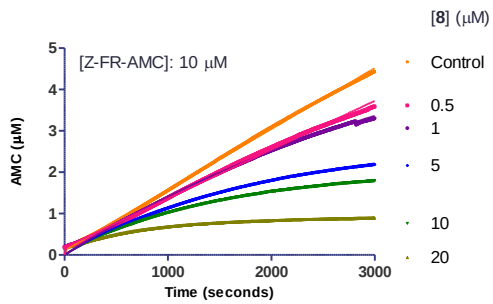


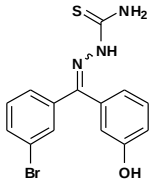




**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | Control | 20 uM | 10 uM | 5 uM | 1 uM | 0.5 uM |
|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | | Ambiguous | Ambiguous | Ambiguous |
| Best-fit values | | | | | | |
| vs | 0.001423 | 0.0003545 | 0.0008308 | 0.0009205 | 0.001256 | 0.001283 |
| vi | ~ 1.833 | 0.002994 | 0.2028 | ~ 0.2786 | ~ 2.064 | ~ 2.577 |
| kobs | ~ 7.446 | 0.003522 | 0.4363 | ~ 0.6231 | ~ 6.293 | ~ 7.384 |
| Std. Error | | | | | | |
| vs | 1.634e-006 | 4.516e-006 | 2.284e-006 | 2.013e-006 | 1.009e-006 | 1.088e-006 |
| vi | ~ 6.361e+007 | 6.051e-005 | 0.06619 | ~ 0.1129 | ~ 1.372e+006 | ~ 3.658e+007 |
| kobs | ~ 2.586e+008 | 0.0001091 | 0.1433 | ~ 0.2538 | ~ 4.187e+006 | ~ 1.049e+008 |
| 95% Confidence Intervals | | | | | | |
| vs | 0.001419 to 0.001426 | 0.0003456 to 0.0003633 | 0.0008263 to 0.0008352 | 0.0009166 to 0.0009244 | 0.001254 to 0.001258 | 0.001281 to 0.001285 |
| vi | (Very wide) | 0.002876 to 0.003113 | 0.07309 to 0.3326 | (Very wide) | (Very wide) | (Very wide) |
| kobs | (Very wide) | 0.003309 to 0.003736 | 0.1556 to 0.7171 | (Very wide) | (Very wide) | (Very wide) |
| Goodness of Fit | | | | | | |
| Degrees of Freedom | 997 | 997 | 997 | 997 | 997 | 997 |
| R square | 0.9988 | 0.9734 | 0.9926 | 0.9953 | 0.9994 | 0.9993 |
| Absolute Sum of Squares | 1.845 | 4.255 | 3.861 | 3.006 | 0.7572 | 0.8543 |
| Sy.x | 0.04302 | 0.06533 | 0.06223 | 0.05491 | 0.02756 | 0.02927 |
| Constraints | | | | | | |
| vs | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | |
| Analyzed | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

A.7 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 25 μM (Preincubation Time: 0 minutes)

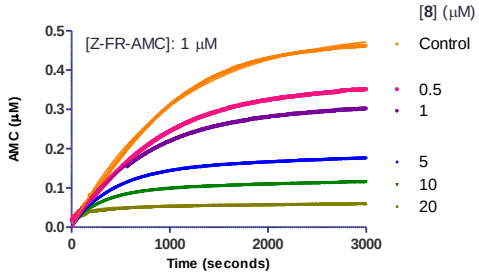


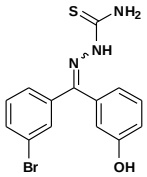




**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | Control | 20 uM | 10 uM | 5 uM | 1 uM | 0.5 uM |
|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | | | | Ambiguous |
| Best-fit values | | | | | | |
| vs | 0.001469 | 8.407e-005 | 0.0001831 | 0.0002328 | 2.674e-014 | 0.001182 |
| vi | ~ 0.8307 | 0.001703 | 0.001567 | 0.001577 | 0.001591 | ~ 0.3222 |
| kobs | ~ 7.910 | 0.002457 | 0.001048 | 0.0008123 | 0.0002494 | ~ 1.800 |
| Std. Error | | | | | | |
| vs | 1.337e-006 | 2.671e-006 | 1.011e-005 | 1.706e-005 | 1.705e-005 | 2.006e-006 |
| vi | ~ 1.068e+008 | 1.739e-005 | 1.018e-005 | 9.726e-006 | 3.808e-006 | ~ 3.825 |
| kobs | ~ 1.019e+009 | 4.219e-005 | 2.609e-005 | 2.747e-005 | 4.326e-006 | ~ 21.46 |
| 95% Confidence Intervals | | | | | | |
| vs | 0.001466 to 0.001472 | 7.884e-005 to 8.931e-005 | 0.0001633 to 0.0002030 | 0.0001994 to 0.0002663 | 0.0 to 3.343e-005 | 0.001178 to 0.001186 |
| vi | (Very wide) | 0.001669 to 0.001737 | 0.001547 to 0.001587 | 0.001558 to 0.001596 | 0.001583 to 0.001598 | (Very wide) |
| kobs | (Very wide) | 0.002375 to 0.002540 | 0.0009972 to 0.001100 | 0.0007585 to 0.0008662 | 0.0002409 to 0.0002579 | (Very wide) |
| Goodness of Fit | | | | | | |
| Degrees of Freedom | 997 | 997 | 997 | 997 | 997 | 997 |
| R square | 0.9993 | 0.9793 | 0.9956 | 0.9967 | 0.9982 | 0.9972 |
| Absolute Sum of Squares | 1.063 | 0.7811 | 0.9939 | 1.161 | 1.663 | 2.991 |
| Sy.x | 0.03265 | 0.02799 | 0.03157 | 0.03412 | 0.04084 | 0.05477 |
| Constraints | | | | | | |
| vs | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | |
| Analyzed | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

A.8 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 10 μM (Preincubation Time: 0 minutes)

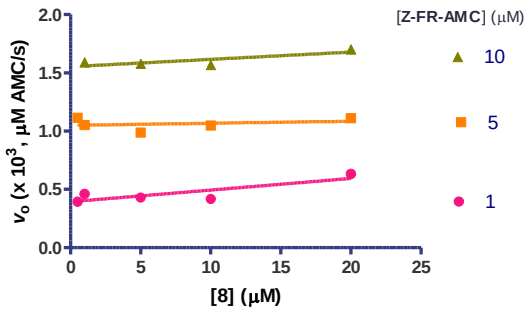



**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | Control | 20 uM | 10 uM | 5 uM | 1 uM | 0.5 uM |
|---|---|---|---|---|---|---|
| Progress Curve | Hit constraint | | | | | Hit constraint |
| Best-fit values | | | | | | |
| vs | ~ 2.162e-016 | 4.460e-006 | 7.883e-006 | 1.078e-005 | 1.211e-005 | ~ 1.304e-016 |
| vi | 0.0004800 | 0.0004655 | 0.0003572 | 0.0003810 | 0.0004115 | 0.0004014 |
| kobs | 0.0009615 | 0.009451 | 0.003719 | 0.002538 | 0.001476 | 0.001099 |
| Std. Error | | | | | | |
| vs | | 9.250e-008 | 1.565e-007 | 2.425e-007 | 5.294e-007 | |
| vi | 1.409e-006 | 7.037e-006 | 2.326e-006 | 1.689e-006 | 1.144e-006 | 7.453e-007 |
| kobs | 7.352e-006 | 0.0001589 | 3.283e-005 | 1.817e-005 | 9.933e-006 | 5.596e-006 |
| 95% Confidence Intervals | | | | | | |
| vs | | 4.279e-006 to 4.642e-006 | 7.576e-006 to 8.190e-006 | 1.030e-005 to 1.125e-005 | 1.107e-005 to 1.315e-005 | |
| vi | 0.0004772 to 0.0004828 | 0.0004517 to 0.0004793 | 0.0003527 to 0.0003618 | 0.0003776 to 0.0003843 | 0.0004092 to 0.0004137 | 0.0003999 to 0.0004028 |
| kobs | 0.0009471 to 0.0009759 | 0.009139 to 0.009762 | 0.003654 to 0.003783 | 0.002503 to 0.002574 | 0.001457 to 0.001496 | 0.001088 to 0.001110 |
| Goodness of Fit | | | | | | |
| Degrees of Freedom | 997 | 997 | 997 | 997 | 997 | 997 |
| R square | 0.9970 | 0.9333 | 0.9885 | 0.9953 | 0.9986 | 0.9991 |
| Absolute Sum of Squares | 0.05083 | 0.004174 | 0.005518 | 0.006891 | 0.008179 | 0.007771 |
| Sy.x | 0.007140 | 0.002046 | 0.002353 | 0.002629 | 0.002864 | 0.002792 |
| Constraints | | | | | | |
| vs | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | |
| Analyzed | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

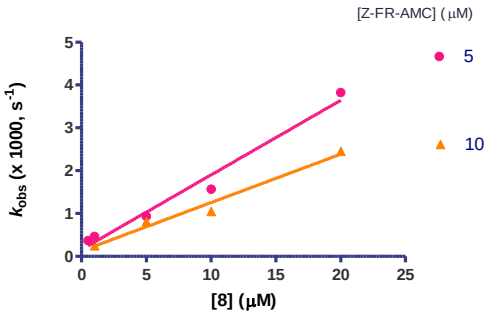
A.9 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 1 μM (Preincubation Time: 0 minutes)




| | [I] | 1 uM [S] | 5 uM [S] | 10 uM [I] |
|---|---|---|---|---|
| | 20.000 | 0.633 | 1.111 | 1.703 |
| | 10.000 | 0.418 | 1.048 | 1.567 |
| | 5.000 | 0.430 | 0.988 | 1.577 |
| | 1.000 | 0.461 | 1.053 | 1.591 |
| | 0.500 | 0.394 | 1.114 | |

| | 1 uM [S] | 5 uM [S] | 10 uM [I] |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.009989 ± 0.003717 | 0.001762 ± 0.003591 | 0.006257 ± 0.003161 |
| Y-intercept when X=0.0 | 0.3943 ± 0.03813 | 1.050 ± 0.03684 | 1.553 ± 0.03625 |
| X-intercept when Y=0.0 | -39.47 | -595.8 | -248.2 |
| 1/slope | 100.1 | 567.5 | 159.8 |
| 95% Confidence Intervals | | | |
| Slope | -0.001837 to 0.02182 | -0.009663 to 0.01319 | -0.007346 to 0.01986 |
| Y-intercept when X=0.0 | 0.2729 to 0.5156 | 0.9327 to 1.167 | 1.397 to 1.709 |
| X-intercept when Y=0.0 | -infinity to -13.46 | -infinity to -73.01 | -infinity to -71.83 |
| Goodness of Fit | | | |
| R square | 0.7066 | 0.07432 | 0.6620 |
| Sy.x | 0.05991 | 0.05787 | 0.04493 |
| Is slope significantly non-zero? | | | |
| F | 7.223 | 0.2409 | 3.918 |
| DFn, DFd | 1.000, 3.000 | 1.000, 3.000 | 1.000, 2.000 |
| P value | 0.0746 | 0.6572 | 0.1863 |
| Deviation from zero? | Not Significant | Not Significant | Not Significant |
| Data | | | |
| Number of X values | 5 | 5 | 4 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 5 | 5 | 4 |
| Number of missing values | 0 | 0 | 1 |

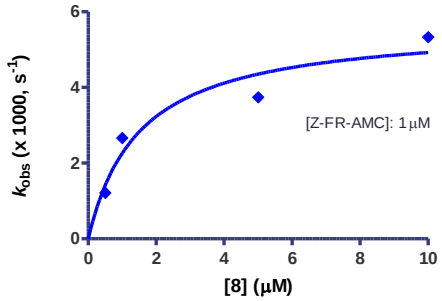
A.10 Plot of $v_o$, vs [**8**] Using Linear Regression. (Preincubation Time: 0 minutes)




| [I] | 5 uM [S] | 10 uM [S] |
|---|---|---|
| 20.000 | 3.824 | 2.457 |
| 10.000 | 1.566 | 1.048 |
| 5.000 | 0.935 | 0.812 |
| 1.000 | 0.465 | 0.249 |
| 0.500 | 0.364 | |

| | 5 uM [S] | 10 uM [S] |
|---|---|---|
| Best-fit values | | |
| Slope | 0.1741 ± 0.01543 | 0.1130 ± 0.01250 |
| Y-intercept when X=0.0 | 0.1598 ± 0.1583 | 0.1245 ± 0.1433 |
| X-intercept when Y=0.0 | -0.9180 | -1.101 |
| 1/slope | 5.743 | 8.848 |
| 95% Confidence Intervals | | |
| Slope | 0.1250 to 0.2232 | 0.05924 to 0.1668 |
| Y-intercept when X=0.0 | -0.3438 to 0.6634 | -0.4923 to 0.7412 |
| X-intercept when Y=0.0 | -4.917 to 1.662 | -11.37 to 3.249 |
| Goodness of Fit | | |
| R square | 0.9770 | 0.9761 |
| Sy.x | 0.2487 | 0.1777 |
| Is slope significantly non-zero? | | |
| F | 127.4 | 81.76 |
| DFn, DFd | 1.000, 3.000 | 1.000, 2.000 |
| P value | 0.0015 | 0.0120 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 5 | 4 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 5 | 4 |
| Number of missing values | 0 | 1 |

A.11 Plot of $k_{obs}$ vs [**8**] Using Linear Regression. ([S]: 5 and 10 μM) (Preincubation Time: 0 minutes)

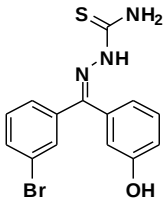



| [I] | 1 uM [S] |
|---|---|
| 20.000 | |
| 10.000 | 5.331 |
| 5.000 | 3.741 |
| 1.000 | 2.667 |
| 0.500 | 1.218 |

| | 1 uM [S] |
|---|---|
| Michaelis-Menten | |
| Best-fit values | |
| Vmax | 5.668 |
| Km | 1.511 |
| Std. Error | |
| Vmax | 0.8430 |
| Km | 0.7593 |
| 95% Confidence Intervals | |
| Vmax | 2.040 to 9.295 |
| Km | 0.0 to 4.778 |
| Goodness of Fit | |
| Degrees of Freedom | 2 |
| R square | 0.9177 |
| Absolute Sum of Squares | 0.7442 |
| Sy.x | 0.6100 |
| Constraints | |
| Km | Km > 0.0 |
| Number of points | |
| Analyzed | 4 |

A.12 Plot of $k_{obs}$ vs [**8**] Using Non-linear Regression. ([S]: 1 µM) (Preincubation Time: 0 minutes)

# RESULTS




**GDK-II-98 4a**
**KGP94**

| % Inhibition |
|---|
| 7.900000 |
| 8.400000 |
| 3.200000 |

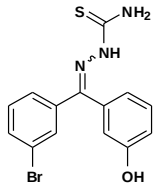
| | % Inhibition |
|---|---|
| Number of values | 3 |
| Mean | 6.500 |
| Std. Deviation | 2.869 |
| Std. Error | 1.656 |
| Sum | 19.50 |

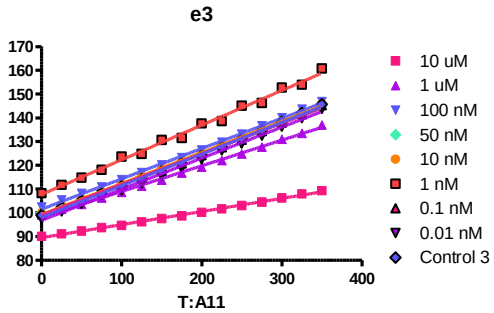
| | EDTA | DTT | BRIJ | NaOAc | DMSO | CAT L | Z-FR-AMC |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | µM |
| | 1 | 3 | 0.01 | 100 | 2 | 1 | 50 |

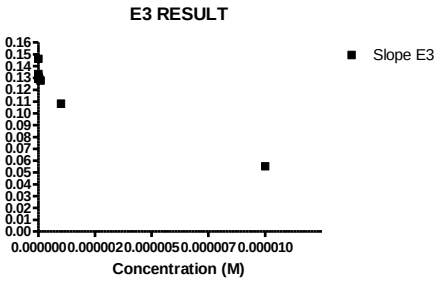
**Pre-incubation time: 0 min**

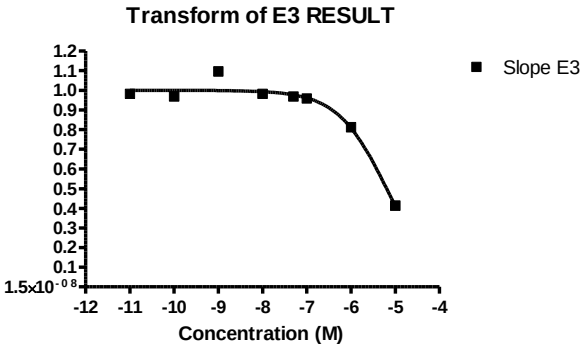
A.13 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 0 minutes)

# GDK-II-98 4a



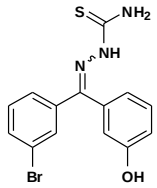

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.055100 |
| 0.000001 | 0.108100 |
| 1.000000e-007 | 0.127600 |
| 5.000000e-008 | 0.128900 |
| 1.000000e-008 | 0.130800 |
| 1.000000e-009 | 0.146000 |
| 1.000000e-010 | 0.129100 |
| 1.000000e-011 | 0.130800 |
| 0.000000 | 0.133100 |

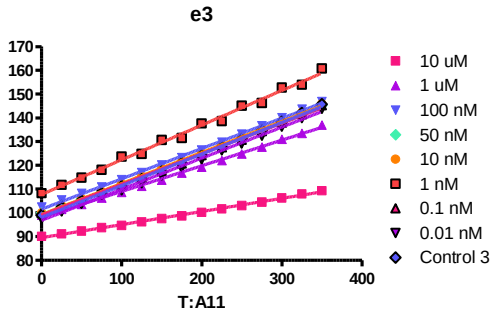
| Concentration (M) | Slope E3 |
|---|---|
| -5.000 | 0.414 |
| -6.000 | 0.812 |
| -7.000 | 0.959 |
| -7.301 | 0.968 |
| -8.000 | 0.983 |
| -9.000 | 1.097 |
| -10.000 | 0.970 |
| -11.000 | 0.983 |
| | 1.000 |

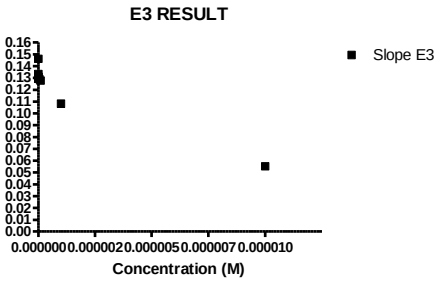
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -5.192 |
| HillSlope | -0.7760 |
| EC50 | 6.420e-006 |
| Std. Error | |
| LogEC50 | 0.08515 |
| HillSlope | 0.1225 |
| 95% Confidence Intervals | |
| LogEC50 | -5.401 to -4.984 |
| HillSlope | -1.076 to -0.4764 |
| EC50 | 3.974e-006 to 1.037e-005 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9644 |
| Absolute Sum of Squares | 0.01102 |
| Sy.x | 0.04285 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

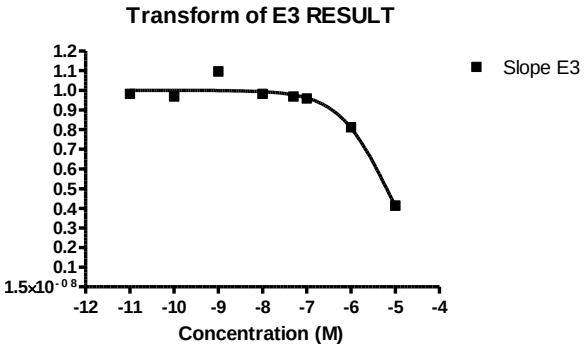
A.14 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 1 minute)

# GDK-II-98 4a



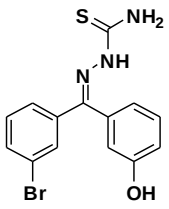

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.055100 |
| 0.000001 | 0.108100 |
| 1.000000e-007 | 0.127600 |
| 5.000000e-008 | 0.128900 |
| 1.000000e-008 | 0.130800 |
| 1.000000e-009 | 0.146000 |
| 1.000000e-010 | 0.129100 |
| 1.000000e-011 | 0.130800 |
| 0.000000 | 0.133100 |

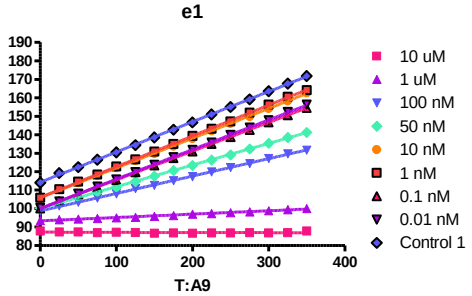
| Concentration (M) | Slope E3 |
|---|---|
| -5.000 | 0.414 |
| -6.000 | 0.812 |
| -7.000 | 0.959 |
| -7.301 | 0.968 |
| -8.000 | 0.983 |
| -9.000 | 1.097 |
| -10.000 | 0.970 |
| -11.000 | 0.983 |
| | 1.000 |

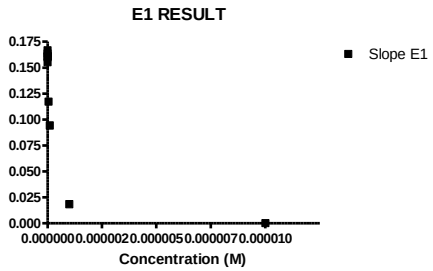
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -5.192 |
| HillSlope | -0.7760 |
| EC50 | 6.420e-006 |
| Std. Error | |
| LogEC50 | 0.08515 |
| HillSlope | 0.1225 |
| 95% Confidence Intervals | |
| LogEC50 | -5.401 to -4.984 |
| HillSlope | -1.076 to -0.4764 |
| EC50 | 3.974e-006 to 1.037e-005 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9644 |
| Absolute Sum of Squares | 0.01102 |
| Sy.x | 0.04285 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

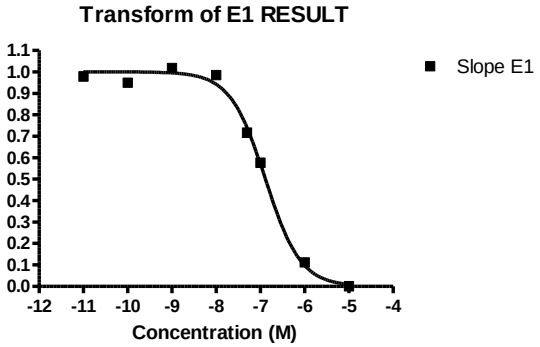
A.15 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 1 minute)

# GDK-II-98 4a



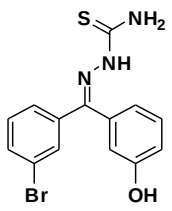

| Concentration (M) | Slope E1 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.018250 |
| 1.000000e-007 | 0.094290 |
| 5.000000e-008 | 0.117200 |
| 1.000000e-008 | 0.161100 |
| 1.000000e-009 | 0.166600 |
| 1.000000e-010 | 0.155300 |
| 1.000000e-011 | 0.160100 |
| 0.000000 | 0.163600 |

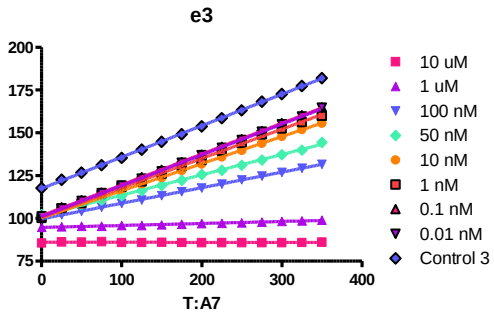
| Concentration (M) | Slope E1 |
|---|---|
| -5.000 | 0.000 |
| -6.000 | 0.112 |
| -7.000 | 0.576 |
| -7.301 | 0.716 |
| -8.000 | 0.985 |
| -9.000 | 1.018 |
| -10.000 | 0.949 |
| -11.000 | 0.979 |
| | 1.000 |

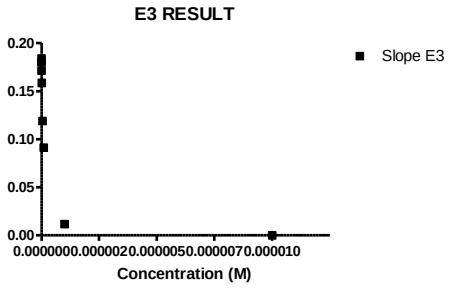
| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -6.884 |
| HillSlope | -1.091 |
| EC50 | 1.305e-007 |
| Std. Error | |
| LogEC50 | 0.04297 |
| HillSlope | 0.1169 |
| 95% Confidence Intervals | |
| LogEC50 | -6.990 to -6.779 |
| HillSlope | -1.377 to -0.8046 |
| EC50 | 1.024e-007 to 1.662e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9947 |
| Absolute Sum of Squares | 0.006132 |
| Sy.x | 0.03197 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

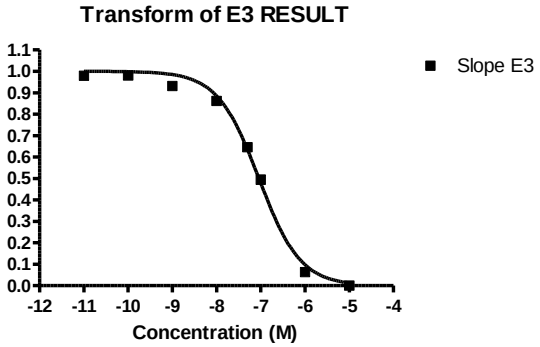
A.16 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 30 minutes)

# GDK-II-98 4a



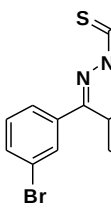

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.011730 |
| 1.000000e-007 | 0.091020 |
| 5.000000e-008 | 0.118900 |
| 1.000000e-008 | 0.158600 |
| 1.000000e-009 | 0.171500 |
| 1.000000e-010 | 0.180400 |
| 1.000000e-011 | 0.180200 |
| 0.000000 | 0.184000 |

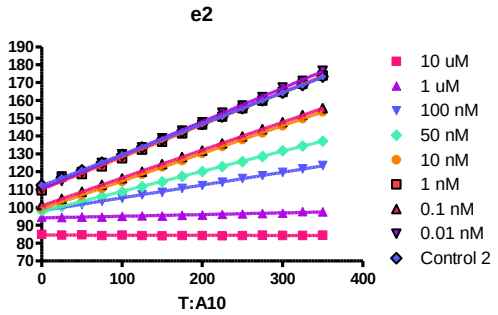
| Concentration (M) | Slope E3 |
|---|---|
| -5.000 | 0.000 |
| -6.000 | 0.064 |
| -7.000 | 0.495 |
| -7.301 | 0.646 |
| -8.000 | 0.862 |
| -9.000 | 0.932 |
| -10.000 | 0.980 |
| -11.000 | 0.979 |
| | 1.000 |

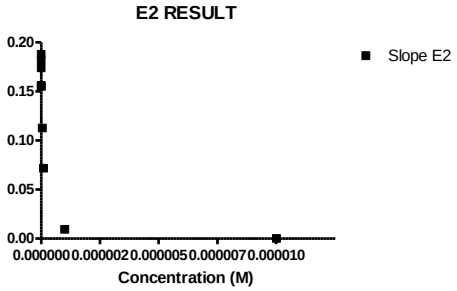
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.040 |
| HillSlope | -0.9265 |
| EC50 | 9.120e-008 |
| Std. Error | |
| LogEC50 | 0.04095 |
| HillSlope | 0.09133 |
| 95% Confidence Intervals | |
| LogEC50 | -7.140 to -6.940 |
| HillSlope | -1.150 to -0.7030 |
| EC50 | 7.240e-008 to 1.149e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9948 |
| Absolute Sum of Squares | 0.005802 |
| Sy.x | 0.03110 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

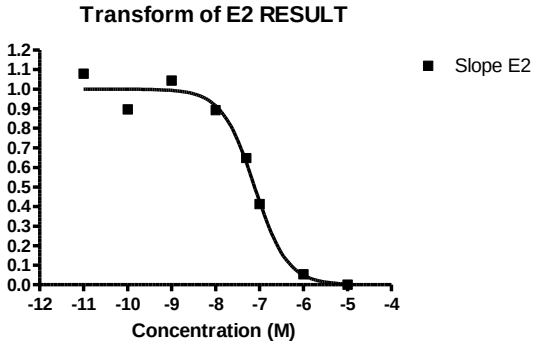
A.17 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 60 minutes)

# GDK-II-98 4a



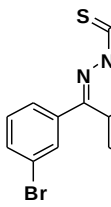

| Concentration (M) | Slope E2 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.009365 |
| 1.000000e-007 | 0.071770 |
| 5.000000e-008 | 0.112700 |
| 1.000000e-008 | 0.155400 |
| 1.000000e-009 | 0.181600 |
| 1.000000e-010 | 0.156100 |
| 1.000000e-011 | 0.188000 |
| 0.000000 | 0.174100 |

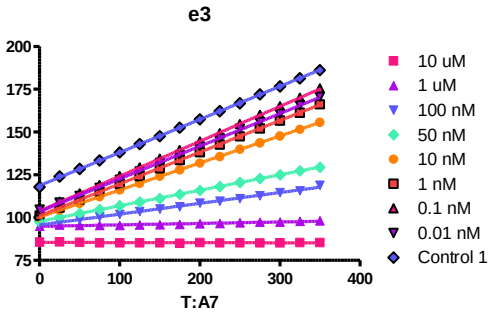
| | Slope E2 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.109 |
| HillSlope | -1.164 |
| EC50 | 7.779e-008 |
| Std. Error | |
| LogEC50 | 0.06374 |
| HillSlope | 0.2564 |
| 95% Confidence Intervals | |
| LogEC50 | -7.265 to -6.953 |
| HillSlope | -1.791 to -0.5363 |
| EC50 | 5.432e-008 to 1.114e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9840 |
| Absolute Sum of Squares | 0.02069 |
| Sy.x | 0.05872 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

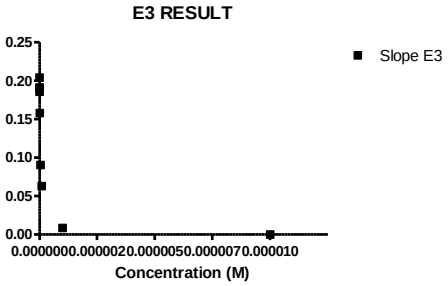
| Concentration (M) | Slope E2 |
|---|---|
| -5.000 | 0.000 |
| -6.000 | 0.054 |
| -7.000 | 0.412 |
| -7.301 | 0.647 |
| -8.000 | 0.893 |
| -9.000 | 1.043 |
| -10.000 | 0.897 |
| -11.000 | 1.080 |
| | 1.000 |

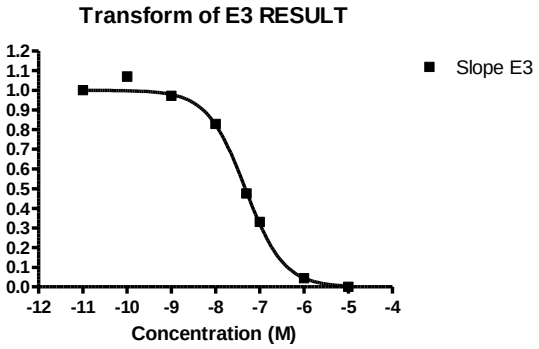
A.18 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 120 minutes)

# GDK-II-98 4a



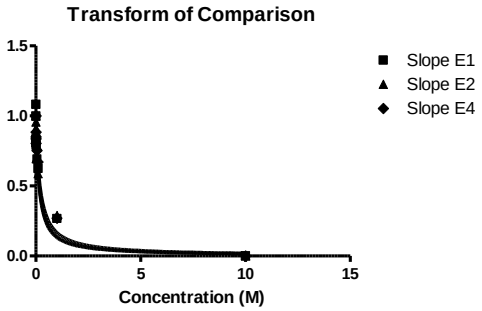




| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.008498 |
| 1.000000e-007 | 0.063020 |
| 5.000000e-008 | 0.090610 |
| 1.000000e-008 | 0.158100 |
| 1.000000e-009 | 0.185500 |
| 1.000000e-010 | 0.204100 |
| 1.000000e-011 | 0.191000 |
| 0.000000 | 0.190700 |

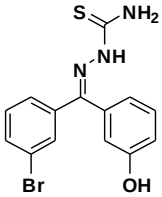
| Concentration (M) | Slope E3 |
|---|---|
| -5.000 | 0.000 |
| -6.000 | 0.045 |
| -7.000 | 0.330 |
| -7.301 | 0.475 |
| -8.000 | 0.829 |
| -9.000 | 0.973 |
| -10.000 | 1.070 |
| -11.000 | 1.002 |
| | 1.000 |

| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.326 |
| HillSlope | -1.000 |
| EC50 | 4.722e-008 |
| Std. Error | |
| LogEC50 | 0.03631 |
| HillSlope | 0.09587 |
| 95% Confidence Intervals | |
| LogEC50 | -7.415 to -7.237 |
| HillSlope | -1.235 to -0.7657 |
| EC50 | 3.849e-008 to 5.794e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9958 |
| Absolute Sum of Squares | 0.005528 |
| Sy.x | 0.03035 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

A.19 Representative $IC_{50}$ Calculation Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 240 minutes)




| Ki (nM) |
|---|
| 3.254 |
| 3.510 |
| 4.390 |



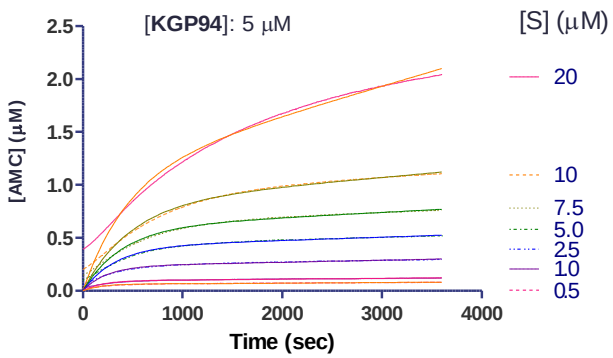

| Concentration (M) | Slope E1 | Slope E2 | Slope E3 | Slope E4 |
|---|---|---|---|---|
| 10.000 | 0.000 | 0.000 | | 0.000 |
| 1.000 | 0.268 | 0.289 | | 0.269 |
| 0.100 | 0.625 | 0.588 | | 0.672 |
| 0.050 | 0.693 | 0.789 | | 0.753 |
| 0.010 | 0.792 | 0.697 | | 0.837 |
| 0.001 | 0.808 | 0.911 | | 0.884 |
| 1.000e-004 | 0.835 | 0.826 | | 0.805 |
| 1.000e-005 | 1.083 | 0.958 | | 0.851 |
| 0.000 | 1.000 | 1.000 | | 1.000 |

| | Ki (nM) |
|---|---|
| Number of values | 3 |
| | |
| | |
| Mean | 3.718 |
| Std. Deviation | 0.5959 |
| Std. Error | 0.3440 |
| | |
| | |
| Sum | 11.15 |

## preincubation: 5 min
## [S]: 50 uM

| | Slope E1 | Slope E2 | Slope E3 | Slope E4 |
|---|---|---|---|---|
| Morrison Ki | | | | |
| Best-fit values | | | | |
| Vo | = 1.000 | = 1.000 | | = 1.000 |
| Et | = 0.0010 | = 0.0010 | | = 0.0010 |
| Ki | 0.003254 | 0.003519 | | 0.004394 |
| S | = 50.00 | = 50.00 | | = 50.00 |
| Km | = 1.100 | = 1.100 | | = 1.100 |
| Std. Error | | | | |
| Ki | 0.001173 | 0.001349 | | 0.001583 |
| 95% Confidence Intervals | | | | |
| Ki | 0.0005496 to 0.005958 | 0.0004078 to 0.006629 | | 0.0007435 to 0.008044 |
| Goodness of Fit | | | | |
| Degrees of Freedom | 8 | 8 | | 8 |
| R square | 0.8821 | 0.8618 | | 0.8836 |
| Absolute Sum of Squares | 0.1123 | 0.1231 | | 0.09893 |
| Sy.x | 0.1185 | 0.1240 | | 0.1112 |
| Constraints | | | | |
| Vo | Vo = 1.000 | Vo = 1.000 | | Vo = 1.000 |
| Et | Et = 0.0010 | Et = 0.0010 | | Et = 0.0010 |
| S | S = 50.00 | S = 50.00 | | S = 50.00 |
| Km | Km = 1.100 | Km = 1.100 | | Km = 1.100 |
| Number of points | | | | |
| Analyzed | 9 | 9 | | 9 |

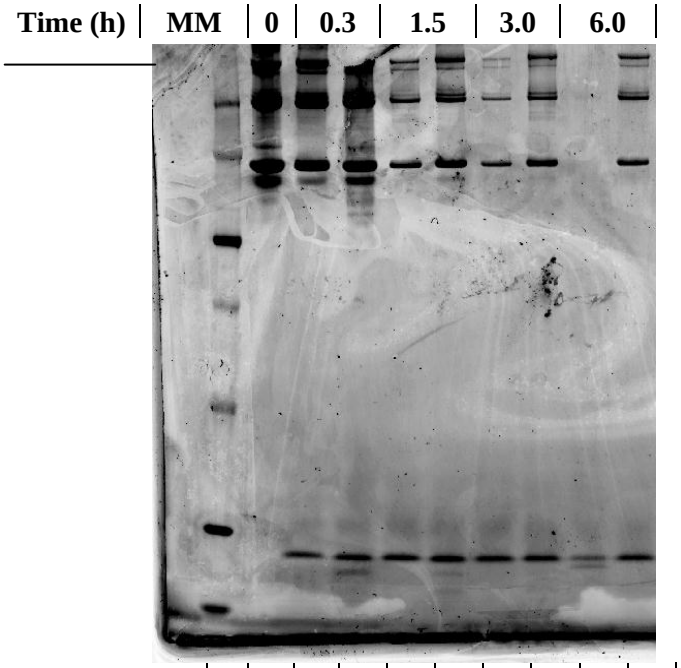
A.20 Calculation of $K_I^{app}$ Using **8** as a Cathepsin L Inhibitor (Preincubation Time: 5 minutes




**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | 20 uM | 10 uM | 7.5 uM | 5.0 uM | 2.5 uM | 1.0 uM | 0.5 uM |
|---|---|---|---|---|---|---|---|
| Progress Curve | | | | | | | |
| Best-fit values | | | | | | | |
| vs | 0.0002770 | 8.329e-005 | 4.884e-005 | 3.055e-005 | 1.919e-005 | 7.727e-006 | 5.059e-006 |
| vi | 0.002723 | 0.001808 | 0.001560 | 0.001302 | 0.001007 | 0.0005189 | 0.0003824 |
| kobs | 0.002219 | 0.002096 | 0.002542 | 0.003058 | 0.004260 | 0.005365 | 0.005975 |
| Std. Error | | | | | | | |
| vs | 5.683e-006 | 2.799e-006 | 1.558e-006 | 8.708e-007 | 4.093e-007 | 1.656e-007 | 1.192e-007 |
| vi | 4.380e-005 | 1.918e-005 | 1.579e-005 | 1.265e-005 | 1.095e-005 | 6.628e-006 | 5.734e-006 |
| kobs | 6.071e-005 | 3.673e-005 | 3.802e-005 | 4.051e-005 | 5.724e-005 | 8.064e-005 | 0.0001035 |
| 95% Confidence Intervals | | | | | | | |
| vs | 0.0002658 to 0.000288 | 7.781e-005 to 8.878e-00 | 4.579e-005 to 5.189e-00 | 2.884e-005 to 3.225e-00 | 1.838e-005 to 1.999e-00 | 7.402e-006 to 8.051e-00 | 4.825e-006 to 5.292e-00 |
| vi | 0.002638 to 0.002809 | 0.001770 to 0.001845 | 0.001529 to 0.001590 | 0.001277 to 0.001327 | 0.0009856 to 0.001028 | 0.0005059 to 0.0005319 | 0.0003712 to 0.0003937 |
| kobs | 0.002100 to 0.002338 | 0.002024 to 0.002168 | 0.002467 to 0.002616 | 0.002979 to 0.003138 | 0.004148 to 0.004372 | 0.005207 to 0.005523 | 0.005772 to 0.006177 |
| Goodness of Fit | | | | | | | |
| Degrees of Freedom | 718 | 718 | 718 | 718 | 718 | 718 | 718 |
| R square | 0.9801 | 0.9836 | 0.9821 | 0.9816 | 0.9770 | 0.9671 | 0.9534 |
| Absolute Sum of Squares | 3.133 | 0.6766 | 0.3023 | 0.1259 | 0.04074 | 0.008050 | 0.004483 |
| Sy.x | 0.06605 | 0.03070 | 0.02052 | 0.01324 | 0.007533 | 0.003348 | 0.002499 |
| Constraints | | | | | | | |
| vs | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | | |
| Analyzed | 721 | 721 | 721 | 721 | 721 | 721 | 721 |

A.21 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [I]: 5 μM (Preincubation Time: 0 minutes)




A.22 Inhibition of collagenase activity of cathepsin L by **8** (Preincubation Time: 120 minutes)

E1

| | 10 uM | Control |
|---|---|---|
| Best-fit values | | |
| Slope | 0.1448 ± 0.002335 | 0.1675 ± 0.003039 |
| Y-intercept when X=0.0 | 168.2 ± 0.4803 | 165.3 ± 0.6249 |
| X-intercept when Y=0.0 | -1162 | -881.9 |
| 1/slope | 6.907 | 5.334 |
| 95% Confidence Intervals | | |
| Slope | 0.1397 to 0.1498 | 0.1809 to 0.1940 |
| Y-intercept when X=0.0 | 167.2 to 169.3 | 164.0 to 166.7 |
| X-intercept when Y=0.0 | -1210 to -1117 | -920.4 to -845.9 |
| Goodness of Fit | | |
| R square | 0.9966 | 0.9996 |
| Sy.x | 0.9769 | 1.271 |
| Is slope significantly non-zero? | | |
| F | 3844 | 3807 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 15 | 15 |
| Number of missing values | 0 | 0 |

E2

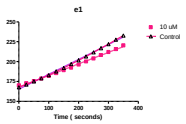
| | 10 uM | Control |
|---|---|---|
| Best-fit values | | |
| Slope | 0.1235 ± 0.002200 | 0.1501 ± 0.003118 |
| Y-intercept when X=0.0 | 162.6 ± 0.4522 | 159.6 ± 0.6409 |
| X-intercept when Y=0.0 | -1317 | -1064 |
| 1/slope | 8.099 | 6.662 |
| 95% Confidence Intervals | | |
| Slope | 0.1187 to 0.1282 | 0.1434 to 0.1568 |
| Y-intercept when X=0.0 | 161.7 to 163.6 | 158.3 to 161.0 |
| X-intercept when Y=0.0 | -1377 to -1262 | -1122 to -1010 |
| Goodness of Fit | | |
| R square | 0.9959 | 0.9944 |
| Sy.x | 0.9202 | 1.304 |
| Is slope significantly non-zero? | | |
| F | 3150 | 2318 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 15 | 15 |
| Number of missing values | 0 | 0 |

E3

| | 10 uM | Control 3 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1246 ± 0.001676 | 0.2257 ± 0.001939 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 156.8 ± 0.3447 | 168.7 ± 0.3986 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1258 | -747.1 | -29806 |
| 1/slope | 8.025 | 4.430 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1210 to 0.1282 | 0.2216 to 0.2299 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 156.1 to 157.6 | 167.8 to 169.5 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1301 to -1218 | -764.7 to -730.2 | -40176 to -23677 |
| Goodness of Fit | | | |
| R square | 0.9977 | 0.9990 | 0.8446 |
| Sy.x | 0.7012 | 0.8110 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 5527 | 13559 | 70.65 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

E4

| | 10 uM | Control |
|---|---|---|
| Best-fit values | | |
| Slope | 0.1242 ± 0.001888 | 0.1459 ± 0.003086 |
| Y-intercept when X=0.0 | 159.8 ± 0.3882 | 164.9 ± 0.6346 |
| X-intercept when Y=0.0 | -1286 | -1130 |
| 1/slope | 8.050 | 6.855 |
| 95% Confidence Intervals | | |
| Slope | 0.1202 to 0.1283 | 0.1392 to 0.1525 |
| Y-intercept when X=0.0 | 158.9 to 160.6 | 163.5 to 166.3 |
| X-intercept when Y=0.0 | -1336 to -1239 | -1193 to -1073 |
| Goodness of Fit | | |
| R square | 0.9970 | 0.9942 |
| Sy.x | 0.7897 | 1.291 |
| Is slope significantly non-zero? | | |
| F | 4330 | 2234 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 15 | 15 |
| Number of missing values | 0 | 0 |

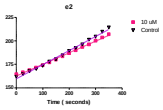


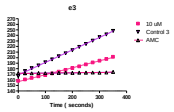

E1

| Time ( seconds) | 10 uM | Control |
|---|---|---|
| 0.00 | 170.224 | 167.788 |
| 25.00 | 172.792 | 171.136 |
| 50.01 | 175.431 | 175.226 |
| 75.00 | 178.748 | 179.070 |
| 100.01 | 182.095 | 182.932 |
| 125.01 | 185.491 | 187.708 |
| 150.01 | 189.086 | 192.038 |
| 175.02 | 192.426 | 196.960 |
| 200.02 | 196.326 | 201.701 |
| 225.01 | 200.219 | 206.583 |
| 250.01 | 203.963 | 211.686 |
| 275.01 | 208.287 | 216.938 |
| 300.02 | 212.060 | 222.022 |
| 325.01 | 215.807 | 227.655 |
| 350.01 | 220.487 | 232.668 |

E2

| Time ( seconds) | 10 uM | Control |
|---|---|---|
| 0.00 | 164.550 | 161.869 |
| 24.90 | 166.615 | 164.729 |
| 49.89 | 168.883 | 167.741 |
| 74.91 | 171.580 | 170.174 |
| 99.90 | 174.100 | 173.623 |
| 124.90 | 177.160 | 177.875 |
| 149.91 | 180.011 | 180.418 |
| 174.91 | 183.490 | 184.704 |
| 199.90 | 186.756 | 189.069 |
| 224.91 | 189.976 | 192.274 |
| 249.90 | 193.417 | 196.247 |
| 274.91 | 196.203 | 201.236 |
| 299.90 | 200.079 | 204.554 |
| 324.90 | 203.567 | 209.794 |
| 349.89 | 207.078 | 214.234 |



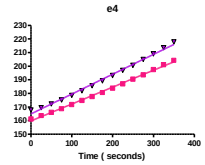




E3

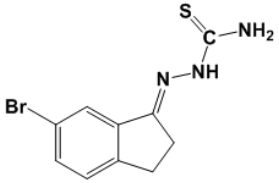
| Time ( seconds) | 10 uM | Control 3 | AMC |
|---|---|---|---|
| 0.00 | 158.248 | 166.572 | 172.054 |
| 24.97 | 160.483 | 175.601 | 171.714 |
| 49.97 | 163.293 | 181.067 | 171.702 |
| 74.97 | 166.161 | 185.722 | 172.106 |
| 99.97 | 168.468 | 191.755 | 172.016 |
| 124.97 | 171.677 | 196.864 | 171.991 |
| 149.98 | 174.696 | 202.323 | 172.678 |
| 174.97 | 178.000 | 207.922 | 172.718 |
| 199.97 | 181.221 | 213.440 | 172.473 |
| 224.98 | 184.366 | 219.246 | 172.947 |
| 249.97 | 187.801 | 224.703 | 172.975 |
| 274.97 | 191.329 | 230.391 | 173.145 |
| 299.98 | 194.574 | 236.762 | 173.009 |
| 324.98 | 197.740 | 241.841 | 173.311 |
| 349.97 | 201.356 | 248.124 | 174.260 |

E4

| Time ( seconds) | 10 uM | Control |
|---|---|---|
| 0.00 | 161.213 | 167.801 |
| 24.97 | 163.676 | 169.518 |
| 49.98 | 166.094 | 172.247 |
| 74.98 | 168.718 | 175.253 |
| 99.97 | 171.742 | 178.653 |
| 124.98 | 174.815 | 181.935 |
| 149.98 | 177.628 | 185.622 |
| 174.97 | 180.665 | 189.325 |
| 199.98 | 183.859 | 193.160 |
| 224.98 | 187.045 | 196.944 |
| 249.97 | 190.422 | 200.635 |
| 274.98 | 193.654 | 204.998 |
| 299.98 | 197.476 | 209.069 |
| 324.97 | 201.072 | 213.539 |
| 349.97 | 204.330 | 217.859 |




| | % Inhibition |
|---|---|
| Number of values | 3 |
| Mean | 18.47 |
| Std. Deviation | 4.005 |
| Std. Error | 2.313 |
| Sum | 55.40 |




| | E1 | E2 | E3 | E4 |
|---|---|---|---|---|
| Control | 0.000 | 0.000 | | 0.000 |
| 10 uM | 0.228 | 0.177 | | 0.149 |

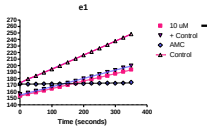
ALJ-I-15
KGP242

| | EDTA | DTT | BRIJ | NaOAc | DMSO | CAT L | Z-FR-AMC |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | µM |
| | 1 | 3 | 0.01 | 100 | 2 | 1 | 50 |

A.23 Representative $IC_{50}$ Calculation Using **83** as a Cathepsin L Inhibitor (Preincubation

**E1**

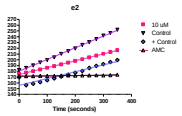
| Best-fit values | 10 uM | Control | + Control |
|---|---|---|---|
| Slope | 0.1146 ± 0.0007484 | 0.2129 ± 0.0008208 | 0.1273 ± 0.001252 |
| Y-intercept when X=0.0 | 153.3 ± 0.1539 | 173.6 ± 0.1688 | 154.9 ± 0.2574 |
| X-intercept when Y=0.0 | -1338 | -815.1 | -1217 |
| 1/slope | 8.723 | 4.697 | 7.857 |
| 95% Confidence Intervals | | | |
| Slope | 0.1130 to 0.1162 | 0.2112 to 0.2147 | 0.1246 to 0.1300 |
| Y-intercept when X=0.0 | 153.0 to 153.7 | 173.2 to 173.9 | 154.3 to 155.4 |
| X-intercept when Y=0.0 | -1359 to -1317 | -823.5 to -806.9 | -1247 to -1188 |
| Goodness of Fit | | | |
| R square | 0.9994 | 0.9998 | 0.9987 |
| Sy.x | 0.3131 | 0.3434 | 0.5237 |
| Is slope significantly non-zero? | | | |
| F | 23463 | 67289 | 10336 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |




| Time (seconds) | 10 uM | Control | + Control |
|---|---|---|---|
| 0.00 | 154.010 | 173.572 | 156.187 |
| 24.94 | 156.381 | 179.626 | 158.332 |
| 49.95 | 159.105 | 184.411 | 161.127 |
| 74.94 | 161.730 | 189.555 | 164.147 |
| 99.94 | 164.726 | 194.585 | 167.119 |
| 124.94 | 167.430 | 199.587 | 170.616 |
| 149.94 | 170.356 | 205.078 | 173.471 |
| 174.94 | 173.151 | 210.693 | 176.667 |
| 199.94 | 175.961 | 216.091 | 180.029 |
| 224.94 | 178.922 | 221.252 | 183.042 |
| 249.94 | 181.806 | 226.610 | 186.720 |
| 274.94 | 184.704 | 231.998 | 190.111 |
| 299.94 | 187.849 | 237.525 | 193.056 |
| 324.94 | 190.843 | 242.949 | 196.794 |
| 349.95 | 194.034 | 248.526 | 199.959 |

**E2**

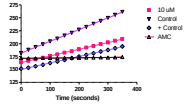
| Best-fit values | 10 uM | Control | + Control | AMC |
|---|---|---|---|---|
| Slope | 0.1188 ± 0.001127 | 0.2024 ± 0.002188 | 0.1182 ± 0.01085 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 174.6 ± 0.2317 | 180.3 ± 0.4499 | 155.7 ± 2.231 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1469 | -890.8 | -1318 | -29806 |
| 1/slope | 8.414 | 4.940 | 8.463 | 173.7 |
| 95% Confidence Intervals | | | | |
| Slope | 0.1164 to 0.1213 | 0.1977 to 0.2072 | 0.09473 to 0.1416 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 174.1 to 175.1 | 179.4 to 181.3 | 150.9 to 160.5 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1503 to -1436 | -916.4 to -868.3 | -1688 to -1070 | -40175 to -23677 |
| Goodness of Fit | | | | |
| R square | 0.9988 | 0.9985 | 0.9012 | 0.8446 |
| Sy.x | 0.4714 | 0.9153 | 4.538 | 0.2865 |
| Is slope significantly non-zero? | | | | |
| F | 11121 | 8561 | 118.6 | 70.65 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant | Significant |
| Data | | | | |
| Number of X values | 15 | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 | 0 |




| Time (seconds) | 10 uM | Control | + Control |
|---|---|---|---|
| 0.00 | 175.769 | 182.379 | 169.764 |
| 24.97 | 177.836 | 185.934 | 156.049 |
| 49.97 | 180.397 | 190.238 | 158.544 |
| 74.97 | 183.262 | 195.566 | 161.485 |
| 99.97 | 186.055 | 199.485 | 164.642 |
| 124.98 | 188.973 | 204.449 | 167.779 |
| 149.97 | 192.090 | 210.524 | 171.146 |
| 174.97 | 195.110 | 215.177 | 174.338 |
| 199.97 | 197.999 | 219.717 | 177.681 |
| 224.97 | 200.989 | 225.775 | 180.729 |
| 249.97 | 204.215 | 230.459 | 185.872 |
| 274.98 | 207.188 | 235.903 | 189.494 |
| 299.97 | 210.475 | 241.964 | 192.476 |
| 324.97 | 213.594 | 246.612 | 196.121 |
| 349.97 | 216.746 | 252.129 | 199.604 |

**E3**

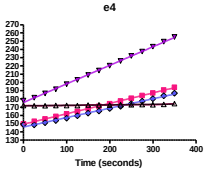
| Best-fit values | 10 uM | Control | + Control | AMC |
|---|---|---|---|---|
| Slope | 0.1298 ± 0.0008498 | 0.2074 ± 0.0007680 | 0.1258 ± 0.001141 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 163.2 ± 0.1747 | 181.8 ± 0.1579 | 150.1 ± 0.2346 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1257 | -799.3 | -1193 | -29805 |
| 1/slope | 7.701 | 4.398 | 7.947 | 173.7 |
| 95% Confidence Intervals | | | | |
| Slope | 0.1280 to 0.1317 | 0.2257 to 0.2291 | 0.1234 to 0.1283 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 162.8 to 163.6 | 181.4 to 182.1 | 149.6 to 150.6 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1277 to -1237 | -808.5 to -792.2 | -1220 to -1166 | -40174 to -23676 |
| Goodness of Fit | | | | |
| R square | 0.9994 | 0.9999 | 0.9989 | 0.8446 |
| Sy.x | 0.3555 | 0.3213 | 0.4772 | 0.2865 |
| Is slope significantly non-zero? | | | | |
| F | 23345 | 87659 | 12166 | 70.66 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant | Significant |
| Data | | | | |
| Number of X values | 15 | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 | 0 |




| Time (seconds) | 10 uM | Control | + Control |
|---|---|---|---|
| 0.00 | 163.955 | 181.764 | 151.345 |
| 24.96 | 166.573 | 188.137 | 153.370 |
| 49.96 | 169.777 | 193.328 | 156.192 |
| 74.96 | 172.723 | 198.856 | 159.054 |
| 99.96 | 176.006 | 204.297 | 162.497 |
| 124.96 | 179.044 | 209.783 | 165.359 |
| 149.96 | 182.607 | 215.634 | 168.429 |
| 174.96 | 185.268 | 221.182 | 171.882 |
| 199.96 | 189.175 | 227.062 | 174.831 |
| 224.97 | 192.292 | 232.550 | 178.324 |
| 249.96 | 195.297 | 238.657 | 181.488 |
| 274.96 | 199.092 | 244.212 | 184.743 |
| 299.96 | 202.416 | 250.027 | 187.979 |
| 324.96 | 205.633 | 255.936 | 191.088 |
| 349.96 | 208.944 | 261.759 | 194.701 |

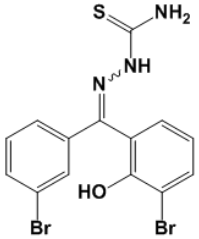
**E4**

| Best-fit values | 10 uM | Control | + Control | AMC |
|---|---|---|---|---|
| Slope | 0.1265 ± 0.0007106 | 0.2250 ± 0.001901 | 0.1142 ± 0.002456 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 149.4 ± 0.1461 | 175.7 ± 0.3809 | 145.9 ± 0.5049 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1181 | -780.7 | -1278 | -29805 |
| 1/slope | 7.908 | 4.444 | 8.759 | 173.7 |
| 95% Confidence Intervals | | | | |
| Slope | 0.1249 to 0.1280 | 0.2209 to 0.2292 | 0.1089 to 0.1195 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 149.0 to 149.7 | 174.9 to 176.5 | 144.9 to 147.0 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1198 to -1165 | -788.6 to -763.5 | -1349 to -1214 | -40172 to -23676 |
| Goodness of Fit | | | | |
| R square | 0.9996 | 0.9991 | 0.9940 | 0.8446 |
| Sy.x | 0.2973 | 0.7952 | 1.027 | 0.2865 |
| Is slope significantly non-zero? | | | | |
| F | 31667 | 14014 | 2161 | 70.67 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant | Significant |
| Data | | | | |
| Number of X values | 15 | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 | 0 |




| Time (seconds) | 10 uM | Control | + Control |
|---|---|---|---|
| 0.00 | 149.893 | 177.913 | 148.954 |
| 24.94 | 152.769 | 181.522 | 148.644 |
| 49.94 | 155.667 | 186.671 | 151.124 |
| 74.94 | 158.854 | 191.686 | 153.752 |
| 99.94 | 161.864 | 197.811 | 156.593 |
| 124.94 | 164.889 | 203.145 | 159.528 |
| 149.94 | 168.114 | 208.948 | 162.805 |
| 174.94 | 171.171 | 214.419 | 165.292 |
| 199.94 | 174.306 | 220.301 | 168.353 |
| 224.95 | 177.465 | 225.983 | 171.085 |
| 249.94 | 180.792 | 232.177 | 174.294 |
| 274.94 | 184.206 | 237.566 | 177.292 |
| 299.94 | 187.370 | 243.434 | 180.298 |
| 324.94 | 190.768 | 249.468 | 183.884 |
| 349.94 | 194.053 | 255.069 | 186.937 |

| | % Inhibition |
|---|---|
| Number of values | 4 |
| Mean | 43.55 |
| Std. Deviation | 2.047 |
| Std. Error | 1.023 |
| Sum | 174.2 |




| | E1 | E2 | E3 | E4 |
|---|---|---|---|---|
| Control | 0.000 | 0.000 | 0.000 | 0.000 |
| 10 uM | 0.462 | 0.413 | 0.429 | 0.438 |
| + Control (50 nM) | 0.402 | 0.416 | 0.447 | 0.492 |

**GDK-III-35 (1 mg)**
**KGP224**

| | EDTA | DTT | BRIJ | NaOAc | DMSO | CAT L | Z-FR-AMC |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | µM |
| | 1 | 3 | 0.01 | 100 | 2 | 1 | 50 |

A.23 Representative $IC_{50}$ Calculation Using **32** as a Cathepsin L Inhibitor

E1

| | 10 uM | Control | + Control |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.2344 ± 0.001800 | 0.2129 ± 0.0008208 | 0.1273 ± 0.001252 |
| Y-intercept when X=0.0 | 177.8 ± 0.3701 | 173.6 ± 0.1888 | 154.9 ± 0.2574 |
| X-intercept when Y=0.0 | -758.3 | -815.1 | -1217 |
| 1/slope | 4.266 | 4.697 | 7.857 |
| 95% Confidence Intervals | | | |
| Slope | 0.2305 to 0.2383 | 0.2112 to 0.2147 | 0.1246 to 0.1300 |
| Y-intercept when X=0.0 | 177.0 to 178.6 | 173.2 to 173.9 | 154.3 to 155.4 |
| X-intercept when Y=0.0 | -774.2 to -743.0 | -823.5 to -806.9 | -1247 to -1188 |
| Goodness of Fit | | | |
| R square | 0.9992 | 0.9998 | 0.9987 |
| Sy.x | 0.7531 | 0.3434 | 0.5237 |
| Is slope significantly non-zero? | | | |
| F | 16951 | 67289 | 10336 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

E2

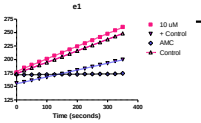
| | 10 uM | Control 2 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1453 ± 0.001092 | 0.2133 ± 0.001033 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 171.2 ± 0.2246 | 170.1 ± 0.2124 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1178 | -797.5 | -29806 |
| 1/slope | 6.882 | 4.688 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1430 to 0.1477 | 0.2111 to 0.2155 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 170.7 to 171.7 | 169.7 to 170.6 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1201 to -1157 | -807.8 to -787.3 | -40175 to -23677 |
| Goodness of Fit | | | |
| R square | 0.9993 | 0.9997 | 0.8446 |
| Sy.x | 0.4570 | 0.4322 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 17695 | 42634 | 70.65 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

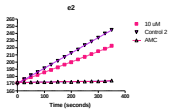
E3

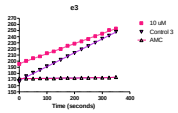
| | 10 uM | Control 3 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1641 ± 0.001633 | 0.2257 ± 0.001940 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 195.9 ± 0.3358 | 168.7 ± 0.3988 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1194 | -747.1 | -29805 |
| 1/slope | 6.093 | 4.430 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1606 to 0.1676 | 0.2216 to 0.2299 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 195.2 to 196.7 | 167.8 to 169.5 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1224 to -1165 | -764.7 to -730.2 | -40174 to -23676 |
| Goodness of Fit | | | |
| R square | 0.9987 | 0.9990 | 0.8446 |
| Sy.x | 0.6832 | 0.8114 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 10096 | 13544 | 70.66 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

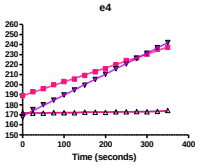
E4

| | 10 uM | Control 4 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1383 ± 0.001325 | 0.2087 ± 0.001368 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 189.0 ± 0.2724 | 168.8 ± 0.2812 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1367 | -808.8 | -29805 |
| 1/slope | 7.233 | 4.790 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1354 to 0.1411 | 0.2058 to 0.2117 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 188.4 to 189.6 | 168.2 to 169.4 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1400 to -1336 | -823.0 to -795.0 | -40172 to -23676 |
| Goodness of Fit | | | |
| R square | 0.9988 | 0.9994 | 0.8446 |
| Sy.x | 0.5542 | 0.5722 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 10893 | 23292 | 70.67 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |




E1

| Time (seconds) | 10 uM | Control | + Control |
|---|---|---|---|
| 0.00 | 175.970 | 173.572 | 156.187 |
| 24.94 | 184.909 | 179.626 | 158.332 |
| 49.95 | 190.199 | 184.411 | 161.127 |
| 74.94 | 195.946 | 189.555 | 164.147 |
| 99.94 | 201.231 | 194.585 | 167.119 |
| 124.94 | 207.169 | 199.587 | 170.616 |
| 149.94 | 212.695 | 205.078 | 173.471 |
| 174.94 | 218.721 | 210.693 | 176.667 |
| 199.94 | 224.257 | 216.091 | 180.029 |
| 224.94 | 230.105 | 221.252 | 183.042 |
| 249.94 | 235.713 | 226.610 | 186.720 |
| 274.94 | 241.764 | 231.998 | 190.111 |
| 299.94 | 247.945 | 237.525 | 193.056 |
| 324.94 | 254.298 | 242.949 | 196.794 |
| 349.95 | 260.466 | 248.526 | 199.959 |

E2

| Time (seconds) | 10 uM | Control 2 | AMC |
|---|---|---|---|
| 0.00 | 171.148 | 169.766 | 172.054 |
| 24.97 | 175.834 | 176.059 | 171.714 |
| 49.97 | 178.886 | 181.630 | 171.702 |
| 74.97 | 182.014 | 185.975 | 172.106 |
| 99.97 | 185.655 | 191.341 | 172.016 |
| 124.98 | 188.899 | 196.345 | 171.991 |
| 149.97 | 192.532 | 201.651 | 172.678 |
| 174.97 | 196.353 | 207.242 | 172.718 |
| 199.97 | 199.760 | 212.745 | 172.473 |
| 224.97 | 203.721 | 217.588 | 172.947 |
| 249.97 | 207.275 | 223.537 | 172.975 |
| 274.98 | 211.130 | 228.678 | 173.145 |
| 299.97 | 215.204 | 233.997 | 173.009 |
| 324.97 | 218.518 | 240.072 | 173.311 |
| 349.97 | 222.758 | 244.950 | 174.260 |

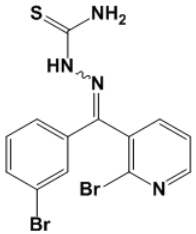
E3

| Time (seconds) | 10 uM | Control 3 | AMC |
|---|---|---|---|
| 0.00 | 195.577 | 166.572 | 172.054 |
| 24.96 | 200.464 | 175.601 | 171.714 |
| 49.96 | 204.920 | 181.067 | 171.702 |
| 74.96 | 207.820 | 185.722 | 172.106 |
| 99.96 | 212.929 | 191.755 | 172.016 |
| 124.96 | 215.998 | 196.864 | 171.991 |
| 149.96 | 219.923 | 202.323 | 172.678 |
| 174.96 | 225.175 | 207.922 | 172.718 |
| 199.96 | 228.661 | 213.440 | 172.473 |
| 224.97 | 232.496 | 219.246 | 172.947 |
| 249.96 | 236.037 | 224.703 | 172.975 |
| 274.96 | 240.791 | 230.391 | 173.145 |
| 299.96 | 244.852 | 236.762 | 173.009 |
| 324.96 | 250.873 | 241.841 | 173.311 |
| 349.96 | 253.246 | 248.124 | 174.260 |

E4

| Time (seconds) | 10 uM | Control 4 | AMC |
|---|---|---|---|
| 0.00 | 189.083 | 167.585 | 172.054 |
| 24.94 | 192.902 | 175.232 | 171.714 |
| 49.94 | 196.066 | 180.001 | 171.702 |
| 74.94 | 199.887 | 184.662 | 172.106 |
| 99.94 | 203.099 | 189.728 | 172.016 |
| 124.94 | 205.601 | 194.787 | 171.991 |
| 149.94 | 209.141 | 199.569 | 172.678 |
| 174.94 | 212.880 | 205.131 | 172.718 |
| 199.94 | 215.982 | 210.431 | 172.473 |
| 224.95 | 219.899 | 215.820 | 172.947 |
| 249.94 | 224.061 | 220.648 | 172.975 |
| 274.94 | 226.752 | 226.260 | 173.145 |
| 299.94 | 230.295 | 231.390 | 173.009 |
| 324.94 | 235.211 | 236.986 | 173.311 |
| 349.94 | 237.226 | 242.101 | 174.260 |

| | % Inhibition |
|---|---|
| Number of values | 3 |
| Mean | 30.97 |
| Std. Deviation | 3.301 |
| Std. Error | 1.906 |
| Sum | 92.90 |




| | E2 | E3 | E4 |
|---|---|---|---|
| Control | 0.000 | 0.000 | 0.000 |
| 10 uM | 0.319 | 0.273 | 0.337 |

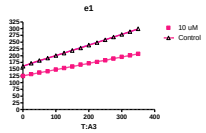
**GDK-III-52 (1.5 mg)**
**KGP233**

| | EDTA | DTT | BRIJ | NaOAc | DMSO | CAT L | Z-FR-AMC |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | µM |
| | 1 | 3 | 0.01 | 100 | 2 | 1 | 50 |

A.23 Representative $IC_{50}$ Calculation Using **38** as a Cathepsin L Inhibitor

E1

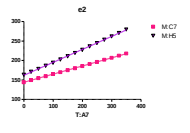
| | 10 uM | Control |
|---|---|---|
| Best-fit values | | |
| Slope | 0.2333 ± 0.001095 | 0.3927 ± 0.001924 |
| Y-intercept when X=0.0 | 124.8 ± 0.2251 | 161.0 ± 0.3955 |
| X-intercept when Y=0.0 | -535.1 | -409.9 |
| 1/slope | 4.287 | 2.546 |
| 95% Confidence Intervals | | |
| Slope | 0.2309 to 0.2357 | 0.3885 to 0.3969 |
| Y-intercept when X=0.0 | 124.4 to 125.3 | 160.1 to 161.8 |
| X-intercept when Y=0.0 | -542.5 to -527.9 | -416.2 to -403.6 |
| Goodness of Fit | | |
| R square | 0.9997 | 0.9997 |
| Sy.x | 0.4580 | 0.8047 |
| Is slope significantly non-zero? | | |
| F | 45397 | 41667 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 15 | 15 |
| Number of missing values | 0 | 0 |




| T:A3 | 10 uM | Control |
|---|---|---|
| 0.00 | 124.407 | 160.577 |
| 24.94 | 131.353 | 171.757 |
| 49.94 | 136.827 | 181.863 |
| 74.94 | 141.905 | 190.668 |
| 99.94 | 148.506 | 200.202 |
| 124.94 | 153.616 | 209.539 |
| 149.93 | 159.428 | 219.230 |
| 174.94 | 166.433 | 228.573 |
| 199.93 | 171.366 | 238.893 |
| 224.93 | 176.835 | 248.422 |
| 249.93 | 182.591 | 258.410 |
| 274.93 | 188.888 | 269.054 |
| 299.93 | 195.100 | 278.818 |
| 324.93 | 200.735 | 289.130 |
| 349.93 | 206.783 | 299.724 |

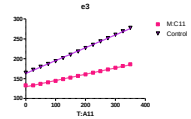
E2

| T:A7 | M:C7 | M:H5 |
|---|---|---|
| 0.00 | 144.293 | 163.562 |
| 24.93 | 149.530 | 171.093 |
| 49.93 | 154.826 | 178.690 |
| 74.92 | 159.739 | 186.654 |
| 99.93 | 165.358 | 194.372 |
| 124.91 | 169.882 | 202.754 |
| 149.91 | 175.291 | 210.988 |
| 174.92 | 180.258 | 219.381 |
| 199.91 | 185.456 | 227.515 |
| 224.89 | 190.986 | 236.274 |
| 249.89 | 196.197 | 244.773 |
| 274.89 | 201.924 | 253.403 |
| 299.89 | 206.964 | 261.959 |
| 324.89 | 212.387 | 270.994 |
| 349.88 | 217.927 | 279.549 |




E3

| T:A11 | M:C11 | Control |
|---|---|---|
| 0.00 | 132.898 | 164.913 |
| 24.92 | 133.039 | 172.469 |
| 49.92 | 136.586 | 179.750 |
| 74.92 | 140.518 | 186.872 |
| 99.92 | 144.192 | 194.147 |
| 124.91 | 148.467 | 201.709 |
| 149.90 | 152.331 | 210.322 |
| 174.92 | 156.580 | 218.039 |
| 199.91 | 160.538 | 226.805 |
| 224.90 | 164.584 | 235.001 |
| 249.89 | 168.706 | 243.521 |
| 274.89 | 172.857 | 251.944 |
| 299.87 | 177.162 | 260.516 |
| 324.87 | 181.662 | 269.109 |
| 349.86 | 186.139 | 277.618 |




E4

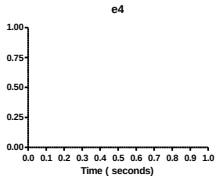
E2

| | M:C7 | M:H5 |
|---|---|---|
| Best-fit values | | |
| Slope | 0.2096 ± 0.0008027 | 0.3328 ± 0.002048 |
| Y-intercept when X=0.0 | 144.1 ± 0.1650 | 161.9 ± 0.4211 |
| X-intercept when Y=0.0 | -687.4 | -486.4 |
| 1/slope | 4.771 | 3.004 |
| 95% Confidence Intervals | | |
| Slope | 0.2079 to 0.2113 | 0.3284 to 0.3373 |
| Y-intercept when X=0.0 | 143.7 to 144.4 | 161.0 to 162.8 |
| X-intercept when Y=0.0 | -694.6 to -680.2 | -495.5 to -477.6 |
| Goodness of Fit | | |
| R square | 0.9998 | 0.9995 |
| Sy.x | 0.3357 | 0.8567 |
| Is slope significantly non-zero? | | |
| F | 68184 | 26406 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 15 | 15 |
| Number of missing values | 0 | 0 |

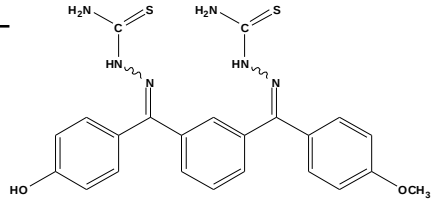
E3

| | M:C11 | Control |
|---|---|---|
| Best-fit values | | |
| Slope | 0.1587 ± 0.002885 | 0.3235 ± 0.002882 |
| Y-intercept when X=0.0 | 129.3 ± 0.5931 | 162.9 ± 0.5923 |
| X-intercept when Y=0.0 | -814.9 | -503.6 |
| 1/slope | 6.301 | 3.091 |
| 95% Confidence Intervals | | |
| Slope | 0.1525 to 0.1649 | 0.3173 to 0.3297 |
| Y-intercept when X=0.0 | 128.0 to 130.6 | 161.7 to 164.2 |
| X-intercept when Y=0.0 | -855.6 to -777.3 | -517.1 to -490.6 |
| Goodness of Fit | | |
| R square | 0.9957 | 0.9990 |
| Sy.x | 1.207 | 1.205 |
| Is slope significantly non-zero? | | |
| F | 3025 | 12605 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant |
| Data | | |
| Number of X values | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 |
| Total number of values | 15 | 15 |
| Number of missing values | 0 | 0 |

You have not entered the right kind of data for this analysis.




| | % Inhibition |
|---|---|
| Number of values | 3 |
| Mean | 42.83 |
| Std. Deviation | 7.214 |
| Std. Error | 4.165 |
| Sum | 128.5 |




| | E1 | E2 | E3 |
|---|---|---|---|
| Control | 0.000 | 0.000 | 0.000 |
| 10 uM | 0.406 | 0.370 | 0.509 |

SJL-III-124-3
KGP271

E4

| | EDTA | DTT | BRIJ | NaOAc | DMSO | CAT L | Z-FR-AMC |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | µM |
| | 1 | 3 | 0.01 | 100 | 2 | 1 | 50 |

A.23 Representative $IC_{50}$ Calculation Using **166** as a Cathepsin L Inhibitor

E1

| | 10 uM | Control | + Control |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1636 ± 0.002270 | 0.2129 ± 0.0008208 | 0.1273 ± 0.001252 |
| Y-intercept when X=0.0 | 153.3 ± 0.4666 | 173.6 ± 0.1688 | 154.9 ± 0.2574 |
| X-intercept when Y=0.0 | -937.3 | -815.1 | -1217 |
| 1/slope | 6.114 | 4.697 | 7.857 |
| 95% Confidence Intervals | | | |
| Slope | 0.1587 to 0.1685 | 0.2112 to 0.2147 | 0.1246 to 0.1300 |
| Y-intercept when X=0.0 | 152.3 to 154.3 | 173.2 to 173.9 | 154.3 to 155.4 |
| X-intercept when Y=0.0 | -971.8 to -904.7 | -823.5 to -806.9 | -1247 to -1188 |
| Goodness of Fit | | | |
| R square | 0.9975 | 0.9998 | 0.9987 |
| Sy.x | 0.9494 | 0.3434 | 0.5237 |
| Is slope significantly non-zero? | | | |
| F | 5194 | 67289 | 10336 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

E2

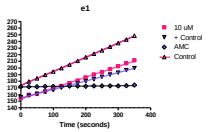
| | 10 uM | Control 2 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1826 ± 0.0007815 | 0.2133 ± 0.001033 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 163.7 ± 0.1607 | 170.1 ± 0.2124 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -896.5 | -797.5 | -29806 |
| 1/slope | 5.476 | 4.688 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1809 to 0.1843 | 0.2111 to 0.2155 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 163.4 to 164.1 | 169.7 to 170.6 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -906.5 to -886.6 | -807.8 to -787.3 | -40175 to -23677 |
| Goodness of Fit | | | |
| R square | 0.9998 | 0.9997 | 0.8446 |
| Sy.x | 0.3269 | 0.4322 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 54608 | 42634 | 70.65 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

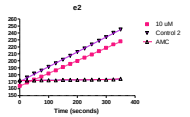
E3

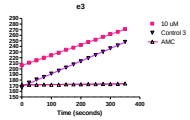
| | 10 uM | Control 3 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1851 ± 0.0007986 | 0.2257 ± 0.001940 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 206.0 ± 0.1642 | 168.7 ± 0.3988 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1113 | -747.1 | -29805 |
| 1/slope | 5.403 | 4.430 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1833 to 0.1868 | 0.2216 to 0.2299 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 205.6 to 206.3 | 167.8 to 169.5 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1125 to -1101 | -764.7 to -730.2 | -40174 to -23676 |
| Goodness of Fit | | | |
| R square | 0.9998 | 0.9990 | 0.8446 |
| Sy.x | 0.3341 | 0.8114 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 53701 | 13544 | 70.66 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

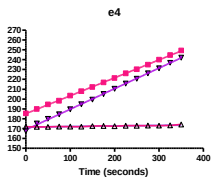
E4

| | 10 uM | Control 4 | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1821 ± 0.0007426 | 0.2087 ± 0.001368 | 0.005757 ± 0.0006849 |
| Y-intercept when X=0.0 | 185.2 ± 0.1527 | 168.8 ± 0.2812 | 171.6 ± 0.1408 |
| X-intercept when Y=0.0 | -1017 | -808.8 | -29805 |
| 1/slope | 5.491 | 4.790 | 173.7 |
| 95% Confidence Intervals | | | |
| Slope | 0.1805 to 0.1837 | 0.2058 to 0.2117 | 0.004278 to 0.007237 |
| Y-intercept when X=0.0 | 184.8 to 185.5 | 168.2 to 169.4 | 171.3 to 171.9 |
| X-intercept when Y=0.0 | -1027 to -1006 | -823.0 to -795.0 | -40172 to -23676 |
| Goodness of Fit | | | |
| R square | 0.9998 | 0.9994 | 0.8446 |
| Sy.x | 0.3107 | 0.5722 | 0.2865 |
| Is slope significantly non-zero? | | | |
| F | 60144 | 23292 | 70.67 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | < 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |




E1

| Time (seconds) | 10 uM | Control | + Control |
|---|---|---|---|
| 0.00 | 154.585 | 173.572 | 156.187 |
| 24.94 | 159.546 | 179.626 | 158.332 |
| 49.95 | 160.940 | 184.411 | 161.127 |
| 74.94 | 164.510 | 189.555 | 164.147 |
| 99.94 | 168.915 | 194.585 | 167.119 |
| 124.94 | 172.896 | 199.587 | 170.616 |
| 149.94 | 177.112 | 205.078 | 173.471 |
| 174.94 | 181.061 | 210.693 | 176.667 |
| 199.94 | 185.744 | 216.091 | 180.029 |
| 224.94 | 189.965 | 221.252 | 183.042 |
| 249.94 | 193.931 | 226.610 | 186.720 |
| 274.94 | 198.277 | 231.998 | 190.111 |
| 299.94 | 202.622 | 237.525 | 193.056 |
| 324.94 | 207.084 | 242.949 | 196.794 |
| 349.95 | 211.487 | 248.526 | 199.959 |

E2

| Time (seconds) | 10 uM | Control 2 | AMC |
|---|---|---|---|
| 0.00 | 164.068 | 169.766 | 172.054 |
| 24.97 | 168.760 | 176.059 | 171.714 |
| 49.97 | 172.968 | 181.630 | 171.702 |
| 74.97 | 177.286 | 185.975 | 172.106 |
| 99.97 | 181.783 | 191.341 | 172.016 |
| 124.98 | 186.256 | 196.345 | 171.991 |
| 149.97 | 190.995 | 201.651 | 172.678 |
| 174.97 | 195.519 | 207.242 | 172.718 |
| 199.97 | 199.852 | 212.745 | 172.473 |
| 224.97 | 204.600 | 217.588 | 172.947 |
| 249.97 | 208.946 | 223.537 | 172.975 |
| 274.98 | 213.781 | 228.678 | 173.145 |
| 299.97 | 218.747 | 233.997 | 173.009 |
| 324.97 | 223.571 | 240.072 | 173.311 |
| 349.97 | 228.006 | 244.950 | 174.260 |

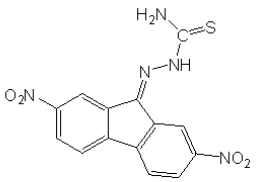
E3

| Time (seconds) | 10 uM | Control 3 | AMC |
|---|---|---|---|
| 0.00 | 206.471 | 166.572 | 172.054 |
| 24.96 | 210.950 | 175.601 | 171.714 |
| 49.96 | 215.312 | 181.067 | 171.702 |
| 74.96 | 219.712 | 185.722 | 172.106 |
| 99.96 | 224.094 | 191.755 | 172.016 |
| 124.96 | 228.697 | 196.864 | 171.991 |
| 149.96 | 233.900 | 202.323 | 172.678 |
| 174.96 | 237.830 | 207.922 | 172.718 |
| 199.96 | 242.597 | 213.440 | 172.473 |
| 224.97 | 247.799 | 219.246 | 172.947 |
| 249.96 | 252.071 | 224.703 | 172.975 |
| 274.96 | 257.015 | 230.391 | 173.145 |
| 299.96 | 261.885 | 236.762 | 173.009 |
| 324.96 | 266.135 | 241.841 | 173.311 |
| 349.96 | 270.960 | 248.124 | 174.260 |

E4

| Time (seconds) | 10 uM | Control 4 | AMC |
|---|---|---|---|
| 0.00 | 185.592 | 167.585 | 172.054 |
| 24.94 | 189.928 | 175.232 | 171.714 |
| 49.94 | 194.513 | 180.001 | 171.702 |
| 74.94 | 198.314 | 184.662 | 172.106 |
| 99.94 | 203.402 | 189.728 | 172.016 |
| 124.94 | 207.915 | 194.787 | 171.991 |
| 149.94 | 212.229 | 199.569 | 172.678 |
| 174.94 | 216.950 | 205.131 | 172.718 |
| 199.94 | 221.301 | 210.431 | 172.473 |
| 224.95 | 226.122 | 215.820 | 172.947 |
| 249.94 | 230.288 | 220.648 | 172.975 |
| 274.94 | 234.971 | 226.260 | 173.145 |
| 299.94 | 240.026 | 231.390 | 173.009 |
| 324.94 | 244.628 | 236.986 | 173.311 |
| 349.94 | 249.348 | 242.101 | 174.260 |

| | % Inhibition |
|---|---|
| Number of values | 3 |
| Mean | 15.03 |
| Std. Deviation | 2.706 |
| Std. Error | 1.562 |
| Sum | 45.10 |




| | E2 | E3 | E4 |
|---|---|---|---|
| Control | 0.000 | 0.000 | 0.000 |
| 10 uM | 0.144 | 0.180 | 0.127 |

**Sample Reference:SJL-II-138**
**KGP236**

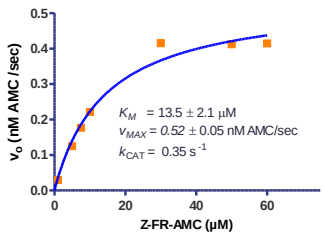
Stock sln: bright orange
Precipitated at 35% DMSO inhibitor dilution

| | EDTA | DTT | BRIJ | NaOAc | DMSO | CAT L | Z-FR-AMC |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | µM |
| | 1 | 3 | 0.01 | 100 | 2 | 1 | 50 |

A.23 Representative $IC_{50}$ Calculation Using **82** as a Cathepsin L Inhibitor

APPENDIX B

*In Vitro* Evaluations of Thiosemicarbazones as Inhibitors of Human Cathepsin K

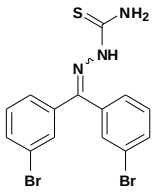


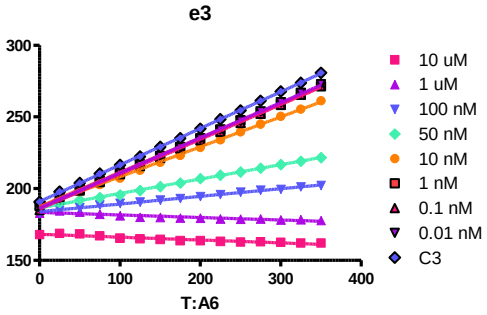

| | Substrate Concentration (uM) | |
|---|---|---|
| | 1.00000 | 0.030 |
| | 5.00000 | 0.125 |
| | 7.50000 | 0.177 |
| | 10.00000 | 0.221 |
| | 30.00000 | 0.416 |
| | 50.00000 | 0.413 |
| | 60.00000 | 0.415 |
| | | 0.372 |

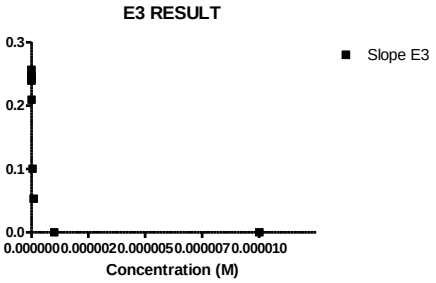
| | |
|---|---|
| Michaelis-Menten | |
| Best-fit values | |
| Vmax | 0.5422 |
| Km | 14.34 |
| Std. Error | |
| Vmax | 0.03844 |
| Km | 2.914 |
| 95% Confidence Intervals | |
| Vmax | 0.4433 to 0.6410 |
| Km | 6.845 to 21.83 |
| Goodness of Fit | |
| Degrees of Freedom | 5 |
| R square | 0.9779 |
| Absolute Sum of Squares | 0.003341 |
| Sy.x | 0.02585 |
| Constraints | |
| Km | Km > 0.0 |
| Number of points | |
| Analyzed | 7 |

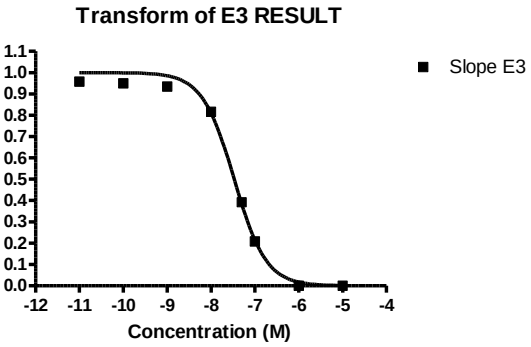
B.1 Calculation of Michaelis-Menten Constant for Cathepsin K using Z-FR-AMC as a Substrate

# LJ-I-006



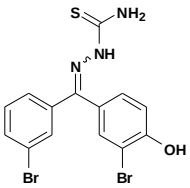

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.053130 |
| 5.000000e-008 | 0.100400 |
| 1.000000e-008 | 0.209100 |
| 1.000000e-009 | 0.239400 |
| 1.000000e-010 | 0.243100 |
| 1.000000e-011 | 0.245300 |
| 0.000000 | 0.256200 |

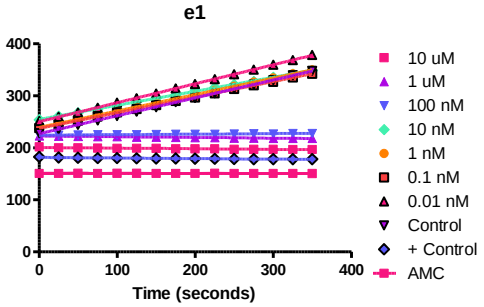
| Concentration (M) | Slope E3 |
|---|---|
| -5.000 | 0.000 |
| -6.000 | 0.000 |
| -7.000 | 0.207 |
| -7.301 | 0.392 |
| -8.000 | 0.816 |
| -9.000 | 0.934 |
| -10.000 | 0.949 |
| -11.000 | 0.957 |
| | 1.000 |

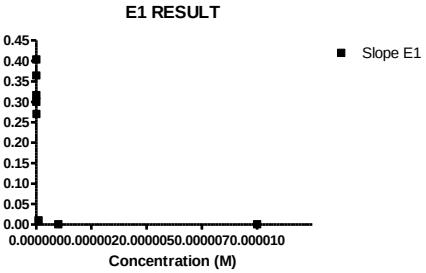
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.471 |
| HillSlope | -1.193 |
| EC50 | 3.381e-008 |
| Std. Error | |
| LogEC50 | 0.03995 |
| HillSlope | 0.1255 |
| 95% Confidence Intervals | |
| LogEC50 | -7.569 to -7.373 |
| HillSlope | -1.500 to -0.8858 |
| EC50 | 2.699e-008 to 4.234e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9943 |
| Absolute Sum of Squares | 0.007339 |
| Sy.x | 0.03497 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

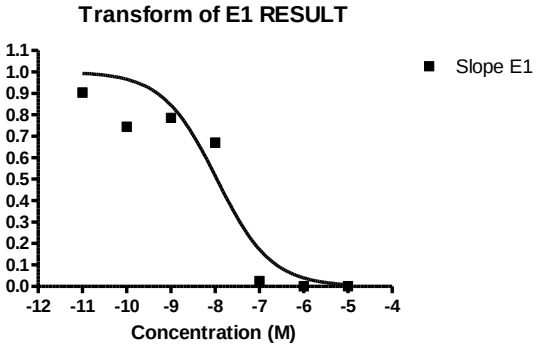
B.2 Representative $IC_{50}$ Calculation Using **1** as a Cathepsin K Inhibitor

# GDK-III-29




| Concentration (M) | Slope E1 |
|---|---|
| 1.00e-005 | 0.000000 |
| 1.00e-006 | 0.000000 |
| 1.00e-007 | 0.009867 |
| 1.00e-008 | 0.270200 |
| 1.00e-009 | 0.316500 |
| 1.00e-010 | 0.300100 |
| 1.00e-011 | 0.364400 |
| 0.00 | 0.403300 |

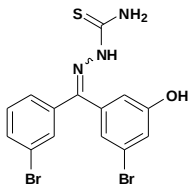
| | Concentration (M) | Slope E1 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.024 |
| | -8.000 | 0.670 |
| | -9.000 | 0.785 |
| | -10.000 | 0.744 |
| | -11.000 | 0.904 |
| | | 1.000 |

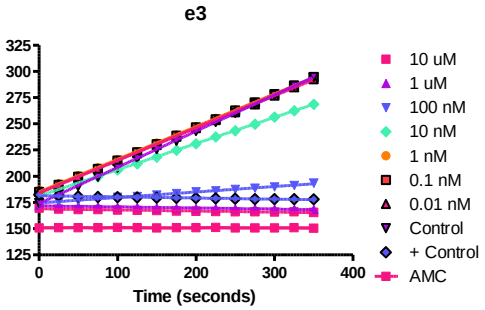
| | E1 |
|---|---|
| Control | 0.000 |
| + Control (100 nM) | 0.882 |

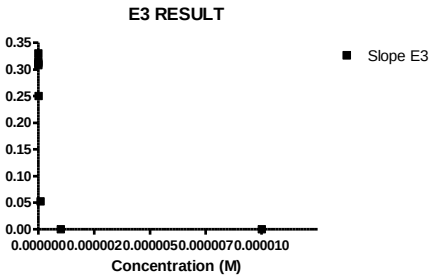
| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.960 |
| HillSlope | -0.7084 |
| EC50 | 1.095e-008 |
| Std. Error | |
| LogEC50 | 0.2811 |
| HillSlope | 0.2893 |
| 95% Confidence Intervals | |
| LogEC50 | -8.683 to -7.238 |
| HillSlope | -1.452 to 0.03548 |
| EC50 | 2.075e-009 to 5.784e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 5 |
| R square | 0.8962 |
| Absolute Sum of Squares | 0.1078 |
| Sy.x | 0.1468 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 7 |

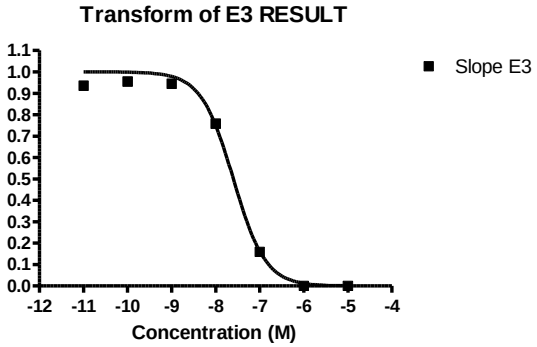
B.3 Representative $IC_{50}$ Calculation using **31** as a Cathepsin K Inhibitor

# GDK-III-35




| Concentration (M) | Slope E3 |
|---|---|
| 1.00e-005 | 0.000000 |
| 1.00e-006 | 0.000000 |
| 1.00e-007 | 0.052280 |
| 1.00e-008 | 0.250000 |
| 1.00e-009 | 0.311200 |
| 1.00e-010 | 0.314700 |
| 1.00e-011 | 0.308200 |
| 0.00 | 0.329600 |

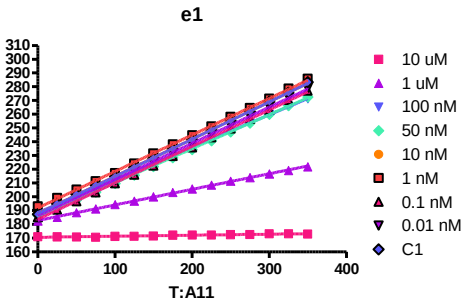
| | Concentration (M) | Slope E3 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.159 |
| | -8.000 | 0.758 |
| | -9.000 | 0.944 |
| | -10.000 | 0.955 |
| | -11.000 | 0.935 |
| | | 1.000 |

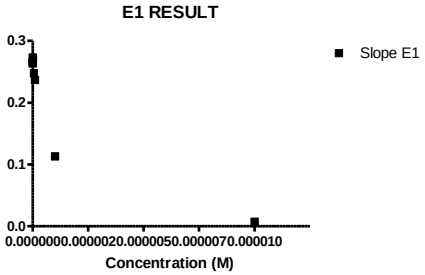
| | E3 |
|---|---|
| Control | 0.000 |
| + Control (100 nM) | 0.881 |

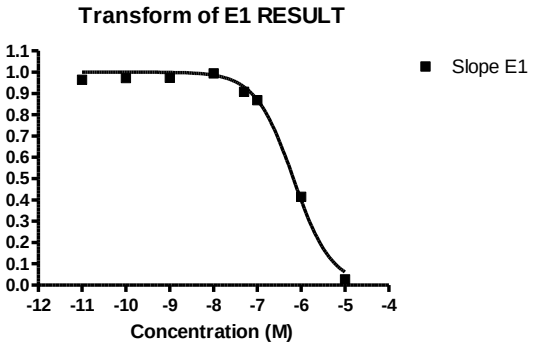
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.598 |
| HillSlope | -1.190 |
| EC50 | 2.524e-008 |
| Std. Error | |
| LogEC50 | 0.06148 |
| HillSlope | 0.1464 |
| 95% Confidence Intervals | |
| LogEC50 | -7.756 to -7.440 |
| HillSlope | -1.567 to -0.8136 |
| EC50 | 1.754e-008 to 3.632e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 5 |
| R square | 0.9940 |
| Absolute Sum of Squares | 0.007570 |
| Sy.x | 0.03891 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 7 |

B.4 Representative $IC_{50}$ Calculation Using **34** as a Cathepsin K Inhibitor

# GDK-II-63



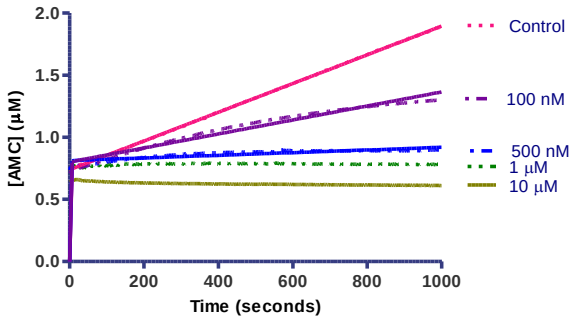




| Concentration (M) | Slope E1 |
|---|---|
| 0.000010 | 0.007452 |
| 0.000001 | 0.113100 |
| 1.000000e-007 | 0.236700 |
| 5.000000e-008 | 0.247500 |
| 1.000000e-008 | 0.270900 |
| 1.000000e-009 | 0.265500 |
| 1.000000e-010 | 0.264900 |
| 1.000000e-011 | 0.263000 |
| 0.000000 | 0.272800 |

| Concentration (M) | Slope E1 |
|---|---|
| -5.000 | 0.027 |
| -6.000 | 0.415 |
| -7.000 | 0.868 |
| -7.301 | 0.907 |
| -8.000 | 0.993 |
| -9.000 | 0.973 |
| -10.000 | 0.971 |
| -11.000 | 0.964 |
| | 1.000 |

| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -6.177 |
| HillSlope | -1.001 |
| EC50 | 6.650e-007 |
| Std. Error | |
| LogEC50 | 0.04391 |
| HillSlope | 0.08421 |
| 95% Confidence Intervals | |
| LogEC50 | -6.285 to -6.070 |
| HillSlope | -1.207 to -0.7945 |
| EC50 | 5.193e-007 to 8.517e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9945 |
| Absolute Sum of Squares | 0.004800 |
| Sy.x | 0.02828 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

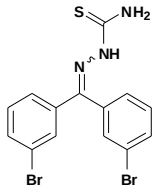
B.5 Representative $IC_{50}$ Calculation Using **18** as a Cathepsin K Inhibitor

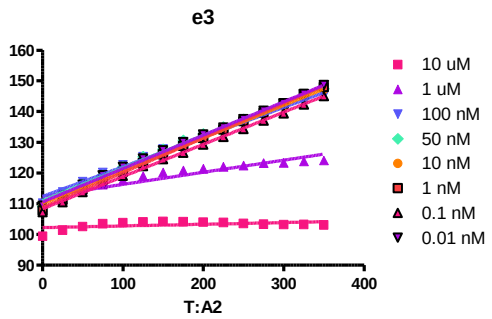


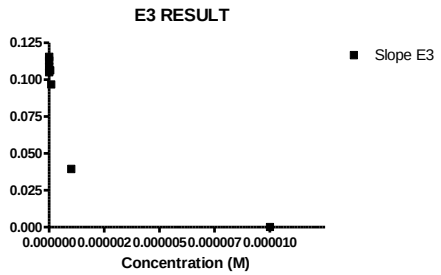

| | Control | Control | 10 uM | 1 uM | 500 nM | 100 nM |
|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | Not converged | Not converged | | |
| Best-fit values | | | | | | |
| vs | 0.001157 | | | | 0.0001072 | 0.0005641 |
| vi | ~ 1.386 | | | | 0.7186 | 0.5464 |
| kobs | ~ 1.873 | | | | 0.8844 | 0.6809 |
| Std. Error | | | | | | |
| vs | 7.686e-006 | | | | 8.494e-006 | 9.655e-006 |
| vi | ~ 3.500 | | | | 0.2047 | 0.1242 |
| kobs | ~ 4.736 | | | | 0.2525 | 0.1554 |
| 95% Confidence Intervals | | | | | | |
| vs | 0.001142 to 0.001172 | | | | 9.059e-005 to 0.0001239 | 0.0005452 to 0.0005830 |
| vi | (Very wide) | | | | 0.3174 to 1.120 | 0.3030 to 0.7898 |
| kobs | (Very wide) | | | | 0.3895 to 1.379 | 0.3762 to 0.9855 |
| Goodness of Fit | | | | | | |
| Degrees of Freedom | 331 | | | | 331 | 331 |
| R square | 0.9857 | | | | -0.4317 | 0.9101 |
| Absolute Sum of Squares | 0.5367 | | | | 0.6537 | 0.8423 |
| Sy.x | 0.04027 | | | | 0.04444 | 0.05044 |
| Constraints | | | | | | |
| vs | vs > 0.0 | | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | |
| Analyzed | 334 | | 334 | 334 | 334 | 334 |

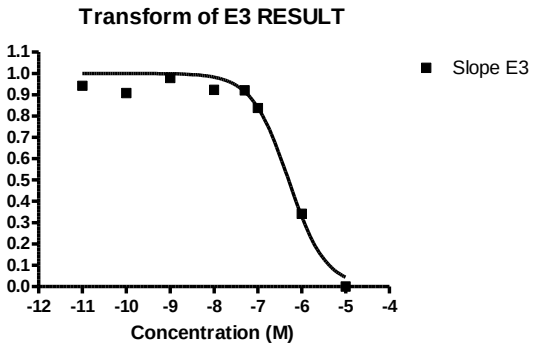
B.6 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 50 µM (Preincubation Time: 0 minutes)

# LJ-I-006



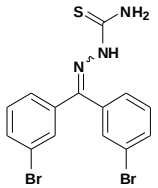

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.039420 |
| 1.000000e-007 | 0.096690 |
| 5.000000e-008 | 0.106300 |
| 1.000000e-008 | 0.106600 |
| 1.000000e-009 | 0.113000 |
| 1.000000e-010 | 0.104900 |
| 1.000000e-011 | 0.108800 |
| 0.000000 | 0.115500 |

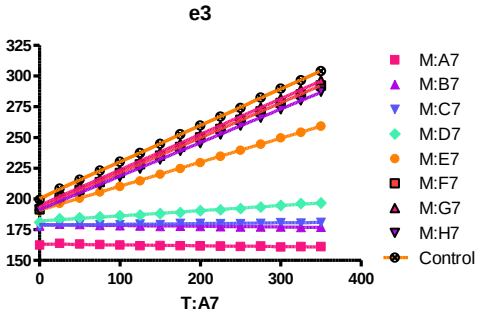
| | Concentration (M) | Slope E3 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.341 |
| | -7.000 | 0.837 |
| | -7.301 | 0.920 |
| | -8.000 | 0.923 |
| | -9.000 | 0.978 |
| | -10.000 | 0.908 |
| | -11.000 | 0.942 |
| | | 1.000 |

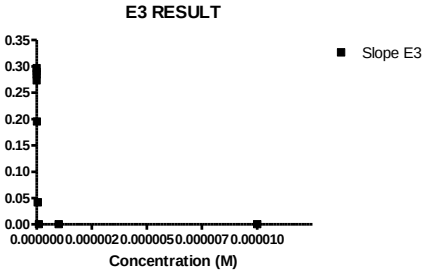
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -6.301 |
| HillSlope | -1.027 |
| EC50 | 4.996e-007 |
| Std. Error | |
| LogEC50 | 0.08417 |
| HillSlope | 0.1550 |
| 95% Confidence Intervals | |
| LogEC50 | -6.507 to -6.095 |
| HillSlope | -1.406 to -0.6477 |
| EC50 | 3.109e-007 to 8.027e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9804 |
| Absolute Sum of Squares | 0.01781 |
| Sy.x | 0.05449 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

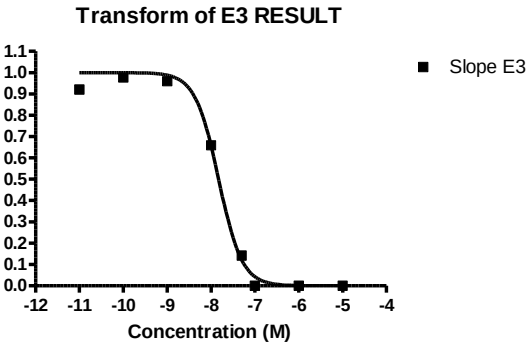
B.7 Representative $IC_{50}$ Calculation Using **1** as a Cathepsin K Inhibitor (Preincubation Time: 0 minutes)

# LJ-I-006



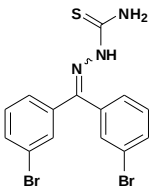

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.000000 |
| 5.000000e-008 | 0.041930 |
| 1.000000e-008 | 0.195400 |
| 1.000000e-009 | 0.284700 |
| 1.000000e-010 | 0.289800 |
| 1.000000e-011 | 0.273300 |
| 0.000000 | 0.296800 |

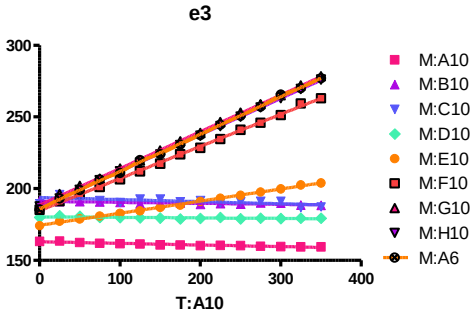
| | Concentration (M) | Slope E3 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.000 |
| | -7.301 | 0.141 |
| | -8.000 | 0.658 |
| | -9.000 | 0.959 |
| | -10.000 | 0.976 |
| | -11.000 | 0.921 |
| | | 1.000 |

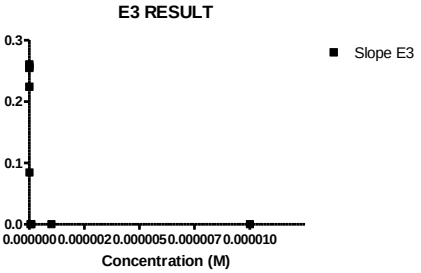
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.826 |
| HillSlope | -1.631 |
| EC50 | 1.494e-008 |
| Std. Error | |
| LogEC50 | 0.04278 |
| HillSlope | 0.2266 |
| 95% Confidence Intervals | |
| LogEC50 | -7.930 to -7.721 |
| HillSlope | -2.185 to -1.076 |
| EC50 | 1.174e-008 to 1.901e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9934 |
| Absolute Sum of Squares | 0.009853 |
| Sy.x | 0.04052 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

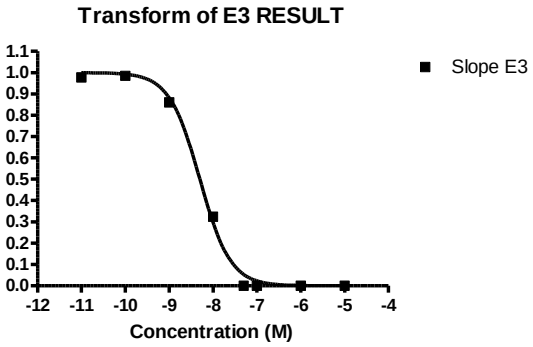
B.8 Representative $IC_{50}$ Calculation Using **1** as a Cathepsin K Inhibitor (Preincubation Time: 15 minutes)

# LJ-I-006



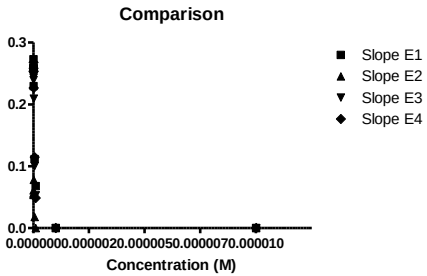




| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.000000 |
| 5.000000e-008 | 0.000000 |
| 1.000000e-008 | 0.084340 |
| 1.000000e-009 | 0.224000 |
| 1.000000e-010 | 0.256500 |
| 1.000000e-011 | 0.254500 |
| 0.000000 | 0.260300 |

| | Concentration (M) | Slope E3 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.000 |
| | -7.301 | 0.000 |
| | -8.000 | 0.324 |
| | -9.000 | 0.861 |
| | -10.000 | 0.985 |
| | -11.000 | 0.978 |
| | | 1.000 |

| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -8.299 |
| HillSlope | -1.250 |
| EC50 | 5.024e-009 |
| Std. Error | |
| LogEC50 | 0.04477 |
| HillSlope | 0.1188 |
| 95% Confidence Intervals | |
| LogEC50 | -8.408 to -8.189 |
| HillSlope | -1.541 to -0.9595 |
| EC50 | 3.904e-009 to 6.466e-009 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9966 |
| Absolute Sum of Squares | 0.005143 |
| Sy.x | 0.02928 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

B.9 Representative $IC_{50}$ Calculation Using **1** as a Cathepsin K Inhibitor (Preincubation Time: 45 minute)

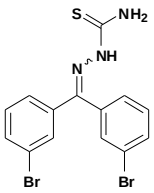


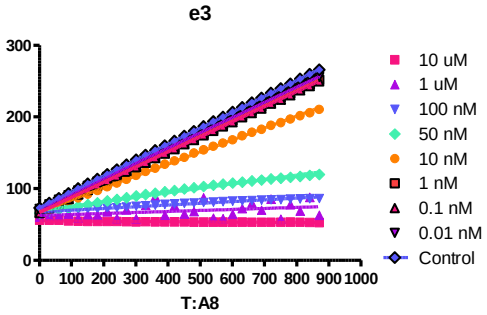

| | Concentration (M) | Slope E1 | Slope E2 | Slope E3 | Slope E4 |
|---|---|---|---|---|---|
| | 10.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 0.100 | 0.248 | 0.000 | 0.207 | 0.182 |
| | 0.050 | 0.409 | 0.295 | 0.392 | 0.428 |
| | 0.010 | 0.841 | 1.272 | 0.816 | 0.846 |
| | 0.001 | 0.946 | 0.921 | 0.934 | 0.944 |
| | 1.000e-004 | 0.948 | 0.891 | 0.949 | 0.948 |
| | 1.000e-005 | 0.949 | 0.876 | 0.957 | 0.964 |
| | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 |

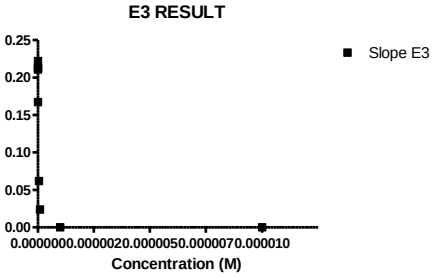
| | Slope E1 | Slope E2 | Slope E3 | Slope E4 |
|---|---|---|---|---|
| Morrison Ki | | | | |
| Best-fit values | | | | |
| Vo | = 1.000 | = 1.000 | = 1.000 | = 1.000 |
| Et | = 0.0015 | = 0.0015 | = 0.0015 | = 0.0015 |
| Ki | 0.006857 | 0.006632 | 0.006050 | 0.006490 |
| S | = 50.00 | = 50.00 | = 50.00 | = 50.00 |
| Km | = 11.80 | = 11.80 | = 11.80 | = 11.80 |
| Std. Error | | | | |
| Ki | 0.0006982 | 0.003949 | 0.0006160 | 0.0008231 |
| 95% Confidence Intervals | | | | |
| Ki | 0.005247 to 0.008467 | -0.002474 to 0.01574 | 0.004630 to 0.007471 | 0.004592 to 0.008388 |
| Goodness of Fit | | | | |
| Degrees of Freedom | 8 | 8 | 8 | 8 |
| R square | 0.9929 | 0.8281 | 0.9932 | 0.9895 |
| Absolute Sum of Squares | 0.01036 | 0.3519 | 0.01008 | 0.01589 |
| Sy.x | 0.03599 | 0.2097 | 0.03550 | 0.04457 |
| Constraints | | | | |
| Vo | Vo = 1.000 | Vo = 1.000 | Vo = 1.000 | Vo = 1.000 |
| Et | Et = 0.0015 | Et = 0.0015 | Et = 0.0015 | Et = 0.0015 |
| S | S = 50.00 | S = 50.00 | S = 50.00 | S = 50.00 |
| Km | Km = 11.80 | Km = 11.80 | Km = 11.80 | Km = 11.80 |
| Number of points | | | | |
| Analyzed | 9 | 9 | 9 | 9 |

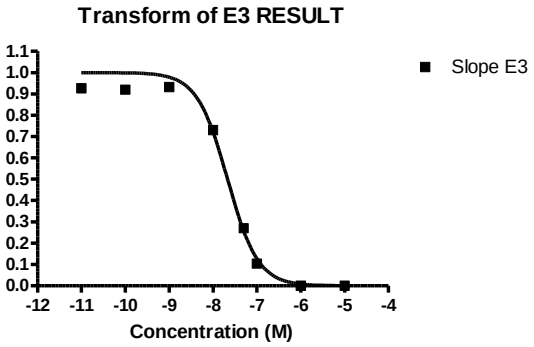
B.10 Calculation of $K_I^{app}$ Using **1** as a Cathepsin K Inhibitor (Preincubation Time: 5 minutes

# LJ-I-006



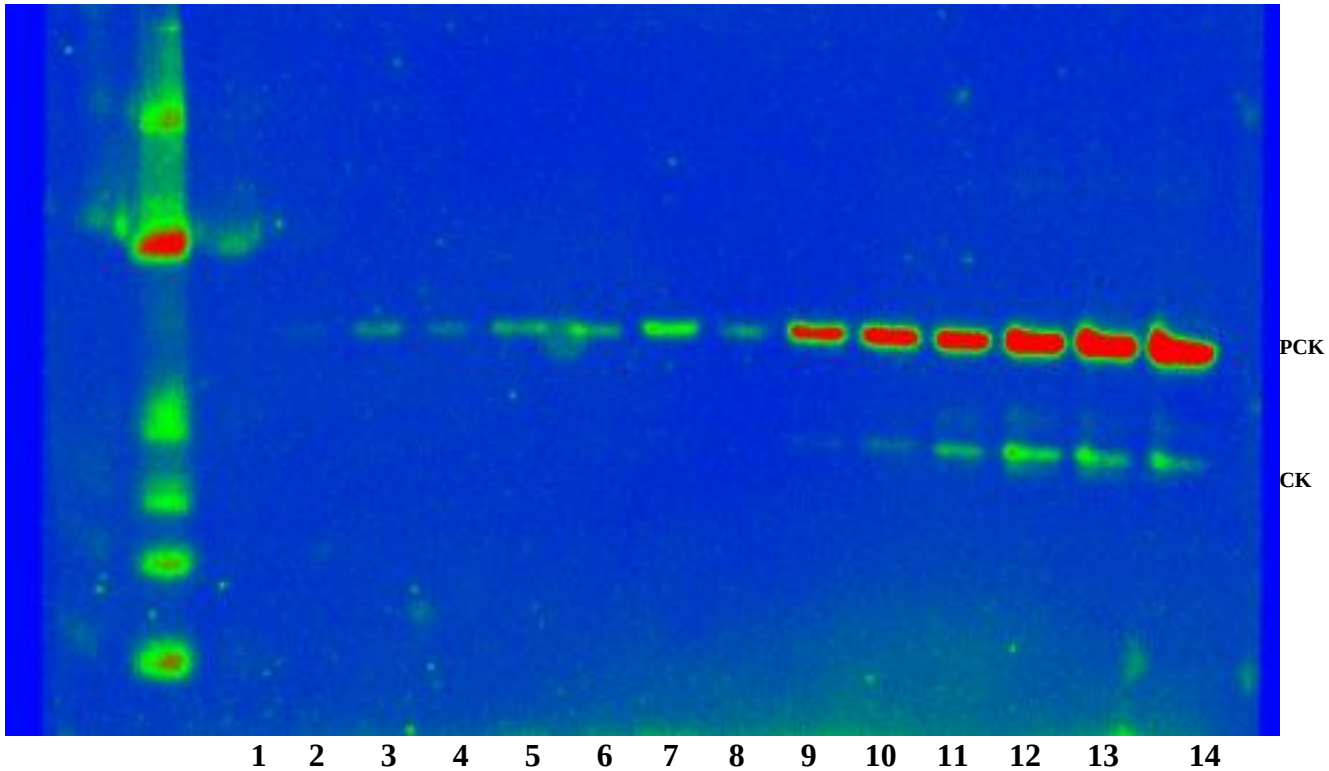




| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.023880 |
| 5.000000e-008 | 0.061840 |
| 1.000000e-008 | 0.167200 |
| 1.000000e-009 | 0.213400 |
| 1.000000e-010 | 0.210800 |
| 1.000000e-011 | 0.212300 |
| 0.000000 | 0.221900 |

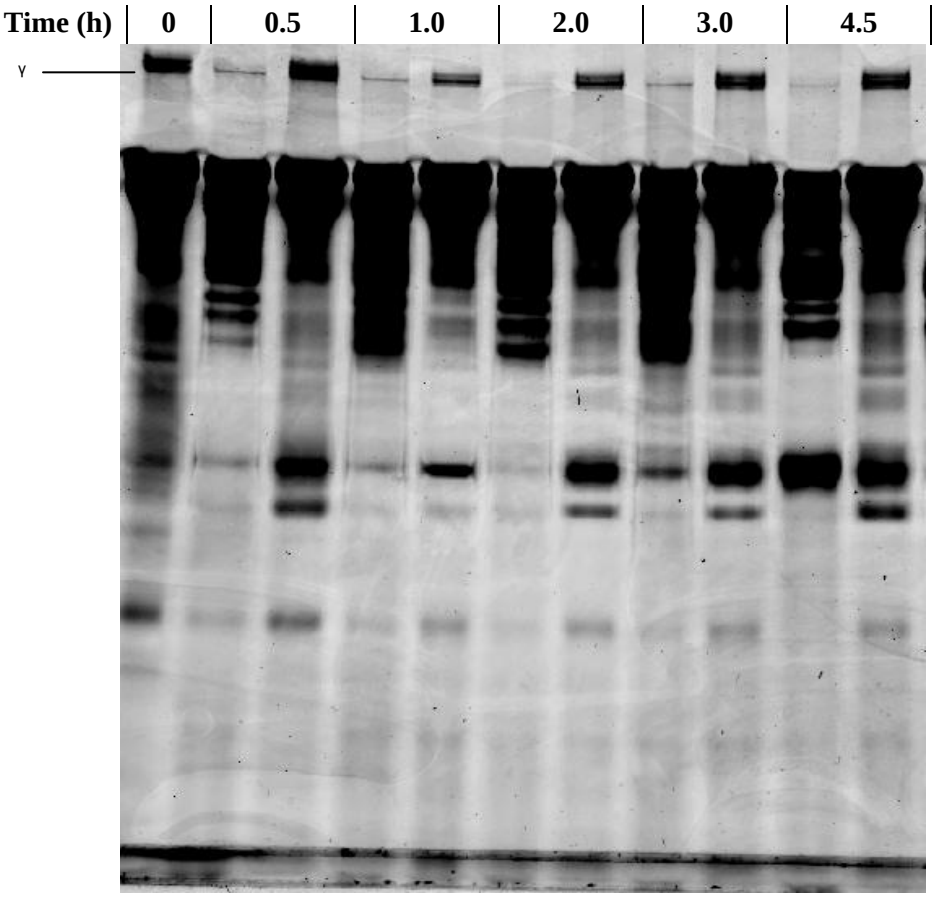
| | Concentration (M) | Slope E3 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.104 |
| | -7.301 | 0.270 |
| | -8.000 | 0.730 |
| | -9.000 | 0.931 |
| | -10.000 | 0.920 |
| | -11.000 | 0.927 |
| | | 0.969 |

| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.668 |
| HillSlope | -1.249 |
| EC50 | 2.148e-008 |
| Std. Error | |
| LogEC50 | 0.05853 |
| HillSlope | 0.1739 |
| 95% Confidence Intervals | |
| LogEC50 | -7.811 to -7.525 |
| HillSlope | -1.675 to -0.8235 |
| EC50 | 1.545e-008 to 2.988e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9888 |
| Absolute Sum of Squares | 0.01461 |
| Sy.x | 0.04934 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

B.11 Representative $IC_{50}$ Calculation Using **1** as a Cathepsin K Inhibitor (Preincubation Time: 5 minutes, [S]: 10 µM)

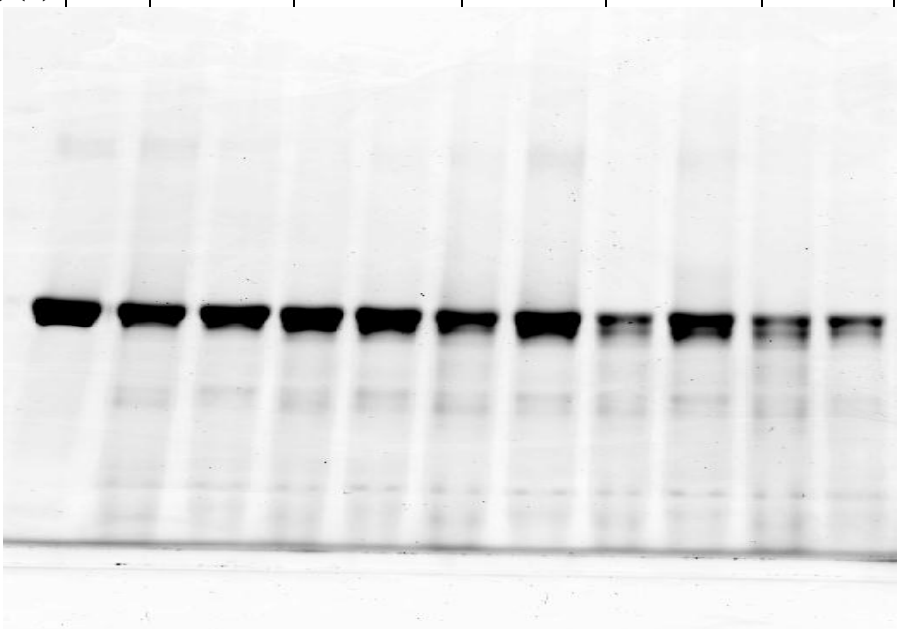



B.12 Western Blotting of the Detection of Procathepsin K. Legend: **1**: 5 ng (not detected); **2**: 10 ng; **3**: 15 ng; **4**: 20 ng; **5**: 25 ng; **6**: 30 ng; **7**: 35 ng; **8**: 40 ng; **9**: 45 ng; **10**: 55 ng; **11**: 65 ng; **12**: 100 ng; **13**: 150 ng; **14**: 200 ng; PCK: procathepsin K (mW~ 43 kDa); CK: active cathepsin K (MW ~29 kDa)




| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Collagen IV** | + | + | + | + | + | + | + | + | + | + | + |
| **Cathepsin K** | - | + | + | + | + | + | + | + | + | + | + |
| **1** | - | - | +- | - | + | - | + | - | + | - | + |
| **C4-S** | + | + | + | + | + | + | + | + | + | + | + |

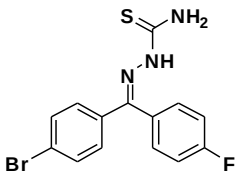
B13. Inhibition of collagenase IV activity of cathepsin K by **1**, preincubation time: 0 hours.

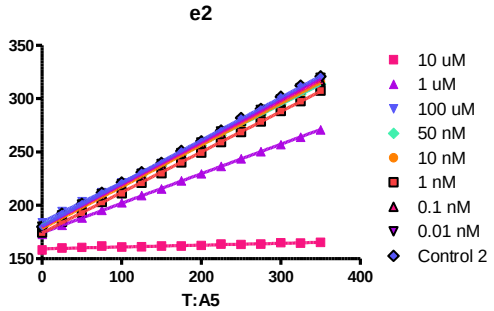


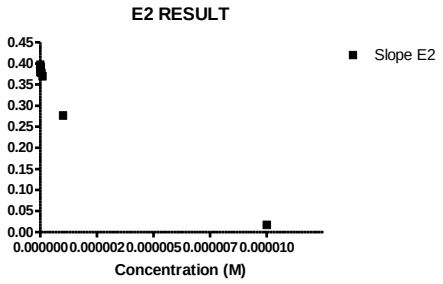

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Tubulin** | + | + | + | + | + | + | + | + | + | + | + |
| **Cathepsin K** | - | + | + | + | + | + | + | + | + | + | + |
| **1** | - | - | +- | - | + | - | + | - | + | - | + |
| **C4-S** | + | + | + | + | + | + | + | + | + | + | + |

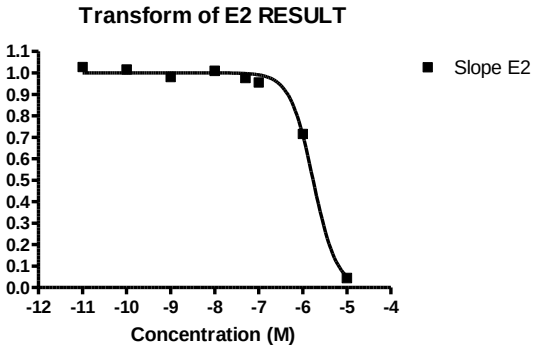
B14. Inhibition of proteolytic activity of cathepsin K by **1**, preincubation time: 0 hours

# GDK-II-33




| Concentration (M) | Slope E2 |
|---|---|
| 0.000010 | 0.017250 |
| 0.000001 | 0.276700 |
| 1.000000e-007 | 0.369400 |
| 5.000000e-008 | 0.377600 |
| 1.000000e-008 | 0.390500 |
| 1.000000e-009 | 0.379400 |
| 1.000000e-010 | 0.392800 |
| 1.000000e-011 | 0.397200 |
| 0.000000 | 0.386700 |

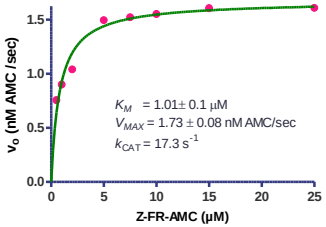
| Concentration (M) | Slope E2 |
|---|---|
| -5.000 | 0.045 |
| -6.000 | 0.716 |
| -7.000 | 0.955 |
| -7.301 | 0.976 |
| -8.000 | 1.010 |
| -9.000 | 0.981 |
| -10.000 | 1.016 |
| -11.000 | 1.027 |
| | 1.000 |

| | Slope E2 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -5.762 |
| HillSlope | -1.649 |
| EC50 | 1.730e-006 |
| Std. Error | |
| LogEC50 | 0.03681 |
| HillSlope | 0.2037 |
| 95% Confidence Intervals | |
| LogEC50 | -5.852 to -5.672 |
| HillSlope | -2.148 to -1.151 |
| EC50 | 1.406e-006 to 2.129e-006 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9959 |
| Absolute Sum of Squares | 0.003222 |
| Sy.x | 0.02317 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

B.15 Representative $IC_{50}$ Calculation Using **48** as a Cathepsin K Inhibitor

APPENDIX C

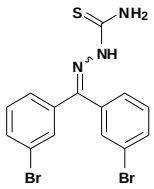
*In Vitro* Evaluations of Thiosemicarbazones as Inhibitors of Cruzain

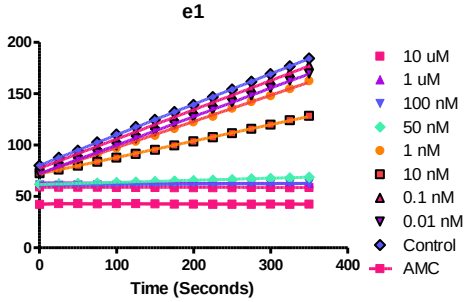


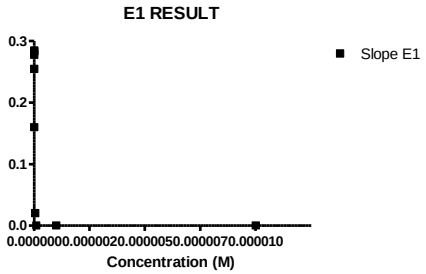

| Substrate Concentration (uM) | 2 A |
|---|---|
| 0.50000 | 0.756 |
| 1.00000 | 0.902 |
| 2.00000 | 1.041 |
| 5.00000 | 1.496 |
| 7.50000 | 1.523 |
| 10.00000 | 1.553 |
| 15.00000 | 1.606 |
| 25.00000 | 1.610 |

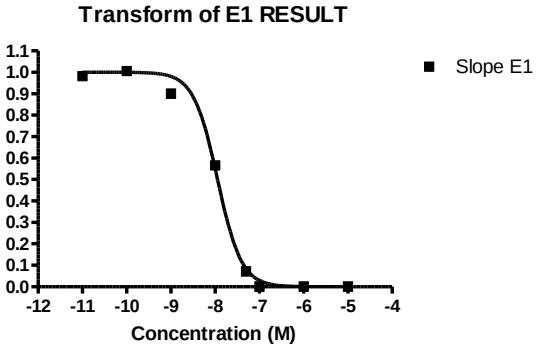
| | 2 A |
|---|---|
| Michaelis-Menten | |
| Best-fit values | |
| Vmax | 1.672 |
| Km | 0.8077 |
| Std. Error | |
| Vmax | 0.05042 |
| Km | 0.1279 |
| 95% Confidence Intervals | |
| Vmax | 1.549 to 1.795 |
| Km | 0.4947 to 1.121 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9530 |
| Absolute Sum of Squares | 0.04052 |
| Sy.x | 0.08218 |
| Constraints | |
| Km | Km > 0.0 |
| Number of points | |
| Analyzed | 8 |

C.1 Calculation of Michaelis-Menten Constant for Cruzain using Z-FR-AMC as a Substrate



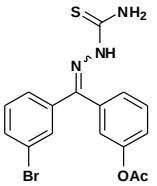



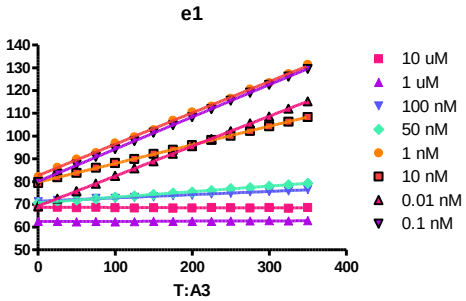

| | |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.000000 |
| 5.000000e-008 | 0.020070 |
| 1.000000e-008 | 0.160000 |
| 1.000000e-009 | 0.254600 |
| 1.000000e-010 | 0.284600 |
| 1.000000e-011 | 0.277700 |
| 0.000000 | 0.283100 |

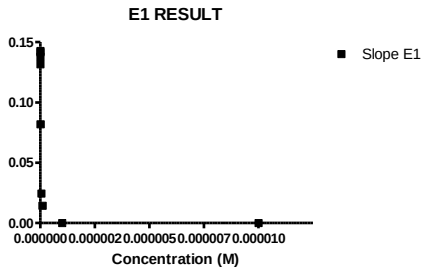
| | Concentration (M) | Slope E1 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.000 |
| | -7.301 | 0.071 |
| | -8.000 | 0.565 |
| | -9.000 | 0.899 |
| | -10.000 | 1.005 |
| | -11.000 | 0.981 |
| | | 1.000 |

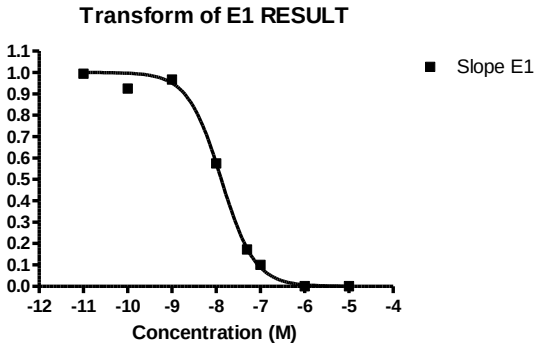
| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.944 |
| HillSlope | -1.611 |
| EC50 | 1.137e-008 |
| Std. Error | |
| LogEC50 | 0.03819 |
| HillSlope | 0.2613 |
| 95% Confidence Intervals | |
| LogEC50 | -8.038 to -7.851 |
| HillSlope | -2.250 to -0.9711 |
| EC50 | 9.165e-009 to 1.409e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9947 |
| Absolute Sum of Squares | 0.008201 |
| Sy.x | 0.03697 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.2 Representative $IC_{50}$ Calculation Using **1** as a Cruzain Inhibitor



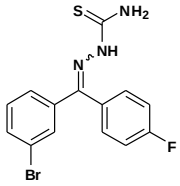



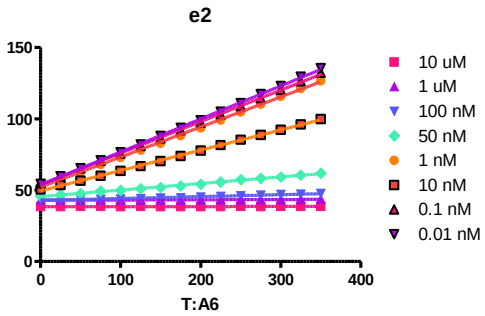

| | |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.014260 |
| 5.000000e-008 | 0.024420 |
| 1.000000e-008 | 0.081800 |
| 1.000000e-009 | 0.137700 |
| 1.000000e-010 | 0.131600 |
| 1.000000e-011 | 0.141500 |
| 0.000000 | 0.142400 |

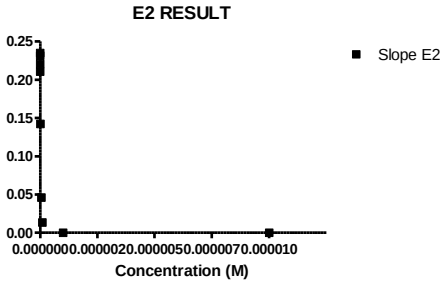
| | Concentration (M) | Slope E1 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.100 |
| | -7.301 | 0.171 |
| | -8.000 | 0.574 |
| | -9.000 | 0.967 |
| | -10.000 | 0.924 |
| | -11.000 | 0.994 |
| | | 1.000 |

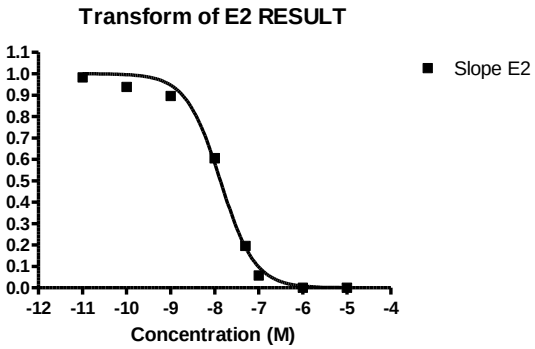
| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.883 |
| HillSlope | -1.146 |
| EC50 | 1.310e-008 |
| Std. Error | |
| LogEC50 | 0.04018 |
| HillSlope | 0.1086 |
| 95% Confidence Intervals | |
| LogEC50 | -7.981 to -7.784 |
| HillSlope | -1.412 to -0.8806 |
| EC50 | 1.045e-008 to 1.643e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9959 |
| Absolute Sum of Squares | 0.005736 |
| Sy.x | 0.03092 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.3 Representative $IC_{50}$ Calculation using **9** as a Cruzain Inhibitor



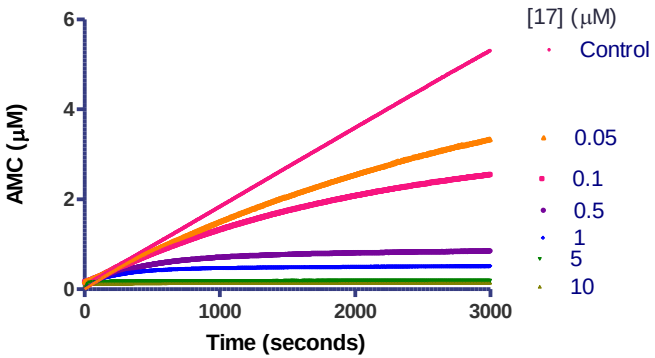



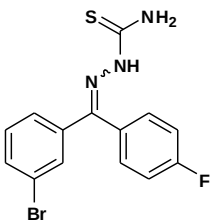

| Concentration (M) | Slope E2 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.013450 |
| 5.000000e-008 | 0.046040 |
| 1.000000e-008 | 0.142400 |
| 1.000000e-009 | 0.210800 |
| 1.000000e-010 | 0.220600 |
| 1.000000e-011 | 0.231300 |
| 0.000000 | 0.235200 |

| | Concentration (M) | Slope E2 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.057 |
| | -7.301 | 0.196 |
| | -8.000 | 0.605 |
| | -9.000 | 0.896 |
| | -10.000 | 0.938 |
| | -11.000 | 0.983 |
| | | 1.000 |

| | Slope E2 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.869 |
| HillSlope | -1.110 |
| EC50 | 1.351e-008 |
| Std. Error | |
| LogEC50 | 0.04964 |
| HillSlope | 0.1253 |
| 95% Confidence Intervals | |
| LogEC50 | -7.991 to -7.748 |
| HillSlope | -1.417 to -0.8037 |
| EC50 | 1.022e-008 to 1.787e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9938 |
| Absolute Sum of Squares | 0.008485 |
| Sy.x | 0.03761 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

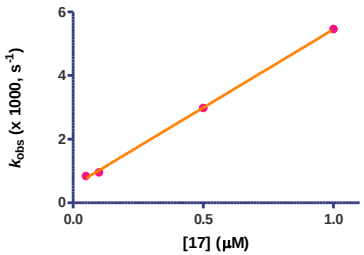
C.4 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor




| | Control | Control | 10 uM | 5 uM | 1 uM | 0.5 uM | 0.1 uM | 0.05 uM |
|---|---|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | | | | | | |
| Best-fit values | | | | | | | | |
| vs | 0.001741 | | 4.533e-006 | 7.658e-006 | 2.832e-005 | 5.803e-005 | 0.0003832 | 0.0007004 |
| vi | ~ 0.7972 | | 0.04930 | 0.01915 | 0.002535 | 0.002149 | 0.001920 | 0.001950 |
| kobs | ~ 7.790 | | 0.3660 | 0.1090 | 0.005666 | 0.003036 | 0.001029 | 0.0009268 |
| Std. Error | | | | | | | | |
| vs | 4.190e-007 | | 1.773e-007 | 2.679e-007 | 9.089e-007 | 1.643e-006 | 9.676e-006 | 1.208e-005 |
| vi | ~ 4.648e+007 | | 0.003839 | 0.0009045 | 2.909e-005 | 1.646e-005 | 9.352e-006 | 9.232e-006 |
| kobs | ~ 4.552e+008 | | 0.02857 | 0.005192 | 7.804e-005 | 3.429e-005 | 2.166e-005 | 2.695e-005 |
| 95% Confidence Intervals | | | | | | | | |
| vs | 0.001740 to 0.001742 | | 4.185e-006 to 4.880e-006 | 7.133e-006 to 8.183e-006 | 2.654e-005 to 3.010e-005 | 5.481e-005 to 6.125e-005 | 0.0003642 to 0.0004022 | 0.0006768 to 0.000724 |
| vi | (Very wide) | | 0.04178 to 0.05683 | 0.01738 to 0.02092 | 0.002478 to 0.002592 | 0.002117 to 0.002181 | 0.001902 to 0.001938 | 0.001932 to 0.001968 |
| kobs | (Very wide) | | 0.3100 to 0.4220 | 0.09886 to 0.1192 | 0.005513 to 0.005819 | 0.002968 to 0.003103 | 0.0009870 to 0.001072 | 0.0008740 to 0.000979 |
| Goodness of Fit | | | | | | | | |
| Degrees of Freedom | 997 | | 997 | 997 | 997 | 997 | 997 | 997 |
| R square | 0.9999 | | 0.2366 | 0.4979 | 0.9524 | 0.9848 | 0.9982 | 0.9989 |
| Absolute Sum of Squares | 0.1309 | | 0.02324 | 0.05176 | 0.2932 | 0.4443 | 0.8554 | 0.9273 |
| Sy.x | 0.01146 | | 0.004828 | 0.007205 | 0.01715 | 0.02111 | 0.02929 | 0.03050 |
| Constraints | | | | | | | | |
| vs | vs > 0.0 | | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | | | |
| Analyzed | 1000 | | 1000 | 1000 | 1000 | 1000 | 1000 | 1000 |

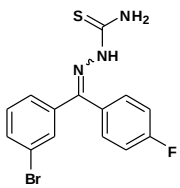
C.5 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [S]: 15 μM (Preincubation Time: 0 minutes)

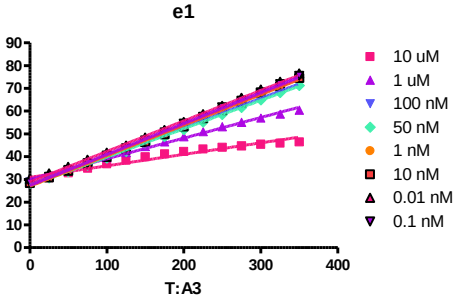


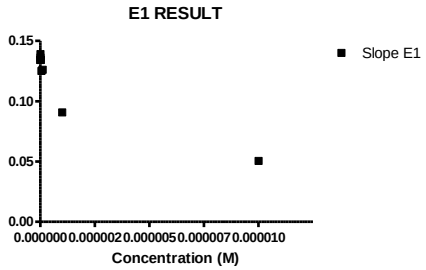

| [I] | kobs | | |
|---|---|---|---|
| | Mean | SEM | N |
| 1.000 | 5.461 | 0.408 | 3 |
| 0.500 | 2.984 | 0.103 | 3 |
| 0.100 | 0.960 | 0.040 | 3 |
| 0.050 | 0.843 | 0.054 | 3 |

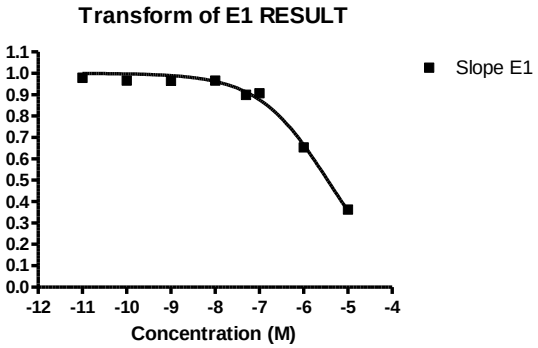
| | kobs |
|---|---|
| Best-fit values | |
| Slope | 4.922 ± 0.08528 |
| Y-intercept when X=0.0 | 0.5318 ± 0.04791 |
| X-intercept when Y=0.0 | -0.1080 |
| 1/slope | 0.2032 |
| 95% Confidence Intervals | |
| Slope | 4.555 to 5.289 |
| Y-intercept when X=0.0 | 0.3256 to 0.7379 |
| X-intercept when Y=0.0 | -0.1593 to -0.06262 |
| Goodness of Fit | |
| R square | 0.9994 |
| Sy.x | 0.06505 |
| Is slope significantly non-zero? | |
| F | 3331 |
| DFn, DFd | 1.000, 2.000 |
| P value | 0.0003 |
| Deviation from zero? | Significant |
| Data | |
| Number of X values | 4 |
| Maximum number of Y replicates | 1 |
| Total number of values | 4 |
| Number of missing values | 0 |

C.6 Plot of $k_{obs}$ vs [**17**] Using Linear Regression. (Preincubation Time: 0 minutes)



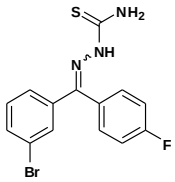



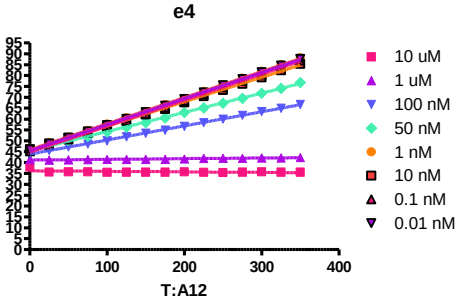

| | |
|---|---|
| 0.000010 | 0.050430 |
| 0.000001 | 0.090750 |
| 1.000000e-007 | 0.126000 |
| 5.000000e-008 | 0.124900 |
| 1.000000e-008 | 0.134200 |
| 1.000000e-009 | 0.134100 |
| 1.000000e-010 | 0.134200 |
| 1.000000e-011 | 0.136000 |
| 0.000000 | 0.139000 |

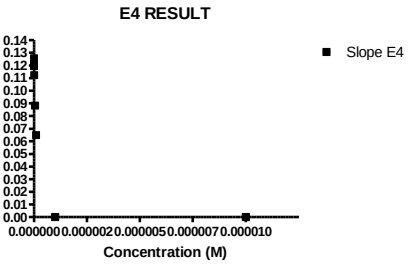
| | Concentration (M) | Slope E1 |
|---|---|---|
| | -5.000 | 0.363 |
| | -6.000 | 0.653 |
| | -7.000 | 0.906 |
| | -7.301 | 0.899 |
| | -8.000 | 0.965 |
| | -9.000 | 0.965 |
| | -10.000 | 0.965 |
| | -11.000 | 0.978 |
| | | 1.000 |

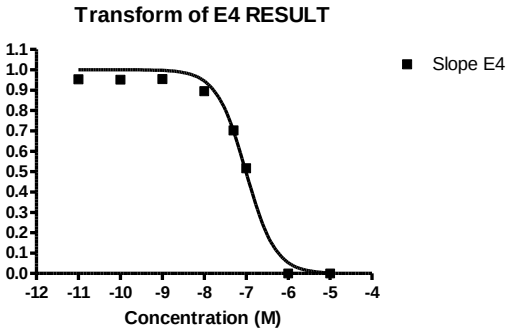
| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -5.457 |
| HillSlope | -0.5508 |
| EC50 | 3.490e-006 |
| Std. Error | |
| LogEC50 | 0.05655 |
| HillSlope | 0.03711 |
| 95% Confidence Intervals | |
| LogEC50 | -5.595 to -5.319 |
| HillSlope | -0.6416 to -0.4600 |
| EC50 | 2.538e-006 to 4.800e-006 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9902 |
| Absolute Sum of Squares | 0.003289 |
| Sy.x | 0.02341 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.7 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor (Preincubation Time: 0 minutes)



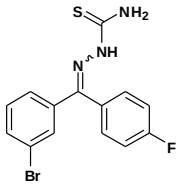



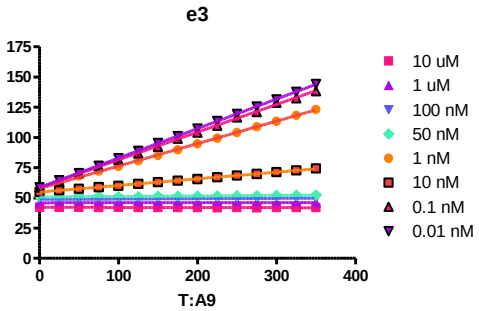

| Concentration (M) | Slope E4 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.064850 |
| 5.000000e-008 | 0.088190 |
| 1.000000e-008 | 0.112300 |
| 1.000000e-009 | 0.119800 |
| 1.000000e-010 | 0.119500 |
| 1.000000e-011 | 0.119700 |
| 0.000000 | 0.125600 |

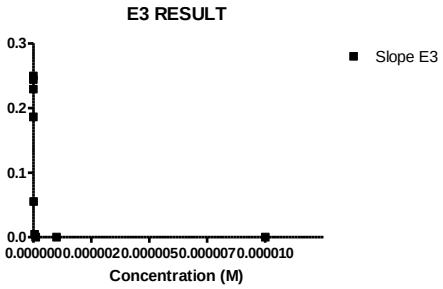
| | Concentration (M) | Slope E4 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.516 |
| | -7.301 | 0.702 |
| | -8.000 | 0.894 |
| | -9.000 | 0.954 |
| | -10.000 | 0.951 |
| | -11.000 | 0.953 |
| | | 1.000 |

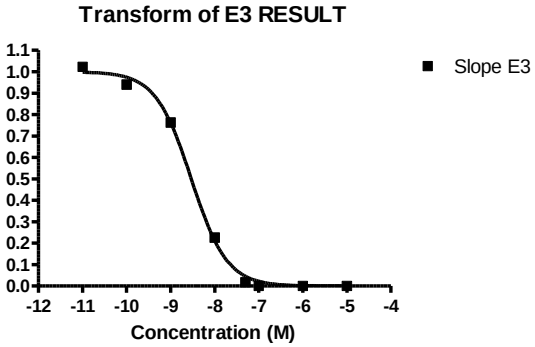
| | Slope E4 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.005 |
| HillSlope | -1.238 |
| EC50 | 9.887e-008 |
| Std. Error | |
| LogEC50 | 0.05165 |
| HillSlope | 0.2217 |
| 95% Confidence Intervals | |
| LogEC50 | -7.131 to -6.879 |
| HillSlope | -1.780 to -0.6950 |
| EC50 | 7.390e-008 to 1.323e-007 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9897 |
| Absolute Sum of Squares | 0.01225 |
| Sy.x | 0.04518 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.8 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor (Preincubation Time: 1 minute)



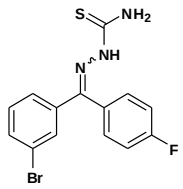



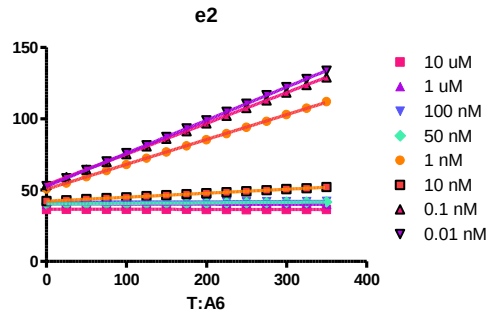

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.000000 |
| 5.000000e-008 | 0.004460 |
| 1.000000e-008 | 0.055080 |
| 1.000000e-009 | 0.186100 |
| 1.000000e-010 | 0.229200 |
| 1.000000e-011 | 0.249400 |
| 0.000000 | 0.243800 |

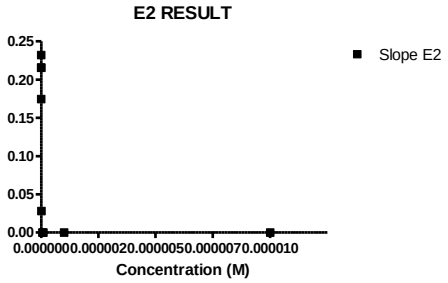
| | Concentration (M) | Slope E3 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.000 |
| | -7.301 | 0.018 |
| | -8.000 | 0.226 |
| | -9.000 | 0.763 |
| | -10.000 | 0.940 |
| | -11.000 | 1.023 |
| | | 1.000 |

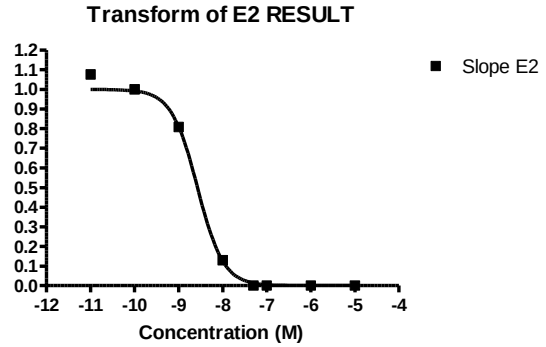
| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -8.530 |
| HillSlope | -1.075 |
| EC50 | 2.955e-009 |
| Std. Error | |
| LogEC50 | 0.03718 |
| HillSlope | 0.07053 |
| 95% Confidence Intervals | |
| LogEC50 | -8.620 to -8.439 |
| HillSlope | -1.247 to -0.9021 |
| EC50 | 2.396e-009 to 3.643e-009 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9978 |
| Absolute Sum of Squares | 0.003238 |
| Sy.x | 0.02323 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.9 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor (Preincubation Time: 30 minutes



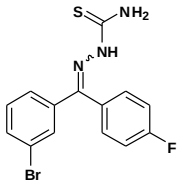



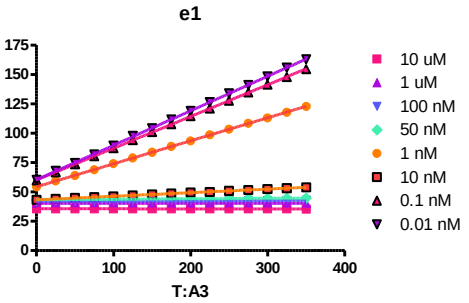

| Concentration (M) | Slope E2 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.000000 |
| 5.000000e-008 | 0.000000 |
| 1.000000e-008 | 0.028020 |
| 1.000000e-009 | 0.174400 |
| 1.000000e-010 | 0.215700 |
| 1.000000e-011 | 0.232200 |
| 0.000000 | 0.215600 |

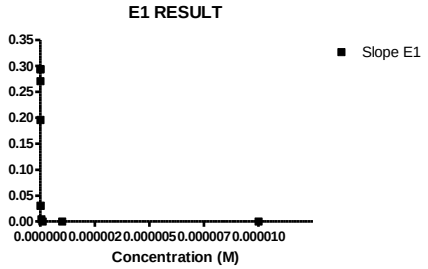
| | Concentration (M) | Slope E2 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.000 |
| | -7.301 | 0.000 |
| | -8.000 | 0.130 |
| | -9.000 | 0.809 |
| | -10.000 | 1.000 |
| | -11.000 | 1.077 |
| | | 1.000 |

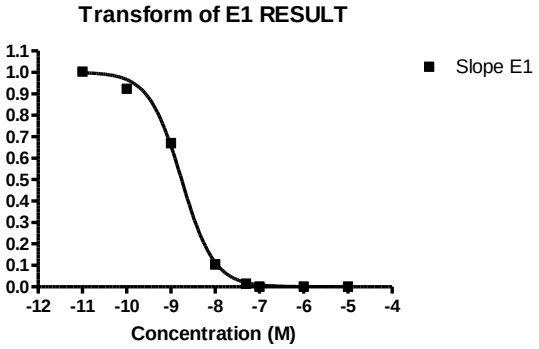
| | Slope E2 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -8.571 |
| HillSlope | -1.474 |
| EC50 | 2.687e-009 |
| Std. Error | |
| LogEC50 | 0.05098 |
| HillSlope | 0.1523 |
| 95% Confidence Intervals | |
| LogEC50 | -8.696 to -8.446 |
| HillSlope | -1.847 to -1.101 |
| EC50 | 2.016e-009 to 3.581e-009 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9963 |
| Absolute Sum of Squares | 0.006256 |
| Sy.x | 0.03229 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.10 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor (Preincubation Time: 60 minutes)



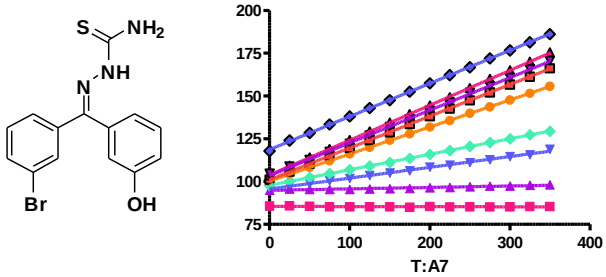

| | |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.000000 |
| 1.000000e-007 | 0.000000 |
| 5.000000e-008 | 0.004140 |
| 1.000000e-008 | 0.030410 |
| 1.000000e-009 | 0.196000 |
| 1.000000e-010 | 0.270400 |
| 1.000000e-011 | 0.293800 |
| 0.000000 | 0.292900 |

| | Concentration (M) | Slope E1 |
|---|---|---|
| | -5.000 | 0.000 |
| | -6.000 | 0.000 |
| | -7.000 | 0.000 |
| | -7.301 | 0.014 |
| | -8.000 | 0.104 |
| | -9.000 | 0.669 |
| | -10.000 | 0.923 |
| | -11.000 | 1.003 |
| | | 1.000 |

| | Slope E1 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -8.758 |
| HillSlope | -1.172 |
| EC50 | 1.744e-009 |
| Std. Error | |
| LogEC50 | 0.02908 |
| HillSlope | 0.07936 |
| 95% Confidence Intervals | |
| LogEC50 | -8.830 to -8.687 |
| HillSlope | -1.366 to -0.9779 |
| EC50 | 1.481e-009 to 2.055e-009 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9984 |
| Absolute Sum of Squares | 0.002223 |
| Sy.x | 0.01925 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

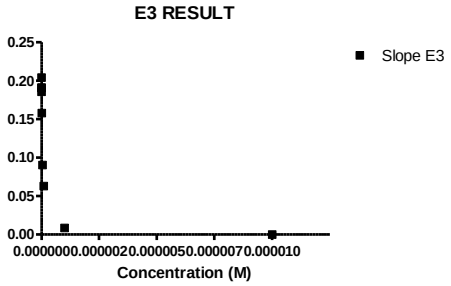
C.11 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor (Preincubation Time: 120 minutes)

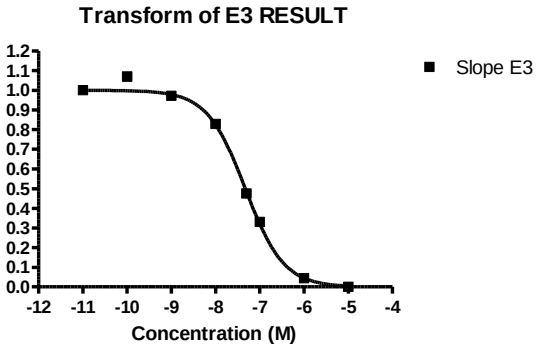
# GDK-II-98 4a









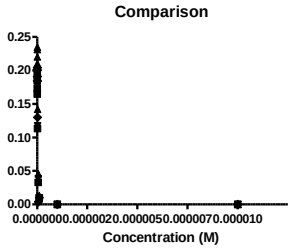



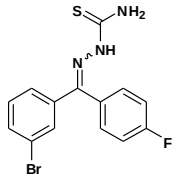

| Concentration (M) | Slope E3 |
|---|---|
| 0.000010 | 0.000000 |
| 0.000001 | 0.008498 |
| 1.000000e-007 | 0.063020 |
| 5.000000e-008 | 0.090610 |
| 1.000000e-008 | 0.158100 |
| 1.000000e-009 | 0.185500 |
| 1.000000e-010 | 0.204100 |
| 1.000000e-011 | 0.191000 |
| 0.000000 | 0.190700 |

| Concentration (M) | Slope E3 |
|---|---|
| -5.000 | 0.000 |
| -6.000 | 0.045 |
| -7.000 | 0.330 |
| -7.301 | 0.475 |
| -8.000 | 0.829 |
| -9.000 | 0.973 |
| -10.000 | 1.070 |
| -11.000 | 1.002 |
| | 1.000 |

| | Slope E3 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -7.326 |
| HillSlope | -1.000 |
| EC50 | 4.722e-008 |
| Std. Error | |
| LogEC50 | 0.03631 |
| HillSlope | 0.09587 |
| 95% Confidence Intervals | |
| LogEC50 | -7.415 to -7.237 |
| HillSlope | -1.235 to -0.7657 |
| EC50 | 3.849e-008 to 5.794e-008 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9958 |
| Absolute Sum of Squares | 0.005528 |
| Sy.x | 0.03035 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.12 Representative $IC_{50}$ Calculation Using **17** as a Cruzain Inhibitor (Preincubation Time: 240 minutes)

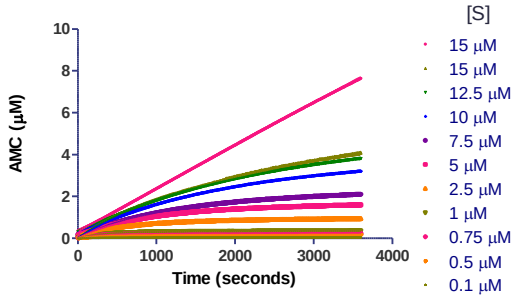


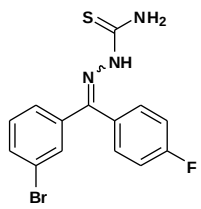




| Concentration (M) | Slope E1 | Slope E2 | Slope E3 | Slope E4 |
|---|---|---|---|---|
| 10.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 1.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 0.100 | 0.048 | 0.057 | 0.052 | 0.051 |
| 0.050 | 0.166 | 0.196 | 0.207 | 0.020 |
| 0.010 | 0.572 | 0.605 | 0.634 | 0.638 |
| 0.001 | 0.832 | 0.896 | 0.962 | 0.904 |
| 1.000e-004 | 0.873 | 0.938 | 0.938 | 0.937 |
| 1.000e-005 | 0.932 | 0.983 | 0.993 | 0.983 |
| 0.000 | 1.000 | 1.000 | 1.000 | 1.000 |

| | Slope E1 | Slope E2 | Slope E3 | Slope E4 |
|---|---|---|---|---|
| Morrison Ki | | | | |
| Best-fit values | | | | |
| Vo | = 1.000 | = 1.000 | = 1.000 | = 1.000 |
| Et | = 1.000e-004 | = 1.000e-004 | = 1.000e-004 | = 1.000e-004 |
| Ki | 0.0006741 | 0.0008138 | 0.0009111 | 0.0006836 |
| S | = 15.00 | = 15.00 | = 15.00 | = 15.00 |
| Km | = 1.010 | = 1.010 | = 1.010 | = 1.010 |
| Std. Error | | | | |
| Ki | 0.0001346 | 8.783e-005 | 0.0001069 | 0.0001633 |
| 95% Confidence Intervals | | | | |
| Ki | 0.0003637 to 0.0009846 | 0.0006113 to 0.001016 | 0.0006646 to 0.001158 | 0.0003070 to 0.001060 |
| Goodness of Fit | | | | |
| Degrees of Freedom | 8 | 8 | 8 | 8 |
| R square | 0.9795 | 0.9942 | 0.9932 | 0.9752 |
| Absolute Sum of Squares | 0.03090 | 0.009465 | 0.01149 | 0.04438 |
| Sy.x | 0.06215 | 0.03440 | 0.03790 | 0.07449 |
| Constraints | | | | |
| Vo | Vo = 1.000 | Vo = 1.000 | Vo = 1.000 | Vo = 1.000 |
| Et | Et = 1.000e-004 | Et = 1.000e-004 | Et = 1.000e-004 | Et = 1.000e-004 |
| S | S = 15.00 | S = 15.00 | S = 15.00 | S = 15.00 |
| Km | Km = 1.010 | Km = 1.010 | Km = 1.010 | Km = 1.010 |
| Number of points | | | | |
| Analyzed | 9 | 9 | 9 | 9 |

C.13 Calculation of $K_I^{app}$ Using **17** as a Cruzain Inhibitor (Preincubation Time: 5 minutes

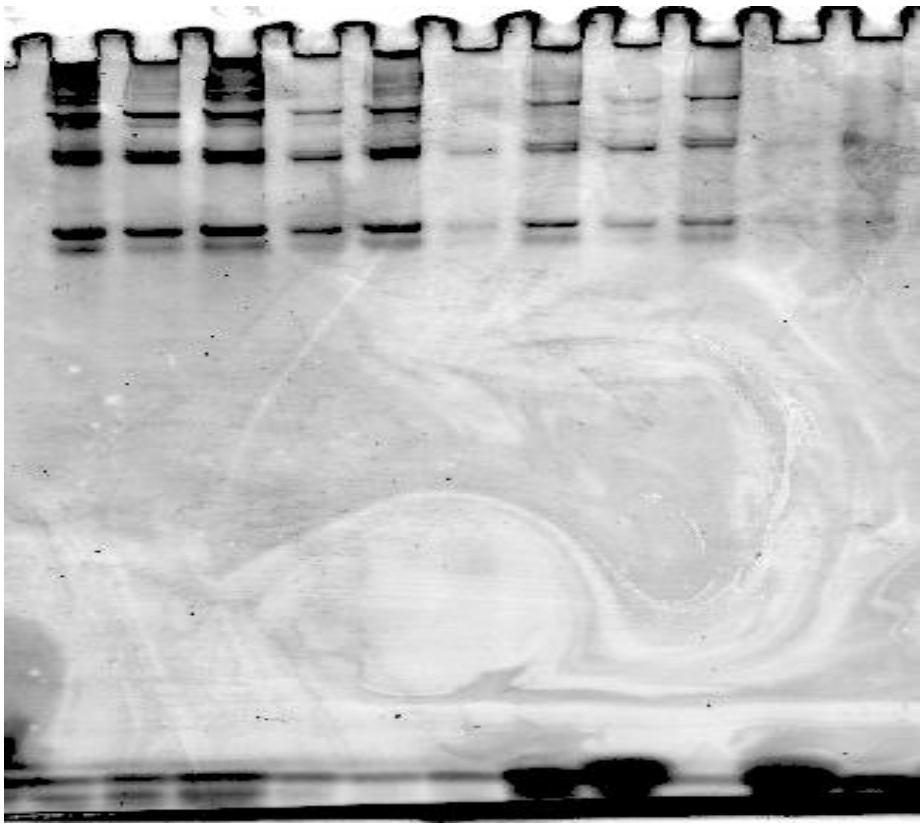



**Y=vs*x+((vi-vs)/kobs)*(1-exp(-kobs*x))**

| | 15 μM | 15 μM | 15 μM | 12.5 μM | 10 μM | 7.5 μM | 5 μM | 2.5 μM | 1 μM | 0.75 μM | 0.5 μM | 0.1 μM |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Progress Curve | Ambiguous | | | | | | | | | | | |
| Best-fit values | | | | | | | | | | | | |
| vs | 0.002052 | | 0.0006798 | 0.0005365 | 0.0003764 | 0.0001708 | 0.0001053 | 5.128e-005 | 1.554e-005 | 1.083e-005 | 8.102e-006 | 2.082e-006 |
| vi | ~ 3.240 | | 0.002796 | 0.002738 | 0.002475 | 0.001988 | 0.001956 | 0.001607 | 0.001021 | 0.0007939 | 0.0006861 | 0.0002676 |
| kobs | ~ 9.855 | | 0.001235 | 0.001114 | 0.001085 | 0.001190 | 0.001494 | 0.002021 | 0.003379 | 0.003859 | 0.004712 | 0.006389 |
| Std. Error | | | | | | | | | | | | |
| vs | 1.823e-006 | | 1.223e-005 | 1.243e-005 | 1.032e-005 | 5.766e-006 | 3.248e-006 | 1.281e-006 | 3.312e-007 | 2.292e-007 | 1.637e-007 | 5.943e-008 |
| vi | ~ 7.775e+008 | | 2.676e-005 | 2.159e-005 | 1.690e-005 | 1.160e-005 | 1.081e-005 | 8.138e-006 | 5.808e-006 | 5.132e-006 | 5.232e-006 | 3.202e-006 |
| kobs | ~ 2.366e+009 | | 3.650e-005 | 2.825e-005 | 2.320e-005 | 1.839e-005 | 1.728e-005 | 1.701e-005 | 2.516e-005 | 3.141e-005 | 4.316e-005 | 8.701e-005 |
| 95% Confidence Intervals | | | | | | | | | | | | |
| vs | 0.002049 to 0.002056 | | 0.0006559 to 0.0007038 | 0.0005121 to 0.0005608 | 0.0003562 to 0.0003966 | 0.0001595 to 0.0001821 | 9.891e-005 to 0.0001116 | 4.877e-005 to 5.379e-005 | 1.489e-005 to 1.619e-005 | 1.038e-005 to 1.127e-005 | 7.781e-006 to 8.423e-006 | 1.966e-006 to 2.199e-006 |
| vi | (Very wide) | | 0.002744 to 0.002849 | 0.002696 to 0.002780 | 0.002442 to 0.002508 | 0.001966 to 0.002011 | 0.001934 to 0.001977 | 0.001591 to 0.001623 | 0.001009 to 0.001032 | 0.0007838 to 0.0008039 | 0.0006758 to 0.0006963 | 0.0002613 to 0.0002738 |
| kobs | (Very wide) | | 0.001164 to 0.001307 | 0.001059 to 0.001169 | 0.001039 to 0.001130 | 0.001154 to 0.001226 | 0.001460 to 0.001528 | 0.001988 to 0.002054 | 0.003330 to 0.003429 | 0.003798 to 0.003921 | 0.004628 to 0.004797 | 0.006219 to 0.006560 |
| Goodness of Fit | | | | | | | | | | | | |
| Degrees of Freedom | 718 | | 718 | 718 | 718 | 718 | 718 | 718 | 718 | 718 | 718 | 718 |
| R square | 0.9998 | | 0.9960 | 0.9966 | 0.9969 | 0.9968 | 0.9962 | 0.9958 | 0.9930 | 0.9906 | 0.9867 | 0.9625 |
| Absolute Sum of Squares | 0.5637 | | 3.326 | 2.499 | 1.585 | 0.6594 | 0.4037 | 0.1312 | 0.02077 | 0.01158 | 0.007120 | 0.001161 |
| Sy.x | 0.02802 | | 0.06806 | 0.05899 | 0.04698 | 0.03031 | 0.02371 | 0.01352 | 0.005378 | 0.004016 | 0.003149 | 0.001271 |
| Constraints | | | | | | | | | | | | |
| vs | vs > 0.0 | | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 | vs > 0.0 |
| vi | vi > 0.0 | | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 | vi > 0.0 |
| kobs | kobs > 0.0 | | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 | kobs > 0.0 |
| Number of points | | | | | | | | | | | | |
| Analyzed | 721 | | 721 | 721 | 721 | 721 | 721 | 721 | 721 | 721 | 721 | 721 |

C.14 Calculation of $k_{obs}$, $v_s$, and $v_o$ Using Nonlinear Regression. [I]: 100 nM (Preincubation Time: 0 minutes)




| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Collagen | + | + | + | + | + | + | + | + | + | + | + |
| Cruzain | - | + | + | + | + | + | + | + | + | + | + |
| **17** | - | - | + | - | + | - | + | - | + | - | + |

C15. Inhibition of collagenase activity of Cruzain by **17**, preincubation time: 0.5 hours.

E1

| | 10 uM | Control | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.2077 ± 0.0004500 | 0.2970 ± 0.0007883 | -0.0009435 ± 0.0005039 |
| Y-intercept when X=0.0 | 66.88 ± 0.09252 | 78.92 ± 0.1621 | 42.82 ± 0.1036 |
| X-intercept when Y=0.0 | -322.0 | -265.8 | 45380 |
| 1/slope | 4.814 | 3.367 | -1060 |
| 95% Confidence Intervals | | | |
| Slope | 0.2067 to 0.2087 | 0.2953 to 0.2987 | -0.002032 to 0.0001448 |
| Y-intercept when X=0.0 | 66.68 to 67.07 | 78.57 to 79.27 | 42.59 to 43.04 |
| X-intercept when Y=0.0 | -324.3 to -319.6 | -268.4 to -263.2 | 21166 to +infinity |
| Goodness of Fit | | | |
| R square | 0.9999 | 0.9999 | 0.2124 |
| Sy.x | 0.1882 | 0.3298 | 0.2108 |
| Is slope significantly non-zero? | | | |
| F | 213079 | 141914 | 3.506 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | 0.0838 |
| Deviation from zero? | Significant | Significant | Not Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

E2

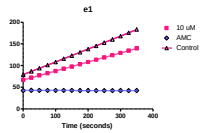
| | 10 uM | Control | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1939 ± 0.0004041 | 0.2824 ± 0.0007757 | -0.002004 ± 0.0003735 |
| Y-intercept when X=0.0 | 62.82 ± 0.08308 | 75.55 ± 0.1595 | 44.21 ± 0.07681 |
| X-intercept when Y=0.0 | -324.0 | -267.5 | 22058 |
| 1/slope | 5.158 | 3.541 | -498.9 |
| 95% Confidence Intervals | | | |
| Slope | 0.1930 to 0.1948 | 0.2807 to 0.2841 | -0.002811 to -0.001198 |
| Y-intercept when X=0.0 | 62.64 to 63.00 | 75.21 to 75.90 | 44.05 to 44.38 |
| X-intercept when Y=0.0 | -326.3 to -321.7 | -270.3 to -264.9 | 15777 to 36801 |
| Goodness of Fit | | | |
| R square | 0.9999 | 0.9999 | 0.6889 |
| Sy.x | 0.1690 | 0.3245 | 0.1563 |
| Is slope significantly non-zero? | | | |
| F | 230251 | 132532 | 28.79 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | 0.0001 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

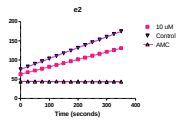
E3

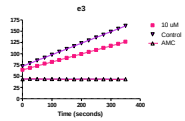
| | 10 uM | Control | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.1778 ± 0.0006521 | 0.2556 ± 0.0004939 | -0.002033 ± 0.0005734 |
| Y-intercept when X=0.0 | 63.95 ± 0.1341 | 72.25 ± 0.1015 | 43.91 ± 0.1179 |
| X-intercept when Y=0.0 | -359.7 | -282.7 | 21603 |
| 1/slope | 5.624 | 3.913 | -492.0 |
| 95% Confidence Intervals | | | |
| Slope | 0.1764 to 0.1792 | 0.2545 to 0.2567 | -0.003271 to -0.0007940 |
| Y-intercept when X=0.0 | 63.66 to 64.24 | 72.03 to 72.47 | 43.66 to 44.17 |
| X-intercept when Y=0.0 | -364.0 to -355.4 | -284.7 to -280.7 | 13490 to 55028 |
| Goodness of Fit | | | |
| R square | 0.9998 | 1.000 | 0.4915 |
| Sy.x | 0.2728 | 0.2066 | 0.2399 |
| Is slope significantly non-zero? | | | |
| F | 74337 | 267822 | 12.57 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | 0.0036 |
| Deviation from zero? | Significant | Significant | Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |

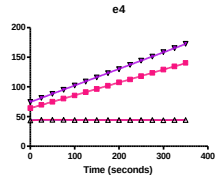
E4

| | 10 uM | Control | AMC |
|---|---|---|---|
| Best-fit values | | | |
| Slope | 0.2169 ± 0.0005517 | 0.2794 ± 0.0006636 | -0.0006939 ± 0.0006337 |
| Y-intercept when X=0.0 | 64.03 ± 0.1134 | 74.38 ± 0.1364 | 44.31 ± 0.1303 |
| X-intercept when Y=0.0 | -295.2 | -266.2 | 63857 |
| 1/slope | 4.611 | 3.579 | -1441 |
| 95% Confidence Intervals | | | |
| Slope | 0.2157 to 0.2181 | 0.2780 to 0.2809 | -0.002063 to 0.0006750 |
| Y-intercept when X=0.0 | 63.79 to 64.28 | 74.09 to 74.68 | 44.03 to 44.59 |
| X-intercept when Y=0.0 | -297.9 to -292.6 | -268.5 to -263.9 | 21596 to +infinity |
| Goodness of Fit | | | |
| R square | 0.9999 | 0.9999 | 0.08443 |
| Sy.x | 0.2308 | 0.2776 | 0.2651 |
| Is slope significantly non-zero? | | | |
| F | 154540 | 177304 | 1.199 |
| DFn, DFd | 1.000, 13.00 | 1.000, 13.00 | 1.000, 13.00 |
| P value | < 0.0001 | < 0.0001 | 0.2934 |
| Deviation from zero? | Significant | Significant | Not Significant |
| Data | | | |
| Number of X values | 15 | 15 | 15 |
| Maximum number of Y replicates | 1 | 1 | 1 |
| Total number of values | 15 | 15 | 15 |
| Number of missing values | 0 | 0 | 0 |




E1

| Time (seconds) | 10 uM | Control | AMC |
|---|---|---|---|
| 0.00 | 67.148 | 79.385 | 42.290 |
| 24.94 | 72.179 | 86.785 | 43.066 |
| 49.95 | 77.404 | 93.897 | 42.870 |
| 74.94 | 82.522 | 101.075 | 42.813 |
| 99.94 | 87.339 | 108.271 | 42.846 |
| 124.94 | 92.704 | 115.845 | 42.993 |
| 149.94 | 97.830 | 123.057 | 42.695 |
| 174.94 | 102.992 | 130.659 | 42.654 |
| 199.94 | 108.242 | 138.004 | 42.506 |
| 224.94 | 113.593 | 145.477 | 42.477 |
| 249.94 | 118.726 | 153.105 | 42.443 |
| 274.94 | 123.992 | 160.480 | 42.444 |
| 299.94 | 129.292 | 167.936 | 42.625 |
| 324.94 | 134.348 | 175.734 | 42.403 |
| 349.95 | 139.895 | 183.446 | 42.648 |

E2

| Time (seconds) | 10 uM | Control | AMC |
|---|---|---|---|
| 0.00 | 62.879 | 75.813 | 44.003 |
| 24.97 | 68.069 | 83.028 | 44.322 |
| 49.97 | 72.526 | 90.086 | 44.131 |
| 74.97 | 77.311 | 96.568 | 44.207 |
| 99.97 | 82.031 | 103.687 | 43.806 |
| 124.98 | 86.922 | 110.455 | 44.205 |
| 149.97 | 91.766 | 117.451 | 43.903 |
| 174.97 | 96.667 | 124.559 | 43.786 |
| 199.97 | 101.556 | 131.802 | 43.921 |
| 224.97 | 106.233 | 139.039 | 43.719 |
| 249.97 | 111.217 | 146.029 | 43.642 |
| 274.98 | 116.081 | 153.131 | 43.525 |
| 299.97 | 121.193 | 160.275 | 43.477 |
| 324.97 | 125.969 | 167.709 | 43.509 |
| 349.97 | 130.790 | 174.794 | 43.759 |

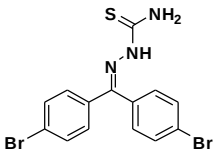
E3

| Time (seconds) | 10 uM | Control | AMC |
|---|---|---|---|
| 0.00 | 64.085 | 72.448 | 43.441 |
| 24.96 | 68.370 | 78.932 | 44.274 |
| 49.96 | 72.793 | 85.136 | 43.941 |
| 74.96 | 78.006 | 91.252 | 43.835 |
| 99.96 | 81.557 | 97.668 | 43.893 |
| 124.96 | 85.831 | 104.227 | 43.823 |
| 149.96 | 90.485 | 110.409 | 43.256 |
| 174.96 | 94.781 | 116.803 | 43.309 |
| 199.96 | 99.373 | 123.122 | 43.538 |
| 224.97 | 103.871 | 129.585 | 43.388 |
| 249.96 | 108.220 | 136.091 | 43.512 |
| 274.96 | 112.764 | 142.338 | 43.348 |
| 299.96 | 117.488 | 148.954 | 43.140 |
| 324.96 | 121.797 | 155.504 | 43.222 |
| 349.96 | 126.452 | 162.084 | 43.408 |

E4

| Time (seconds) | 10 uM | Control | AMC |
|---|---|---|---|
| 0.00 | 64.229 | 74.689 | 43.551 |
| 24.94 | 69.735 | 81.651 | 44.617 |
| 49.94 | 74.941 | 88.439 | 44.540 |
| 74.94 | 80.168 | 95.303 | 44.544 |
| 99.94 | 85.611 | 102.335 | 44.276 |
| 124.94 | 91.071 | 108.879 | 44.219 |
| 149.94 | 96.383 | 116.130 | 44.182 |
| 174.94 | 101.817 | 123.147 | 44.093 |
| 199.94 | 107.515 | 129.734 | 44.314 |
| 224.95 | 112.441 | 137.017 | 44.218 |
| 249.94 | 118.034 | 144.185 | 44.060 |
| 274.94 | 123.547 | 151.134 | 44.090 |
| 299.94 | 129.097 | 158.532 | 43.908 |
| 324.94 | 134.597 | 165.233 | 44.111 |
| 349.94 | 140.438 | 172.561 | 44.076 |

| | % Inhibition |
|---|---|
| Number of values | 3 |
| Mean | 30.63 |
| Std. Deviation | 0.6807 |
| Std. Error | 0.3930 |
| Sum | 91.90 |




GDK-II-38
KGP96

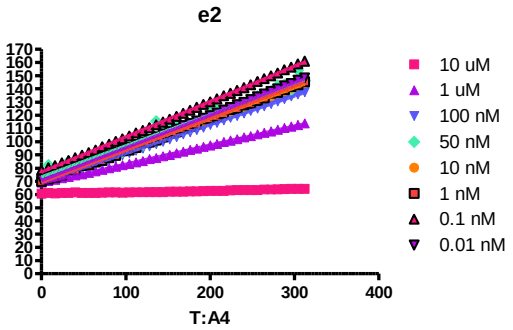
| | E1 | E2 | E3 | E4 |
|---|---|---|---|---|
| Control | 0.000 | 0.000 | 0.000 | 0.000 |
| 10 uM | 0.301 | 0.314 | 0.304 | 0.224 |

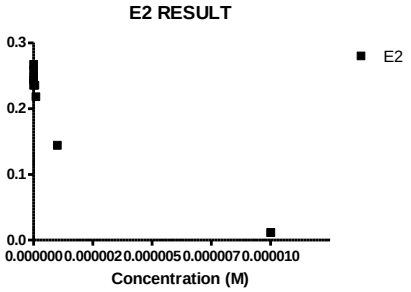
FINAL CONDITIONS

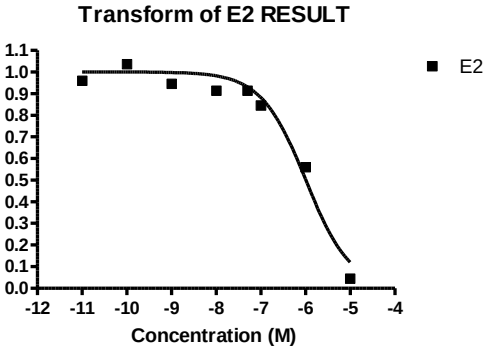
| | EDTA | DTT | BRIJ | NaOAc | DMSO | Cruzain | SUBS |
|---|---|---|---|---|---|---|---|
| | mM | mM | % | mm | % | nM | uM |
| FC | 1 | 2.5 | 0.01 | 100 | 2 | 0.1001 | 15 |

C.16. Representative Analysis of **47** as an Inactive Cruzain Inhibitor.

# GDK-II-33



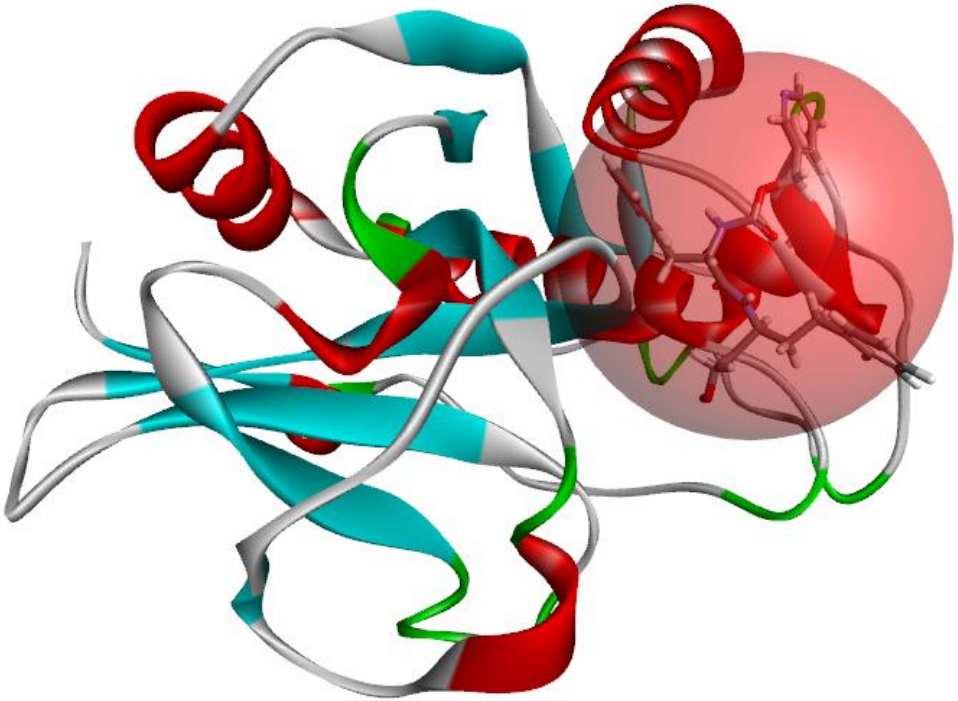




| Concentration (M) | E2 |
|---|---|
| 0.000010 | 0.011490 |
| 0.000001 | 0.144200 |
| 1.000000e-007 | 0.217800 |
| 5.000000e-008 | 0.235400 |
| 1.000000e-008 | 0.235600 |
| 1.000000e-009 | 0.243500 |
| 1.000000e-010 | 0.267000 |
| 1.000000e-011 | 0.247300 |

| | Concentration (M) | E2 |
|---|---|---|
| | -5.000 | 0.045 |
| | -6.000 | 0.559 |
| | -7.000 | 0.845 |
| | -7.301 | 0.913 |
| | -8.000 | 0.914 |
| | -9.000 | 0.945 |
| | -10.000 | 1.036 |
| | -11.000 | 0.959 |

| | E2 |
|---|---|
| Sigmoidal dose-response (variable slope) | |
| Best-fit values | |
| Bottom | = 0.0 |
| Top | = 1.000 |
| LogEC50 | -5.996 |
| HillSlope | -0.8665 |
| EC50 | 1.009e-006 |
| Std. Error | |
| LogEC50 | 0.1002 |
| HillSlope | 0.1484 |
| 95% Confidence Intervals | |
| LogEC50 | -6.241 to -5.751 |
| HillSlope | -1.230 to -0.5033 |
| EC50 | 5.739e-007 to 1.776e-006 |
| Goodness of Fit | |
| Degrees of Freedom | 6 |
| R square | 0.9720 |
| Absolute Sum of Squares | 0.02110 |
| Sy.x | 0.05930 |
| Constraints | |
| Bottom | Bottom = 0.0 |
| Top | Top = 1.000 |
| Number of points | |
| Analyzed | 8 |

C.17. Representative Analysis of **48** as an Inactive Cruzain Inhibitor



| | -CDOCKER_ENERGY | -CDOCKER_INTERACTION_ENERGY | POSE_NUMBER | NumberOfTautomers | StereoisomerIndex | TautomerIndex | Ionization pH | Number( |
|---|---|---|---|---|---|---|---|---|
| | 71.3987 | 69.0827 | 1 | 1 | 4 | 1 | 7.5 | 4 |
| | 71.3014 | 69.3615 | 2 | 1 | 4 | 1 | 7.5 | 4 |
| | 71.0952 | 69.303 | 3 | 1 | 4 | 1 | 7.5 | 4 |
| | 70.986 | 68.9851 | 4 | 1 | 4 | 1 | 7.5 | 4 |
| | 70.6787 | 68.4982 | 5 | 1 | 4 | 1 | 7.5 | 4 |
| | 70.0659 | 68.6286 | 6 | 1 | 4 | 1 | 7.5 | 4 |
| | 69.5473 | 68.1114 | 7 | 1 | 4 | 1 | 7.5 | 4 |
| | 69.1792 | 68.6784 | 8 | 1 | 4 | 1 | 7.5 | 4 |

C18. Validation of Cruzain Crystal Structure docked an Irreversible Inhibitor Using Discovery Studio 3.0

REFERENCES


(1) Copeland, R. A.; Olhava, E. J.; Scott, M. P. Targeting epigenetic enzymes for drug discovery. *Current Opinion in Chemical Biology* **2010**, *14*, 505–510.

(2) HIV and AIDS Activities - Antiretroviral drugs used in the treatment of HIV infection http://www.fda.gov/forconsumers/byaudience/forpatientadvocates/hivandaidsactivities/ucm118915.htm (accessed Jun 21, 2012).

(3) Siles, R. Design, synthesis and biological evaluation of new anti-Cancer nitrogen-containing combretastatins and novel cysteine protease inhibitors for the treatment of Chagas, Baylor University: Waco, Texas, 2005.

(4) Chen, S.-E. Modeling, design, and development of potential inhibitors of γ-glutamylamine cyclotransferase and inhibitors of cruzain as therapeutic agents for Chagas' disease., Baylor University: Waco, Texas, 2008.

(5) Angulo, A. Inhibitors of human cathepsin L and cruzain as therapeutic agents., Waco, Texas, 2008.

(6) Jones, L. M. Design and Synthesis of Functionalized Small-Molecule Inhibitors of Cathepsins L and K, Baylor University: Waco, Texas, 2012.

(7) Song, J. Design, Synthesis, and Biological Evaluation of Non-peptidic Thiosemicarbazone Derivatives as Inhibitors of Cathepsins L and K, Baylor University: Waco, Texas, 2012.

(8) Du, X.; Guo, C.; Hansell, E.; Doyle, P. S.; Caffrey, C. R.; Holler, T. P.; McKerrow, J. H.; Cohen, F. E. Synthesis and structure-activity relationship study of potent trypanocidal thio semicarbazone inhibitors of the trypanosomal cysteine protease cruzain. *J. Med. Chem.* **2002**, *45*, 2695–2707.

(9) Siles, R.; Chen, S.-E.; Zhou, M.; Pinney, K. G.; Trawick, M. L. Design, synthesis, and biochemical evaluation of novel cruzain inhibitors with potential application in the treatment of Chagas' disease. *Bioorganic & Medicinal Chemistry Letters* **2006**, *16*, 4405–4409.

(10) Kishore Kumar, G. D.; Chavarria, G. E.; Charlton-Sevcik, A. K.; Arispe, W. M.; MacDonough, M. T.; Strecker, T. E.; Chen, S.-E.; Siim, B. G.; Chaplin, D. J.; Trawick, M. L.; Pinney, K. G. Design, synthesis, and biological evaluation of potent thiosemicarbazone based cathepsin L inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 1415–1419.

(11) Kumar, G. D. K.; Chavarria, G. E.; Charlton-Sevcik, A. K.; Yoo, G. K.; Song, J.; Strecker, T. E.; Siim, B. G.; Chaplin, D. J.; Trawick, M. L.; Pinney, K. G. Functionalized benzophenone, thiophene, pyridine, and fluorene thiosemicarbazone derivatives as inhibitors of cathepsin L. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 6610–6615.

(12) Song, J.; Jones, L. M.; Kumar, G. D. K.; Conner, E. S.; Bayeh, L.; Chavarria, G. E.; Charlton-Sevcik, A. K.; Chen, S.-E.; Chaplin, D. J.; Trawick, M. L.; Pinney, K. G. Synthesis and Biochemical Evaluation of Thiochromanone Thiosemicarbazone Analogues as Inhibitors of Cathepsin L. *ACS Med. Chem. Lett.* **2012**.

(13) Trawick, M. L.; Chen, S.-E.; Arispe, W. M.; Siles, R.; Zhou, M.; Pinney, K. G. Inhibitors of cruzipain and cathepsin L. In *FASEB J 2006*; San Francisco, USA, 2006; Vol. 20, p. A51.

(14) Cullen, K. J.; Smith, H. S.; Hill, S.; Rosen, N.; Lippman, M. E. Growth Factor Messenger RNA Expression by Human Breast Fibroblasts from Benign and Malignant Lesions. *Cancer Res* **1991**, *51*, 4978–4985.

(15) Jögi, A.; Vaapil, M.; Johansson, M.; Påhlman, S. Cancer cell differentiation heterogeneity and aggressive behavior in solid tumors. *Upsala Journal of Medical Sciences* **2012**, *117*, 217–224.

(16) Duffy, M. J.; Maguire, T. M.; McDermott, E. W.; O'Higgins, N. Urokinase plasminogen activator: A prognostic marker in multiple types of cancer. *Journal of Surgical Oncology* **1999**, *71*, 130–135.

(17) Polednak, A. P.; Flannery, J. T. Brain, other central nervous system, and eye cancer. *Cancer* **1995**, *75*, 330–337.

(18) Diamandopoulos, G. T. Cancer: An historical perspective. *Anticancer Research* **1996**, *16*, 1595–1602.

(19) Gallucci, B. B. Selected concepts of cancer as a disease: from 1900 to oncogenes. *Oncology nursing forum* **1985**, *12*, 69–78.

(20) FASTSTATS - Leading Causes of Death http://www.cdc.gov/nchs/fastats/lcod.htm (accessed May 22, 2012).

(21) Siegel, R.; Naishadham, D.; Jemal, A. Cancer statistics, 2012. *CA: A Cancer Journal for Clinicians* **2012**, *62*, 10–29.

(22) Yabroff, KR Economic Cost of Cancer Mortality Is High in U.S., Regardless of How Cost Is Measured. *JNCI J Natl Cancer Inst* **2008**, *100*, 1741–1741.

(23) Larsson, L.-G. Oncogene- and tumor suppressor gene-mediated suppression of cellular senescence. *Seminars in Cancer Biology* **2011**, *21*, 367–376.

(24) Pylayeva-Gupta, Y.; Grabocka, E.; Bar-Sagi, D. RAS oncogenes: weaving a tumorigenic web. *Nature Reviews Cancer* **2011**, *11*, 761–774.

(25) Nawijn, M. C.; Alendar, A.; Berns, A. For better or for worse: the role of Pim oncogenes in tumorigenesis. *Nature Reviews Cancer* **2010**, *11*, 23–34.

(26) Gastaldi, S.; Comoglio, P. M.; Trusolino, L. The Met oncogene and basal-like breast cancer: another culprit to watch out for? *Breast cancer research : BCR* **2010**, *12*, 208.

(27) Li, W.; Zhao, H.; Qian, W.; Li, H.; Zhang, L.; Ye, Z.; Zhang, G.; Xia, M.; Li, J.; Gao, J.; Li, B.; Kou, G.; Dai, J.; Wang, H.; Guo, Y. Chemotherapy for gastric cancer by finely tailoring anti-Her2 anchored dual targeting immunomicelles. *Biomaterials* **2012**, *33*, 5349–5362.

(28) Nys, K.; Agostinis, P. Bcl-2 family members: Essential players in skin cancer. *Cancer Letters* **2012**, *320*, 1–13.

(29) Thomas, A.; Rajan, A.; Giaccone, G. Tyrosine Kinase Inhibitors in Lung Cancer. *Hematology/Oncology Clinics of North America* **2012**, *26*, 589–605.

(30) Kwon, K. H. PDGFRA promoter polymorphisms are associated with the risk of papillary thyroid cancer. *Molecular Medicine Reports* **2012**.

(31) Garrido, C.; Paco, L.; Romero, I.; Berruguilla, E.; Stefansky, J.; Collado, A.; Algarra, I.; Garrido, F.; Garcia-Lora, A. M. MHC Class I Molecules Act as Tumor Suppressor Genes Regulating the Cell Cycle Gene Expression, Invasion and Intrinsic Tumorigenicity of Melanoma Cells. *Carcinogenesis* **2012**, *33*, 687–693.

(32) Yoshioka, Y.; Shimizu, S.; Ito, T.; Taniguchi, M.; Nomura, M.; Nishida, T.; Sawa, Y. p53 inhibits vascular endothelial growth factor expression in solid tumor. *Journal of Surgical Research* **2012**, *174*, 291–297.

(33) Jafarnejad, S.; Li, G. Regulation of p53 by ING family members in suppression of tumor initiation and progression. *Cancer and Metastasis Reviews* **2012**, *31*, 55–73.

(34) Passaperuma, K.; Warner, E.; Causer, P. A.; Hill, K. A.; Messner, S.; Wong, J. W.; Jong, R. A.; Wright, F. C.; Yaffe, M. J.; Ramsay, E. A.; Balasingham, S.; Verity, L.; Eisen, A.; Curpen, B.; Shumak, R.; Plewes, D. B.; Narod, S. A. Long-term results of screening with magnetic resonance imaging in women with BRCA mutations **2012**.

(35) Lu, C.-L.; Chang, S.-Y.; Tseng, Y.-T.; Wu, B.-R.; Liu, W.-C.; Hsieh, C.-Y.; Wu, P.-Y.; Sun, H.-Y.; Hung, C.-C. Immune reconstitution inflammatory syndrome of Kaposi's sarcoma in an HIV-infected patient. *Journal of Microbiology, Immunology and Infection*.

(36) Izumiya, S.; Mitsuaki, I.; Hodohara, K.; Yoshida, T.; Okabe, H. Epstein-Barr virus-associated lymphoproliferative disorder developed following autologous peripheral blood stem cell transplantation for relapsing Hodgkin's lymphoma. *Oncology Letters* **2012**, *3*, 1203–1206.

(37) Castillo, J. J.; Furman, M.; Beltrán, B. E.; Bibas, M.; Bower, M.; Chen, W.; Díez-Martín, J. L.; Liu, J. J.; Miranda, R. N.; Montoto, S.; Nanaji, N. M.; Navarro, J.-T.; Seegmiller, A. C.; Vose, J. M. Human immunodeficiency virus-associated plasmablastic lymphoma: Poor prognosis in the era of highly active antiretroviral therapy **2012**.

(38) Gyllensten, U.; Gustavsson, I.; Lindell, M.; Wilander, E. Primary high-risk HPV screening for cervical cancer in post-menopausal women. *Gynecologic Oncology* **2012**, *125*, 343–345.

(39) Wikström, A.; Hedblad, M.-A.; Syrjänen, S. Penile intraepithelial neoplasia: histopathological evaluation, HPV typing, clinical presentation and treatment. *Journal of the European Academy of Dermatology and Venereology* **2012**, *26*, 325–330.

(40) Kaminagakura, E.; Villa, L. L.; Andreoli, M. A.; Sobrinho, J. S.; Vartanian, J. G.; Soares, F. A.; Nishimoto, I. N.; Rocha, R.; Kowalski, L. P. High-risk human papillomavirus in oral squamous cell carcinoma of young patients. *International Journal of Cancer* **2012**, *130*, 1726–1732.

(41) Medeiros Vinci, R.; Mestdagh, F.; De Meulenaer, B. Acrylamide formation in fried potato products – Present and future, a critical review on mitigation strategies. *Food Chemistry* **2012**, *133*, 1138–1154.

(42) Fernández, M. I.; López, J. F.; Vivaldi, B.; Coz, F. Long-Term Impact of Arsenic in Drinking Water on Bladder Cancer Health Care and Mortality Rates 20 Years After End of Exposure. *The Journal of Urology* **2012**, *187*, 856–861.

(43) Dai, S.; Ren, D.; Chou, C.-L.; Finkelman, R. B.; Seredin, V. V.; Zhou, Y. Geochemistry of trace elements in Chinese coals: A review of abundances, genetic types, impacts on human health, and industrial utilization. *International Journal of Coal Geology* **2012**, *94*, 3–21.

(44) Kim, S.-I.; Yoon, J.-I.; Tommasi, S.; Besaratinia, A. New Experimental Data Linking Secondhand Smoke Exposure to Lung Cancer in Nonsmokers. *FASEB J* **2012**, *26*, 1845–1854.

(45) Lefemine, V.; Sweetland, H. The role of surgeons in cancer management. *Surgery (Oxford)* **2012**, *30*, 181–185.

(46) Zhou, G.; Niu, L.; Chiu, D.; He, L.; Xu, K. Changes in the expression of serum markers CA242, CA199, CA125, CEA, TNF-α and TSGF after cryosurgery in pancreatic cancer patients. *Biotechnology Letters* 1–7.

(47) Olweny, E. O.; Park, S. K.; Tan, Y. K.; Best, S. L.; Trimmer, C.; Cadeddu, J. A. Radiofrequency Ablation Versus Partial Nephrectomy in Patients with Solitary Clinical T1a Renal Cell Carcinoma: Comparable Oncologic Outcomes at a Minimum of 5 Years of Follow-Up. *European Urology* **2012**, *61*, 1156–1161.

(48) Herrmann, T. R. W.; Liatsikos, E. N.; Nagele, U.; Traxer, O.; Merseburger, A. S. EAU Guidelines on Laser Technologies. *European Urology* **2012**, *61*, 783–795.

(49) Fisher, B.; Costantino, J. P.; Wickerham, D. L.; Redmond, C. K.; Kavanah, M.; Cronin, W. M.; Vogel, V.; Robidoux, A.; Dimitrov, N.; Atkins, J.; Daly, M.; Wieand, S.; Tan-Chiu, E.; Ford, L.; Wolmark, N. Tamoxifen for Prevention of Breast Cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *JNCI J Natl Cancer Inst* **1998**, *90*, 1371–1388.

(50) Perlmutter, M. A.; Lepor, H. Androgen Deprivation Therapy in the Treatment of Advanced Prostate Cancer. *Rev Urol* **2007**, *9*, S3–S8.

(51) Winer, E. P.; Hudis, C.; Burstein, H. J.; Wolff, A. C.; Pritchard, K. I.; Ingle, J. N.; Chlebowski, R. T.; Gelber, R.; Edge, S. B.; Gralow, J.; Cobleigh, M. A.; Mamounas, E. P.; Goldstein, L. J.; Whelan, T. J.; Powles, T. J.; Bryant, J.; Perkins, C.; Perotti, J.; Braun, S.; Langer, A. S.; Browman, G. P.; Somerfield, M. R. American Society of Clinical Oncology Technology Assessment on the Use of Aromatase Inhibitors As Adjuvant Therapy for Postmenopausal Women With Hormone Receptor–Positive Breast Cancer: Status Report 2004. *JCO* **2005**, *23*, 619–629.

(52) Niranjan, A.; Lunsford, L. D.; Emerick, R. L. Stereotactic Radiosurgery for Patients with Metastatic Brain Tumors: Development of a Consensus Radiosurgery Guideline Recommendation. In *Progress in Neurological Surgery*; Kim, D. G.; Lunsford, L. D., Eds.; KARGER: Basel, 2012; Vol. 25, pp. 123–138.

(53) Shirvani, S. M.; Klopp, A. H.; Likhacheva, A.; Jhingran, A.; Soliman, P. T.; Lu, K. H.; Eifel, P. J. Intensity modulated radiation therapy for definitive treatment of paraortic relapse in patients with endometrial cancer. *Practical Radiation Oncology*.

(54) Kim, M.; Ghate, A.; Phillips, M. H. A stochastic control formalism for dynamic biologically conformal radiation therapy. *European Journal of Operational Research* **2012**, *219*, 541–556.

(55) Nascimento, A.; Falzon, P. Producing effective treatment, enhancing safety: Medical physicists' strategies to ensure quality in radiotherapy. *Applied Ergonomics* **2012**, *43*, 777–784.

(56) Hussain, M.; Javeed, A.; Ashraf, M.; Al-Zaubai, N.; Stewart, A.; Mukhtar, M. M. Non-steroidal anti-inflammatory drugs, tumour immunity and immunotherapy. *Pharmacological Research* **2012**, *66*, 7–18.

(57) Wang, B. X.; Rahbar, R.; Fish, E. N. Interferon: Current Status and Future Prospects in Cancer Therapy. *Journal of Interferon & Cytokine Research* **2011**, *31*, 545–552.

(58) Laberge, R.-M.; Awad, P.; Campisi, J.; Desprez, P.-Y. Epithelial-Mesenchymal Transition Induced by Senescent Fibroblasts. *Cancer Microenvironment* **2012**, *5*, 39–44.

(59) Mojtahedi, Z.; Khademi, B.; Yehya, A.; Talebi, A.; Fattahi, M. J.; Ghaderi, A. Serum levels of interleukins 4 and 10 in head and neck squamous cell carcinoma. *The Journal of Laryngology & Otology* **2012**, *126*, 175–179.

(60) Patra, S. K.; Arora, S. Integrative role of neuropeptides and cytokines in cancer anorexia–cachexia syndrome. *Clinica Chimica Acta* **2012**, *413*, 1025–1034.

(61) DeRose, P.; Thorpe, P. E.; Gerber, D. E. Development of bavituximab, a vascular targeting agent with immune-modulating properties, for lung cancer treatment. *Immunotherapy* **2011**, *3*, 933–944.

(62) Shi, W.; Han, X.; Yao, J.; Yang, J.; Shi, Y. Combined effect of histone deacetylase inhibitor suberoylanilide hydroxamic acid and anti-CD20 monoclonal antibody rituximab on mantle cell lymphoma cells apoptosis. *Leukemia Research* **2012**, *36*, 749–755.

(63) Adler, M. J.; Dimitrov, D. S. Therapeutic Antibodies against Cancer. *Hematology/Oncology Clinics of North America* **2012**, *26*, 447–481.

(64) Tombal, B. Continuous Improvement Versus Innovation: The Case for Sipuleucel-T. *European Urology* **2012**, *61*, 648–649.

(65) Agarwal, N.; Sonpavde, G.; Sternberg, C. N. Novel Molecular Targets for the Therapy of Castration-Resistant Prostate Cancer. *European Urology* **2012**, *61*, 950–960.

(66) A. McCubrey, J.; S. Steelman, L.; L. Abrams, S.; Misaghian, N.; H. Chappell, W.; Basecke, J.; Nicoletti, F.; Libra, M.; Ligresti, G.; Stivala, F.; Maksimovic-Ivanic, D.; Mijatovic, S.; Montalto, G.; Cervello, M.; Laidler, P.; Bonati, A.; Evangelisti, C.; Cocco, L.; M. Martelli, A. Targeting the Cancer Initiating Cell: The Ultimate Target for Cancer Therapy. *Current Pharmaceutical Design* **2012**, *18*, 1784–1795.

(67) Serrano, C.; Morales, R.; Suárez, C.; Núñez, I.; Valverde, C.; Rodón, J.; Humbert, J.; Padrós, O.; Carles, J. Emerging therapies for urothelial cancer. *Cancer Treatment Reviews* **2012**, *38*, 311–317.

(68) Shojaei, F. Anti-angiogenesis therapy in cancer: Current challenges and future perspectives. *Cancer Letters* **2012**, *320*, 130–137.

(69) Shandala, T.; Shen Ng, Y.; Hopwood, B.; Yip, Y.-C.; Foster, B. K.; Xian, C. J. The role of osteocyte apoptosis in cancer chemotherapy-induced bone loss. *Journal of Cellular Physiology* **2012**, *227*, 2889–2897.

(70) Gerber, D. E. Targeted therapies: a new generation of cancer treatments. *Am Fam Physician* **2008**, *77*, 311–319.

(71) Norris, R. E.; Adamson, P. C. Clinical potency of methotrexate, aminopterin, talotrexin and pemetrexed in childhood leukemias. *Cancer Chemother. Pharmacol.* **2010**, *65*, 1125–1130.

(72) Farley, J. H.; Heathcock, R. B.; Branch, W.; Larsen, W.; Homas, D. Treatment of metastatic gestational choriocarcinoma with oral methotrexate in a combat environment. *Obstet Gynecol* **2005**, *105*, 1250–1254.

(73) DeVita, V. T.; Chu, E. A History of Cancer Chemotherapy. *Cancer Res* **2008**, *68*, 8643–8653.

(74) Freedman, R.; Anders, C. Treatment of Breast Cancer Brain Metastases. *Current Breast Cancer Reports* **2012**, *4*, 1–9.

(75) Kawauchi, T. Cell Adhesion and Its Endocytic Regulation in Cell Migration during Neural Development and Cancer Metastasis. *International Journal of Molecular Sciences* **2012**, *13*, 4564–4590.

(76) Crisafuli, F. A. P.; Cesconetto, E. C.; Ramos, E. B.; Rocha, M. S. DNA–cisplatin interaction studied with single molecule stretching experiments. *Integrative Biology* **2012**, *4*, 568.

(77) Lee, J. W.; Kang, H. J.; Lee, S. H.; Yu, K.-S.; Kim, N. H.; Yuk, Y. J.; Jang, M. K.; Han, E. J.; Kim, H.; Song, S. H.; Park, K. D.; Shin, H. Y.; Jang, I.-J.; Ahn, H. S. Highly Variable Pharmacokinetics of Once-Daily Intravenous Busulfan When Combined with Fludarabine in Pediatric Patients: Phase I Clinical Study for Determination of Optimal Once-Daily Busulfan Dose Using Pharmacokinetic Modeling. *Biology of Blood and Marrow Transplantation* **2012**, *18*, 944–950.

(78) Hatiboglu, M. A.; Kong, L.-Y.; Wei, J.; Wang, Y.; McEnery, K. A.; Fuller, G. N.; Qiao, W.; Davies, M. A.; Priebe, W.; Heimberger, A. B. The tumor microenvironment expression of p-STAT3 influences the efficacy of cyclophosphamide with WP1066 in murine melanoma models. *International Journal of Cancer* **2012**, *131*, 8–17.

(79) Sáez-Ayala, M.; Fernández-Pérez, M. P.; Montenegro, M. F.; Sánchez-del-Campo, L.; Chazarra, S.; Piñero-Madrona, A.; Cabezas-Herrera, J.; Rodríguez-López, J. N. Melanoma coordinates general and cell-specific mechanisms to promote methotrexate resistance. *Experimental Cell Research* **2012**, *318*, 1146–1159.

(80) Octavia, Y.; Tocchetti, C. G.; Gabrielson, K. L.; Janssens, S.; Crijns, H. J.; Moens, A. L. Doxorubicin-induced cardiomyopathy: From molecular mechanisms to therapeutic strategies. *Journal of Molecular and Cellular Cardiology* **2012**, *52*, 1213–1225.

(81) Karkos, P. D.; Leong, S. C.; Sastry, A.; Assimakopoulos, A. D.; Swift, A. C. Evidence-based applications of mitomycin C in the nose. *American Journal of Otolaryngology* **2011**, *32*, 422–425.

(82) Hartwell, D.; Jones, J.; Loveman, E.; Harris, P.; Clegg, A.; Bird, A. Topotecan for relapsed small cell lung cancer: A systematic review and economic evaluation. *Cancer Treatment Reviews* **2011**, *37*, 242–249.

(83) Leu, L.; Solimando, J.; Waddell; J. Aubrey Waddell, P. Cancer Chemotherapy Update - Etoposide, Doxorubicin, Vincristine, Cyclophosphamide, Prednisone (EPOCH) for Treatment of Lymphomas. *Hospital Pharmacy* **2007**, *42*, 1006–1012.

(84) Chirgwin, J.; Chua, S. L. Management of breast cancer with nanoparticle albumin-bound (nab)-paclitaxel combination regimens: A clinical review. *The Breast* **2011**, *20*, 394–406.

(85) Kumar, S.; Flinn, I.; Richardson, P. G.; Hari, P.; Callander, N.; Noga, S. J.; Stewart, A. K.; Turturro, F.; Rifkin, R.; Wolf, J.; Estevam, J.; Mulligan, G.; Shi, H.; Webb, I. J.; Rajkumar, S. V. Randomized, Multicenter, Phase 2 Study (EVOLUTION) of Combinations of Bortezomib, Dexamethasone, Cyclophosphamide, and Lenalidomide in Previously Untreated Multiple Myeloma. *Blood* **2012**, *119*, 4375–4382.

(86) Hafejee, A.; Coulson, I. H. Dysphagia in dermatomyositis secondary to bladder cancer: rapid response to combined immunoglobulin and methylprednisolone. *Clinical and Experimental Dermatology* **2005**, *30*, 93–94.

(87) Morrison, E.; Waddell; J. Aubrey Waddell, P.; Solimando, J. Cancer Chemotherapy Update - Cyclophosphamide, Bortezomib and Dexamethasone (CyBorD) Regimen for Multiple Myeloma. *Hospital Pharmacy* **2011**, *46*, 839–847.

(88) Yerlikaya, A.; Okur, E.; Ulukaya, E. The p53-independent induction of apoptosis in breast cancer cells in response to proteasome inhibitor bortezomib. *Tumor Biology* 1–8.

(89) Geiger, T. R.; Peeper, D. S. Metastasis mechanisms. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **2009**, *1796*, 293–308.

(90) Koblinski, J. E.; Ahram, M.; Sloane, B. F. Unraveling the role of proteases in cancer. *Clinica Chimica Acta* **2000**, *291*, 113–135.

(91) Turk, B.; Turk, D.; Turk, V. Protease signalling: the cutting edge. *The EMBO Journal* **2012**, *31*, 1630–1643.

(92) Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **2012**, *1824*, 68–88.

(93) Hiraiwa, M. Cathepsin A/protective protein: an unusual lysosomal multifunctional protein. *Cellular and Molecular Life Sciences* **1999**, *56*, 894–907.

(94) Burster, T.; Macmillan, H.; Hou, T.; Boehm, B. O.; Mellins, E. D. Cathepsin G: Roles in antigen presentation and beyond. *Molecular Immunology* **2010**, *47*, 658–665.

(95) Huber-Lang, M.; Denk, S.; Fulda, S.; Erler, E.; Kalbitz, M.; Weckbach, S.; Schneider, E. M.; Weiss, M.; Kanse, S. M.; Perl, M. Cathepsin D is released after severe tissue trauma in vivo and is capable of generating C5a in vitro. *Molecular Immunology* **2012**, *50*, 60–65.

(96) Yamamoto, K.; Kawakubo, T.; Yasukochi, A.; Tsukuba, T. Emerging roles of cathepsin E in host defense mechanisms. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **2012**, *1824*, 105–112.

(97) Turk, B.; Turk, D.; Turk, V. Lysosomal cysteine proteases: more than scavengers. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **2000**, *1477*, 98–111.

(98) Heutinck, K. M.; ten Berge, I. J. M.; Hack, C. E.; Hamann, J.; Rowshani, A. T. Serine proteases of the human immune system in health and disease. *Molecular Immunology* **2010**, *47*, 1943–1955.

(99) Sjö, P. Neutrophil elastase inhibitors: recent advances in the development of mechanism-based and nonelectrophilic inhibitors. *Future Medicinal Chemistry* **2012**, *4*, 651–660.

(100) Dunn, B. M. Structure and Mechanism of the Pepsin-Like Family of Aspartic Peptidases. *Chem. Rev.* **2002**, *102*, 4431–4458.

(101) Stoka, V.; Turk, B.; Turk, V. Lysosomal Cysteine Proteases: Structural Features and their Role in Apoptosis. *IUBMB Life* **2005**, *57*, 347–353.

(102) Klein, T.; Bischoff, R. Physiology and pathophysiology of matrix metalloproteases. *Amino Acids* **2011**, *41*, 271–290.

(103) Structural Classification of Proteins http://scop.mrc-lmb.cam.ac.uk/scop/search.cgi?ver=1.75&key=cathepsin (accessed May 25, 2012).

(104) Protein Structure Classification http://www.cathdb.info/cathnode/3.90.70.10 (accessed May 25, 2012).

(105) Berdowska, I. Cysteine proteases as disease markers. *Clinica Chimica Acta* **2004**, *342*, 41–69.

(106) Lecaille, F.; Kaleta, J.; Brömme, D. Human and Parasitic Papain-Like Cysteine Proteases: Their Role in Physiology and Pathology and Recent Developments in Inhibitor Design. *Chem. Rev.* **2002**, *102*, 4459–4488.

(107) Han, S.-R.; Momeni, A.; Strach, K.; Suriyaphol, P.; Fenske, D.; Paprotka, K.; Hashimoto, S. I.; Torzewski, M.; Bhakdi, S.; Husmann, M. Enzymatically Modified LDL Induces Cathepsin H in Human Monocytes Potential Relevance in Early Atherogenesis. *Arterioscler Thromb Vasc Biol* **2003**, *23*, 661–667.

(108) Nägler, D. K.; Kraus, S.; Feierler, J.; Mentele, R.; Lottspeich, F.; Jochum, M.; Faussner, A. A cysteine-type carboxypeptidase, cathepsin X, generates peptide receptor agonists. *International Immunopharmacology* **2010**, *10*, 134–139.

(109) Puzer, L.; Barros, N. M. T.; Paschoalin, T.; Hirata, I. Y.; Tanaka, A. S.; Oliveira, M. C.; Brömme, D.; Carmona, A. K. Cathepsin V, but not cathepsins L, B and K, may release angiostatin-like fragments from plasminogen. *Biol. Chem.* **2008**, *389*, 195–200.

(110) Naour, N.; Rouault, C.; Fellahi, S.; Lavoie, M.-E.; Poitou, C.; Keophiphath, M.; Eberlé, D.; Shoelson, S.; Rizkalla, S.; Bastard, J.-P.; Rabasa-Lhoret, R.; Clément, K.; Guerre-Millo, M. Cathepsins in Human Obesity: Changes in Energy Balance Predominantly Affect Cathepsin S in Adipose Tissue and in Circulation. *JCEM* **2010**, *95*, 1861–1868.

(111) Dauth, S.; Arampatzidou, M.; Rehders, M.; Yu, D. M. T.; Führer, D.; Brix, K. Thyroid Cathepsin K: Roles in Physiology and Thyroid Disease. *Clinical Reviews in Bone and Mineral Metabolism* **2011**, *9*, 94–106.

(112) Halangk, W.; Lerch, M. M.; Brandt-Nedelev, B.; Roth, W.; Ruthenbuerger, M.; Reinheckel, T.; Domschke, W.; Lippert, H.; Peters, C.; Deussing, J. Role of cathepsin B in intracellular trypsinogen activation and the onset of acute pancreatitis. *Journal of Clinical Investigation* **2000**, *106*, 773–781.

(113) Sutton, V. R.; Waterhouse, N. J.; Browne, K. A.; Sedelies, K.; Ciccone, A.; Anthony, D.; Koskinen, A.; Mullbacher, A.; Trapani, J. A. Residual Active Granzyme B in Cathepsin C–Null Lymphocytes Is Sufficient for Perforin-Dependent Target Cell Apoptosis. *J Cell Biol* **2007**, *176*, 425–433.

(114) Bania, J.; Gatti, E.; Lelouard, H.; David, A.; Cappello, F.; Weber, E.; Camosseto, V.; Pierre, P. Human Cathepsin S, but Not Cathepsin L, Degrades Efficiently MHC Class II-Associated Invariant Chain in Nonprofessional APCs. *PNAS* **2003**, *100*, 6664–6669.

(115) Costantino, C. M.; Ploegh, H. L.; Hafler, D. A. Cathepsin S Regulates Class II MHC Processing in Human CD4+ HLA-DR+ T Cells. *J Immunol* **2009**, *183*, 945–952.

(116) Li, C. Y.; Jepsen, K. J.; Majeska, R. J.; Zhang, J.; Ni, R.; Gelb, B. D.; Schaffler, M. B. Mice Lacking Cathepsin K Maintain Bone Remodeling but Develop Bone Fragility Despite High Bone Mass. *Journal of Bone and Mineral Research* **2006**, *21*, 865–875.

(117) Benavides, F.; Starost, M. F.; Flores, M.; Gimenez-Conti, I. B.; Guénet, J.-L.; Conti, C. J. Impaired Hair Follicle Morphogenesis and Cycling with Abnormal Epidermal Differentiation in nackt Mice, a Cathepsin L-Deficient Mutation. *The American Journal of Pathology* **2002**, *161*, 693–703.

(118) Sluka, P.; O'Donnell, L.; Stanton, P. G. Stage-Specific Expression of Genes Associated with Rat Spermatogenesis: Characterization by Laser-Capture Microdissection and Real-Time Polymerase Chain Reaction. *Biol Reprod* **2002**, *67*, 820–828.

(119) Hishita, T.; Tada-Oikawa, S.; Tohyama, K.; Miura, Y.; Nishihara, T.; Tohyama, Y.; Yoshida, Y.; Uchiyama, T.; Kawanishi, S. Caspase-3 Activation by Lysosomal Enzymes in Cytochrome C-Independent Apoptosis in Myelodysplastic Syndrome-Derived Cell Line P39. *Cancer Res* **2001**, *61*, 2878–2884.

(120) Wiederanders, B. The Function of Propeptide Domains of Cysteine Proteinases. In *Cellular Peptidases in Immune Functions and Diseases 2*; Langner, J.; Ansorge, S., Eds.; Advances in Experimental Medicine and Biology; Springer Netherlands, 2002; Vol. 477, pp. 261–270.

(121) Karrer, K. M.; Peiffer, S. L.; DiTomas, M. E. Two Distinct Gene Subfamilies Within the Family of Cysteine Protease Genes. *PNAS* **1993**, *90*, 3063–3067.

(122) Musil, D.; Zucic, D.; Turk, D.; Engh, R. A.; Mayr, I.; Huber, R.; Popovic, T.; Turk Towatari, V. T.; Katunuma, N.; Bode, W. The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: The structural basis for its specificity. *EMBO Journal* **1991**, *10*, 2321–2330.

(123) Fujishima, A.; Imai, Y.; Nomura, T.; Fujisawa, Y.; Yamamoto, Y.; Sugawara, T. The crystal structure of human cathepsin L complexed with E-64. *FEBS Letters* **1997**, *407*, 47–50.

(124) McGrath, M. E.; Klaus, J. L.; Barnes, M. G.; Bromme, D. Crystal structure of human cathepsin K complexed with a potent inhibitor. *Nature Structural Biology* **1997**, *4*, 105–109.

(125) Gunčar, G.; Podobnik, M.; Pungerčar, J.; BorutŠtrukelj; Turk, V.; Turk, D. Crystal structure of porcine cathepsin H determined at 2.1 å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function. *Structure* **1998**, *6*, 51–61.

(126) Gunčar, G.; Klemenčič, I.; Turk, B.; Turk, V.; Karaoglanovic-Carmona, A.; Juliano, L.; Turk, D. Crystal structure of cathepsin X: A flip-flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. *Structure* **2000**, *8*, 305–313.

(127) Somoza, J. R.; Zhan, H.; Bowman, K. K.; Yu, L.; Mortara, K. D.; Palmer, J. T.; Clark, J. M.; McGrath, M. E. Crystal structure of human cathepsin V. *Biochemistry* **2000**, *39*, 12543–12551.

(128) Turk, D.; Janjić, V.; Štern, I.; Podobnik, M.; Lamba, D.; Dahl, S. W.; Lauritzen, C.; Pedersen, J.; Turk, V.; Turk, B. Structure of human dipeptidyl peptidase I (cathepsin C): Exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO Journal* **2001**, *20*, 6570–6582.

(129) Turkenburg, J. P.; Lamers, M. B. A. C.; Brzozowski, A. M.; Wright, L. M.; Hubbard, R. E.; Sturt, S. L.; Williams, D. H. Structure of a Cys25→Ser mutant of human cathepsin S. *Acta Crystallographica Section D: Biological Crystallography* **2002**, *58*, 451–455.

(130) Somoza, J. R.; Palmer, J. T.; Ho, J. D. The crystal structure of human cathepsin F and its implications for the development of novel immunomodulators. *Journal of Molecular Biology* **2002**, *322*, 559–568.

(131) Huang, L.; Brinen, L. S.; Ellman, J. A.; Brinen, L. S.; Hansell, E.; Cheng, J.; Roush, W. R.; McKerrow, J. H.; Fletterick, R. J.; MCGRATH, M. E.; EAKIN, A. E.; ENGEL, J. C.; MCKERROW, J. H.; CRAIK, C. S.; FLETTERICK, R. J. Crystal Structures of Reversible Ketone-based Inhibitors of the Cysteine Protease Cruzain. *Bioorg.Med.Chem.* **2002**, *11*, 21–29.

(132) Schechter, I.; Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.

(133) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; 1st ed.; W. H. Freeman, 1998.

(134) McIntyre, G. F.; Erickson, A. H. Procathepsins L and D Are Membrane-Bound in Acidic Microsomal Vesicles. *J. Biol. Chem.* **1991**, *266*, 15438–15445.

(135) Guay, J.; Falgueyret, J.-P.; Ducret, A.; Percival, M. D.; Mancini, J. A. Potency and selectivity of inhibition of cathepsin K, L and S by their respective propeptides. *European Journal of Biochemistry* **2001**, *267*, 6311–6318.

(136) Turk, D.; Podobnik, M.; Kuhelj, R.; Dolinar, M.; Turk, V. Crystal structures of human procathepsin B at 3.2 and 3.3 Å resolution reveal an interaction motif between a papain-like cysteine protease and its propeptide. *FEBS Letters* **1996**, *384*, 211–214.

(137) Coulombe, R.; Grochulski, P.; Sivaraman, J.; Ménard, R.; Mort, J. S.; Cygler, M. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO J* **1996**, *15*, 5492–5503.

(138) LaLonde, J. M.; Zhao, B.; Janson, C. A.; D'Alessio, K. J.; McQueney, M. S.; Orsini, M. J.; Debouck, C. M.; Smith, W. W. The Crystal Structure of Human Procathepsin K‡. *Biochemistry* **1998**, *38*, 862–869.

(139) Sivaraman, J.; Nägler, D. K.; Zhang, R.; Ménard, R.; Cygler, M. Crystal structure of human procathepsin X: a cysteine protease with the proregion covalently linked to the active site cysteine. *Journal of Molecular Biology* **2000**, *295*, 939–951.

(140) Dahl, S. W.; Halkier, T.; Lauritzen, C.; Dolenc, I.; Pedersen, J.; Turk, V.; Turk, B. Human recombinant pro-dipeptidyl peptidase I (cathepsin C) can be activated by cathepsins L and S but not by autocatalytic processing. *Biochemistry* **2001**, *40*, 1671–1678.

(141) Nishimura, Y.; Kawabata, T.; Furuno, K.; Kato, K. Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. *Archives of Biochemistry and Biophysics* **1989**, *271*, 400–406.

(142) Ménard, R.; Carmona, E.; Takebe, S.; Dufour, É.; Plouffe, C.; Mason, P.; Mort, J. S. Autocatalytic processing of recombinant human procathepsin L: Contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro. *Journal of Biological Chemistry* **1998**, *273*, 4478–4484.

(143) Jerala, R.; Žerovnik, E.; Kidrič, J.; Turk, V. pH-induced conformational transitions of the propeptide of human cathepsin L. A role for a molten globule state in zymogen activation. *Journal of Biological Chemistry* **1998**, *273*, 11498–11504.

(144) Parks, W. C.; Mecham, R. P. *Extracellular Matrix Degradation*; Springer, 2011.

(145) Higgins, W. J.; Fox, D. M.; Kowalski, P. S.; Nielsen, J. E.; Worrall, D. M. Heparin Enhances Serpin Inhibition of the Cysteine Protease Cathepsin L. *J. Biol. Chem.* **2010**, *285*, 3722–3729.

(146) Wallin, H.; Bjarnadottir, M.; Vogel, L. K.; Wassélius, J.; Ekström, U.; Abrahamson, M. Cystatins – Extra- and intracellular cysteine protease inhibitors: High-level secretion and uptake of cystatin C in human neuroblastoma cells. *Biochimie* **2010**, *92*, 1625–1634.

(147) Lenarcic, B.; Bevec, T. Thyropins--new structurally related proteinase inhibitors. *Biol. Chem.* **1998**, *379*, 105–111.

(148) Lalmanach, G.; Naudin, C.; Lecaille, F.; Fritz, H. Kininogens: More than cysteine protease inhibitors and kinin precursors. *Biochimie* **2010**, *92*, 1568–1579.

(149) Marques, E. F.; Bueno, M. A.; Duarte, P. D.; Silva, L. R. S. P.; Martinelli, A. M.; Dos Santos, C. Y.; Severino, R. P.; Brömme, D.; Vieira, P. C.; Corrêa, A. G. Evaluation of synthetic acridones and 4-quinolinones as potent inhibitors of cathepsins L and v. *European Journal of Medicinal Chemistry* **2012**, *54*, 10–21.

(150) Lorenz, J. C.; Busacca, C. A.; Feng, X.; Grinberg, N.; Haddad, N.; Johnson, J.; Kapadia, S.; Lee, H.; Saha, A.; Sarvestani, M.; Spinelli, E. M.; Varsolona, R.; Wei, X.; Zeng, X.; Senanayake, C. H. Large-scale asymmetric synthesis of a cathepsin S inhibitor. *Journal of Organic Chemistry* **2010**, *75*, 1155–1161.

(151) Frlan, R.; Gobec, S. Inhibitors of Cathepsin B. *Current Medicinal Chemistry* **2006**, *13*, 2309–2327.

(152) Wang, D.; Li, W.; Pechar, M.; Kopečková, P.; Brömme, D.; Kopeček, J. Cathepsin K inhibitor–polymer conjugates: potential drugs for the treatment of osteoporosis and rheumatoid arthritis. *International Journal of Pharmaceutics* **2004**, *277*, 73–79.

(153) Matsumoto, K.; Yamamoto, D.; Ohishi, H.; Tomoo, K.; Ishida, T.; Inoue, M.; Sadatome, T.; Kitamura, K.; Mizuno, H. Mode of binding of E-64-c, a potent thiol protease inhibitor, to papain as determined by X-ray crystal analysis of the complex. *FEBS Letters* **1989**, *245*, 177–180.

(154) Matsumoto, K.; Murata, M.; Sumiya, S.; Mizoue, K.; Kitamura, K.; Ishida, T. X-Ray crystal structure of papain complexed with cathepsin B-specific covalent-type inhibitor: substrate specificity and inhibitory activity. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1998**, *1383*, 93–100.

(155) KATUNUMA, N. Structure-based development of specific inhibitors for individual cathepsins and their medical applications. *Proc Jpn Acad Ser B Phys Biol Sci* **2011**, *87*, 29–39.

(156) McConnell, R. M.; York, J. L.; Frizzell, D.; Ezell, C. Inhibition studies of some serine and thiol proteinases by new leupeptin analogs. *J. Med. Chem.* **1993**, *36*, 1084–1089.

(157) Kalinowski, D. S.; Quach, P.; Richardson, D. R. Thiosemicarbazones: the new wave in cancer treatment. *Future Med Chem* **2009**, *1*, 1143–1151.

(158) Costa, A. G.; Cusano, N. E.; Silva, B. C.; Cremers, S.; Bilezikian, J. P. Cathepsin K: its skeletal actions and role as a therapeutic target in osteoporosis. *Nature Reviews Rheumatology* **2011**, *7*, 447–456.

(159) Ikeda, Y.; Ikata, T.; Mishiro, T.; Nakano, S.; Ikebe, M.; Yasuoka, S. Cathepsins B and L in synovial fluids from patients with rheumatoid arthritis and the effect of cathepsin B on the activation of pro-urokinase. *J. Med. Invest.* **2000**, *47*, 61–75.

(160) Hou, W.-S.; Li, W.; Keyszer, G.; Weber, E.; Levy, R.; Klein, M. J.; Gravallese, E. M.; Goldring, S. R.; Brömme, D. Comparison of cathepsins K and S expression within the rheumatoid and osteoarthritic synovium. *Arthritis Rheum.* **2002**, *46*, 663–674.

(161) Deschamps, K.; Cromlish, W.; Weicker, S.; Lamontagne, S.; Huszar, S. L.; Gauthier, J. Y.; Mudgett, J. S.; Guimond, A.; Romand, R.; Frossard, N.; Percival, M. D.; Slipetz, D.; Tan, C. M. Genetic and pharmacological evaluation of cathepsin s in a mouse model of asthma. *Am. J. Respir. Cell Mol. Biol.* **2011**, *45*, 81–87.

(162) Tolosa, E.; Li, W.; Yasuda, Y.; Wienhold, W.; Denzin, L. K.; Lautwein, A.; Driessen, C.; Schnorrer, P.; Weber, E.; Stevanovic, S.; Kurek, R.; Melms, A.; Brömme, D. Cathepsin V is involved in the degradation of invariant chain in human thymus and is overexpressed in myasthenia gravis. *J Clin Invest* **2003**, *112*, 517–526.

(163) Liu, J.; Ma, L.; Yang, J.; Ren, A.; Sun, Z.; Yan, G.; Sun, J.; Fu, H.; Xu, W.; Hu, C.; Shi, G.-P. Increased serum cathepsin S in patients with atherosclerosis and diabetes. *Atherosclerosis* **2006**, *186*, 411–419.

(164) Guo, J.; Bot, I.; De Nooijer, R.; Hoffman, S. J.; Stroup, G. B.; Biessen, E. A. L.; Benson, G. M.; Groot, P. H. E.; Van Eck, M.; Berkel, V.; J.c, T. Leucocyte Cathepsin K Affects Atherosclerotic Lesion Composition and Bone Mineral Density in Low-Density Lipoprotein Receptor Deficient Mice. *Cardiovasc Res* **2009**, *81*, 278–285.

(165) Eley, B. M.; Cox, S. W. Cathepsin B- and L-like activities at local gingival sites of chronic periodontitis patients. *J. Clin. Periodontol.* **1991**, *18*, 499–504.

(166) Minematsu, N.; Nakamura, H.; Furuuchi, M.; Nakajima, T.; Takahashi, S.; Tsujimura, S.; Tateno, H.; Ishizaka, A. Common functional polymorphisms in the cathepsin S promoter in Japanese subjects: possible contribution to pulmonary emphysema. *Respirology* **2008**, *13*, 498–504.

(167) Taggart, C. C.; Lowe, G. J.; Greene, C. M.; Mulgrew, A. T.; O'Neill, S. J.; Levine, R. L.; McElvaney, N. G. Cathepsin B, L, and S Cleave and Inactivate Secretory Leucoprotease Inhibitor. *J. Biol. Chem.* **2001**, *276*, 33345–33352.

(168) Kukor, Z.; Mayerle, J.; Krüger, B.; Tóth, M.; Steed, P. M.; Halangk, W.; Lerch, M. M.; Sahin-Tóth, M. Presence of Cathepsin B in the Human Pancreatic Secretory Pathway and Its Role in Trypsinogen Activation During Hereditary Pancreatitis. *J. Biol. Chem.* **2002**, *277*, 21389–21396.

(169) Schönefuss, A.; Wendt, W.; Schattling, B.; Schulten, R.; Hoffmann, K.; Stuecker, M.; Tigges, C.; Lübbert, H.; Stichel, C. Upregulation of cathepsin S in psoriatic keratinocytes. *Exp. Dermatol.* **2010**, *19*, e80–88.

(170) Bylaite, M.; Moussali, H.; Marciukaitiene, I.; Ruzicka, T.; Walz, M. Expression of cathepsin L and its inhibitor hurpin in inflammatory and neoplastic skin diseases. *Exp. Dermatol.* **2006**, *15*, 110–118.

(171) Kawada, A.; Hara, K.; Kominami, E.; Hiruma, M.; Noguchi, H.; Ishibashi, A. Processing of cathepsins L, B and D in psoriatic epidermis. *Archives of Dermatological Research* **1997**, *289*, 87–93.

(172) Hart, T. C.; Hart, P. S.; Michalec, M. D.; Zhang, Y.; Firatli, E.; Van Dyke, T. E.; Stabholz, A.; Zlotogorski, A.; Shapira, L.; Soskolne, W. A.; Zlorogorski, A. Haim-Munk syndrome and Papillon-Lefèvre syndrome are allelic mutations in cathepsin C. *J. Med. Genet.* **2000**, *37*, 88–94.

(173) Moran, M. T.; Schofield, J. P.; Hayman, A. R.; Shi, G.-P.; Young, E.; Cox, T. M. Pathologic Gene Expression in Gaucher Disease: Up-Regulation of Cysteine Proteinases Including Osteoclastic Cathepsin K. *Blood* **2000**, *96*, 1969–1978.

(174) Cimerman, N.; Mesko Brguljan, P.; Krasovec, M.; Suskovic, S.; Kos, J. Serum concentration and circadian profiles of cathepsins B, H and L, and their inhibitors, stefins A and B, in asthma. *Clin. Chim. Acta* **2001**, *310*, 113–122.

(175) Steinfeld, S.; Maho, A.; Chaboteaux, C.; Daelemans, P.; Pochet, R.; Appelboom, T.; Kiss, R. Prolactin Up-Regulates Cathepsin B and D Expression in Minor Salivary Glands of Patients with Sj|[ouml]|gren's Syndrome. *Laboratory Investigation* **2000**, *80*, 1711–1720.

(176) Chemistry, R. B. S. P. D. O. *Organic Chemistry of Enzyme-Catalyzed Reactions, Revised Edition, Second Edition*; 2nd ed.; Academic Press, 2002.

(177) Copeland, R. A. *Evaluation of Enzyme Inhibitors in Drug Discovery: A Guide for Medicinal Chemists and Pharmacologists*; 1st ed.; Wiley-Interscience, 2005.

(178) Gowen, M.; Lazner, F.; Dodds, R.; Kapadia, R.; Feild, J.; Tavaria, M.; Bertoncello, I.; Drake, F.; Zavarselk, S.; Tellis, I.; Hertzog, P.; Debouck, C.; Kola, I. Cathepsin K knockout mice develop osteopetrosis due to a deficit in matrix degradation but not demineralization. *J. Bone Miner. Res.* **1999**, *14*, 1654–1663.

(179) McCullough, A. R. Four-Year Review of Sildenafil Citrate. *Rev Urol* **2002**, *4*, S26–S38.

(180) Zhang, H.; Vrang, L.; Bäckbro, K.; Unge, T.; Noréen, R.; Oberg, B. Enzymatic properties and sensitivity to inhibitors of human immunodeficiency virus type 1 (HIV-1) reverse transcriptase with Glu-138-->Arg and Tyr-188-->His mutations. *Antiviral Res.* **1994**, *24*, 43–57.

(181) Staker, B. L.; Hjerrild, K.; Feese, M. D.; Behnke, C. A.; Burgin, A. B., Jr; Stewart, L. The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15387–15392.

(182) Matsumoto, K.; Mizoue, K.; Kitamura, K.; Tse, W. C.; Huber, C. P.; Ishida, T. Structural basis of inhibition of cysteine proteases by E-64 and its derivatives. *Biopolymers* **1999**, *51*, 99–107.

(183) Lanter, J.; Zhang, X.; Sui, Z. Chapter sixteen - Medicinal Chemistry Inspired Fragment-Based Drug Discovery. In *Fragment-Based Drug Design Tools, Practical Approaches, and Examples*; Academic Press, 2011; Vol. Volume 493, pp. 421–445.

(184) Thompson, M. J.; Louth, J. C.; Ferrara, S.; Sorrell, F. J.; Irving, B. J.; Cochrane, E. J.; Meijer, A. J. H. M.; Chen, B. Structure–Activity Relationship Refinement and Further Assessment of Indole-3-glyoxylamides as a Lead Series against Prion Disease. *ChemMedChem* **2010**, *6*, 115–130.

(185) Xue, F.; Huang, J.; Ji, H.; Fang, J.; Li, H.; Martásek, P.; Roman, L. J.; Poulos, T. L.; Silverman, R. B. Structure-based design, synthesis, and biological evaluation of lipophilic-tailed monocationic inhibitors of neuronal nitric oxide synthase. *Bioorganic & Medicinal Chemistry* **2010**, *18*, 6526–6537.

(186) Neubig, R. R.; Spedding, M.; Kenakin, T.; Christopoulos, A. International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on Terms and Symbols in Quantitative Pharmacology. *Pharmacol Rev* **2003**, *55*, 597–606.

(187) Cer, R. Z.; Mudunuri, U.; Stephens, R.; Lebeda, F. J. IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of enzyme activity and ligand binding. *Nucleic Acids Research* **2009**, *37*, W441–W445.

(188) Goličnik, M.; Stojan, J. Slow-binding inhibition: A theoretical and practical course for students. *Biochemistry and Molecular Biology Education* **2006**, *32*, 228–235.

(189) Wei, M.; Wynn, R.; Hollis, G.; Liao, B.; Margulis, A.; Reid, B. G.; Klabe, R.; Liu, P. C. C.; Becker-Pasha, M.; Rupar, M.; Burn, T. C.; McCall, D. E.; Li, Y. High-throughput determination of mode of inhibition in lead identification and optimization. *J Biomol Screen* **2007**, *12*, 220–228.

(190) Murphy, D. J. Determination of accurate KI values for tight-binding enzyme inhibitors: an in silico study of experimental error and assay design. *Anal. Biochem.* **2004**, *327*, 61–67.

(191) IUBMB http://www.chem.qmul.ac.uk/iubmb/enzyme/EC3/4/22/15.html.

(192) Shuja, S.; Sheahan, K.; Murnane, M. J. Cysteine endopeptidase activity levels in normal human tissues, colorectal adenomas and carcinomas. *International Journal of Cancer* **2006**, *49*, 341–346.

(193) The Comprehensive Enzyme Information System http://www.brenda-enzymes.org/php/result_flat.php4?ecno=3.4.22.15.

(194) KIRSCHKE, H.; LANGER, J.; WIEDERANDERS, B.; ANSORGE, S.; BOHLEY, P. Cathepsin L. *European Journal of Biochemistry* **1977**, *74*, 293–301.

(195) Chauhan, S. S.; Popescu, N. C.; Ray, D.; Fleischmann, R.; Gottesman, M. M.; Troen, B. R. Cloning, Genomic Organization, and Chromosomal Localization of Human Cathepsin L. *J. Biol. Chem.* **1993**, *268*, 1039–1045.

(196) Palungwachira, P.; Kakuta, M.; Yamazaki, M.; Yaguchi, H.; Tsuboi, R.; Takamori, K.; Ogawa, H. Immunohistochemical localization of cathepsin L and cystatin A in normal skin and skin tumors. *J. Dermatol.* **2002**, *29*, 573–579.

(197) Yokota, S.; Nishimura, Y.; Kato, K. Localization of cathepsin L in rat kidney revealed by immunoenzyme and immunogold techniques. *Histochemistry and Cell Biology* **1988**, *90*, 277–283.

(198) Uchiyama, Y.; Watanabe, T.; Watanabe, M.; Ishii, Y.; Matsuba, H.; Waguri, S.; Kominami, E. Immunocytochemical localization of cathepsins B, H, L, and T4 in follicular cells of rat thyroid gland. *J. Histochem. Cytochem.* **1989**, *37*, 691–696.

(199) Ryvnyak, V. V.; Ryvnyak, E. I.; Tudos, R. V. Electron histochemical localization of cathepsin L in the liver. *Bull. Exp. Biol. Med.* **2004**, *137*, 90–91.

(200) Meng, W.; Yonenobu, K.; Ariga, K.; Nakase, T.; Okuda, S.; Obata, K.; Yoshikawa, H. Localization of cathepsins G and L in spontaneous resorption of intervertebral discs in a rat experimental model. *J Musculoskelet Neuronal Interact* **2001**, *2*, 171–176.

(201) Kiyoshima, T.; Kido, M. A.; Nishimura, Y.; Himeno, M.; Tsukuba, T.; Tashiro, H.; Yamamoto, K.; Tanaka, T. Immunocytochemical localization of cathepsin L in the synovial lining cells of the rat temporomandibular joint. *Archives of Oral Biology* **1994**, *39*, 1049–1056.

(202) Goulet, B.; Baruch, A.; Moon, N.-S.; Poirier, M.; Sansregret, L. L.; Erickson, A.; Bogyo, M.; Nepveu, A. A Cathepsin L Isoform that Is Devoid of a Signal Peptide Localizes to the Nucleus in S Phase and Processes the CDP/Cux Transcription Factor. *Molecular Cell* **2004**, *14*, 207–219.

(203) Sullivan, S.; Tosetto, M.; Kevans, D.; Coss, A.; Wang, L.; O'Donoghue, D.; Hyland, J.; Sheahan, K.; Mulcahy, H.; O'Sullivan, J. Localization of nuclear cathepsin L and its association with disease progression and poor outcome in colorectal cancer. *Int. J. Cancer* **2009**, *125*, 54–61.

(204) Yasothornsrikul, S.; Greenbaum, D.; Medzihradszky, K. F.; Toneff, T.; Bundey, R.; Miller, R.; Schilling, B.; Petermann, I.; Dehnert, J.; Logvinova, A.; Goldsmith, P.; Neveu, J. M.; Lane, W. S.; Gibson, B.; Reinheckel, T.; Peters, C.; Bogyo, M.; Hook, V. Cathepsin L in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9590–9595.

(205) Röcken, C.; Menard, R.; Bühling, F.; Vöckler, S.; Raynes, J.; Stix, B.; Krüger, S.; Roessner, A.; Kähne, T. Proteolysis of serum amyloid A and AA amyloid proteins by cysteine proteases: cathepsin B generates AA amyloid proteins and cathepsin L may prevent their formation. *Ann. Rheum. Dis.* **2005**, *64*, 808–815.

(206) Liu, Y.; Li, X.; Peng, D.; Tan, Z.; Liu, H.; Qing, Y.; Xue, Y.; Shi, G.-P. Usefulness of serum cathepsin L as an independent biomarker in patients with coronary heart disease. *Am. J. Cardiol.* **2009**, *103*, 476–481.

(207) Nagaya, T.; Murata, Y.; Yamaguchi, S.; Nomura, Y.; Ohmori, S.; Fujieda, M.; Katunuma, N.; Yen, P. M.; Chin, W. W.; Seo, H. Intracellular Proteolytic Cleavage of 9-Cis-Retinoic Acid Receptor A by Cathepsin L-Type Protease Is a Potential Mechanism for Modulating Thyroid Hormone Action. *J. Biol. Chem.* **1998**, *273*, 33166–33173.

(208) Boudreau, F.; Lussier, C. R.; Mongrain, S.; Darsigny, M.; Drouin, J. L.; Doyon, G.; Suh, E. R.; Beaulieu, J.-F.; Rivard, N.; Perreault, N. Loss of Cathepsin L Activity Promotes Claudin-1 Overexpression and Intestinal Neoplasia. *FASEB J* **2007**, *21*, 3853–3865.

(209) Kasai, M.; Shirasawa, T.; Kitamura, M.; Ishido, K.; Kominami, E.; Hirokawa, K. Proenzyme Form of Cathepsin L Produced by Thymic Epithelial Cells Promotes Proliferation of Immature Thymocytes in the Presence of IL-1, IL-7, and Anti-CD3 Antibody. *Cellular Immunology* **1993**, *150*, 124–136.

(210) Laha, T. T.; Hawley, M.; Rock, K. L.; Goldberg, A. L. Gamma-interferon causes a selective induction of the lysosomal proteases, cathepsins B and L, in macrophages. *FEBS Letters* **1995**, *363*, 85–89.

(211) Roth, W.; Deussing, J.; Botchkarev, V. A.; Pauly-Evers, M.; Saftig, P.; Hafner, A.; Schmidt, P.; Schmahl, W.; Scherer, J.; Anton-Lamprecht, I.; Von Figura, K.; Paus, R.; Peters, C. Cathepsin L deficiency as molecular defect of furless: hyperproliferation of keratinocytes and pertubation of hair follicle cycling. *FASEB J.* **2000**, *14*, 2075–2086.

(212) Charron, M.; Chern, J.-Y.; Wright, W. W. The cathepsin L first intron stimulates gene expression in rat sertoli cells. *Biol. Reprod.* **2007**, *76*, 813–824.

(213) Wright, W. W.; Smith, L.; Kerr, C.; Charron, M. Mice that express enzymatically inactive cathepsin L exhibit abnormal spermatogenesis. *Biol. Reprod.* **2003**, *68*, 680–687.

(214) Oksjoki, S.; Söderström, M.; Vuorio, E.; Anttila, L. Differential Expression Patterns of Cathepsins B, H, K, L and S in the Mouse Ovary. *Mol. Hum. Reprod.* **2001**, *7*, 27–34.

(215) Divya; Chhikara, P.; Mahajan, V. S.; Datta Gupta, S.; Chauhan, S. S. Differential activity of cathepsin L in human placenta at two different stages of gestation. *Placenta* **2002**, *23*, 59–64.

(216) Zhang, J.; Wang, P.; Huang, Y.-B.; Li, J.; Zhu, J.; Luo, X.; Shi, H.-M.; Li, Y. Plasma cathepsin L and its related pro/antiangiogenic factors play useful roles in predicting rich coronary collaterals in patients with coronary heart disease. *J. Int. Med. Res.* **2010**, *38*, 1389–1403.

(217) Nguyen, Q.; Mort, J. S.; Roughley, P. J. Cartilage proteoglycan aggregate is degraded more extensively by cathepsin L than by cathepsin B. *Biochem. J.* **1990**, *266*, 569–573.

(218) Kitamoto, S.; Sukhova, G. K.; Sun, J.; Yang, M.; Libby, P.; Love, V.; Duramad, P.; Sun, C.; Zhang, Y.; Yang, X.; Peters, C.; Shi, G.-P. Cathepsin L deficiency reduces diet-induced atherosclerosis in low-density lipoprotein receptor-knockout mice. *Circulation* **2007**, *115*, 2065–2075.

(219) Yang, M.; Zhang, Y.; Pan, J.; Sun, J.; Liu, J.; Libby, P.; Sukhova, G. K.; Doria, A.; Katunuma, N.; Peroni, O. D.; Guerre-Millo, M.; Kahn, B. B.; Clement, K.; Shi, G.-P. Cathepsin L activity controls adipogenesis and glucose tolerance. *Nat. Cell Biol.* **2007**, *9*, 970–977.

(220) Guinec, N.; Dalet-Fumeron, V.; Pagano, M. “In vitro” study of basement membrane degradation by the cysteine proteinases, cathepsins B, B-like and L. Digestion of collagen IV, laminin, fibronectin, and release of gelatinase activities from basement membrane fibronectin. *Biol. Chem. Hoppe-Seyler* **1993**, *374*, 1135–1146.

(221) Page, A. E.; Hayman, A. R.; Andersson, L. M.; Chambers, T. J.; Warburton, M. J. Degradation of bone matrix proteins by osteoclast cathepsins. *Int. J. Biochem.* **1993**, *25*, 545–550.

(222) Lecaille, F.; Chowdhury, S.; Purisima, E.; Brömme, D.; Lalmanach, G. The S2 subsites of cathepsins K and L and their contribution to collagen degradation. *Protein Sci* **2007**, *16*, 662–670.

(223) Maciewicz, R. A.; Wotton, S. F.; Etherington, D. J.; Duance, V. C. Susceptibility of the cartilage collagens types II, IX and XI to degradation by the cysteine proteinases, cathepsins B and L. *FEBS Lett.* **1990**, *269*, 189–193.

(224) Lodish, H.; Zipursky, SL Collagen: The Fibrous Proteins of the Matrix - Molecular Cell Biology - NCBI Bookshelf http://www.ncbi.nlm.nih.gov/books/NBK21582/ (accessed Jun 17, 2012).

(225) Choe, Y.; Leonetti, F.; Greenbaum, D. C.; Lecaille, F.; Bogyo, M.; Brömme, D.; Ellman, J. A.; Craik, C. S. Substrate profiling of cysteine proteases using a combinatorial peptide library identifies functionally unique specificities. *J. Biol. Chem.* **2006**, *281*, 12824–12832.

(226) Ritonja, A.; Coetzer, T. H.; Pike, R. N.; Dennison, C. The amino acid sequences, structure comparisons and inhibition kinetics of sheep cathepsin L and sheep stefin B. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* **1996**, *114*, 193–198.

(227) Wada, K.; Takai, T.; Tanabe, T. Amino acid sequence of chicken liver cathepsin L. *Eur. J. Biochem.* **1987**, *167*, 13–18.

(228) Guncar, G.; Pungercic, G.; Klemencic, I.; Turk, V.; Turk, D. Crystal structure of MHC class II-associated p41 Ii fragment bound to cathepsin L reveals the structural basis for differentiation between cathepsins L and S. *EMBO J* **1999**, *18*, 793–803.

(229) Shenoy, R. T.; Sivaraman, J. PDB ID: 3OF9 http://www.rcsb.org/pdb/explore/explore.do?structureId=3OF9.

(230) Shenoy, R. T.; Sivaraman, J. Structural basis for reversible and irreversible inhibition of human cathepsin L by their respective dipeptidyl glyoxal and diazomethylketone inhibitors. *Journal of Structural Biology* **2011**, *173*, 14–19.

(231) PDB ID 1CS8 http://www.pdb.org/pdb/explore/explore.do?structureId=1CS8.

(232) Reilly, J. J., Jr; Mason, R. W.; Chen, P.; Joseph, L. J.; Sukhatme, V. P.; Yee, R.; Chapman, H. A., Jr Synthesis and processing of cathepsin L, an elastase, by human alveolar macrophages. *Biochem. J.* **1989**, *257*, 493–498.

(233) Sever, S.; Altintas, M. M.; Nankoe, S. R.; Möller, C. C.; Ko, D.; Wei, C.; Henderson, J.; del Re, E. C.; Hsing, L.; Erickson, A.; Cohen, C. D.; Kretzler, M.; Kerjaschki, D.; Rudensky, A.; Nikolic, B.; Reiser, J. Proteolytic processing of dynamin by cytoplasmic cathepsin L is a mechanism for proteinuric kidney disease. *Journal of Clinical Investigation* **2007**, *117*, 2095–2104.

(234) Ishidoh, K.; Kominami, E. Multi-step processing of procathepsin L in vitro. *FEBS Letters* **1994**, *352*, 281–284.

(235) Hibbetts, K.; Hines, B.; Williams, D. An Overview of Proteinase Inhibitors. *Journal of Veterinary Internal Medicine* **2008**, *13*, 302–308.

(236) Cheng, T.; Hitomi, K.; Vlijmen-Willems, V.; J, I. M. J.; Jongh, D.; J, G.; Yamamoto, K.; Nishi, K.; Watts, C.; Reinheckel, T.; Schalkwijk, J.; Zeeuwen, P. L. J. M. Cystatin M/E Is a High Affinity Inhibitor of Cathepsin V and Cathepsin L by a Reactive Site That Is Distinct from the Legumain-Binding Site A NOVEL CLUE FOR THE ROLE OF CYSTATIN M/E IN EPIDERMAL CORNIFICATION. *J. Biol. Chem.* **2006**, *281*, 15893–15899.

(237) Žerovnik, E. *Human Stefins And Cystatins*; Nova Publishers, 2006.

(238) Kos, J.; Krašovec, M.; Cimerman, N.; Nielsen, H. J.; Christensen, I. J.; Brünner, N. Cysteine Proteinase Inhibitors Stefin A, Stefin B, and Cystatin C in Sera from Patients with Colorectal Cancer: Relation to Prognosis. *Clin Cancer Res* **2000**, *6*, 505–511.

(239) Hsieh, W.-T.; Fong, D.; Sloane, B. F.; Golembieski, W.; Smith, D. I. Mapping of the gene for human cysteine proteinase inhibitor stefin A, STF1, to chromosome 3cen-q21. *Genomics* **1991**, *9*, 207–209.

(240) Henskens, Y. M. C.; Veerman, E. C. I.; Mantel, M. S.; Van Der Velden, U.; Nieuw Amerongen, A. V. Cystatins S and C in Human Whole Saliva and in Glandular Salivas in Periodontal Health and Disease. *J DENT RES* **1994**, *73*, 1606–1614.

(241) Welss, T.; Sun, J.; Irving, J. A.; Blum, R.; Smith, A. I.; Whisstock, J. C.; Pike, R. N.; von Mikecz, A.; Ruzicka, T.; Bird, P. I.; Abts, H. F. Hurpin is a selective inhibitor of lysosomal cathepsin L and protects keratinocytes from ultraviolet-induced apoptosis. *Biochemistry* **2003**, *42*, 7381–7389.

(242) Schirmeister, T.; Kaeppler, U. Non-Peptidic Inhibitors of Cysteine Proteases. *Mini-Reviews in Medicinal Chemistry* **2003**, *3*, 361–373.

(243) Shenoy, R. T.; Chowdhury, S. F.; Kumar, S.; Joseph, L.; Purisima, E. O.; Sivaraman, J. A Combined Crystallographic and Molecular Dynamics Study of Cathepsin L Retrobinding Inhibitors. *J. Med. Chem.* **2009**, *52*, 6335–6346.

(244) Shah, P. P.; Wang, T.; Kaletsky, R. L.; Myers, M. C.; Purvis, J. E.; Jing, H.; Huryn, D. M.; Greenbaum, D. C.; Smith, A. B.; Bates, P.; Diamond, S. L. A Small-Molecule Oxocarbazate Inhibitor of Human Cathepsin L Blocks Severe Acute Respiratory Syndrome and Ebola Pseudotype Virus Infection into Human Embryonic Kidney 293T Cells. *Mol Pharmacol* **2010**, *78*, 319–324.

(245) Takahashi, K.; Ueno, T.; Tanida, I.; Minematsu-Ikeguchi, N.; Murata, M.; Kominami, E. Characterization of CAA0225, a Novel Inhibitor Specific for Cathepsin L, as a Probe for Autophagic Proteolysis. *Biological and Pharmaceutical Bulletin* **2009**, *32*, 475–479.

(246) Asaad, N.; Bethel, P. A.; Coulson, M. D.; Dawson, J. E.; Ford, S. J.; Gerhardt, S.; Grist, M.; Hamlin, G. A.; James, M. J.; Jones, E. V.; Karoutchi, G. I.; Kenny, P. W.; Morley, A. D.; Oldham, K.; Rankine, N.; Ryan, D.; Wells, S. L.; Wood, L.; Augustin, M.; Krapp, S.; Simader, H.; Steinbacher, S. Dipeptidyl nitrile inhibitors of Cathepsin L. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 4280–4283.

(247) Myers, M. C.; Shah, P. P.; Beavers, M. P.; Napper, A. D.; Diamond, S. L.; Smith III, A. B.; Huryn, D. M. Design, synthesis, and evaluation of inhibitors of cathepsin L: Exploiting a unique thiocarbazate chemotype. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 3646–3651.

(248) Wang, S.-Q.; Du, Q.-S.; Zhao, K.; Li, A.-X.; Wei, D.-Q.; Chou, K.-C. Virtual screening for finding natural inhibitor against cathepsin-L for SARS therapy. *Amino Acids* **2007**, *33*, 129–135.

(249) Marquis, R. W.; James, I.; Zeng, J.; Trout, R. E. L.; Thompson, S.; Rahman, A.; Yamashita, D. S.; Xie, R.; Ru, Y.; Gress, C. J.; Blake, S.; Lark, M. A.; Hwang, S.-M.; Tomaszek, T.; Offen, P.; Head, M. S.; Cummings, M. D.; Veber, D. F. Azepanone-Based Inhibitors of Human Cathepsin L. *J. Med. Chem.* **2005**, *48*, 6870–6878.

(250) Lynas, J. F.; Hawthorne, S. J.; Walker, B. Development of peptidyl α-keto-β-aldehydes as new inhibitors of cathepsin L - Comparisons of potency and selectivity profiles with cathepsin B. *Bioorganic and Medicinal Chemistry Letters* **2000**, *10*, 1771–1773.

(251) Spano, D.; Heck, C.; De Antonellis, P.; Christofori, G.; Zollo, M. Molecular networks that regulate cancer metastasis. *Seminars in Cancer Biology* **2012**, *22*, 234–249.

(252) Lankelma, J. M.; Voorend, D. M.; Barwari, T.; Koetsveld, J.; Van der Spek, A. H.; De Porto, A. P. N. A.; Van Rooijen, G.; Van Noorden, C. J. F. Cathepsin L, target in cancer treatment? *Life Sci.* **2010**, *86*, 225–233.

(253) Stoka, V.; Turk, B.; Schendel, S. L.; Kim, T. H.; Cirman, T.; Snipas, S. J.; Ellerby, L. M.; Bredesen, D.; Freeze, H.; Abrahamson, M.; Bromme, D.; Krajewski, S.; Reed, J. C.; Yin, X. M.; Turk, V.; Salvesen, G. S. Lysosomal protease pathways to apoptosis. Cleavage of bid, not pro-caspases, is the most likely route. *J. Biol. Chem.* **2001**, *276*, 3149–3157.

(254) Di Piazza, M.; Mader, C.; Geletneky, K.; Herrero Y Calle, M.; Weber, E.; Schlehofer, J.; Deleu, L.; Rommelaere, J. Cytosolic Activation of Cathepsins Mediates Parvovirus H-1-Induced Killing of Cisplatin and TRAIL-Resistant Glioma Cells. *J. Virol.* **2007**, *81*, 4186–4198.

(255) Gocheva, V.; Zeng, W.; Ke, D.; Klimstra, D.; Reinheckel, T.; Peters, C.; Hanahan, D.; Joyce, J. A. Distinct roles for cysteine cathepsin genes in multistage tumorigenesis. *Genes Dev* **2006**, *20*, 543–556.

(256) O'Reilly, M. S.; Boehm, T.; Shing, Y.; Fukai, N.; Vasios, G.; Lane, W. S.; Flynn, E.; Birkhead, J. R.; Olsen, B. R.; Folkman, J. Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. *Cell* **1997**, *88*, 277–285.

(257) Felbor, U.; Dreier, L.; Bryant, R. A. R.; Ploegh, H. L.; Olsen, B. R.; Mothes, W. Secreted cathepsin L generates endostatin from collagen XVIII. *EMBO J* **2000**, *19*, 1187–1194.

(258) Joyce, J. A.; Baruch, A.; Chehade, K.; Meyer-Morse, N.; Giraudo, E.; Tsai, F.-Y.; Greenbaum, D. C.; Hager, J. H.; Bogyo, M.; Hanahan, D. Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. *Cancer Cell* **2004**, *5*, 443–453.

(259) Urbich, C.; Heeschen, C.; Aicher, A.; Sasaki, K.; Bruhl, T.; Farhadi, M. R.; Vajkoczy, P.; Hofmann, W. K.; Peters, C.; Pennacchio, L. A.; Abolmaali, N. D.; Chavakis, E.; Reinheckel, T.; Zeiher, A. M.; Dimmeler, S. Cathepsin L is required for endothelial progenitor cell-induced neovascularization. *Nat. Med.* **2005**, *11*, 206–213.

(260) Colella, R.; Casey, S. F. Decreased activity of cathepsins L + B and decreased invasive ability of PC3 prostate cancer cells. *Biotech Histochem* **2003**, *78*, 101–108.

(261) Svatek, R. S.; Karam, J.; Karakiewicz, P. I.; Gallina, A.; Casella, R.; Roehrborn, C. G.; Shariat, S. F. Role of urinary cathepsin B and L in the detection of bladder urothelial cell carcinoma. *J. Urol.* **2008**, *179*, 478–484; discussion 484.

(262) Kayser, K.; Richter, N.; Hufnagl, P.; Kayser, G.; Kos, J.; Werle, B. Expression, proliferation activity and clinical significance of cathepsin B and cathepsin L in operated lung cancer. *Anticancer Res.* **2003**, *23*, 2767–2772.

(263) de la Taille, A.; Buttyan, R.; Katz, A. E.; McKiernan, J.; Burchardt, M.; Burchardt, T.; Chopin, D. K.; Sawczuk, I. S. Biomarkers of renal cell carcinoma: Past and future considerations. *Urologic Oncology: Seminars and Original Investigations* **2000**, *5*, 139–148.

(264) Stabuc, B.; Mrevlje, Z.; Markovic, J.; Stabuc-Silih, M. Expression and prognostic significance of Cathepsin L in early cutaneous malignant melanoma. *Neoplasma* **2006**, *53*, 259–262.

(265) Ishibashi, O.; Mori, Y.; Kurokawa, T.; Kumegawa, M. Breast cancer cells express cathepsins B and L but not cathepsins K or H. *Cancer Biochem. Biophys.* **1999**, *17*, 69–78.

(266) Nishida, Y.; Kohno, K.; Kawamata, T.; Morimitsu, K.; Kuwano, M.; Miyakawa, I. Increased cathepsin L levels in serum in some patients with ovarian cancer: comparison with CA125 and CA72-4. *Gynecol. Oncol.* **1995**, *56*, 357–361.

(267) Hsu, K.-F.; Wu, C.-L.; Huang, S.-C.; Wu, C.-M.; Hsiao, J.-R.; Yo, Y.-T.; Chen, Y.-H.; Shiau, A.-L.; Chou, C.-Y. Cathepsin L mediates resveratrol-induced autophagy and apoptotic cell death in cervical cancer cells. *Autophagy* **2009**, *5*, 451–460.

(268) Xu, X.; Yuan, G.; Liu, W.; Zhang, Y.; Chen, W. Expression of cathepsin L in nasopharyngeal carcinoma and its clinical significance. *Exp. Oncol.* **2009**, *31*, 102–105.

(269) Saleh, Y.; Siewiński, M.; Kielan, W.; Ziółkowski, P.; Gryboś, M.; Rybka, J. Regulation of cathepsin B and L expression in vitro in gastric cancer tissues by egg cystatin. *J. Exp. Ther. Oncol.* **2003**, *3*, 319–324.

(270) Levi|[ccaron]|ar, N.; Dewey, R. A.; Daley, E.; Bates, T. E.; Davies, D.; Kos, J.; Pilkington, G. J.; Lah, T. T. Selective suppression of cathepsin L by antisense cDNA impairs human brain tumor cell invasion in vitro and promotes apoptosis. *Cancer Gene Therapy* **2003**, *10*, 141–151.

(271) Thomssen, C.; Oppelt, P.; Jänicke, F.; Ulm, K.; Harbeck, N.; Höfler, H.; Kuhn, W.; Graeff, H.; Schmitt, M. Identification of low-risk node-negative breast cancer patients by tumor biological factors PAI-1 and cathepsin L. *Anticancer Res.* **1998**, *18*, 2173–2180.

(272) Zajc, I.; Sever, N.; Bervar, A.; Lah, T. T. Expression of cysteine peptidase cathepsin L and its inhibitors stefins A and B in relation to tumorigenicity of breast cancer cell lines. *Cancer Letters* **2002**, *187*, 185–190.

(273) Foekens, J. A.; Kos, J.; Peters, H. A.; Krasovec, M.; Look, M. P.; Cimerman, N.; Meijer-Van Gelder, M. E.; Henzen-Logmans, S. C.; Van Putten, W. L.; Klijn, J. G. Prognostic Significance of Cathepsins B and L in Primary Human Breast Cancer. *JCO* **1998**, *16*, 1013–1021.

(274) Shaw, L. M. Tumor cell invasion assays. *Methods Mol. Biol.* **2005**, *294*, 97–105.

(275) Hulkower, K. I.; Herber, R. L. Cell Migration and Invasion Assays as Tools for Drug Discovery. *Pharmaceutics* **2011**, *3*, 107–124.

(276) Albini, A.; Noonan, D. M. The "chemoinvasion" assay, 25 years and still going strong: the use of reconstituted basement membranes to study cell invasion and angiogenesis. *Current Opinion in Cell Biology* **2010**, *22*, 677–689.

(277) BD BIosciences Matrigel http://www.bdbiosciences.com/cellculture/ecm/ecmtypes/index.jsp.

(278) Lu, P.; Weaver, V. M.; Werb, Z. The Extracellular Matrix: A Dynamic Niche in Cancer Progression. *J Cell Biol* **2012**, *196*, 395–406.

(279) Hanein, D.; Horwitz, A. R. The structure of cell–matrix adhesions: the new frontier. *Current Opinion in Cell Biology* **2012**, *24*, 134–140.

(280) Huang, G.; Greenspan, D. S. ECM roles in the function of metabolic tissues. *Trends in Endocrinology & Metabolism* **2012**, *23*, 16–22.

(281) Yates, C.; Shepard, C. R.; Papworth, G.; Dash, A.; Beer Stolz, D.; Tannenbaum, S.; Griffith, L.; Wells, A. Novel Three-Dimensional Organotypic Liver Bioreactor to Directly Visualize Early Events in Metastatic Progression. In *Advances in Cancer Research*; Academic Press, 2007; Vol. Volume 97, pp. 225–246.

(282) Brandt, B.; Heyder, C.; Gloria-Maercker, E.; Hatzmann, W.; Rötger, A.; Kemming, D.; Zänker, K. S.; Entschladen, F.; Dittmar, T. 3D-extravasation model – selection of highly motile and metastatic cancer cells. *Seminars in Cancer Biology* **2005**, *15*, 387–395.

(283) Wolf, K.; Friedl, P. Extracellular matrix determinants of proteolytic and non-proteolytic cell migration. *Trends in Cell Biology* **2011**, *21*, 736–744.

(284) Eccles, S. A.; Box, C.; Court, W. Cell migration/invasion assays and their application in cancer drug discovery. In *Biotechnology Annual Review*; Elsevier, 2005; Vol. Volume 11, pp. 391–421.

(285) Chatterjee, S. K.; Zetter, B. R. Cancer biomarkers: knowing the present and predicting the future. *Future Oncology* **2005**, *1*, 37–50.

(286) Blum, G.; Degenfeld, G. von; Merchant, M. J.; Blau, H. M.; Bogyo, M. Noninvasive optical imaging of cysteine protease activity using fluorescently quenched activity-based probes. *Nature Chemical Biology* **2007**, *3*, 668–677.

(287) Mason, S. D.; Joyce, J. A. Proteolytic networks in cancer. *Trends Cell Biol.* **2011**, *21*, 228–237.

(288) Hamilton, G.; Colbert, J. D.; Schuettelkopf, A. W.; Watts, C. Cystatin F is a cathepsin C-directed protease inhibitor regulated by proteolysis. *EMBO J* **2008**, *27*, 499–508.

(289) Jokimaa, V.; Oksjoki, S.; Kujari, H.; Vuorio, E.; Anttila, L. Expression patterns of cathepsins B, H, K, L and S in the human endometrium. *Mol. Hum. Reprod.* **2001**, *7*, 73–78.

(290) Mohamed, M. M.; Sloane, B. F. Cysteine cathepsins: multifunctional enzymes in cancer. *Nature Reviews Cancer* **2006**, *6*, 764–775.

(291) Davydova, O.; Yakovlev, A. Protease-activated receptors and neuroplasticity: Protease-activated receptors as a possible target for cathepsin B. *Neurochemical Journal* **2010**, *4*, 1–7.

(292) Zheng, X.; Chu, F.; Mirkin, B. L.; Sudha, T.; Mousa, S. A.; Rebbaa, A. Role of the proteolytic hierarchy between cathepsin L, cathepsin D and caspase-3 in regulation of cellular susceptibility to apoptosis and autophagy. *Biochim. Biophys. Acta* **2008**, *1783*, 2294–2300.

(293) Turk, V.; Turk, B.; Turk, D. Lysosomal cysteine proteases: facts and opportunities. *The EMBO Journal* **2001**, *20*, 4629–4633.

(294) Hauck, C. R.; Hsia, D. A.; Schlaepfer, D. D. The focal adhesion kinase--a regulator of cell migration and invasion. *IUBMB Life* **2002**, *53*, 115–119.

(295) Hsieh, Y.-H.; Wu, T.-T.; Huang, C.-Y.; Hsieh, Y.-S.; Hwang, J.-M.; Liu, J.-Y. P38 Mitogen-Activated Protein Kinase Pathway Is Involved in Protein Kinase Cα–Regulated Invasion in Human Hepatocellular Carcinoma Cells. *Cancer Res* **2007**, *67*, 4320–4327.

(296) Hood, J. D.; Cheresh, D. A. Role of integrins in cell invasion and migration. *Nat. Rev. Cancer* **2002**, *2*, 91–100.

(297) Schlaepfer, D. D.; Mitra, S. K. Multiple connections link FAK to cell motility and invasion. *Current Opinion in Genetics & Development* **2004**, *14*, 92–101.

(298) Desgrosellier, J. S.; Cheresh, D. A. Integrins in cancer: biological implications and therapeutic opportunities. *Nature Reviews Cancer* **2010**, *10*, 9–22.

(299) Alexander, S.; Friedl, P. Cancer invasion and resistance: interconnected processes of disease progression and therapy failure. *Trends in Molecular Medicine* **2012**, *18*, 13–26.

(300) Steeg, P. S. Tumor metastasis: mechanistic insights and clinical challenges. *Nature Medicine* **2006**, *12*, 895–904.

(301) Mundy, G. R. Metastasis: Metastasis to bone: causes, consequences and therapeutic opportunities. *Nature Reviews Cancer* **2002**, *2*, 584–593.

(302) Roato, I.; D'Amelio, P.; Gorassini, E.; Grimaldi, A.; Bonello, L.; Fiori, C.; Delsedime, L.; Tizzani, A.; De Libero, A.; Isaia, G.; Ferracini, R. Osteoclasts Are Active in Bone Forming Metastases of Prostate Cancer Patients. *PLoS ONE* **2008**, *3*, e3627.

(303) Guise, T. A.; Yin, J. J.; Mohammad, K. S. Role of endothelin-1 in osteoblastic bone metastases. *Cancer* **2003**, *97*, 779–784.

(304) Guise, T. A. The vicious cycle of bone metastases. *Journal of Musculoskeletal Neuronal Interactions* **2002**, *2*, 570–572.

(305) Tanaka, S.; Nakamura, K.; Takahasi, N.; Suda, T. Role of RANKL in physiological and pathological bone resorption and therapeutics targeting the RANKL-RANK signaling system. *Immunol. Rev.* **2005**, *208*, 30–49.

(306) Wolf, A.; Couttet, P.; Dong, M.; Grenet, O.; Heron, M.; Junker, U.; Laengle, U.; Ledieu, D.; Marrer, E.; Nussher, A.; Persohn, E.; Pognan, F.; Rivière, G.-J.; Roth, D. R.; Trendelenburg, C.; Tsao, J.; Roman, D. Imatinib does not induce cardiotoxicity at clinically relevant concentrations in preclinical studies. *Leuk. Res.* **2010**, *34*, 1180–1188.

(307) Rennert, G. Bisphosphonates: Beyond Prevention of Bone Metastases. *JNCI J Natl Cancer Inst* **2011**.

(308) Drake, J. M.; Danke, J. R.; Henry, M. D. Bone-specific growth inhibition of prostate cancer metastasis by atrasentan. *Cancer Biol. Ther.* **2010**, *9*, 607–614.

(309) Smith, M. R.; Saad, F.; Coleman, R.; Shore, N.; Fizazi, K.; Tombal, B.; Miller, K.; Sieber, P.; Karsh, L.; Damião, R.; Tammela, T. L.; Egerdie, B.; Van Poppel, H.; Chin, J.; Morote, J.; Gómez-Veiga, F.; Borkowski, T.; Ye, Z.; Kupic, A.; Dansey, R.; Goessl, C. Denosumab and bone-metastasis-free survival in men with castration-resistant prostate cancer: results of a phase 3, randomised, placebo-controlled trial. *Lancet* **2012**, *379*, 39–46.

(310) Leto, G.; Sepporta, M. V.; Crescimanno, M.; Flandina, C.; Tumminello, F. M. Cathepsin L in metastatic bone disease: therapeutic implications. *Biol. Chem.* **2010**, *391*, 655–664.

(311) Damiens, C.; Grimaud, E.; Rousselle, A. V.; Charrier, C.; Fortun, Y.; Heymann, D.; Padrines, M. Cysteine protease production by human osteosarcoma cells (MG63, SAOS2) and its modulation by soluble factors. *Cytokine* **2000**, *12*, 539–542.

(312) Furuyama, N.; Fujisawa, Y. Regulation of collagenolytic protease secretion through c-Src in osteoclasts. *Biochem. Biophys. Res. Commun.* **2000**, *272*, 116–124.

(313) Katunuma, N.; Tsuge, H.; Nukatsuka, M.; Asao, T.; Fukushima, M. Structure-based design of specific cathepsin inhibitors and their application to protection of bone metastases of cancer cells. *Arch. Biochem. Biophys.* **2002**, *397*, 305–311.

(314) Zheng, X.; Chu, F.; Chou, P. M.; Gallati, C.; Dier, U.; Mirkin, B. L.; Mousa, S. A.; Rebbaa, A. Cathepsin L inhibition suppresses drug resistance in vitro and in vivo: a putative mechanism. *Am. J. Physiol., Cell Physiol.* **2009**, *296*, C65–74.

(315) Abboud-Jarrous, G.; Rangini-Guetta, Z.; Aingorn, H.; Atzmon, R.; Elgavish, S.; Peretz, T.; Vlodavsky, I. Site-directed mutagenesis, proteolytic cleavage, and activation of human proheparanase. *Journal of Biological Chemistry* **2005**, *280*, 13568–13575.

(316) Keerthivasan, S.; Keerthivasan, G.; Mittal, S.; Chauhan, S. S. Transcriptional upregulation of human cathepsin L by VEGF in glioblastoma cells. *Gene* **2007**, *399*, 129–136.

(317) Colella, R.; Lu, G.; Glazewski, L.; Korant, B.; Matlapudi, A.; England, M. R.; Craft, C.; Frantz, C. N.; Mason, R. W. Induction of cell death in neuroblastoma by inhibition of cathepsins B and L. *Cancer Lett* **2010**, *294*, 195–203.

(318) Cailhier, J.-F.; Sirois, I.; Laplante, P.; Lepage, S.; Raymond, M.-A.; Brassard, N.; Prat, A.; Iozzo, R. V.; Pshezhetsky, A. V.; Hébert, M.-J. Caspase-3 Activation Triggers Extracellular Cathepsin L Release and Endorepellin Proteolysis. *J. Biol. Chem.* **2008**, *283*, 27220–27229.

(319) Margulis, A.; Zhang, W.; Alt-Holland, A.; Crawford, H. C.; Fusenig, N. E.; Garlick, J. A. E-Cadherin Suppression Accelerates Squamous Cell Carcinoma Progression in Three-Dimensional, Human Tissue Constructs. *Cancer Res* **2005**, *65*, 1783–1791.

(320) *Lab Ref, Volume 1: A Handbook of Recipes, Reagents, and Other Reference Tools for Use at the Bench*; Roskams, J.; Rodgers, L., Eds.; Cold Spring Harbor Laboratory Press, 2002.

(321) Cloutier, S. M.; Kündig, C.; Felber, L. M.; Fattah, O. M.; Chagas, J. R.; Gygi, C. M.; Jichlinski, P.; Leisinger, H.-J.; Deperthes, D. Development of recombinant inhibitors specific to human kallikrein 2 using phage-display selected substrates. *European Journal of Biochemistry* **2004**, *271*, 607–613.

(322) Baeriswyl, V.; Rapley, H.; Pollaro, L.; Stace, C.; Teufel, D.; Walker, E.; Chen, S.; Winter, G.; Tite, J.; Heinis, C. Bicyclic Peptides with Optimized Ring Size Inhibit Human Plasma Kallikrein and its Orthologues While Sparing Paralogous Proteases. *ChemMedChem* **2012**.

(323) Wu, G. *Assay Development: Fundamentals and Practices*; John Wiley & Sons, 2010.

(324) Setti, E. L.; Davis, D.; Chung, T.; McCarter, J. 3,4-Disubstituted azetidinones as selective inhibitors of the cysteine protease cathepsin K. Exploring P2 elements for selectivity. *Bioorganic & Medicinal Chemistry Letters* **2003**, *13*, 2051–2053.

(325) Harris, A. L.; Gregory, J. S.; Maycock, A. L.; Graybill, T. L.; Osifo, I. K.; Schmidt, S. J.; Dolle, R. E. Characterization of a continuous fluorogenic assay for calpain I. Kinetic evaluation of peptide aldehydes, halomethyl ketones and (acyloxy)methyl ketones as inhibitors of the enzyme. *Bioorganic & Medicinal Chemistry Letters* **1995**, *5*, 393–398.

(326) Shah, P. P.; Myers, M. C.; Beavers, M. P.; Purvis, J. E.; Jing, H.; Grieser, H. J.; Sharlow, E. R.; Napper, A. D.; Huryn, D. M.; Cooperman, B. S.; Smith, A. B., 3rd; Diamond, S. L. Kinetic characterization and molecular docking of a novel, potent, and selective slow-binding inhibitor of human cathepsin L. *Mol. Pharmacol.* **2008**, *74*, 34–41.

(327) Mallari, J. P.; Shelat, A.; Kosinski, A.; Caffrey, C. R.; Connelly, M.; Zhu, F.; McKerrow, J. H.; Guy, R. K. Discovery of trypanocidal thiosemicarbazone inhibitors of rhodesain and TbcatB. *Bioorg Med Chem Lett* **2008**, *18*, 2883–2885.

(328) Dash, C.; Vathipadiekal, V.; George, S. P.; Rao, M. Slow-tight binding inhibition of xylanase by an aspartic protease inhibitor: kinetic parameters and conformational changes that determine the affinity and selectivity of the bifunctional nature of the inhibitor. *J. Biol. Chem.* **2002**, *277*, 17978–17986.

(329) Hang, J. Q.; Yang, Y.; Harris, S. F.; Leveque, V.; Whittington, H. J.; Rajyaguru, S.; Ao-Ieong, G.; McCown, M. F.; Wong, A.; Giannetti, A. M.; Le Pogam, S.; Talamás, F.; Cammack, N.; Nájera, I.; Klumpp, K. Slow binding inhibition and mechanism of resistance of non-nucleoside polymerase inhibitors of hepatitis C virus. *J. Biol. Chem.* **2009**, *284*, 15517–15529.

(330) Vogel, K. G.; Paulsson, M.; Heinegård, D. Specific inhibition of type I and type II collagen fibrillogenesis by the small proteoglycan of tendon. *Biochem. J.* **1984**, *223*, 587–597.

(331) Zhang, Z.; Li, G.; Shi, B. Physicochemical properties of collagen, gelatin and collagen hydrolysate derived from bovine limed split wastes. *Journal of the Society of Leather Technologies and Chemists* **2006**, *90*, 23–28.

(332) Imeson, A. *Thickening and Gelling Agents for Food*; Springer, 1997.

(333) Parenteau-Bareil, R.; Gauvin, R.; Berthod, F. Collagen-Based Biomaterials for Tissue Engineering Applications. *Materials* **2010**, *3*, 1863–1887.

(334) Telis, V.; Wolf, K.; Sobral, P. Characterizations of Collagen Fibers for Biodegradable Films Production. In; EDP Sciences, 2006.

(335) PDB ID: 1ICF http://www.pdb.org/pdb/explore/explore.do?structureId=1ICF.

(336) Shoji-Kasai, Y.; Senshu, M.; Iwashita, S.; Imahori, K. Thiol protease-specific inhibitor E-64 arrests human epidermoid carcinoma A431 cells at mitotic metaphase. *Proc Natl Acad Sci U S A* **1988**, *85*, 146–150.

(337) Goodyear, S.; Sharma, M. C. Roscovitine regulates invasive breast cancer cell (MDA-MB231) proliferation and survival through cell cycle regulatory protein cdk5. *Experimental and Molecular Pathology* **2007**, *82*, 25–32.

(338) Enzyme Nomenclature http://www.chem.qmul.ac.uk/iubmb/enzyme/EC34/3422.html#2238 (accessed Jan 6, 2012).

(339) Lecaille, F.; Brömme, D.; Lalmanach, G. Biochemical properties and regulation of cathepsin K activity. *Biochimie* **2008**, *90*, 208–226.

(340) Nägler, D. K.; Ménard, R. Family C1 cysteine proteases: biological diversity or redundancy? *Biol. Chem.* **2003**, *384*, 837–843.

(341) Tezuka, K.; Tezuka, Y.; Maejima, A.; Sato, T.; Nemoto, K.; Kamioka, H.; Hakeda, Y.; Kumegawa, M. Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts. *Journal of Biological Chemistry* **1994**, *269*, 1106–9.

(342) Bossard, M. J.; Tomaszek, T. A.; Thompson, S. K.; Amegadzie, B. Y.; Hanning, C. R.; Jones, C.; Kurdyla, J. T.; McNulty, D. E.; Drake, F. H.; Gowen, M.; Levy, M. A. Proteolytic activity of human osteoclast cathepsin K. Expression, purification, activation, and substrate identification. *J. Biol. Chem.* **1996**, *271*, 12517–12524.

(343) Bühling, F.; Reisenauer, A.; Gerber, A.; Krüger, S.; Weber, E.; Brömme, D.; Roessner, A.; Ansorge, S.; Welte, T.; Röcken, C. Cathepsin K--a marker of macrophage differentiation? *J. Pathol.* **2001**, *195*, 375–382.

(344) Haeckel, C.; Krueger, S.; Buehling, F.; Broemme, D.; Franke, K.; Schuetze, A.; Roese, I.; Roessner, A. Expression of cathepsin K in the human embryo and fetus. *Developmental Dynamics* **1999**, *216*, 89–95.

(345) Quintanilla-Dieck, M. J.; Codriansky, K.; Keady, M.; Bhawan, J.; Rünger, T. M. Expression and regulation of cathepsin K in skin fibroblasts. *Exp. Dermatol.* **2009**, *18*, 596–602.

(346) Zhao, Q.; Jia, Y.; Xiao, Y. Cathepsin K: A therapeutic target for bone diseases. *Biochemical and Biophysical Research Communications* **2009**, *380*, 721–723.

(347) Brömme, D.; Okamoto, K. Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution. *Biol. Chem. Hoppe-Seyler* **1995**, *376*, 379–384.

(348) R|[uuml]|nger, T. M.; Quintanilla-Dieck, M. J.; Bhawan, J. Role of Cathepsin K in the Turnover of the Dermal Extracellular Matrix during Scar Formation. *Journal of Investigative Dermatology* **2007**, *127*, 293–297.

(349) Hou, W.-S.; Li, Z.; Büttner, F. H.; Bartnik, E.; Brömme, D. Cleavage site specificity of cathepsin K toward cartilage proteoglycans and protease complex formation. *Biol. Chem.* **2003**, *384*, 891–897.

(350) Codriansky, K. A.; Quintanilla-Dieck, M. J.; Gan, S.; Keady, M.; Bhawan, J.; Rünger, T. M. Intracellular degradation of elastin by cathepsin K in skin fibroblasts--a possible role in photoaging. *Photochem. Photobiol.* **2009**, *85*, 1356–1363.

(351) Garnero, P.; Borel, O.; Byrjalsen, I.; Ferreras, M.; Drake, F. H.; McQueney, M. S.; Foged, N. T.; Delmas, P. D.; Delaissé, J.-M. The Collagenolytic Activity of Cathepsin K Is Unique Among Mammalian Proteinases. *J. Biol. Chem.* **1998**, *273*, 32347–32352.

(352) Dejica, V. M.; Mort, J. S.; Laverty, S.; Percival, M. D.; Antoniou, J.; Zukor, D. J.; Poole, A. R. Cleavage of type II collagen by cathepsin K in human osteoarthritic cartilage. *Am. J. Pathol.* **2008**, *173*, 161–169.

(353) Golovatch, P.; Mercer, B. A.; Lemaître, V.; Wallace, A.; Foronjy, R. F.; D'Armiento, J. Role for Cathepsin K in emphysema in smoke-exposed guinea pigs. *Exp Lung Res* **2009**, *35*, 631–645.

(354) Quintanilla-Dieck, M. J.; Codriansky, K.; Keady, M.; Bhawan, J.; R|[uuml]|nger, T. M. Cathepsin K in Melanoma Invasion. *Journal of Investigative Dermatology* **2008**, *128*, 2281–2288.

(355) Etherington, D. J. Collagen degradation. *Ann Rheum Dis* **1977**, *36*, 14–17.

(356) Troeberg, L.; Nagase, H. Proteases involved in cartilage matrix degradation in osteoarthritis. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **2012**, *1824*, 133–145.

(357) Selent, J.; Kaleta, J.; Li, Z.; Lalmanach, G.; Brömme, D. Selective Inhibition of the Collagenase Activity of Cathepsin K. *J. Biol. Chem.* **2007**, *282*, 16492–16501.

(358) Atley, L. .; Mort, J. .; Lalumiere, M.; Eyre, D. . Proteolysis of human bone collagen by cathepsin K: characterization of the cleavage sites generating the cross-linked N-telopeptide neoepitope. *Bone* **2000**, *26*, 241–247.

(359) Lecaille, F.; Weidauer, E.; Juliano, M. A.; Brömme, D.; Lalmanach, G. Probing cathepsin K activity with a selective substrate spanning its active site. *Biochem. J.* **2003**, *375*, 307–312.

(360) Eastoe, J. E. The amino acid composition of mammalian collagen and gelatin. *Biochem J* **1955**, *61*, 589–600.

(361) Troen, B. R. The role of cathepsin K in normal bone resorption. *Drug News Perspect.* **2004**, *17*, 19–28.

(362) Brömme, D.; Okamoto, K.; Wang, B. B.; Biroc, S. Human Cathepsin O2, a Matrix Protein-Degrading Cysteine Protease Expressed in Osteoclasts FUNCTIONAL EXPRESSION OF HUMAN CATHEPSIN O2 IN SPODOPTERA FRUGIPERDA AND CHARACTERIZATION OF THE ENZYME. *J. Biol. Chem.* **1996**, *271*, 2126–2132.

(363) Bühling, F.; Peitz, U.; Krüger, S.; Küster, D.; Vieth, M.; Gebert, I.; Roessner, A.; Weber, E.; Malfertheiner, P.; Wex, T. Cathepsins K, L, B, X and W are differentially expressed in normal and chronically inflamed gastric mucosa. *Biological Chemistry* **2004**, *385*, 439–445.

(364) Lee, C. H.; Singla, A.; Lee, Y. Biomedical applications of collagen. *International Journal of Pharmaceutics* **2001**, *221*, 1–22.

(365) Meena, C.; Mengi, S.; Deshpande, S. Biomedical and industrial applications of collagen. *Journal of Chemical Sciences* **1999**, *111*, 319–329.

(366) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. Crystal and molecular structure of a collagen-like peptide at 1.9 A resolution. *Science* **1994**, *266*, 75–81.

(367) Vanacore, R. M.; Shanmugasundararaj, S.; Friedman, D. B.; Bondar, O.; Hudson, B. G.; Sundaramoorthy, M. The alpha1.alpha2 network of collagen IV. Reinforced stabilization of the noncollagenous domain-1 by noncovalent forces and the absence of Met-Lys cross-links. *J.Biol.Chem.* **2004**, *279*, 44723–44730.

(368) Fields, G. B. Synthesis and biological applications of collagen-model triple-helical peptides. *Organic & Biomolecular Chemistry* **2010**, *8*, 1237.

(369) Doi, M.; Imori, K.; Sakaguchi, N.; Asano, A. Boc–Pro–Hyp–Gly–OBzl and Boc–Ala–Hyp–Gly–OBzl, two repeating triplets found in collagen. *Acta Crystallographica Section C Crystal Structure Communications* **2006**, *62*, o577–o580.

(370) Khoshnoodi, J.; Pedchenko, V.; Hudson, B. G. Mammalian collagen IV. *Microscopy Research and Technique* **2008**, *71*, 357–370.

(371) Li, Z.; Hou, W.-S.; Escalante-Torres, C. R.; Gelb, B. D.; Bromme, D. Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate. *J. Biol. Chem.* **2002**, *277*, 28669–28676.

(372) Li, Z.; Yasuda, Y.; Li, W.; Bogyo, M.; Katz, N.; Gordon, R. E.; Fields, G. B.; Brömme, D. Regulation of collagenase activities of human cathepsins by glycosaminoglycans. *J. Biol. Chem.* **2004**, *279*, 5470–5479.

(373) Li, Z.; Kienetz, M.; Cherney, M. M.; James, M. N. G.; Brömme, D. The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex. *J. Mol. Biol.* **2008**, *383*, 78–91.

(374) Theriot, C. M.; Tove, S. R.; Grunden, A. M. Chapter 3 Biotechnological Applications of Recombinant Microbial Prolidases. In *Advances in Applied Microbiology*; Academic Press, 2009; Vol. Volume 68, pp. 99–132.

(375) Reiser, J.; Adair, B.; Reinheckel, T. Specialized roles for cysteine cathepsins in health and disease. *J. Clin. Invest.* **2010**, *120*, 3421–3431.

(376) Rankovic, Z.; Cai, J.; Kerr, J.; Fradera, X.; Robinson, J.; Mistry, A.; Hamilton, E.; McGarry, G.; Andrews, F.; Caulfield, W.; Cumming, I.; Dempster, M.; Waller, J.; Scullion, P.; Martin, I.; Mitchell, A.; Long, C.; Baugh, M.; Westwood, P.; Kinghorn, E.; Bruin, J.; Hamilton, W.; Uitdehaag, J.; Zeeland, M. V.; Potin, D.; Saniere, L.; Fouquet, A.; Chevallier, F.; Deronzier, H.; Dorleans, C.; Nicolai, E. Design and optimization of a series of novel 2-cyano-pyrimidines as cathepsin K inhibitors. *Bioorg.Med.Chem.Lett.* **2009**, -.

(377) Protein Data Bank. Cathepsin K http://www.pdb.org/pdb/explore/explore.do?structureId=3KX1.

(378) Frizler, M.; Lohr, F.; Lülsdorff, M.; Gütschow, M. Facing the gem-Dialkyl Effect in Enzyme Inhibitor Design: Preparation of Homocycloleucine-Based Azadipeptide Nitriles. *Chemistry – A European Journal* **2011**, *17*, 11419–11423.

(379) Isabel, E.; Mellon, C.; Boyd, M. J.; Chauret, N.; Deschênes, D.; Desmarais, S.; Falgueyret, J.-P.; Gauthier, J. Y.; Khougaz, K.; Lau, C. K.; Léger, S.; Levorse, D. A.; Li, C. S.; Massé, F.; David Percival, M.; Roy, B.; Scheigetz, J.; Thérien, M.; Truong, V. L.; Wesolowski, G.; Young, R. N.; Zamboni, R.; Cameron Black, W. Difluoroethylamines as an amide isostere in inhibitors of cathepsin K. *Bioorganic & Medicinal Chemistry Letters* **2011**, *21*, 920–923.

(380) Rankovic, Z.; Cai, J.; Kerr, J.; Fradera, X.; Robinson, J.; Mistry, A.; Finlay, W.; McGarry, G.; Andrews, F.; Caulfield, W.; Cumming, I.; Dempster, M.; Waller, J.; Arbuckle, W.; Anderson, M.; Martin, I.; Mitchell, A.; Long, C.; Baugh, M.; Westwood, P.; Kinghorn, E.; Jones, P.; Uitdehaag, J. C. M.; van Zeeland, M.; Potin, D.; Saniere, L.; Fouquet, A.; Chevallier, F.; Deronzier, H.; Dorleans, C.; Nicolai, E. Optimisation of 2-cyano-pyrimidine inhibitors of cathepsin K: Improving selectivity over hERG. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 6237–6241.

(381) Rankovic, Z.; Cai, J.; Fradera, X.; Dempster, M.; Mistry, A.; Mitchell, A.; Long, C.; Hamilton, E.; King, A.; Boucharens, S.; Jamieson, C.; Gillespie, J.; Cumming, I.; Uitdehaag, J.; Zeeland, M. van Dioxo-triazines as a novel series of cathepsin K inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 1488–1490.

(382) Robichaud, J.; Black, W. C.; Thérien, M.; Paquet, J.; Oballa, R. M.; Bayly, C. I.; McKay, D. J.; Wang, Q.; Isabel, E.; Léger, S.; Mellon, C.; Kimmel, D. B.; Wesolowski, G.; Percival, M. D.; Massé, F.; Desmarais, S.; Falgueyret, J.-P.; Crane, S. N. Identification of a Nonbasic, Nitrile-Containing Cathepsin K Inhibitor (MK-1256) that is Efficacious in a Monkey Model of Osteoporosis. *J. Med. Chem.* **2008**, *51*, 6410–6420.

(383) Teno, N.; Irie, O.; Miyake, T.; Gohda, K.; Horiuchi, M.; Tada, S.; Nonomura, K.; Kometani, M.; Iwasaki, G.; Betschart, C. New chemotypes for cathepsin K inhibitors. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 2599–2603.

(384) Kim, S. H.; Jhon, D. J.; Song, J. H.; No, J. S.; Kang, N. S. Effect of novel N-cyano-tetrahydro-pyridazine compounds, a class of cathepsin K inhibitors, on the bone resorptive activity of mature osteoclasts. *Bioorganic & Medicinal Chemistry Letters* **2008**, *18*, 3988–3991.

(385) Robichaud, J.; Bayly, C. I.; Black, W. C.; Desmarais, S.; Léger, S.; Massé, F.; McKay, D. J.; Oballa, R. M.; Pâquet, J.; Percival, M. D.; Truchon, J.-F.; Wesolowski, G.; Crane, S. N. β-Substituted cyclohexanecarboxamide cathepsin K inhibitors: Modification of the 1,2-disubstituted aromatic core. *Bioorganic & Medicinal Chemistry Letters* **2007**, *17*, 3146–3151.

(386) Cameron Black, W.; David Percival, M. Chapter 8 Cathepsin K Inhibitors. In *Annual Reports in Medicinal Chemistry*; Academic Press, 2007; Vol. Volume 42, pp. 111–127.

(387) Duong, L. T. Therapeutic inhibition of cathepsin K[mdash]reducing bone resorption while maintaining bone formation. *BoneKEy Reports* **2012**, *1*.

(388) Podgorski, I. Future of anticathepsin K drugs: dual therapy for skeletal disease and atherosclerosis? *Future Med Chem* **2009**, *1*, 21.

(389) Rodan, S. B.; Duong, L. T. Cathepsin K - A new molecular target for osteoporosis. *IBMS BoneKEy* **2008**, *5*, 16–24.

(390) Eastell, R.; Nagase, S.; Ohyama, M.; Small, M.; Sawyer, J.; Boonen, S.; Spector, T.; Kuwayama, T.; Deacon, S. Safety and efficacy of the cathepsin K inhibitor ONO-5334 in postmenopausal osteoporosis: the OCEAN study. *J. Bone Miner. Res.* **2011**, *26*, 1303–1312.

(391) Beaulieu, C.; Isabel, E.; Fortier, A.; Massé, F.; Mellon, C.; Méthot, N.; Ndao, M.; Nicoll-Griffith, D.; Lee, D.; Park, H.; Black, W. C. Identification of potent and reversible cruzipain inhibitors for the treatment of Chagas disease. *Bioorganic & Medicinal Chemistry Letters* **2010**, *20*, 7444–7449.

(392) Gauthier, J. Y.; Chauret, N.; Cromlish, W.; Desmarais, S.; Duong, L. T.; Falgueyret, J.-P.; Kimmel, D. B.; Lamontagne, S.; Léger, S.; LeRiche, T.; Li, C. S.; Massé, F.; McKay, D. J.; Nicoll-Griffith, D. A.; Oballa, R. M.; Palmer, J. T.; Percival, M. D.; Riendeau, D.; Robichaud, J.; Rodan, G. A.; Rodan, S. B.; Seto, C.; Thérien, M.; Truong, V.-L.; Venuti, M. C.; Wesolowski, G.; Young, R. N.; Zamboni, R.; Black, W. C. The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 923–928.

(393) Pérez-Castrillón, J. L.; Pinacho, F.; De Luis, D.; Lopez-Menendez, M.; Dueñas Laita, A. Odanacatib, a New Drug for the Treatment of Osteoporosis: Review of the Results in Postmenopausal Women. *Journal of Osteoporosis* **2010**, *2010*, 1–5.

(394) Eisman, J. A.; Bone, H. G.; Hosking, D. J.; McClung, M. R.; Reid, I. R.; Rizzoli, R.; Resch, H.; Verbruggen, N.; Hustad, C. M.; DaSilva, C.; Petrovic, R.; Santora, A. C.; Ince, B. A.; Lombardi, A. Odanacatib in the treatment of postmenopausal women with low bone mineral density: three-year continued therapy and resolution of effect. *J. Bone Miner. Res.* **2011**, *26*, 242–251.

(395) Canalis, E. New Treatment Modalities in Osteoporosis. *Endocrine Practice* **2010**, *16*, 855–863.

(396) Brown, J. P.; Prince, R. L.; Deal, C.; Recker, R. R.; Kiel, D. P.; de Gregorio, L. H.; Hadji, P.; Hofbauer, L. C.; Álvaro-Gracia, J. M.; Wang, H.; Austin, M.; Wagman, R. B.; Newmark, R.; Libanati, C.; San Martin, J.; Bone, H. G. Comparison of the Effect of Denosumab and Alendronate on BMD and Biochemical Markers of Bone Turnover in Postmenopausal Women With Low Bone Mass: A Randomized, Blinded, Phase 3 Trial*. *Journal of Bone and Mineral Research* **2009**, *24*, 153–161.

(397) Ochi, Y.; Yamada, H.; Mori, H.; Nakanishi, Y.; Nishikawa, S.; Kayasuga, R.; Kawada, N.; Kunishige, A.; Hashimoto, Y.; Tanaka, M.; Sugitani, M.; Kawabata, K. Effects of ONO-5334, a novel orally-active inhibitor of cathepsin K, on bone metabolism. *Bone* **2011**, *49*, 1351–1356.

(398) Brömme, D.; Lecaille, F. Cathepsin K inhibitors for osteoporosis and potential off-target effects. *Expert Opin Investig Drugs* **2009**, *18*, 585–600.

(399) Troen, B. R. The regulation of cathepsin K gene expression. *Ann. N. Y. Acad. Sci.* **2006**, *1068*, 165–172.

(400) Kim, H.-S.; Suh, K. S.; Sul, D.; Kim, B.-J.; Lee, S. K.; Jung, W.-W. The inhibitory effect and the molecular mechanism of glabridin on RANKL-induced osteoclastogenesis in RAW264.7 cells. *Int. J. Mol. Med.* **2012**, *29*, 169–177.

(401) Zaidi, M.; Davies, T. F.; Zallone, A.; Blair, H. C.; Iqbal, J.; Moonga, S. S.; Mechanick, J.; Sun, L. Thyroid-stimulating hormone, thyroid hormones, and bone loss. *Curr Osteoporos Rep* **2009**, *7*, 47–52.

(402) Väänänen, H. K.; Zhao, H.; Mulari, M.; Halleen, J. M. The cell biology of osteoclast function. *J. Cell. Sci.* **2000**, *113 ( Pt 3)*, 377–381.

(403) Nakamura, I.; Takahashi, N.; Jimi, E.; Udagawa, N.; Suda, T. Regulation of osteoclast function. *Mod Rheumatol* **2012**, *22*, 167–177.

(404) Gallagher, J. C.; Sai, A. J. Molecular biology of bone remodeling: implications for new therapeutic targets for osteoporosis. *Maturitas* **2010**, *65*, 301–307.

(405) Littlewood-Evans, A.; Kokubo, T.; Ishibashi, O.; Inaoka, T.; Wlodarski, B.; Gallagher, J. A.; Bilbe, G. Localization of cathepsin K in human osteoclasts by in situ hybridization and immunohistochemistry. *Bone* **1997**, *20*, 81–86.

(406) Kiviranta, R.; Morko, J.; Uusitalo, H.; Aro, H. T.; Vuorio, E.; Rantakokko, J. Accelerated turnover of metaphyseal trabecular bone in mice overexpressing cathepsin K. *J. Bone Miner. Res.* **2001**, *16*, 1444–1452.

(407) Gelb, B. D.; Shi, G. P.; Chapman, H. A.; Desnick, R. J. Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science* **1996**, *273*, 1236–1238.

(408) Konttinen, Y. T.; Mandelin, J.; Li, T.-F.; Salo, J.; Lassus, J.; Liljeström, M.; Hukkanen, M.; Takagi, M.; Virtanen, I.; Santavirta, S. Acidic cysteine endoproteinase cathepsin K in the degeneration of the superficial articular hyaline cartilage in osteoarthritis. *Arthritis Rheum.* **2002**, *46*, 953–960.

(409) Liu, J.; Sukhova, G. K.; Sun, J.-S.; Xu, W.-H.; Libby, P.; Shi, G.-P. Lysosomal cysteine proteases in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **2004**, *24*, 1359–1366.

(410) Lutgens, E.; Lutgens, S. P. M.; Faber, B. C. G.; Heeneman, S.; Gijbels, M. M. J.; de Winther, M. P. J.; Frederik, P.; van der Made, I.; Daugherty, A.; Sijbers, A. M.; Fisher, A.; Long, C. J.; Saftig, P.; Black, D.; Daemen, M. J. A. P.; Cleutjens, K. B. J. M. Disruption of the cathepsin K gene reduces atherosclerosis progression and induces plaque fibrosis but accelerates macrophage foam cell formation. *Circulation* **2006**, *113*, 98–107.

(411) Samokhin, A. O.; Wong, A.; Saftig, P.; Brömme, D. Role of cathepsin K in structural changes in brachiocephalic artery during progression of atherosclerosis in apoE-deficient mice. *Atherosclerosis* **2008**, *200*, 58–68.

(412) Wilson, T. J.; Singh, R. K. Proteases as modulators of tumor-stromal interaction: primary tumors to bone metastases. *Biochim. Biophys. Acta* **2008**, *1785*, 85–95.

(413) Yan, X.; Takahara, M.; Xie, L.; Oda, Y.; Nakahara, T.; Uchi, H.; Takeuchi, S.; Tu, Y.; Moroi, Y.; Furue, M. Stromal expression of cathepsin K in squamous cell carcinoma. *J Eur Acad Dermatol Venereol* **2011**, *25*, 362–365.

(414) Brubaker, K. D.; Vessella, R. L.; True, L. D.; Thomas, R.; Corey, E. Cathepsin K mRNA and protein expression in prostate cancer progression. *J. Bone Miner. Res.* **2003**, *18*, 222–230.

(415) Sturge, J.; Caley, M. P.; Waxman, J. Bone metastasis in prostate cancer: emerging therapeutic strategies. *Nat Rev Clin Oncol* **2011**, *8*, 357–368.

(416) Le Gall, C.; Bellahcène, A.; Bonnelye, E.; Gasser, J. A.; Castronovo, V.; Green, J.; Zimmermann, J.; Clézardin, P. A cathepsin K inhibitor reduces breast cancer induced osteolysis and skeletal tumor burden. *Cancer Res.* **2007**, *67*, 9894–9902.

(417) Chen, Y.-C.; Sosnoski, D. M.; Mastro, A. M. Breast cancer metastasis to the bone: mechanisms of bone loss. *Breast Cancer Research* **2010**, *12*, 215.

(418) Herrmann, M.; Seibel, M. The amino- and carboxyterminal cross-linked telopeptides of collagen type I, NTX-I and CTX-I: A comparative review. *Clinica Chimica Acta* **2008**, *393*, 57–75.

(419) Seibel, M. J. Clinical use of markers of bone turnover in metastatic bone disease. *Nat Clin Pract Oncol* **2005**, *2*, 504–517; quiz 1 p following 533.

(420) Tamada, T.; Sone, T.; Tomomitsu, T.; Jo, Y.; Tanaka, H.; Fukunaga, M. Biochemical markers for the detection of bone metastasis in patients with prostate cancer: diagnostic efficacy and the effect of hormonal therapy. *J. Bone Miner. Metab.* **2001**, *19*, 45–51.

(421) Voorzanger-Rousselot, N.; Juillet, F.; Mareau, E.; Zimmermann, J.; Kalebic, T.; Garnero, P. Association of 12 serum biochemical markers of angiogenesis, tumour invasion and bone turnover with bone metastases from breast cancer: a crossectional and longitudinal evaluation. *Br. J. Cancer* **2006**, *95*, 506–514.

(422) Johnson, G. V.; Litersky, J. M.; Whitaker, J. N. Proteolysis of microtubule-associated protein 2 and tubulin by cathepsin D. *J. Neurochem.* **1991**, *57*, 1577–1583.

(423) Cherney, M. M.; Lecaille, F.; Kienitz, M.; Nallaseth, F. S.; Li, Z.; James, M. N. G.; Brömme, D. Structure-Activity Analysis of Cathepsin K/Chondroitin 4-Sulfate Interactions. *J. Biol. Chem.* **2011**, *286*, 8988–8998.

(424) Monaco, S.; Sparano, V.; Gioia, M.; Sbardella, D.; Di Pierro, D.; Marini, S.; Coletta, M. Enzymatic processing of collagen IV by MMP-2 (gelatinase A) affects neutrophil migration and it is modulated by extracatalytic domains. *Protein Sci* **2006**, *15*, 2805–2815.

(425) Parasitic Diseases http://www.nlm.nih.gov/medlineplus/parasiticdiseases.html (accessed Jun 11, 2012).

(426) McKerrow, J. H.; Caffrey, C.; Kelly, B.; Loke, P.; Sajid, M. Proteases in parasitic diseases. *Annu Rev Pathol* **2006**, *1*, 497–536.

(427) Quezada, L. A. L.; Sajid, M.; Lim, K. C.; McKerrow, J. H. A Blood Fluke Serine Protease Inhibitor Regulates an Endogenous Larval Elastase. *J. Biol. Chem.* **2012**, *287*, 7074–7083.

(428) Lannes-Vieira, J.; de Araújo-Jorge, T. C.; Soeiro, M. de N. C.; Gadelha, P.; Corrêa-Oliveira, R. The Centennial of the Discovery of Chagas Disease: Facing the Current Challenges. *PLoS Negl Trop Dis* **2010**, *4*, e645.

(429) Prevention, C.-C. for D. C. and CDC - Chagas Disease http://www.cdc.gov/parasites/chagas/ (accessed Jun 11, 2012).

(430) WHO | Chagas disease (American trypanosomiasis) http://www.who.int/mediacentre/factsheets/fs340/en/index.html (accessed Jun 11, 2012).

(431) Pérez-Molina, J.; Norman, F.; López-Vélez, R. Chagas Disease in Non-Endemic Countries: Epidemiology, Clinical Presentation and Treatment. *Current Infectious Disease Reports* **2012**, *14*, 263–274.

(432) Rassi, A., Jr; Rassi, A.; Marin-Neto, J. A. Chagas disease. *Lancet* **2010**, *375*, 1388–1402.

(433) Bern, C.; Kjos, S.; Yabsley, M. J.; Montgomery, S. P. Trypanosoma Cruzi and Chagas' Disease in the United States. *Clin. Microbiol. Rev.* **2011**, *24*, 655–681.

(434) Hughes, A. L.; Piontkivska, H. Phylogeny of Trypanosomatidae and Bodonidae (Kinetoplastida) Based on 18S rRNA: Evidence for Paraphyly of Trypanosoma and Six Other Genera. *Mol Biol Evol* **2003**, *20*, 644–652.

(435) Noireau, F.; Diosque, P.; Jansen, A. M. Trypanosoma cruzi: adaptation to its vectors and its hosts. *Vet Res* **2009**, *40*.

(436) Clayton, J. Chagas disease 101. *Nature* **2010**, *465*, S4–S5.

(437) Hemmige, V.; Tanowitz, H.; Sethi, A. Trypanosoma cruzi infection: a review with emphasis on cutaneous manifestations. *International Journal of Dermatology* **2012**, *51*, 501–508.

(438) Macedo, A. M.; Oliveira, R. P.; Pena, S. D. J. Chagas disease: role of parasite genetic variation in pathogenesis. *Expert Rev Mol Med* **2002**, *4*, 1–16.

(439) Benjamin, R. J.; Stramer, S. L.; Leiby, D. A.; Dodd, R. Y.; Fearon, M.; Castro, E. Trypanosoma cruzi infection in North America and Spain: evidence in support of transfusion transmission. *Transfusion* **2012**, no–no.

(440) Skolnick, A. Does Influx From Endemic Areas Mean More Transfusion-Associated Chagas' Disease? *JAMA: The Journal of the American Medical Association* **1989**, *262*, 1433–1433.

(441) Pereira, K. S.; Schmidt, F. L.; Guaraldo, A. M. A.; Franco, R. M. B.; Dias, V. L.; Passos, L. A. C. Chagas' disease as a foodborne illness. *J. Food Prot.* **2009**, *72*, 441–446.

(442) Salas, N. A.; Cot, M.; Schneider, D.; Mendoza, B.; Santalla, J. A.; Postigo, J.; Chippaux, J. P.; Brutus, L. Risk factors and consequences of congenital Chagas disease in Yacuiba, south Bolivia. *Trop. Med. Int. Health* **2007**, *12*, 1498–1505.

(443) Antas, P. R.; Medrano-Mercado, N.; Torrico, F.; Ugarte-Fernandez, R.; Gómez, F.; Correa Oliveira, R.; Chaves, A. C.; Romanha, A. J.; Araújo-Jorge, T. C. Early, intermediate, and late acute stages in Chagas' disease: a study combining anti-galactose IgG, specific serodiagnosis, and polymerase chain reaction analysis. *Am. J. Trop. Med. Hyg.* **1999**, *61*, 308–314.

(444) Chagas Disease (American Trypanosomiasis). **2012**.

(445) Almeida, E. A.; Lima, J. N.; Lages-Silva, E.; Guariento, M. E.; Aoki, F. H.; Torres-Morales, A. E.; Pedro, R. J. Chagas' disease and HIV co-infection in patients without effective antiretroviral therapy: prevalence, clinical presentation and natural history. *Trans. R. Soc. Trop. Med. Hyg.* **2010**, *104*, 447–452.

(446) Bern, C.; Martin, D. L.; Gilman, R. H. Acute and congenital Chagas disease. *Adv. Parasitol.* **2011**, *75*, 19–47.

(447) Rassi, A.; Rassi, A.; Franco-Paredes, C. A Latin American Man with Palpitations, Dizziness, Episodes of Nonsustained Ventricular Tachycardia, and an Apical Aneurysm. *PLoS Negl Trop Dis* **2011**, *5*.

(448) Lofstedt, A. R. S., James A. Roth, Jane Galyon, Jeanne *Emerging and Exotic Diseases of Animals*; CFSPH Iowa State University, 2010.

(449) Solari, A.; Ortíz, S.; Soto, A.; Arancibia, C.; Campillay, R.; Contreras, M.; Salinas, P.; Rojas, A.; Schenone, H. Treatment of Trypanosoma Cruzi-Infected Children with Nifurtimox: A 3 Year Follow-up by PCR. *J. Antimicrob. Chemother.* **2001**, *48*, 515–519.

(450) Malan, A. K.; Avelar, E.; Litwin, S. E.; Hill, H. R.; Litwin, C. M. Serological diagnosis of Trypanosoma cruzi: evaluation of three enzyme immunoassays and an indirect immunofluorescent assay. *J Med Microbiol* **2006**, *55*, 171–178.

(451) Amato Neto, V.; Amato, V. S.; Tuon, F. F.; Gakiya, E.; de Marchi, C. R.; de Souza, R. M.; Furucho, C. R. False-positive results of a rapid K39-based strip test and Chagas disease. *Int. J. Infect. Dis.* **2009**, *13*, 182–185.

(452) Wincker, P.; Bosseno, M. F.; Britto, C.; Yaksic, N.; Cardoso, M. A.; Morel, C. M.; Brenière, S. F. High correlation between Chagas' disease serology and PCR-based detection of Trypanosoma cruzi kinetoplast DNA in Bolivian children living in an endemic area. *FEMS Microbiol. Lett.* **1994**, *124*, 419–423.

(453) Croft, A. M.; Baker, D.; von Bertele, M. J. An evidence-based vector control strategy for military deployments: the British Army experience. *Med Trop (Mars)* **2001**, *61*, 91–98.

(454) Jackson, Y.; Alirol, E.; Getaz, L.; Wolff, H.; Combescure, C.; Chappuis, F. Tolerance and Safety of Nifurtimox in Patients with Chronic Chagas Disease. *Clin Infect Dis.* **2010**, *51*, e69–e75.

(455) Garcia, S.; Ramos, C. O.; Senra, J. F. V.; Vilas-Boas, F.; Rodrigues, M. M.; Campos-De-Carvalho, A. C.; Ribeiro-Dos-Santos, R.; Soares, M. B. P. Treatment with Benznidazole During the Chronic Phase of Experimental Chagas' Disease Decreases Cardiac Alterations. *Antimicrob. Agents Chemother.* **2005**, *49*, 1521–1528.

(456) Castro, J. A.; Diaz de Toranzo, E. G. Toxic effects of nifurtimox and benznidazole, two drugs used against American trypanosomiasis (Chagas' disease). *Biomed. Environ. Sci.* **1988**, *1*, 19–33.

(457) Maya, J. D.; Cassels, B. K.; Iturriaga-Vásquez, P.; Ferreira, J.; Faúndez, M.; Galanti, N.; Ferreira, A.; Morello, A. Mode of action of natural and synthetic drugs against Trypanosoma cruzi and their interaction with the mammalian host. *Comp. Biochem. Physiol., Part A Mol. Integr. Physiol.* **2007**, *146*, 601–620.

(458) Enzyme Database http://www.brenda-enzymes.info/php/result_flat.php4?ecno=3.4.22.51.

(459) Itow, S.; Camargo, E. P. Proteolytic activites in cell extracts of Trypanosoma cruzi. *J. Protozool.* **1977**, *24*, 591–595.

(460) Eakin, A. E.; Mills, A. A.; Harth, G.; McKerrow, J. H.; Craik, C. S. The sequence, organization, and expression of the major cysteine protease (cruzain) from Trypanosoma cruzi. *J. Biol. Chem.* **1992**, *267*, 7411–7420.

(461) Engel, J. C.; Torres, C.; Hsieh, I.; Doyle, P. S.; McKerrow, J. H.; Garcia, C. T. Upregulation of the secretory pathway in cysteine protease inhibitor-resistant Trypanosoma cruzi. *J. Cell. Sci.* **2000**, *113 ( Pt 8)*, 1345–1354.

(462) Engel, J. C.; Doyle, P. S.; Palmer, J.; Hsieh, I.; Bainton, D. F.; McKerrow, J. H. Cysteine protease inhibitors alter Golgi complex ultrastructure and function in Trypanosoma cruzi. *J. Cell. Sci.* **1998**, *111 ( Pt 5)*, 597–606.

(463) Labriola, C.; Cazzulo, J. J.; Parodi, A. J. Retention of glucose units added by the UDP-GLC:glycoprotein glucosyltransferase delays exit of glycoproteins from the endoplasmic reticulum. *J. Cell Biol.* **1995**, *130*, 771–779.

(464) Souto-Padron, T.; Campetella, O. E.; Cazzulo, J. J.; Souza, W. de Cysteine proteinase in Trypanosoma cruzi: immunocytochemical localization and involvement in parasite-host cell interaction. *J Cell Sci* **1990**, *96*, 485–490.

(465) Nascimento, A. High resolution localization of cruzipain and Ssp4 in Trypanosoma cruzi by replica staining label fracture. *Biology of the Cell* **1996**, *86*, 53–58.

(466) Lima, A. P.; dos Reis, F. C.; Serveau, C.; Lalmanach, G.; Juliano, L.; Ménard, R.; Vernet, T.; Thomas, D. Y.; Storer, A. C.; Scharfstein, J. Cysteine protease isoforms from Trypanosoma cruzi, cruzipain 2 and cruzain, present different substrate preference and susceptibility to inhibitors. *Mol. Biochem. Parasitol.* **2001**, *114*, 41–52.

(467) Aparicio, I. M.; Scharfstein, J.; Lima, A. P. C. A. A New Cruzipain-Mediated Pathway of Human Cell Invasion by Trypanosoma cruzi Requires Trypomastigote Membranes. *Infect Immun* **2004**, *72*, 5892–5902.

(468) Giordanengo, L.; Maldonado, C.; Rivarola, H. W.; Iosa, D.; Girones, N.; Fresno, M.; Gea, S. Induction of antibodies reactive to cardiac myosin and development of heart alterations in cruzipain-immunized mice and their offspring. *Eur. J. Immunol.* **2000**, *30*, 3181–3189.

(469) Giordanengo, L.; Fretes, R.; Díaz, H.; Cano, R.; Bacile, A.; Vottero-Cima, E.; Gea, S. Cruzipain induces autoimmune response against skeletal muscle and tissue damage in mice. *Muscle Nerve* **2000**, *23*, 1407–1413.

(470) Zhang, S.; Kim, C. C.; Batra, S.; McKerrow, J. H.; Loke, P. Delineation of diverse macrophage activation programs in response to intracellular parasites and cytokines. *PLoS Negl Trop Dis* **2010**, *4*, e648.

(471) Doyle, P. S.; Zhou, Y. M.; Hsieh, I.; Greenbaum, D. C.; McKerrow, J. H.; Engel, J. C. The Trypanosoma cruzi protease cruzain mediates immune evasion. *PLoS Pathog.* **2011**, *7*, e1002139.

(472) Stempin, C.; Giordanengo, L.; Gea, S.; Cerbán, F. Alternative activation and increase of Trypanosoma cruzi survival in murine macrophages stimulated by cruzipain, a parasite antigen. *J. Leukoc. Biol.* **2002**, *72*, 727–734.

(473) Del Nery, E.; Juliano, M. A.; Lima, A. P.; Scharfstein, J.; Juliano, L. Kininogenase activity by the major cysteinyl proteinase (cruzipain) from Trypanosoma cruzi. *J. Biol. Chem.* **1997**, *272*, 25713–25718.

(474) Scharfstein, J.; Schmitz, V.; Morandi, V.; Capella, M. M.; Lima, A. P.; Morrot, A.; Juliano, L.; Müller-Esterl, W. Host cell invasion by Trypanosoma cruzi is potentiated by activation of bradykinin B(2) receptors. *J. Exp. Med.* **2000**, *192*, 1289–1300.

(475) Aoki, M. D. P.; Cano, R. C.; Pellegrini, A. V.; Tanos, T.; Guiñazú, N. L.; Coso, O. A.; Gea, S. Different signaling pathways are involved in cardiomyocyte survival induced by a Trypanosoma cruzi glycoprotein. *Microbes Infect.* **2006**, *8*, 1723–1731.

(476) Tomas, A. M.; Miles, M. A.; Kelly, J. M. Overexpression of cruzipain, the major cysteine proteinase of Trypanosoma cruzi, is associated with enhanced metacyclogenesis. *Eur. J. Biochem.* **1997**, *244*, 596–603.

(477) Tomás, A. M.; Kelly, J. M. Stage-regulated expression of cruzipain, the major cysteine protease of Trypanosoma cruzi is independent of the level of RNA1. *Mol. Biochem. Parasitol.* **1996**, *76*, 91–103.

(478) Huete-Pérez, J. A.; Engel, J. C.; Brinen, L. S.; Mottram, J. C.; McKerrow, J. H. Protease Trafficking in Two Primitive Eukaryotes Is Mediated by a Prodomain Protein Motif. *J. Biol. Chem.* **1999**, *274*, 16249–16256.

(479) Gillmor, S. A.; Craik, C. S.; Fletterick, R. J. Structural determinants of specificity in the cysteine protease cruzain. *Protein Sci.* **1997**, *6*, 1603–1611.

(480) Piña-Vázquez, C.; Reyes-López, M.; Ortíz-Estrada, G.; de la Garza, M.; Serrano-Luna, J. Host-Parasite Interaction: Parasite-Derived and -Induced Proteases That Degrade Human Extracellular Matrix. *Journal of Parasitology Research* **2012**, *2012*, 1–24.

(481) Duaso, J.; Yanez, E.; Castillo, C.; Galanti, N.; Cabrera, G.; Corral, G.; Maya, J. D.; Zulantay, I.; Apt, W.; Kemmerling, U. Reorganization of extracellular matrix in placentas from women with asymptomatic chagas disease: mechanism of parasite invasion or local placental defense? *J Trop Med* **2012**, *2012*, 758357.

(482) Jose Cazzulo, J.; Stoka, V.; Turk, V. The major cysteine proteinase of Trypanosoma cruzi: a valid target for chemotherapy of Chagas disease. *Curr. Pharm. Des.* **2001**, *7*, 1143–1156.

(483) McGrath, M. E.; Eakin, A. E.; Engel, J. C.; McKerrow, J. H.; Craik, C. S.; Fletterick, R. J. The crystal structure of cruzain: a therapeutic target for Chagas' disease. *J. Mol. Biol.* **1995**, *247*, 251–259.

(484) Reis, F. C. G.; Costa, T. F. R.; Sulea, T.; Mezzetti, A.; Scharfstein, J.; Brömme, D.; Ménard, R.; Lima, A. P. C. A. The propeptide of cruzipain--a potent selective inhibitor of the trypanosomal enzymes cruzipain and brucipain, and of the human enzyme cathepsin F. *FEBS J.* **2007**, *274*, 1224–1234.

(485) Lalmanach, G.; Lecaille, F.; Chagas, J. R.; Authié, E.; Scharfstein, J.; Juliano, M. A.; Gauthier, F. Inhibition of Trypanosomal Cysteine Proteinases by Their Propeptides. *J. Biol. Chem.* **1998**, *273*, 25112–25116.

(486) Monteiro, A. C. S.; Abrahamson, M.; Lima, A. P. C. A.; Vannier-Santos, M. A.; Scharfstein, J. Identification, characterization and localization of chagasin, a tight-binding cysteine protease inhibitor in Trypanosoma cruzi. *J Cell Sci* **2001**, *114*, 3933–3942.

(487) Bevec, T.; Stoka, V.; Pungercic, G.; Cazzulo, J. J.; Turk, V. A fragment of the major histocompatibility complex class II-associated p41 invariant chain inhibits cruzipain, the major cysteine proteinase from Trypanosoma cruzi. *FEBS Lett.* **1997**, *401*, 259–261.

(488) Chiyanzu, I.; Hansell, E.; Gut, J.; Rosenthal, P. J.; McKerrow, J. H.; Chibale, K. Synthesis and evaluation of isatins and thiosemicarbazone derivatives against cruzain, falcipain-2 and rhodesain. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3527–3530.

(489) Carvalho, S. A.; Feitosa, L. O.; Soares, M.; Costa, T. E. M. M.; Henriques, M. G.; Salomão, K.; de Castro, S. L.; Kaiser, M.; Brun, R.; Wardell, J. L.; Wardell, S. M. S. V.; Trossini, G. H. G.; Andricopulo, A. D.; da Silva, E. F.; Fraga, C. A. M. Design and synthesis of new (E)-cinnamic N-acylhydrazones as potent antitrypanosomal agents. *European Journal of Medicinal Chemistry*.

(490) Mott, B. T.; Ferreira, R. S.; Simeonov, A.; Jadhav, A.; Ang, K. K.-H.; Leister, W.; Shen, M.; Silveira, J. T.; Doyle, P. S.; Arkin, M. R.; McKerrow, J. H.; Inglese, J.; Austin, C. P.; Thomas, C. J.; Shoichet, B. K.; Maloney, D. J. Identification and Optimization of Inhibitors of Trypanosomal Cysteine Proteases: Cruzain, Rhodesain, and TbCatB. *J. Med. Chem.* **2009**, *53*, 52–60.

(491) Bryant, C.; Kerr, I. D.; Debnath, M.; Ang, K. K. H.; Ratnam, J.; Ferreira, R. S.; Jaishankar, P.; Zhao, D.; Arkin, M. R.; McKerrow, J. H.; Brinen, L. S.; Renslo, A. R. Novel non-peptidic vinylsulfones targeting the S2 and S3 subsites of parasite cysteine proteases. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 6218–6221.

(492) Pizzo, C.; Faral-Tello, P.; Salinas, G.; Fló, M.; Robello, C.; Wipf, P.; Graciela Mahler, S. Selenosemicarbazones as potent cruzipain inhibitors and their antiparasitic properties against Trypanosoma cruzi. *MedChemComm* **2012**, *3*, 362.

(493) Brak, K.; Kerr, I. D.; Barrett, K. T.; Fuchi, N.; Debnath, M.; Ang, K.; Engel, J. C.; McKerrow, J. H.; Doyle, P. S.; Brinen, L. S.; Ellman, J. A. Nonpeptidic Tetrafluorophenoxymethyl Ketone Cruzain Inhibitors as Promising New Leads for Chagas Disease Chemotherapy. *J. Med. Chem.* **2010**, *53*, 1763–1773.

(494) Pizzo, C.; Saiz, C.; Talevi, A.; Gavernet, L.; Palestro, P.; Bellera, C.; Blanch, L. B.; Benítez, D.; Cazzulo, J. J.; Chidichimo, A.; Wipf, P.; Mahler, S. G. Synthesis of 2-Hydrazolyl-4-Thiazolidinones Based on Multicomponent Reactions and Biological Evaluation against Trypanosoma Cruzi. *Chemical Biology & Drug Design* **2011**, *77*, 166–172.

(495) Clayton, J. Chagas disease: pushing through the pipeline. *Nature* **2010**, *465*, S12–S15.

(496) Renslo, A. R.; McKerrow, J. H. Drug discovery and development for neglected parasitic diseases. *Nature Chemical Biology* **2006**, *2*, 701–710.

(497) Ascenzi, P.; Bocedi, A.; Visca, P.; Antonini, G.; Gradoni, L. Catalytic properties of cysteine proteinases from Trypanosoma cruzi and Leishmania infantum: a pre-steady-state and steady-state study. *Biochem. Biophys. Res. Commun.* **2003**, *309*, 659–665.

(498) Salvati, L.; Mattu, M.; Polticelli, F.; Tiberi, F.; Gradoni, L.; Venturini, G.; Bolognesi, M.; Ascenzi, P. Modulation of the catalytic activity of cruzipain, the major cysteine proteinase from Trypanosoma cruzi, by temperature and pH. *Eur. J. Biochem.* **2001**, *268*, 3253–3258.

(499) Eakin, A. E.; McGrath, M. E.; McKerrow, J. H.; Fletterick, R. J.; Craik, C. S. Production of crystallizable cruzain, the major cysteine protease from Trypanosoma cruzi. *J. Biol. Chem.* **1993**, *268*, 6115–6118.

(500) Shoichet, B. K. Interpreting steep dose-response curves in early inhibitor discovery. *J. Med. Chem.* **2006**, *49*, 7274–7277.