Characterization of glycosaminoglycan interaction sites and exosite inhibitors of cathepsin K

by

HAORAN CUI

B.Sc.H, Shandong University, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Craniofacial Science)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

February 2014

Abstract

Cysteine cathepsins belong to the papain-like family and are critically involved in the pathogenesis of various diseases, such as osteoporosis and cardiovascular diseases. These disorders are associated with failures in the degradation of extracellular matrix proteins, collagens and elastins. Based on enzymatic and structural studies, the active site mechanism has been well elucidated and many active site inhibitors have been identified, especially inhibitors for cathepsin K. Cathepsin K is primarily expressed in osteoclasts and the enzyme is responsible for most of the bone resorption, and thus an important pharmaceutical target for the treatment of osteoporosis. However, as cathepsin K is expressed not only in osteoclasts but also has functions outside the skeletal system, active site inhibitors may lead to severe side effects when used for treatment. In cathepsins, there are certain binding sites that are distinct from the active site, termed exosites, which have been defined as important for the degradation of extracellular matrix proteins. The hypothesis is that exosites could block the degradation of matrix proteins but would not interfere with other biological functions of the target protease such as cathepsin K. My thesis is focused on the characterization of these exosites and the identification of selective exosite inhibitors.

In Project 1, protein-GAG interaction sites were studied for their involvement in the collagenase activity. According to the X-ray structures of cathepsin K -chondroitin sulfate (CS) complex, two mechanistic models were built: cathepsin K/CS tetramer and dimer. The hypothesis is that the cathepsin K/CS dimer complex solubilizes collagen fibers into tropocollagen and the cathepsin K/CS tetramer complex further degrades tropocollagen into peptides. The objective is to elucidate the mechanism of cathepsin K/CS complex-mediated collagen degradation by generating GAG binding sites mutants.

In project 2, potential exosite inhibitors of cathepsin K were identified by drug screening assays. Compounds which have been reported to be effective in osteoporosis animal models were studied. The hypothesis is that exosite inhibitors for cathepsins specifically inhibit the degradation of collagen but not that of other biological substrates. The objective is to determine IC_{50} values of the compounds for the inhibition of the degradation of collagen.

Preface

Dr. Dieter Brömme was the principle investigator of this research project. Under his general supervision, I was responsible for performing the research, analysis of the research data, and manuscript preparation for all the work described in this thesis. Dr. Adleke Aguda and Dr. Xin Du developed the cathepsin K/GAG dimer and tetramer models.

Table of Contents

List of Tables

List of Figures

Abbreviations

AFM: atomic force microscopy

BMGY: buffered glycerol complex medium

BMMY: buffered methanol complex medium

CS: chondroitin sulfate

C4-S: chondroitin sulfate A

DDT: dithioerythreitol

DMSO: dimethyl sulfoxide

E-64: L-3-carboxy-trans-2-3-epoxypropionyl-leucylamido-(4guanidino)-butane

EDTA: ethylenediaminetetraacetic acid

GAG: glycosaminoglycan

IgG: immunoglobulin G

kDa: kilodalton

MD: minimal dextrose

MM: minimal methanol

MMP: matrix metalloprotease

MW: molecular weight

PBS: phosphate buffered saline

PCR: polymerase chain reaction

rpm: revolutions per minute

SAR: structure-activity relationship study

SDS-PAGE: sodium dodecyl sulfate polyacrylamide electrophoresis

SEM: scanning electron microscopy

TGF-β1: transforming growth factor beta 1

Z-FR-MCA: carbobenzoxy-phenylalanine-arginine-4-methylcoumarin-7-amide

Acknowledgements

First of all, I would like to show my gratitude to the Department of Dentistry and particularly to Dr. Dieter Brömme, who gave me excellent guidance. It is his endless support, inspiration and encouragement that motivated me to finish my projects.

I am greatly thankful to my committee members, Dr. Francois Jean and Dr. Lari Häkkinen, for their support and assistance.

I would also like to thank my colleagues. I offer my appreciation to Dr. Xin Du for his assistance and instructions. I really appreciate Shinako Takimoto, who entered the laboratory with me at the same time. She taught me a lot not only about research, but also English language.

Finally and most importantly, I thank my parents for their unconditional love and financial support.

Chapter 1: Introduction

1.1 Proteases

Proteases (also termed peptidases) are enzymes hydrolysing peptide bonds to complete protein catabolism. Proteases can be categorized into subgroups according to their substrate specificity and mechanisms of action. Based on the catalytic active sites, they are divided into 6 groups, which are serine proteases, threonine proteases, cysteine proteases, aspartate proteases, glutamic proteases, and metalloproteases (1). Cysteine proteases are named according to their common features of exploiting a cysteine residue in their active site. A histidine residue serves as nucleophile and general base. Cysteine proteases are divided into different families based on their sequence homology and structural similarities (2). In the following section, papain-like cysteine proteases, which are known as cathepsin cysteine proteases, will be discussed (3).

1.2 Cathepsins

Cysteine cathepsins belong to the papain-like protease family. There are 11 cysteine cathepsins in humans (cathepsins B, H, L, S, C, K, O, F, V, X, and W). They are known as lysosomal proteases and are involved in the intracellular protein degradation as housekeeping proteases (Figure 1). However, cathepsins K, V, S, and F are also found to participate in specific biological processes (4). Previous studies have shown that cathepsin K plays a major role in osteoporosis. Osteoporosis is a bone disease that leads to increased rate of fracture and the incidence of it increases when people age (5). This disease is related to an unbalanced bone remodelling. Bone remodelling contains two phases, one is bone resorption mediated by osteoclasts and the other one is bone formation mediated by osteoblasts. Usually in healthy

individuals, these two phases are balanced and the bone is constantly renewed. Bone formation begins after several weeks when the bone resorption forms the “bone remodelling unit” (6). Bone resorption consists of two steps: mineral solubilization and focalized matrix degradation (7). The former step is relied on the acid secretion by osteoclasts and the latter step is relied on the activity of a cysteine protease, cathepsin K (8). When bone resorption predominates over bone formation, loss of skeletal mass and increased risk of fractures are observed, causing osteoporosis (6). Cathepsin K plays a major role in bone remodelling activity, and consequently is highly expressed in osteoclasts (9). Cathepsin K exerts a collagenase activity at acidic pH (10) that can cleave triple helical collagens at multiple sites (11). By secreting protons into the resorption lacuna, the osteoclast first dissolves the mineralized component of bone (6). In the second step, the osteoclast secretes cathepsin K into the acidified environment underneath the osteoclast and degrades the organic collagen matrix.

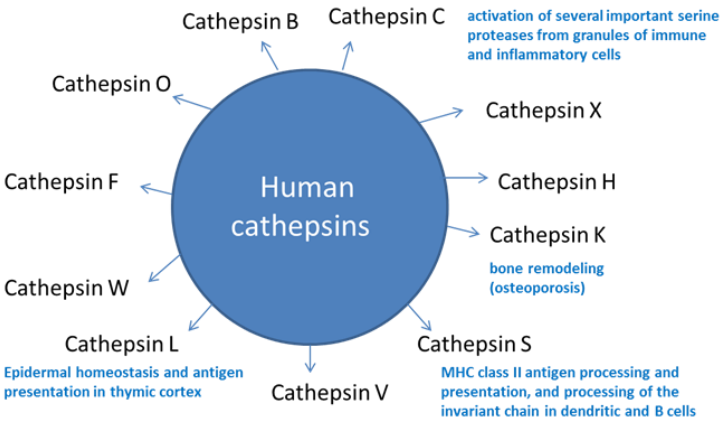


Figure 1. Human Cathepsins and Their Functions.

There are 11 cysteine cathepsins in the human genome (B, H, L, S, C, K, O, F, V, X, and W). Among them, cathepsin B, H, L, C, O, X, and F are ubiquitously expressed cathepsins, while cathepsin S, K, V, and W are tissue specific cathepsins. They have various functions in humans (12).

The substrate binding site of cathepsins can be divided into subsites S1-S4 and S1'-S3' according to the Schechter-Berger model (Figure 2) (13). S1-S4 refer to the sites N-terminally and S1'-S4' to the sites C-terminally to the scissile bond. The appropriate amino acids in the peptide substrate are named as P to P'. According to previous structural studies, the most significant subsites in cysteine proteases are S2-S1'. However, recent studies showed that S3 subsite of cathepsin K is important for determining the substrate specificity as well (4). The main subsite pocket, S2, allows the binding of aliphatic amino acids such as Leu, Ile, and Val. S2 and S1' subsites are emphasized for substrate binding. The S1 subsite helps facilitating catalysis (4, 14, 15). In contrast to all other cathepsins, cathepsin K also accepts Pro in the S2 subsite (16). At the P1' and P3 position, cathepsin K prefers Gly (17, 84).

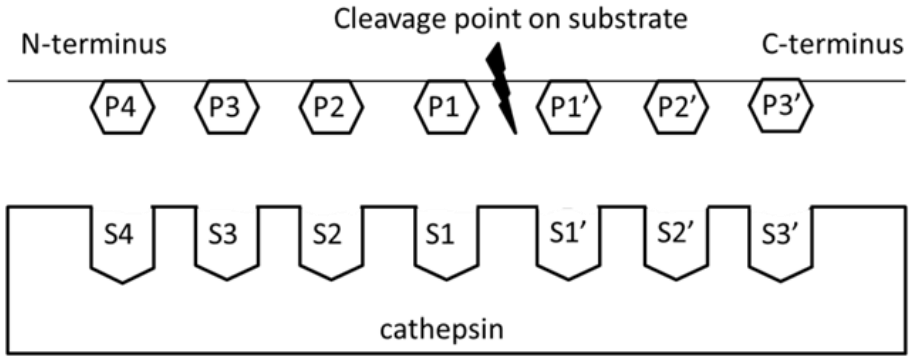


Figure 2. Schechter and Berger Subsite Model of Cathepsins.

S stands for subsites and P stands for amino acid residues in the peptide substrate. The cleavage site on substrate is between P1 and P1' residues. (adapted from Schechter, I. and A. Berger (1967). "On the size of the active site in proteases. I. Papain." Biochem Biophys Res Commun 27(2): 157-162.)

1.3 Catalytic Mechanism of Cathepsin K

Mature cathepsin K is ~24 kDa in size. The full length preproenzyme (~36kDa) contains 329-residue which includes a 15-amino acid signal sequence and a 99-residue long propeptide. Active mature cathepsin K has two domains (Figure 3). Between the two domains, there is the "V"-shaped active site cleft. The right domain is mainly composed of β-barrel motifs, while the left one has α- helices. The active site of cathepsin K and that of all other cysteine cathepsins is formed by a catalytic triad. This triad is composed of Cys25, His162 and Asn182. Cys25 and His162 are thought to exist as the thiolate-imidazolium ion pair that is stabilized by Asn182 via a hydrogen bond with His159 (18). Using E-64, an irreversible inhibitor of cathepsin K, we know that the Cys25 acts as a nucleophile. In this case, Cys25 forms a covalent bond with the C2 of the epoxy-ring of E-64 (19).

This process is started by the attack of the active cysteine SH group from the backside of E-64 epoxy O1 atom and leads to the ring opening and covalent bond formation. Additionally, hydrogen bonds or electrostatic short contacts are formed by the inhibitor with polar atoms of substrate binding site residues (19). There are three stronger interactions and one weak interaction of hydrogen bonds. For cathepsin K, the leucyl group of E-64 is buried in the S2 subsite. Met-68, Ala-134 and Leu-160 are responsible for forming the S2 subsites in cathepsin K (19).

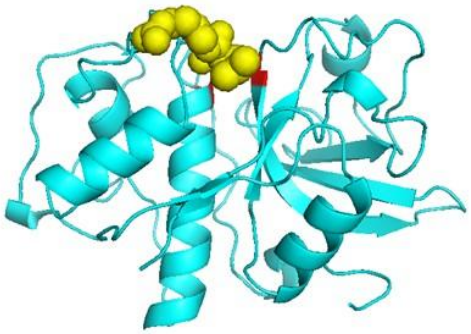


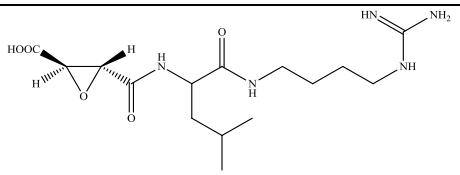
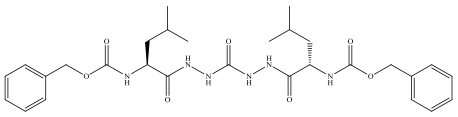
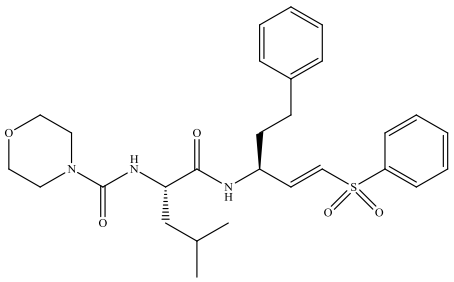
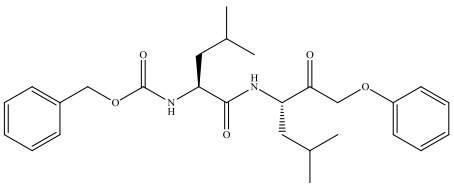
Figure 3. Crystal Structure of Cathepsin K and E64

Cathepsin K monomer is in cyan. Cathepsin K contains two domains. The left is rich in α helix elements and the right domain is mainly composed of β sheets. Active residues (Cys25, His162 and Asn182) are in red. E64 is in yellow.

1.4 Active Site Inhibitors of Cathepsin K

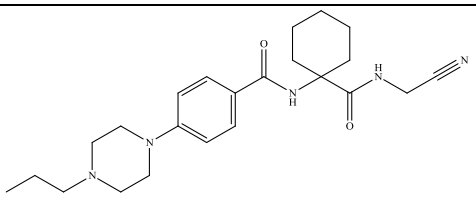
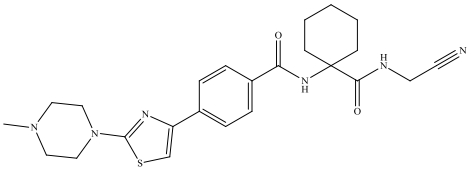
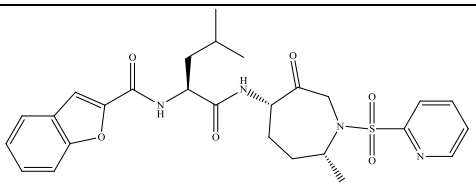
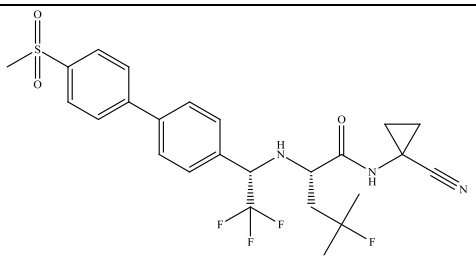
According to previous studies, there are many known low molecular weight inhibitors of cathepsin K. They are divided into irreversible and reversible inhibitors (Table 1 and 2). Irreversible inhibitors are not well suited for long-term clinical usage. As they form covalent adducts, they may form novel haptens for antigens. This could potentially lead to autoimmune diseases. Moreover, due to their covalent action, they will also inactivate related and even less related proteins over time. Although irreversible inhibitors are in general not favoured clinically, they are still very useful for laboratory use, such as in cathepsin assays (22). Compared to irreversible inhibitors, reversible inhibitors are more appropriate for clinical applications. However, reversible active site-directed inhibitors also have the disadvantage that they block the active site of the target protease and thus would inhibit the degradation of any substrate equally.

Table 1: Irreversible Inhibitors of Cathepsin K.

Irreversible Inhibitor (reference)	Chemical structure	Activity
E-64 (L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane) Epoxysuccinyl chemical (19)		IC_{50}: 1.4 nM
Diacyl-bis hydrazide (12)		K_i: 0.7 nM
K-17 (Morpholino-leucine-homophenylalanine-vinylsulfone) Vinyl sulfone (20)		NA
Dipeptidyl ketone (12)		K_i: 8.0 nM

Chemical structures, K_i values and IC_{50} values of inhibitors are listed.

Table 2: Reversible Inhibitors of Cathepsin K.

Reversible Inhibitor	Chemical structure	Activity
Cystatin C (natural inhibitor in blood) (21)	NA	K_i:4.1 pM
Balicatib Nitril-based inhibitor (22)		IC_{50}: 1.4 nM
L-006235 Nitril-based inhibitor (23)		K_i:0.2 nM
Relacatib Non-basic 7-methyl-substituted azepanone analogue (22)		K_i:0.041 nM
Odanacatib Nitrile-based inhibitor (22)		IC_{50}:0.041nM
ONO-5334 (24)	NA	K_i: 0.1 nM
MV061194 Ketone based inhibitor (22)	NA	K_i: 2.5 nM

Chemical structures, K_i values, and IC_{50} values of inhibitors.

1.5 Importance of Exosite Inhibitors

If an enzyme such as cathepsin K had only the collagenase function in osteoclasts, the blocking of its active site would not be a problem. However, there are now several studies showing that the expression of cathepsin K is not limited to osteoclasts and that many other physiologically relevant substrates exist. For example, chondrocytes, synovial fibroblasts, and macrophages also express this protease (25) and Zhang et al. identified that cathepsin K has been implicated in the degradation of growth factors such as TGF-ß1. Hence, blocking the active sites of cathepsins could lead to the loss of functions in other cells. Some cathepsin K inhibitors, which were developed for treating osteoporosis such as balicatib and relacatib, were terminated due to causing severe side effects (summarized in review paper 22). Thus, from our point of view, it is not wise to block the active sites of cathepsins. We believe that a new strategy should be developed to selectively block existing exosites of cathepsins, which are required for the degradation of extracellular matrix proteins. In this way, bone resorption mediated by cathepsin K in osteoclasts, i.e. the degradation of collagen, could be inhibited without affecting other activities in the skin and lungs.

Interestingly, cathepsin K-deficient mice have severe learning and memory impairments (26). It has been shown that this protease is also expressed in the thyroid epithelium. It is possible that cathepsin K participates in thyroid-globulin processing (27). This further illustrates that cathepsin K has other non-collagenous substrates. If cathepsin K deficiency causes learning and memory impairments in mice, long-term treatment with cathepsin K active site-directed inhibitors of osteoporosis patients is of concern.

In addition, cathepsin K has a kininase activity, where it cleaves the regulatory peptide, bradykinin. Bradykinin can lead to bronchial constriction and hyper-responsiveness in

asthmatic patients (28). Once cathepsin K activity is totally blocked, it could not cleave bradykinin and thus may increase the susceptibility toward asthmatic attacks. All these examples demonstrate that exosite-targeting inhibitors would be beneficial as they would not block these essential activities.

1.6 Current Understanding of Exosite

Substrates do not only bind the active site of a target enzyme but could also bind to sites on other surface domains that are distant from the active site. These sites are called exosites (29). To date, research in exosite biology is still very limited. Some progress has been made in serine and metalloproteases (29, 30).

An example of putative exosite use is found in matrix metalloproteinases (MMP). It has been shown that the degradation of the connective tissue and proteolysis of the non-matrix proteins could result in physical diseases, such as inflammatory diseases. MMPs have the function of degrading extracellular matrix proteins as well as the processing of many bioactive molecules (30). In MMPs, with the exception of the binding of the substrate to the active sites, the secondary binding occurs through specialized secondary substrate binding sites on discrete substrate binding domains, or smaller functional modules, located outside the active site. These secondary specific binding sites are termed exosites (31). The exosites have the function of increasing the affinity of the proteinases to the substrate and modifying the function of the catalytic domain. These different binding properties could alter the substrate preference of the MMP and may increase its ability of degrading particular substrates. For example, exosites for native collagen are found on the hemopexin C domain of collagenases MMP-1, 8 and 13 (32). Exosites may play a critical role in substrate preparation prior to cleavage. In MMPs, one location of the exosites is on the hemopexin C

domain. In MT-MMP1, the hemopexin C domain, which acts as the exosite, binds to native collagen and it is thought that this interaction is necessary for the unwinding of the triple helical structure of collagens (30).

While exosites of gelatinase, MMP-2, are on three-fibronectin type II modules, which form an alternative collagen binding domain, the hemopexin C domain of MMP-2 is unable to bind native collagen. The triple fibronectin type II modules are involved in binding denatured collagen, which defines MMP-2 as a gelatinase and not a collagenase (30). Hence, the various exosites of MMPs modulate the various functions of the proteinases (30).

Cathepsins have been shown to have exosites needed for the binding of specific substrates such as elastin. Cathepsin V shares 78% amino acid sequence identity and similar 3D crystal structure with cathepsin L (33, 34). Cathepsin V is a highly effective elastase, however, cathepsin L only has minimal elastolytic activity (85). This suggests there are exosites interacting with elastin existing in cathepsin V. To determine the locations of these exosites, 11 chimeras were generated by exchanging analogous sequences of cathepsin L into cathepsin V. By evaluating their elastolytic activities, two exosites in cathepsin V were found (35). Exosite 1 contains amino acid residues V_{92}AVDEICKYRPEN$_{104}$ and exosite 2 covers the region of T_{113}VVAPG$_{119}$. The hydrophobicity of these two exosites is critical for binding elastin, which mainly consists of neutral and hydrophobic amino acid residues (36). In addition to the active site, the exosites contribute to the elastase activity of cathepsin V. These sites can be exploited for the binding of inhibitors competing with these substrates or interfering with special ligand binding required for matrix protein cleavage.

1.7 Collagen

Collagen is mostly found in connective tissues. There are 28 types of collagen in mammalian species, but not all have a helical structure. The most abundant collagens are type I and II collagens (37). They are characterized by their triple helix content, which make them resistant against general proteolysis. (38). Each parallel chain will form a left-handed, polyproline II-type helical coil. (41) The packing structure of polyproline II-type helices results in an identical sequence to GlyXaaYaa (Xaa and Yaa refer to any amino acid) (39). The most common motifs in collagen are GlyXaaHyp and GlyProYaa (40). The individual triple helix is called tropocollagen and forms collagen fibrils by covalent cross-linking with other tropocollagen molecules. Multiple collagen fibrils then will form collagen fibers (41). Type I collagen is the most abundant collagen representing over 90% of the organic bone matrix. The major collagen in cartilage is type II collagen (42). The most effective mammalian collagenases are matrix metalloproteases and lysosomal cathepsin K. The former one has one target cleaving site generating a ¼ and ¾ fragments each and the latter one cleaves at multiple sites and releases a range of fragments (9).

1.8 Glycosaminoglycans (GAGs)

GAGs are known as a component of proteoglycans, which are glycosylated proteins. They are mostly found in the extracellular matrix (ECM) and have multiple functions including hydration, storage, regulation, and tissue buffering (44). GAGs are polysaccharides which contain repeating disaccharide units. The repeating disaccharides are composed of N-acetylglucosamine or N-acetylgalacosamine paired with a uronic sugar (44). GAGs are divided into different classes based on the disaccharide type (Table 4): chondroitin sulfate (CS), heparan/heparan sulfate (HS), dermatan sulfate (DS), keratan sulfate (KS) and

hyaluronic acid (HA) (46). CS is the most common GAG existing in cartilage, bone, skin and tendons. Each monosaccharide in CS can be unsulfated, sulfated once or sulfated twice, which makes CS to exert different properties (45). Hence, CS is divided into CSA (C-4S), CSC (C-6S), CSD (C-2,6S) and CSE (C4, 6S) due to the type of sulfation. It has been demonstrated that CS has a beneficial effect on the metabolism of chondrocytes, synoviocytes, and on subchondral bone (45).

Table 3: Different Types of GAGs and Their Disaccharide Types

GAG	Disaccharide type	U	H	Modifications
HS	$U_{2X\ (\alpha/\beta1,4)}\ H_{NY,\ 3X,\ 6X\ (\alpha1,4)}$	IdoA/ GlcA	Glucosaminine	X- sulfated Y- accetylated/sulfated
CS/ DS	$U_{2X\ (\alpha/\beta1,3)}\ H_{NAc,\ 4X,\ 6X\ (\beta1,4)}$	IdoA/ GlcA	Galactosamine	X- sulfated
KS	$Gal_{6X\ (\beta1,4)}\ H_{NAc,\ 6X(\beta1,3)}$	Gal	Glucosamine	X- sulfated
HA	$GlcA_{\ (\beta1,3)}\ H_{NAc\ (\beta1,4)}$	GlcA	Glucosamine	None

Abbreviations used in the table are as follows: IdoA or I stands for α-L-iduronic acid; GlcA or G stands for β-D-glucuronic acid; Gal stands for β-D-galactose; Ac stands for Acetylation. Depending on the GAG class, H is either GlcNAc (α/ β-D-glucosamine) or GalNAc (β-D-galactosamine) (47).

1.9 Bone Related Diseases

Bone remodelling contains two phases, bone resorption and bone formation. It is important to achieve a balance between these two phases. Once the balance is destroyed, different bone diseases will arise, such as osteoporosis and pycnodysostosis.

1.9.1 Osteoporosis

Osteoporosis is the common bone disease found in elderly patients that is characterized by low bone mineral density and a high risk of fracture (5). The disease is more frequent in women due to the lack of estrogen after menopause. Estrogen downregulates osteoclast formation and activity (86). Today, approximately two hundred million people worldwide suffer from osteoporosis (48). The mechanism of the onset of osteoporosis is that osteoclasts mediate excessive bone loss that is not rectified by the bone formation mediated by osteoblasts (4). Clinically, osteoporosis can be defined by low bone mineral density (BMD) and a history of fragility fractures.

Cathepsin K is responsible for the degradation of bone matrix proteins. Excessive bone degradation in osteoporosis has been linked to excessive cathepsin K activity (4). Current therapies for osteoporosis include hormone therapy, bisphosphonates, selective estrogen receptor modulators, and human parathyroid hormone therapy. Various side effects were observed in patients upon taking these medications. Hormone replacement therapy was reported to increase risks in breast cancer, heart attacks and blot clot formation (50). Bisphosphonates were indicated to induce apoptosis of osteoclasts and suppress the osteoclast-mediated bone resorption (13). Jaw and hip necrosis occurred in patients on oral osteoporosis therapy of bisphosphonates (22). Raloxifene, one of the selective estrogen receptor modulators, had adverse effects related to blood clots, strokes and developmental abnormalities (51).

Taking these side effects and cathepsin K's predominant collagenase activity into consideration, alternative therapy inhibiting cathepsin K were considered. Various pharmaceutical companies have put efforts into developing cathepsin K inhibitors. However,

none of the compounds have yet been approved by the FDA and some of them failed due to their intolerability and extraskeletal side effects, such as Balicatib (52). Odanacatib, a nitrile peptidomimetic derivative developed by Merck & Co, is the most promising drug candidate so far, which successfully passed a large patient phase III clinical trial (49). However, there are still safety issues open before Odanacatib can get approved.

1.9.2 Pycnodysostosis

Pycnodysostosis is a bone disease caused by cathepsin K deficiency. The disease phenotype was first described in 1962 by Maroteaux and Lamy (53). The main features of this disease are short stature, cranial dysplasia, an increase in the bone density of long bones, stubby hands and feet with dystrophic nails and pathological fractures with poor healing. A potential pycnodysostosis gene was localized on chromosome 1q21 and later confirmed as locus for cathepsin K (54). When cathepsin K's expression is affected, the collagenase activity is abolished, which results in no degradation of the bone matrix. So far, there are 33 mutations on the CTSK gene that are reported to cause pycnodysostosis (54).

In the following, I will describe the two projects of my thesis, which 1) are aimed to characterize GAG-protein interaction sites in cathepsin K, and 2) to characterize and identify specific exosite inhibitors of this protease.

Chapter 2: Characterization of GAG Binding Sites in Tetramer and Dimer Cathepsin K Complexes

2.1 Introduction

Two cathepsin K/ GAG complex models have been proposed in our laboratory based on X-ray structure; the dimer and tetramer models (Figure 4). Site directed mutagenesis were performed at protein-protein binding sites and protein-GAG binding sites in our laboratory. The aim of my project is to characterize the effects of protein-GAG binding sites in both models on collagen degradation. In Figure 4, protein residues interacting with GAGs are highlighted.

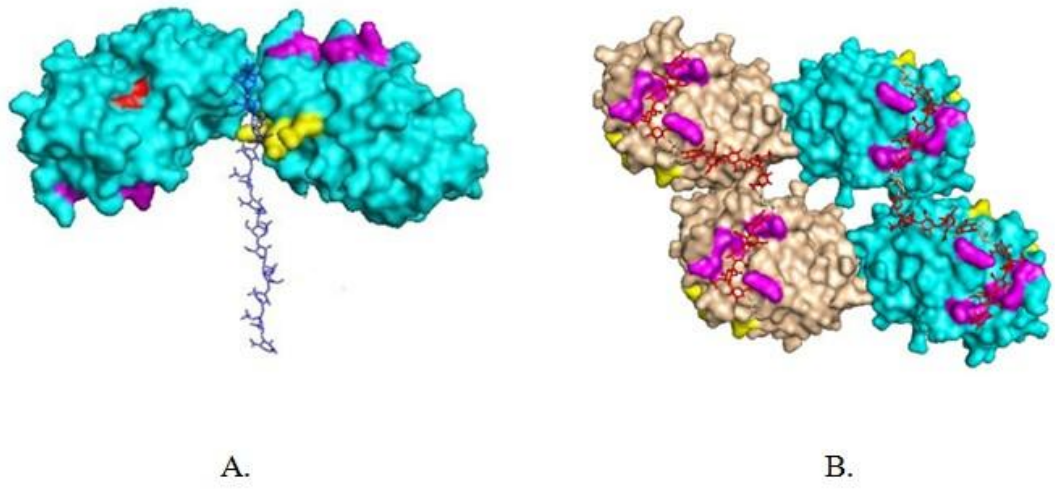


Figure 4. Cathepsin K Dimer and Tetramer Models Derived from Crystal Structures of Cathepsin K/GAG Complexes.

A. Cathepsin K/GAG dimer model B. Cathepsin K/ GAG tetramer model. For the cathepsin K/ GAG dimer model, cathepsin K molecules are in cyan, GAG binding sites are in yellow, GAG is in blue, and active sites are in red. For the cathepsin K/GAG tetramer model, cathepsin K molecules are in wheat and cyan, GAG binding sites are in magenta and GAG is in red. The active site is hidden as it is located on the reverse side of the complex.

Mutations of GAG binding sites were made for both cathepsin K dimer and tetramer complexes (Table 4). In the dimer model, amino acid residues K119, K122, R123 and R127 were mutated to alanine and D, M, E and T (cathepsin L like replacement), respectively. These amino acid residues were identified as being involved in the GAG binding, according to our X-ray structures (unpublished data). Alanine scanning is widely used to determine the contribution of individual amino acid residues on the activity and function of proteins as alanine only contains a small methyl group as side chain and thus side chain interactions are limited. Since cathepsin L does not show collagenolytic activity in the presence or in the absence of GAGs, GAG binding residues in cathepsin K, lysine and arginine residues were substituted with those present in cathepsin L (D, M, E and T) (11). Lysine and arginine residues are positively charged, which could interact with negatively charged GAG. Residues D and E provide negative charges which could disrupt the interaction of cathepsin K with GAGs. In case of the tetramer model, amino acid residues K9, I171, Q172, N190, K191 and L195 were mutated to alanine. These residues have been shown to interact directly with CS in the X-ray structure (11).

Earlier studies on the cathepsin K/ C4-S structure revealed that each cathepsin K molecule interacts with 6 hexasaccharide units of C4-S (55). X-ray structure analysis of the complex identified more than 12 amino acid residues in cathepsin K that are involved in the C4-S binding. Our laboratory has previously substituted 6 of them that were specific for cathepsin K binding with the appropriate residues present in cathepsin L and a 60% inhibition of the collagenolytic activity was obtained (11). In my project, these same residues were mutated into alanines to determine their individual role in collagenolytic activity.

Table 4: Mutations of Cathepsin K Based on Cathepsin K/ GAG Dimer and Tetramer Models.

Based model	Tetramer model	Dimer model
GAG Binding residues	K9, I171, Q172, N190, K191, L195	K119, K122, R123, R127
Mutants	Alanine scan (M3)	Alanine scan (M1), Cathepsin L like replacement (M2)

2.2 Hypothesis

Specific cathepsin K/GAG complexes are needed to degrade collagen. The dimer model and tetramer model, which were derived from crystal structures, have been proposed by our laboratory. Cathepsin GAG interaction sites are located at various surfaces in both models. The dimer complex can degrade collagen fibers into soluble tropocollagen molecules, which can further be degraded by the tetramer. These complexes are not required for non-collagenase activities of cathepsin K. **My hypothesis is that mutating the GAG binding sites in the dimer complexes selectively inhibits the degradation of insoluble collagen fibers by cathepsin K. While mutating those sites in the tetramer complexes primarily affects the degradation of soluble tropocollagen.**

2.3 Objectives

Mutations at cathepsin K/GAG binding sites will be generated in both dimer and tetramer models. Enzyme kinetic and stability parameters of these mutants will be determined.

Collagenolytic, gelatinolytic, and elastinolytic activities will be determined to evaluate the role of cathepsin K/GAG binding sites in these activities.

2.4 Material and Methods

2.4.1 Materials

Benzyloxycarbonyl-Phe-Arg-7-amido-4-methylcoumarin (Z-FR-MCA) was from Enzyme System Products (Dublin, USA). Soluble calf-skin type I collagen was purchased from Affymetrix, Inc. (Santa Clara, CA). Rat-tail type I collagen, chondroitin sulfate A (C4-S) were from Sigma-Aldrich (St. Louis, MO). The cathepsin inhibitor E-64 (L-3-carboxy-trans-2-3-epoxypropionyl-leucylamido-(4guanidino)-butane) was purchased from Bio Basic Inc. (Markham, Canada).

2.4.2 Construction of Cathepsin K Mutants

Wild-type human cathepsin K in pPIC9 vector (Invitrogen, Burlington, ON) was used as a template for site directed mutagenesis. *Pfu*-polymerase (Fermentas, Burlington, ON) was used in PCR. PCR system contains 10 ng of template plasmid, 125 ng of each primer, 250 µM dNTP, 1 U pfu DNA polymerase, 5 µL of 10x Buffer and H_2O to 50 µL total volume. The conditions of PCR site-directed mutagenesis are listed in Table 5 and Table 6. Primers used are shown in Table 7 (Integrated DNA Technology, Coralville, Iowa). The mutated cDNA was digested by *EcoRI* and *NotI* restriction enzymes (NEB, CANADA), that then ligated into pPIC9 vector. Mutations were confirmed by DNA sequencing (Genewiz, South Plainfield, NJ). The construct was linearized with *SacI* restriction enzyme, which went through electroporation into *Pichia pastoris* GS115 (Invitrogen protocol).

Table 5: The Conditions of the First Step PCR for Cathepsin K Mutants.

1. Initial denaturing	95 °C 45s
2. Denaturing	95 °C 30s
3. Annealing	58 °C 30s
4. Extension	72 °C 1 min
Repeat step 2-4 for 35 times	
5. Final extension	72 °C 10 min

Table 6: The Conditions of the Second Step PCR for Cathepsin K Mutants.

1. Initial denaturing	95 °C 3 min
2. Denaturing	95 °C 15s
3. Annealing	55 °C 40s
4. Extension	72 °C 1 min 30s
Repeat step 2-4 for 26 times	
5. Final extension	72 °C 10 min

Table 7: PCR Primers for Cathepsin K Mutants.

Mutants	Primer sequences
M1 (K119, K122, R123, R127 to A)	
Forward	5’ ggg aat gag gca gcc ctg gcg gcg gca gtg gcc gca gtg 3’
Reverse	5’ cac tgc ggc cac tgc cgc cgc cag ggc tgc ctc att ccc 3’
M2 (K119, K122, R123, R127 to D, M, E, T)	
Forward	5’ ggg aat gag gac gcc ctg atg ggg gca gtg gcc aca gtg 3’
Reverse	5’ cac tgt ggc cac tgc ccc cat cag ggc gtc ctc att ccc 3’
M3 (K9, I171, Q172, N190, K191, L195 to alanine)	
Forward	K9: 5’ gtc gac tat cga gcg aaa gga tat gtt 3’ I171, Q172: 5’ gtg gga tat gga gcc gcg aag gga aac aag 3’ N190, K191, L195: 5’ gaa aac tgg gga gcc gca gga

Mutants	Primer sequences
	tat atc gcc atg gct cga 3’
Reverse	K9: 5’ aac ata tcc ttt cgc tcg ata gtc gac 3’ I171, Q172: 5’ ctt gtt tcc ctt cgc ggc tcc ata tcc cac 3’ N190, K191, L195: 5’ tcg agc tac ggc gat ata tcc tgc ggc tcc cca gtt ttc 3’

A summary of all the primers used in constructing cathepsin K mutants. In M1, K119, K122, R123, R127 were replaced to alanines. M2 replaces K119, K122, R123, R127 to D, M, E, T (cathepsin L like replacement). M3 replaces K9, I171, Q172, N190, K191, L195 to alanines. The substituted nucleotides are underlined.

2.4.3 Protein Expression

After cathepsin K mutants’ cDNAs were transformed into GS115 yeast cells, minimal dextrose (MD) and minimal methanol (MM) agar plates were used to screen mutants. Since the pPIC9 vector has the promoter for alcohol oxidase, only positive colonies (colonies with pPIC9 in the genome) can grow on MM plates. Positive clones were picked up to be cultured in 5mL BMGY media at 30°C until the culture reached an O.D.600 of 2.0-6.0. BMGY is composed of 1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 1.34% yeast nitrogen base, 4x 10-15% biotin, and 1% glycerol. The cells were spun down and supernatant was discarded. 5 mL of BMMY media were added into cells. BMMY contains 1% yeast extract, 2% peptone, 100 mM potassium phosphate (pH 6.0), 1.34% yeast nitrogen base, 4x 10-15% biotin, and 1% methanol. 0.5% methanol was added to the yeast culture every 24 hours to induce protein expression. At day 4 of methanol induction, usually maximum

protein production is reached. After being activated by pepsin (as described below), the yeast culture supernatants were tested for activity using the synthetic substrate, Z-FR-MCA. Positive clones, which were His+ Mut+, contained detectable peptidolytic activity. The clone that had the highest activity was chosen for large scale expression. 5 mL of MD media was used to grow the clone at 30 °C for 24 hours, which was inoculated into 50 mL of BMGY media for 12 hours. After the cells were spun down and supernatant was discarded, 500 mL BMGY was added into pellet, which was cultured until O.D.600 was 6.0. Then cells were transferred into 1.5 L BMMY media, followed by 0.5% methanol induction every 24 hours. 500µL of sample were used to monitor the Z-FR-MCA hydrolyzing activity of the culture.

It was particularly important to safeguard against bacterial contamination by evaluating the culture supernatants microscopically every day. The cells were ready to be harvested at day 4 post methanol induction and the supernatant was concentrated using an Amicon Ultrafiltration membrane (EMD Millipore, Darmstadt, Germany) to a final volume of 30 mL at 4 °C. This material was used for the subsequent pepsin-mediated protease activation.

2.4.4 Protein Activation and Purification

0.6 mg/mL pepsin (Sigma-Aldrich, St. Louis, MO) was used to activate the cathepsin K precursor at pH of 4. The pH was lowered to pH 4.0 with 20% glacial acetic acid (Fisher Scientific, Ottawa, ON). The activation mixture was incubated at 37°C and the cathepsin K enzyme activity was monitored with Z-FR-MCA in a standard activity buffer. Activity buffer is composed of 100 mM sodium acetate, pH 5.5, containing 2.5 mM EDTA, and 2.5 mM DTT.

It is necessary to raise the pH to 5.5 to stop pepsin activity when maximum activity is reached as further incubation will lead to a progressive degradation of the cathepsin K activity. For the purification process I added ammonium sulfate to the activated enzyme to 2M and concentrated to 50 mL by centrifugation in a centricon tube (cut-off: 10kDa) at 4000 g at 4 °C. Supernatant was kept and loaded onto a n-butyl sepharose column (GE Healthcare, Little Chalfont, U.K.). Here the loading buffer was used to wash the column. The loading buffer contains 100 mM sodium acetate, pH 5.0, 2 M ammonium sulfate, 0.5 mM EDTA, and 0.5 mM DTT. Using an ammonium sulfate linear gradient, the protein was eluted. Z-FR-MCA was used as substrate to monitor the elution of active protease. Active fractions eluted at 0.5 M ammonium sulfate and were combined and concentrated using Amicon ultra concentrator (EMD Millipore, Darmstadt, Germany). Buffer exchange was performed using 100 mM sodium acetate buffer (pH 5.5, 0.5 mM EDTA and 0.5 mM DTT). Protein was analyzed on SDS-PAGE gels to monitor the purity of the preparation and samples were stored at -80 °C till further use.

2.4.5 Enzyme Concentration Determination and Kinetic Studies

E-64 active site titration was used to determine the enzyme concentrations using the spectro-fluorometer Perkin-Elmer LB50 (16). The excitation wavelength was 380 nm and emission wavelength was 460 nm. 5 µL diluted enzyme was mixed with 5 µL of various concentrations of E-64 solutions and incubated for 30 min on ice in 90 µl of activity buffer. 5µL of the mixture were used to measure the residual activity by Z-FR-MCA hydrolysis. Activity levels were plotted against E-64 concentrations and the enzyme concentrations were calculated.

Non-linear regression analysis was used to determine Michaelis-Menten constants (k_{cat} and K_m). Enzymes were first diluted to 1-5 nM and activity was determined at variable Z-FR-MCA concentrations (1-20 μM) in activity buffer (as described above).

2.4.6 Stability Assay

Wild-type cathepsin K and mutants were incubated at different temperatures in standard activity buffer (room temperature/22°C, 28°C, and 37°C). The enzyme activity was determined at the following time points: 0, 0.5, 1, 2, 3, and 4 hours using the substrate Z-FR-MCA. Residual activities were plotted against time.

2.4.7 Protein Degradation Assay with Soluble Tropocollagen

In the 50 μL reaction volume, 0.6 mg/mL soluble calf-skin type I collagen was incubated with 400 nM wild-type cathepsin K or mutants, and 200 nM of C4-S in activity buffer for 4 hours at 28°C. After this, 1 μL of 100 μM E-64 was added to each tube to inhibit the residual enzyme activity. All reactions were kept overnight at 4°C and then ran on 8% SDS-PAGE gels which were stained with Coomassie Blue. To characterize the collagen degradation, the density of collagen α1 band of each reaction was quantified using GeneSnap program (Syngene Inc. Frederick, MD). To monitor the degree of degradation, densities of collagen α1 bands of all experimental groups were standardized against collagen control group.

2.4.8 Protein Degradation Assay with Insoluble Collagen

1 mg insoluble rat-tail type I collagen fibrils isolated from Sprague Dawley rats were mixed with wild-type cathepsin K or mutants (final concentration 400nM) in activity buffer and then incubated for 4 hours at 28˚C. 1 μL of 100 μM E64 was added to inhibit cathepsin K's activity. Before running samples on SDS-PAGE, samples were centrifuged for 20 min and

supernatants were loaded onto the gel. Only fragments released from the fibers will be visible on the polyacrylamide gel. After Coomassie Blue staining, densities of collagen α bands were measured and standardized against wild-type cathepsin K group. Released α1 bands visible on SDS-PAGE gels reveal the degradation of insoluble collagen.

2.4.9 Gelatinase Assay

0.6 mg/mL gelatin was generated by heat denaturation of soluble calf-skin type I collagen at 95°C for 20 minutes. Then, gelatin was incubated with 5 nM wild-type cathepsin K or mutants in activity buffer for 30 min at 28°C. Subsequently, 1 µL of 100 µM E-64 was added to stop the reaction. After overnight storage at 4°C, all samples were run on 8% SDS-PAGE gels and stained with Coomassie Blue.

2.4.10 Elastase Assay

1mg of elastin congo red (Sigma, St. Louis, MO) was incubated with 1 µM wild-type cathepsin K or mutants in standard activity buffer at pH 5.5 and incubated overnight at 37°C. All samples were centrifuged at 13.2×10^5 rpm for 5 min by using a bench-top AccuSpin microcentrifuge (Fisher Scientific, Ottawa, ON). Supernatants were measured at 490 nm using Vmax kinetic microplate reader (Molecular Device, Sunnyvale, CA).

2.4.11 Western Blot

Protein concentrations were determined by Bicinchoninic acid assay and 20 µg of each sample were loaded onto a 12% SDS-PAGE gel, which were transferred after electrophoretic separation to a nitrocellulose membrane at 200 mA for 2 hours using transfer buffer. Transfer buffer contained 25 mM Tris, 192 mM glycine, and 10% methanol. After transferring samples to Nitrocellulose membrane, 3% defatted milk dissolved in PBS buffer was used for

blocking of non-specific binding for 1h at room temperature. PBS buffer contained 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4, at pH 7.4. MS2 antibody, a rabbit polyclonal anti-human cathepsin K antibody (previously made in our laboratory), was used at the dilution of 1:200 as primary antibody overnight at 4°C. After washing out the primary antibody, an anti IgG HRP conjugated antibody (SCBT) was used as secondary antibody for 1 h at room temperature with a dilution of 1:1000. The membrane was washed with both PBST buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, and 1 mL tween-20, at pH 7.4) and PBS buffer thoroughly. For developing, ECL substrate (Biorad) was used and samples were detected. ECL substrate is an enhanced chemiluminescent substrate for detection of horseradish peroxidase activity from antibodies.

2.4.12 Statistics

Results are expressed as mean ± SD. The significance of differences of the mean values was calculated using one-way ANOVA (t-test). A p-value of less than 0.05 was considered significant.

2.5 Results

2.5.1 Purification

M1 and M2 mutant cathepsin K variants were successfully expressed and purified. Figure 5 shows single bands of purified mutants variants, M1 and M2, after n-butyl Sepharose chromatography.

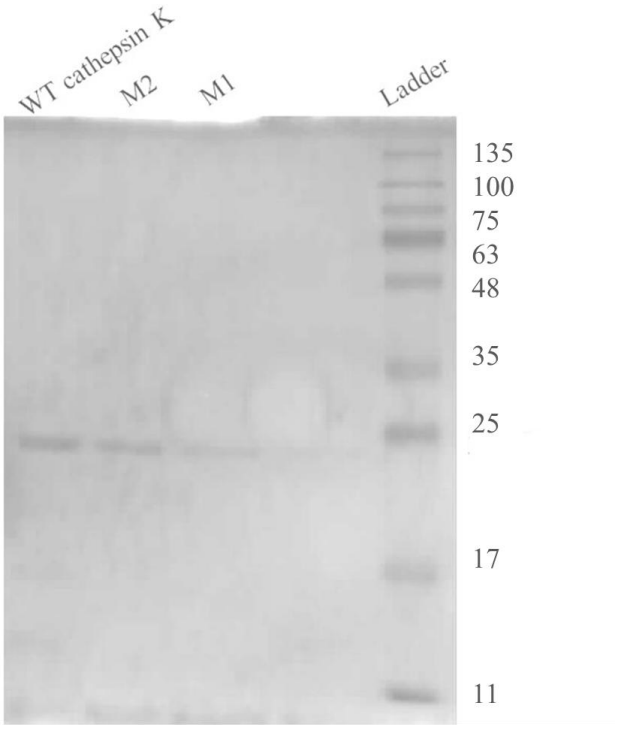


Figure 5. SDS-PAGE Gel of Purified Proteins (Coomassie Staining)

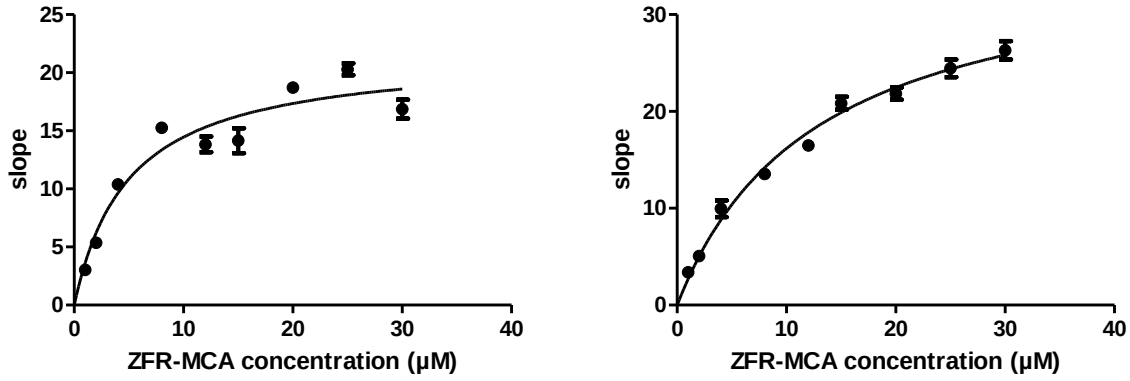
2.5.2 Enzyme Kinetics

M1 and M2 mutants were expressed and evaluated for their kinetic parameters. The kinetic parameters did not reveal major alterations when compared to wild-type cathepsin K. Both, the k_{cat}, K_m and k_{cat}/K_m values were comparable to the parameters obtained for the wild-type protease, indicating that mutations did not interfere with the integrity of the active site of these two mutants. The data are displayed in Table 8 and Figure 6.

Table 8: Kinetic Values for the Peptidase Activity of Cathepsin K.

Enzyme name	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (10^6 $s^{-1}M^{-1}$)
Wild-type cathepsin K	12.1±2.0	17.5±0.9	1.45±0.7
M1	5.9±2.2	9.2±1.0	1.55±1.2
M2	12.8±1.8	18.9±0.5	1.48±1.5

The constants K_m, k_{cat}, and k_{cat}/K_m were determined by using Michaelis-Menten kinetics. Different concentrations of Z-FR-MCA as fluorogenic substrate were used. The K_m is measured in μM and k_{cat} in s^{-1}. M1 contains alanine replacements for residues K119, K122, K123 and R127. In M2, K119, K122, K123 and R127 were replaced by D, M, E, and T. Data represent mean ± SD values of three independent experiments.



A. B.

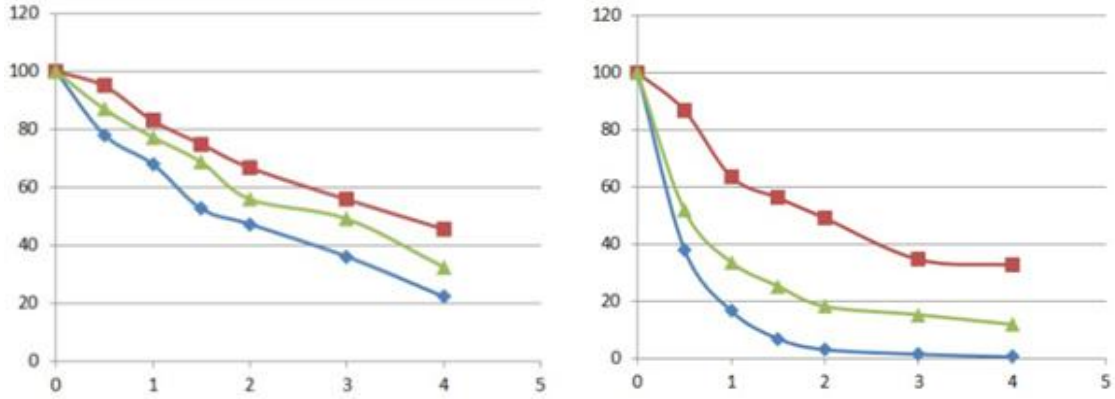
Figure 6. Plots for the Determination of K_m and k_{cat} according to Michaelis-Menten Kinetics.

A. Michaelis-Menten plot of M1 B. Michaelis-Menten plot of M2

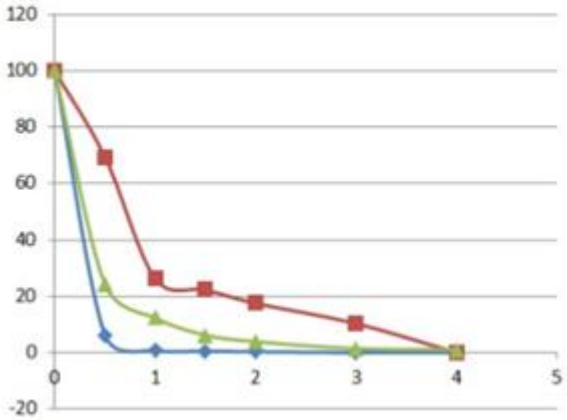
2.5.3 Stability of Mutants

Stability of both mutants was determined at different temperatures, including room temperature (22ºC), 28ºC, and 37ºC (Figure 7). At room temperature, wild-type cathepsin K retained 20% enzyme activity after 4 hours incubation at pH 5.5. Mutants M1 and M2

revealed to be more stable than wild-type cathepsin K, which was reflected by their higher residual activities at each individual time point. After four hours, M1 retained 45% residual activity, while M2 had 30% residual activity. At 28ºC, wild-type cathepsin K lost all activity after 4 hours at pH 5.5, while M1 and M2 retained significant residual activities after 4 hours (35% and 15%, respectively). At 37ºC, all enzymes lost activity after 4 hours at pH 5.5. From time point 0 to 4 hours, mutants kept higher residue activity than wild-type cathepsin K.



(a) (b)



(c)

Figure 7. Stability of Wide-type Cathepsin K and the Mutants

Stability of mutants and wild-type cathepsin K under different temperatures were tested against the substrate Z-FR-MCA (1mM). Samples were tested at room temperature (22ºC), 28ºC, and 37ºC at the time points of 30min, 60min, 90min, 2h, 3h, and 4h. The residual activities were plotted against time. (a) Residual activities at room temperature, (b) residual activities at 28ºC, (c) residual activities at 37ºC. Wild-type cathepsin K is in blue, M1 is in red, and M2 is in green. M1 has replaced amino acid residues K119, K122, R123 and R127 into alanine residues. M2 replaces K119, K122, R123 and R127 into D, M, E and T. Values were representatives of one independent experiment.

2.5.4 Gelatinase Activity

Gelatinase assay is used to determine whether mutations led to an alteration of the general protease activity of the variant cathepsin K proteases. As revealed by SDS-PAGE analysis,

the gelatin control line showed gelatin bands whereas gelatin was totally degraded by variants M1 and M2, and wild-type cathepsin K (not shown) at 28ºC after half an hour incubation indicating that the gelatinase activities of the mutant variants were comparable to that of wild-type cathepsin K (Figure 8).

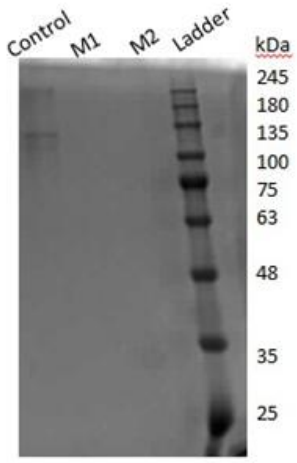


Figure 8. Gelatinase Assay.

Gelatinase assay of variants M1 and M2. Gelatin was prepared by heating calf-skin collagen at 100ºC for 30 minutes. 0.6 mg/mL gelatin and 5 nM enzyme were used in the assay at 28ºC for 30min. Values were representatives of three independent experiments.

2.5.5 Collagenase Activity

Soluble type I tropocollagen and insoluble collagen fibers were used to determine the collagenase activities of the M1 and M2 variants and wild-type cathepsin K. All samples were analysed on SDS-PAGE gels. After Coomassie Blue staining, the densities of collagen α1 bands were quantified to determine collagen degradation (Figure 9A and B).

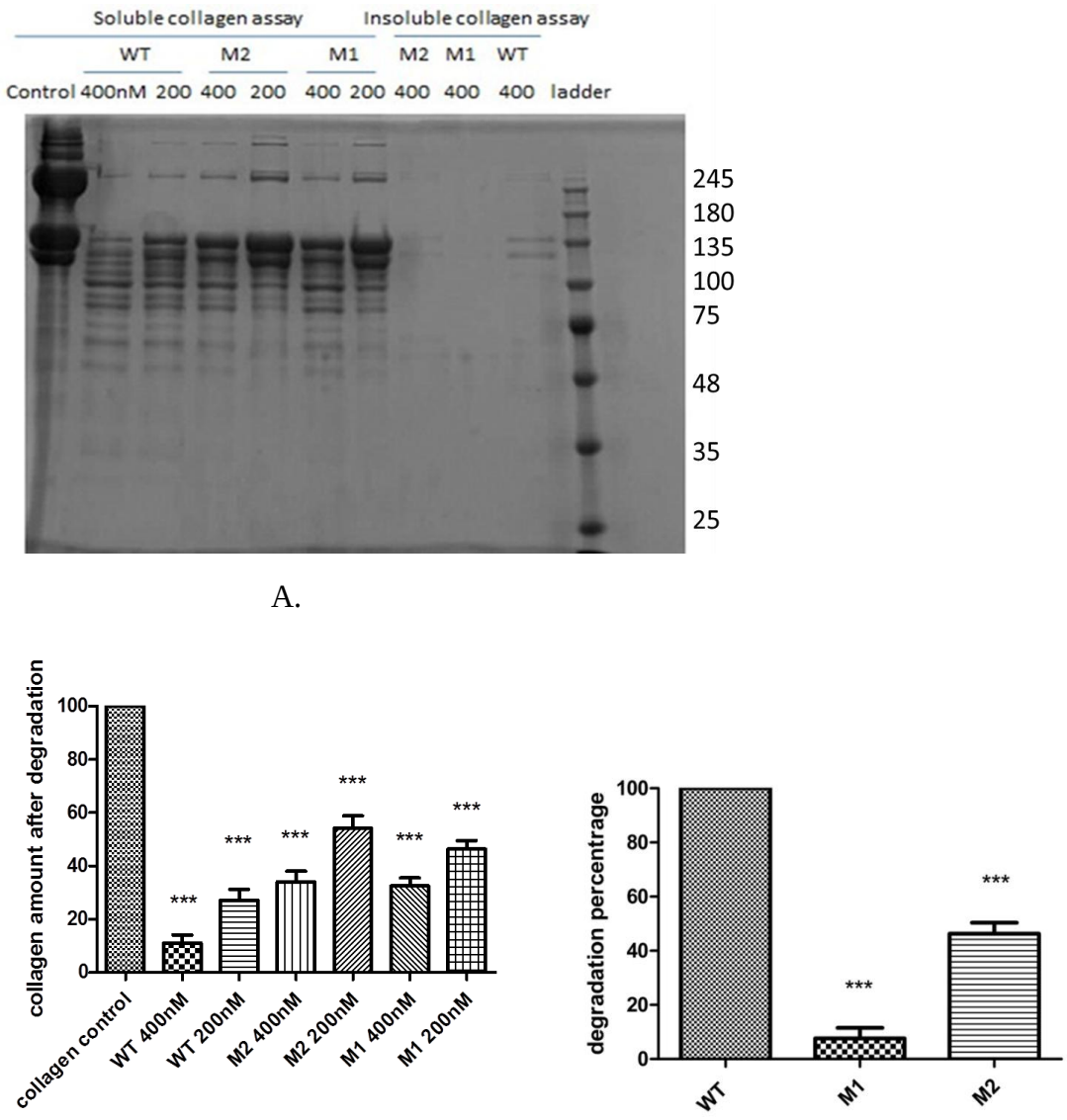


Figure 9. Collagenase Activity of Wild-type Cathepsin K, M1 and M2.

A. Collagenase assay of wild-type cathepsin K, M1 and M2 mutants in the presence of C4-S. 200 nM and 400 nM of wild-type cathepsin K, M1 and M2 were used with 100 nM and200 nM C4-S, respectively. 0.6 mg/mL calf-skin tropocollagen was used for soluble collagen degradation assay. 1 mg rat tail collagen fibers were used for the insoluble collagen degradation assay. All assays were performed at 28ºC for 4 hours and repeated three times. B. Corresponding collagenolytic activity quantifications. Densitometry of α1-bands on SDS-PAGE gels was used to determine the collagenolytic activity. (a) Quantifications of soluble collagen degradation assays (b) Quantifications of insoluble collagen degradation assays.

Experiments were repeated 3 times. *P*-value represents unpaired *t*-test statistical analysis relative to control, *** indicates $P < 0.001$.

Wild-type cathepsin K in the presence of C4-S was tested at two different protease concentrations for collagenase activity: 400 nM and 200 nM and revealed the expected collagenolytic activity as shown by the appearance of lower molecular weight degradation products (Figure 9). In contrast, variants M1 and M2 had both a reduced collagenase activity. At 200 nM variant concentration, M1 lost 19±5% activity, while the activity of M2 was decreased by 27±8% when compared to wild-type cathepsin K. The results at 400 nM variant concentrations were comparable as well. M1 lost 22±5% activity and M2 lost 23±7% activity when compared to wild- type cathepsin K.

M1 and M2 mutants were also tested against insoluble rat tail collagen fibers. When collagen fibers are degraded by cathepsin K, soluble tropocollagen will be released and can be quantified SDS-PAGE analysis. Variants M1 and M2 exhibited a reduction on degrading collagen fibrils with residual activities of 8±7% and 46±7% left, respectively. The measurements are based on density measurements of collagen α1 bands and were repeated 3 times.

2.5.6 Elastase Activity

Besides its potent collagenase activity, cathepsin K also has a significant elastin-degrading capability. Elastase activities of variants were characterized using Congo-Red elastin as substrate. Wild-type cathepsin K showed an optical density of 0.6 of Congo-Red release, whereas M1 and M2 revealed much higher optical densities of 1.2 and 1.1, respectively (Figure 10). This suggests that the elastase activity of M1 and M2 mutants was not affected

by these mutations. To the contrary, M1 and M2 possessed more elastase activity than wild-type cathepsin K, which is likely related to the increased enzyme stability as shown in Figure 7.

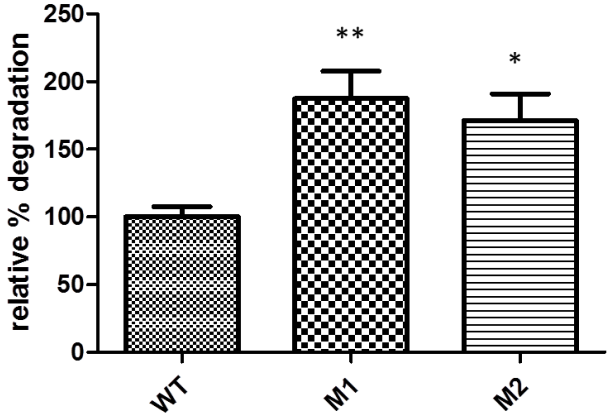


Figure 10. Elastase Activities of Wild-type Cathepsin K and Variants M1 and M2.

Elastase assay of wild-type cathepsin K, M1 and M2 variants. Assays were performed at 1 μM enzyme, 1 mg of insoluble elastin congo red as substrate at 37ºC overnight. Experiments were repeated 3 times. *P*-value represents unpaired *t*-test statistical analysis relative to WT, ** indicates $P < 0.01$, * indicates $P < 0.05$.

2.6. Characterization of M3 mutations

For the M3 mutant, no clones expressing a protease activity (which are His+ Mut+ and need to possess detectable peptidolytic activity against synthetic substrate Z-FR-MCA) could be obtained. To verify whether there was protein expressed for M3, a western blot was performed (Figure 11). Samples prior pepsin activation and after activation at different time points at day 4 were examined with wild-type cathepsin K as control.

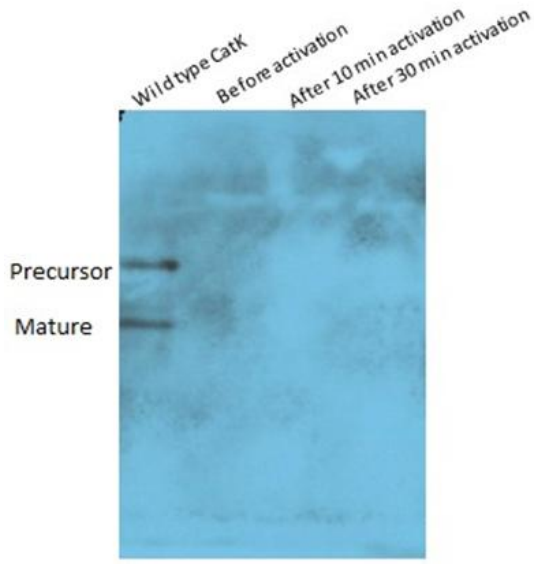


Figure 11. Western Blot of Wild-type Cathepsin K and M3 Mutant

MS2 rabbit polyclonal anti-human cathepsin K antibody was used as primary antibody and anti-rabbit IgG HRP conjugated antibody was used as secondary antibody. Purified wild-type cathepsin K was used as control. M3 samples were collected on day 4 after induction. M3 samples were checked both prior to pepsin activation and after pepsin activation (10 min and 30 min). In M3, amino acid residues K9, I171, Q172, N190, K191 and L195 were replaced by alanine residues.

Figure 11 reveals that wild-type cathepsin K control sample had two bands; the upper one was the precursor protein and the lower one was the active protein. This shows that western blotting allows the identification of the two protein forms of the enzyme. In contrast, no protein was detected for the M3 mutant, neither before nor after pepsin activation. This illustrated that no stable enzyme was obtained from yeast culture, which explains that no enzyme activity against Z-FR-MCA after pepsin activation was detected.

2.7 Discussion

To act as a collagenase, cathepsin K needs to bind with GAGs to form oligomeric complexes (9). However, the mechanism of collagenase activity of cathepsin K/GAG is not yet completely understood. Based on X-ray diffraction crystallography data, cathepsin K/GAG complex dimer and tetramer models were hypothesized. My project focused on the individual protein-GAG interaction sites present in these two models. Two mutants involved in the dimer complex, M1 and M2, were successfully generated and examined for the effect of their substitutions on the proteolytic and specifically on the collagenolytic activities. Mutations in M3, which is involved in the tetramer complex, led to protein destabilization. Potential reasons will be discussed later in this chapter.

2.7.1 Effects of the Activities of M1 and M2 Variants

The dimer model revealed binding sites on cathepsin K for GAGs, which include amino acid residues K119, K122, R123, and R127. In M1, these sites were mutated into alanine residues and in M2 mutated into D, M, E, and T residues, representing a cathepsin L like replacement. Since the mutated residues were remote from the active site of cathepsin K and the structure of cathepsin K was not altered by the complex formation (55), the specificity constants k_{cat}/K_m for these two variants were expected to be comparable to that of wild-type cathepsin K. The K_m value is the substrate concentration at which the reaction rate is half of V_{max}. In my project, the substrate refers to Z-FR-MCA. Wild-type cathepsin K and M2 reach their half of V_{max} at 18 μM of Z-FR-MCA. M1 could get the same reaction rate when 9 μM of Z-FR-MCA was applied. k_{cat} is the "turnover number" which tells how many Z-FR-MCA molecules are cleaved in one second. Wild-type cathepsin K and M2 turn over 12 Z-FR-MCA molecules into product per second, which is two folds more than that of M1. The

accuracy of the individual kinetic parameters, k_{cat} and K_m depends whether a full saturation of the Michaelis/Menten curves was reached. In contrast, k_{cat}/K_m values are independent from this and are in general a better parameter to compare enzyme activities. k_{cat}/K_m values for variants, M1 and M2, are about 1.5×10^6 $s^{-1}M^{-1}$, which is comparable to that of wild-type cathepsin K. This suggests that the introduced substitutions did not alter the catalytic and substrate binding efficacies of variants.

Furthermore, the gelatin degradation assay was not affected by the amino acid substitutions, as the overall turnover efficacy of the mutant variants was comparable to that of wild-type cathepsin K.

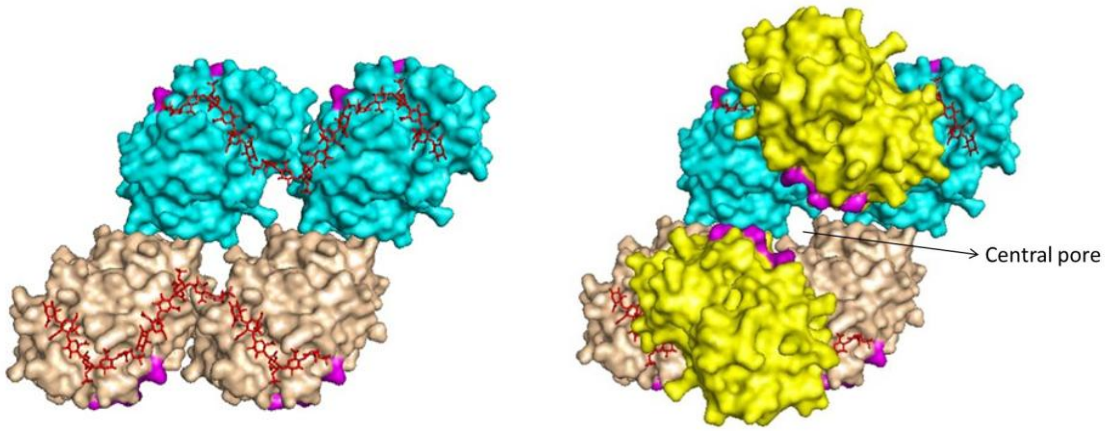
However, both of the variants influenced the collagenolytic activity of cathepsin K towards tropocollagen and insoluble collagen fibers to various degrees. Excluding the influence factor of enzyme stability, we can confirm that the specific binding of cathepsin K to GAG at the sites identified in the X-ray structure is critical for collagen degradation. M1 and M2 variants inhibited the degradation of insoluble collagen fibers by 92% and 54%, respectively.

Stability issues are unlikely the cause of this inhibition as the M1 mutant was even more stable that the M2 mutant. According to our hypothesis, the dimer complex should solubilize insoluble collagen fibers into tropocollagens where the GAG binding sites are required for the docking of the dimer to the collagen (submitted for publication, Brömme lab). Considering that the degradation of insoluble collagen fibers was inhibited, it can be assumed that the dimer complex or its binding to collagen fibers was disrupted by the substitution of GAG binding residues. Interestingly, the M1 variant showed a more potent inhibitory effect on the degradation of insoluble collagen fibers than the M2 variant at 28°C for 4 hours. The

potential reasons that the mutation in M1 resulted in a more effective inhibition are not clear. It would have been expected that the M2 variant which contains two negatively charged amino acid residues, D and E would interfere with the GAG bind more profoundly as these residues would repel the SO_4^{2-} groups in CS more effectively.

Regarding the tropocollagen degradation, the M1 and M2 variants revealed a 20%-30% inhibitory effect, compared to wild-type cathepsin K under the same conditions. Our laboratory has shown that two clusters of positively charged lysine and arginine residues contribute to collagenolytic activity (56). Among these lysine and arginine residues, K122 and R127, revealed approximately 20% inhibition of the collagenase activity when they were substituted by alanine residues (56). Since K122 and R127 are two of the residues mutated in M1 and M2, it is not surprising that M1 and M2 revealed a 20%-30% reduction of the collagenolytic activity. Our hypothesis is that the dimer complex could solubilize collagen fibers into tropocollagen-sized fragments and the tetramer complex would degrade these tropocollagen fragments into smaller peptides. Based on this, both variants are supposed to have no or little effect on the collagenolytic activity towards soluble collagen. X-ray structure of cathepsin K/GAG complex previously solved revealed 6 specific residues involved in the GAG binding (10). The substitution of these residues into those of cathepsin L still maintained the interactions of GAG and the variant, but altered the binding mode which led to the reduction of collagenolytic activity of the variant (11). When K119, K122, R123, and R127 were substituted by alanine, it was quite possible that these mutations did not prevent the binding of GAG to the variants, but lead to significant alteration of binding mode as well, resulting in the reduction of collagenolytic activity. Moreover, there is possibility that tropocollagen is the common substrate of the cathepsin K dimer as well as of the tetramer

complexes. The potential mechanism of the degradation of collagen fibers might be that cathepsin K dimer first degrades collagen fibers into tropocollagen, then cathepsin K dimer and tetramer cooperate to unwind and degrade tropocollagen. Another explanation is that cathepsin K dimer complex is responsible for degradation of collagen fibers, while tetramer complex is not the only one functionalized to degrade tropocollagen. The crystal structures of collagenolytically inactive cathepsin K with GAG form a "beads-on-a-strand" like organization. To obtain the cathepsin K tetramer model from this X-ray structure, two cathepsin K molecules binding on GAG are removed for simplification. By adding these two cathepsin K monomers back to tetramer complexes, a hexamer model was proposed (Figure 12).



A. B.

Figure 12. Cathepsin K/GAG Tetramer Model and Hexamer Model Derived from Crystal Structure.

A. Cathepsin K/GAG tetramer model, cathepsin K molecules are in wheat and cyan, GAG binding sites of dimer model are in magenta, GAG and active sites are in red. B. Cathepsin K/GAG hexamer model. Cathepsin K molecules which are the same as the ones in tetramer model are in wheat and cyan. Two additional cathepsin K molecules are in yellow. GAG binding sites of dimer model are in magenta, GAGs and active sites are in red.

In the cathepsin K tetramer model, it is assumed that there is a central pore in the middle of these four cathepsin K monomers to unwind triple helical strands of tropocollagen and further degrade them into small pieces. In the cathepsin K hexamer model, two additional cathepsin K molecules bind to the GAGs from the tetramer model and the putative GAG binding sites mutated in the dimer model are now facing the potential tropocollagen docking site which will guide the collagen molecule through the central pore. These additional two cathepsin K molecules in the hexamer model, if required for the collagen binding, could affect the guiding of the triple helical collagen into the pore. Hence, the tetramer model revealed some limitations based on my findings which are better explained by the hexamer model.

Elastin and collagen, two major fibrous proteins in the extracellular matrix, are highly resistant to proteolytic degradation. Only a very few selected metallo and cysteine proteases are capable of hydrolyzing them. Cathepsin K is one of the proteases functioning as both collagenase and elastase. There are two different mechanisms for collagenase activity and elastase activity. As stated above, GAG binding to cathepsin K is necessary for the collagenolytic degradation. The elastase activity of cathepsin K and other elastolytic cathepsins depends on the presence of at least two exosites which are required for elastin binding (35). Both exosites are on the surface of cathepsin K and distant from the active site (35). Considering that the residues mutated are remote from the active site and the two exosites, it is not surprising that M1 and M2 variants maintained their elastase activity in comparison to the wild-type protease. Interestingly, the elastase activities of these variants were even higher than that of wild-type cathepsin K. This could be explained by the increased protease stability of the two mutant variants (see Figure 7).

2.7.2 Destabilizing Mutations

The characterization of M3 variant failed due to the instability of the mutant protein. In the M3 variant, residues K9, I171, Q172, N190, K191, and L195 were mutated into alanines. It should be noted that a mutant (M5) substituting these 6 residues with those present in cathepsin L was generated and examined for proteolytic and collagenolytic activities by our laboratory before (11). Even though M5 and M3 have common mutation sites, it is quite possible that alanine replacement and cathepsin L-like replacement have significantly different effects on enzyme activities. The potential reasons why no active protein could be obtained are as follows: Firstly, mutation sites might directly destabilize the active site of cathepsin K so that the enzyme was inactive. Secondly, mutation sites might be involved in mutation sites known in pycnodysostosis, a disease which is caused by the deficiency of cathepsin K activity. It is well known that mutants related to pycnodysostosis fail to produce a stable protein (54). Thirdly, mutation sites might be highly conserved among cysteine cathepsins and thus mutations would be likely essential for the protein stability and activity.

The X-ray structure of the cathepsin K monomer provides the information about locations of its active site and the M3 mutation sites (Figure 13). Active site residue, Cys25, is located in the left domain, while His162 is in the right domain. M3 mutation sites are all in the right domain. Among them, the residues having potential effects on His162, are N190 and K191. N190 and K191 are located at the loop connecting to the ß-sheet where His162 is located. Introducing alanine residues into these two sites could disrupt the active site. This could be a potential reason that no stable protein could be expressed for the M3 variant.

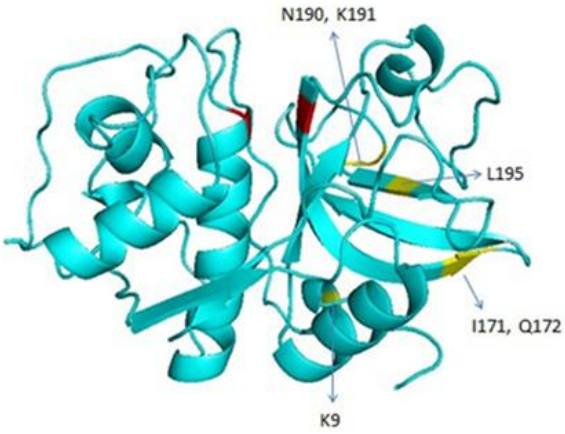


Figure 13. X-ray Structure of Cathepsin K Monomer.

Cathepsin K molecule is in cyan. Active sites Cys25 (left domain) and His162 (right domain) are in red. Amino acid residues that have mutated (K9, I171, Q172, N190, K191 and L195) are in yellow.

Taking pycnodysostosis causing mutations into consideration, 33 mutation sites on cathepsin K from 159 patients have been confirmed to date (54). Mutation sites are highlighted in Figure 14 A. By comparing mutation sites related to pycnodysostosis to mutation sites in M3, I found L195 is one of these mutation sites. Using pymol software, I found that the distances between the carbon atom in the side chain of L195 and the carbon atoms in the side chains of H177, I179 and V149 are 4.0, 4.1 and 4.2 Å, respectively (Figure 14 B). When L195 is mutated into alanine, the distances between the carbon atom in the side chain of alanine and the carbon atom in the side chains of V149, H177 and I179 are changed to 4.5, 4.9 and 5.4 Å, respectively (Figure 14 C). Thus, when L195 is substituted by alanine, the distances between L195 and other residues increase, which may disrupt van der Waals interactions. This may explain why no stable protein was obtained.

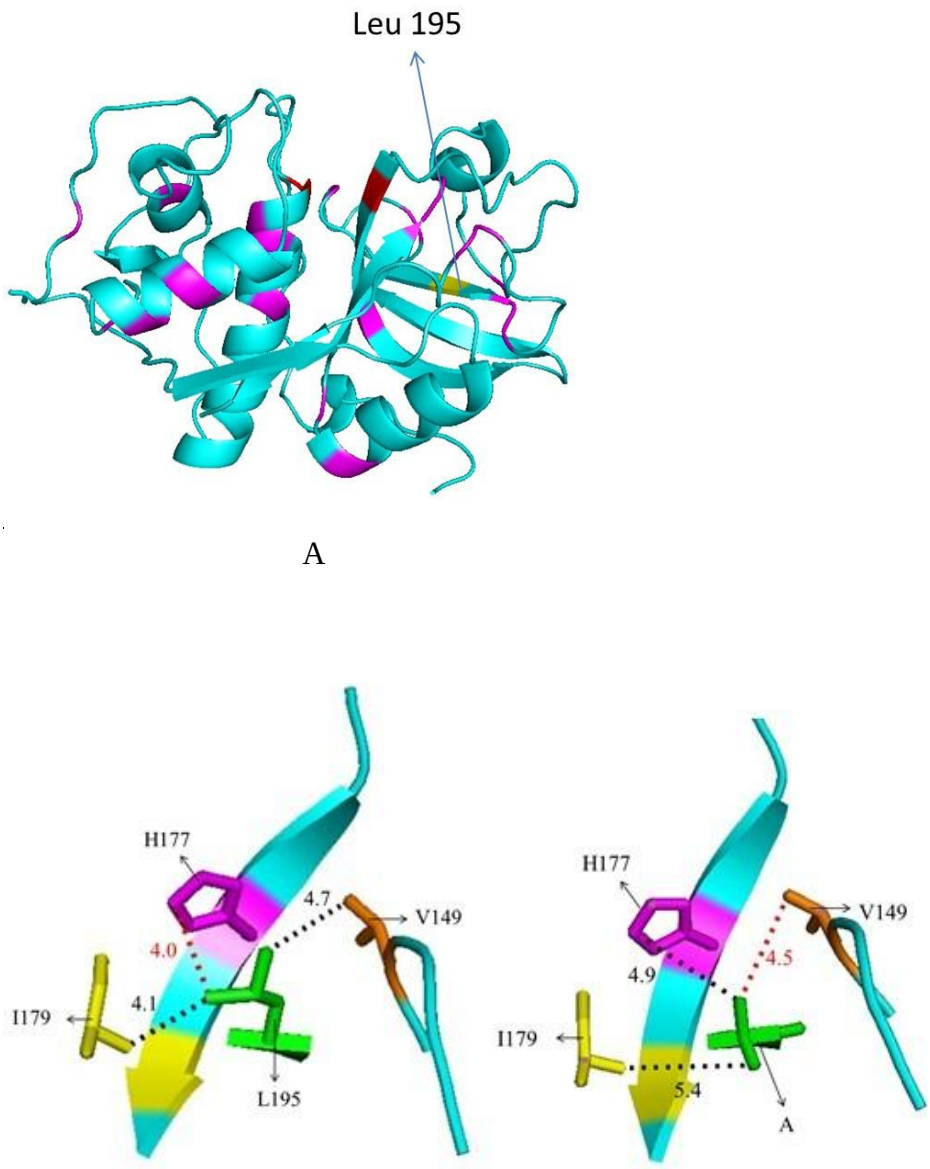
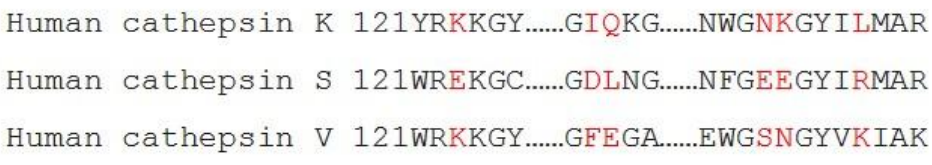


Figure 14. Crystal Structure of Cathepsin K Monomer.

A. Cathepsin K monomer is in cyan. Amino acid residues involved in pycnodysostosis are in magenta. Overlapping residue L195 involved in pycnodysostosis and GAG binding sites in tetramer model is in yellow. M3 replaced K9, I171, Q172, N190, K191 and L195 to alanine. B. Cathepsin K monomer is in cyan. L195 is in green. V149 is in orange. H177 is in magenta. I179 is in yellow. Distances between L195 and V149, H177, I179 are labeled, respectively. C. Cathepsin K monomer is in cyan. Alanine (substitute L195) is in green. V149 is in orange. H177 is in magenta. I179 is in yellow. Distances between alanine and V149, H177, I179 are labeled, respectively.

Furthermore, sequence alignment analysis was used to examine whether the mutated residues were highly conserved among cathepsins (Figure 15). The sequence analysis revealed that none of the mutations residues of M3 variant were conserved and thus sequence conservation is unlikely the reason for the observed protease instability.



A

B

Figure 15. Sequence Alignment of K9, I171, Q172, N190, K191, and L195

A. Protein sequence alignment of thee human cathepsins shows that the C4-S interacting residues in cathepain K are not conserved in related cathepsins S and V. B. Residues K9, Q172, N190, and L195 are conserved among the mammalian orthologues of cathepsin K. Residues I171, and K191 show some variations among the mammalian orthologues of cathepsin K. The change of I171 to M is mostly conservative (56). K191 is only altered into Y in African bush elephant. MEROPS http://merops.sanger.ac.uk/

In summary, considering that the collagenase activity of cathepsin K relies on protein-GAG interactions besides protein-protein interactions, disrupting these interactions could be a novel way to selectively inhibit its collagenase activity without influencing the enzyme integrity and overall hydrolytic activity. This could be developed as a potential therapy for future osteoporosis treatment.

Chapter 3: Characterization of Complex Formation Inhibitors (Exosite Inhibitors) of Cathepsin K

3.1 Introduction

For osteoporosis treatment, many active site inhibitors of cathepsin K have been synthesized and characterized. However, none of them has been approved by the FDA (22). Some of them were stopped in different phases of clinical trials due to off-target effects, while some of them are still going through clinical testing (Table 9).

Table 9: Active Site Inhibitors of Cathepsin K in Clinical Development

Name	Phase of Development	Expected Benefits and/or Limitations
Balicatib	Phase II (discontinued)	Benefits: antiresorptive effect without affecting bone formation Limitations: dermatologic adverse effects (57)
Relacatib	Phase I (discontinued)	Benefits: reduction of bone resorption Limitations: drug-drug interactions with paracetamol, ibuprofen and atorvastatin (49)
ONO-5334	Phase II	Benefits: less impact on the suppression of bone formation Limitations: adverse events, such as hypertension and dyspepsia (24)
Odanacatib	Phase III	Benefits: highly selective for cathepsin K, a long half-life, well tolerated, no effects on bone formation. It perturbs osteoclast-mediated bone resorption by preventing cathepsin K degradation of collagen (49)

Considering the side effects of active site inhibitors of cathepsin K, the discovery of complex formation inhibitors (synonymous to exosite inhibitors) could be beneficial.

Our laboratory screened a library of 1280 compounds by using a high throughput fluorescence polarization assay and identified several potential inhibitors disrupting complex formation without affecting the active site of cathepsin K. Among them, dihydrotanshinone and (-)-epigallocatechin gallate are compounds extracted from Chinese herbal plants including from green tea. My project is to extend previous studies of our laboratory to characterize exosite inhibitors from Traditional Chinese Medicine.

Traditional Chinese Medicine (TCM), an empirical system of multicomponent therapeutics, has been applied to treat various diseases for centuries (58). It includes various forms of herbal medicine, food therapy, acupuncture, massage and therapeutic exercise (59). For herbal medicine, naturally occurring herbs could be used individually or as a formula in disease treatment. The theories of TCM are to study the relationship between the physiologic activities and the pathologic changes according to the phenomena and laws of nature (60). Most TCM therapies are based on documentations recorded in ancient books and they have been used for thousands of years with the traditional belief that they are safe and effective for treating certain diseases (61). A wide range of traditional Chinese medicines are well known to have effects for skeletal diseases (62). The effects of these herbal plants and some of their effective constituents will be discussed next:

Compounds tanshinone IIA, tanshinone I, *ortho*-dihydrotanshinone (*o*-dihydrotanshinone) and cryptotanshinone are diterpenoids extracted from *Salvia miltiorrhiza.* Extracts of *Salvia miltiorrhiza, named* Danshen in Mandarin, have been shown to be effective in preventing bone resorption (63). Tanshinone IIA, tanshinone I, *o*-dihydrotanshinone and cryptotanshinone have been shown to have the inhibitory effects on osteoclast generation in vitro (64).

Berberine is an isoquinoline alkaloid and an active constituent of *berberis aquifolium* (65). *Berberis aquifolium* is a clinically important medicinal plant which has therapeutic potential for rheumatoid arthritis in the collagen-induced arthritis mouse model (66). Berberine also significantly decreased the number of osteoclasts when it was used in glucocorticoid-induced osteoporosis rats (67).

Icariin is a flavonol and a primary active constituent of *Epimedium sagittatum* which is another effective herbal component of Er-xian decoction, the formula for osteoporosis treatment widely used in China (68). Icariin treatment on ovariectomized rats suppressed osteoclastogenesis (68), which illustrated the potential therapeutic effects on osteoporosis. In mice with collagen-induced arthritis, icariin was able to reduce the serum levels of the bone resorption marker, C terminal telopeptide I of type I collagen (83).

Osthole, is an O-methylated coumarin, which is extracted from *Cnidium monnieri. Cnidium monnieri* has been shown to possess antiosteoporotic effects by inhibiting the formation and differentiation of osteoclasts (69). As the effective constituent of *Cnidium monnieri*, osthole, has also been shown to stimulate osteoblast differentiation and bone formation (70), osthole reveals a dual activity on bone formation and bone resorption.

Curculigoside is produced by *Curculigo orchinoides,* which is another component in Er-xian decoction (68). Curculigoside has been demonstrated to increase osteoblast proliferation and alkaline phosphatase activity (71).

Betulinic acid is a naturally occurring triterpenoid existing in a variety of plants and fruits (73). It is abundant in the plants of the genus *Sambucus*, including *Sambucus williamsii*, which is well known as a source of anti-inflammatory agents and as a remedy for

osteoporosis and bone fractures (72). Previous studies illustrated that betulinic acid is able to promote osteoblast differentiation *in vitro* to improve bone formation (72).

Myricetin, a flavonoid compound, is a naturally occurring compound extracted from *Myrica crifera,* which is used for bone fracture treatment, dental problems, stomach ache, and constipation in TCM (82). It is indicated that myricetin can promote bone formation by stimulating osteoblast differentiation *in vitro* (74).

Tetracycline is a widely used polyketide antibiotic, which is produced by the *Streptomyces* genus of Actinobacteria. It is of interest that tetracycline was shown to inhibit collagenase MMPs, which inspired people to investigate its effect on bone diseases (75). It has been demonstrated that tetracycline can improve bone formation, which was tested in osteopenic diabetic rats (76) and ovariectomized rats (77).

These herbs and compounds described above have either effects on bone resorption or bone formation or on both. Taking the less side effects of traditional Chinese medicine into consideration, it is worthwhile to evaluate these herbs for exosite inhibitors of cathepsin K.

3.2 Hypothesis

Selective exosite inhibitors can be found in compounds present in Chinese medicinal plants which specifically may inhibit cathepsin K's collagenase activity without affecting the other proteolytic functions of the enzyme.

3.3 Objective

Collagenase assay of cathepsin K will be used to identify and characterize potential exosite inhibitors. Compounds which inhibit the collagenolytic activity of cathepsin K will be further

examined for their effects on the gelatinolytic and Z-FR-MCA cleaving activity to exclude active site-targeting inhibitors.

3.4 Material and Methods

3.4.1 Material

The following is the list of inhibitors evaluated in this project. It includes inhibitors of known anti-resorptive activity: tanshinone I, cryptotanshinone, tanshinlactone, *p*-dihydrotanshinone, *o*-dihydrotanthinone, berberine, icariin; compounds affecting bone formation: curculigoside, betulinic acid, osthole, myricetin, tetracycline. HCl, tetracycline; as well as two active site-directed control inhibitors: antipain and odanacatib. Antipain, betulinic acid, and tetracycline free base were purchased from Sigma-Aldrich (St. Louis, MO). Tetracycline hydrochloride was kindly provided by Dr. Beatty (Professor, UBC, Vancouver). Tanshinone I, cryptotanshinone, tanshinlactone, and one of the *p*-dihydrotanshinones were purchased from ChemFaces (Wuhan, China). Two other dihydrotanshinones were purchased from MAYA company (China). Odanacatib, osthole, berberine, icariin and curculigoside were purchased from Selleckchem Company (United States). The active site cathepsin inhibitor E-64 (L-3-carboxy-trans-2-3-epoxypropionyl-leucylamido-(4guanidino)-butane) was purchased from Bio Basic Inc. (Markham, ON). The fluorogenic cathepsin substrate, Z-FR-MCA, was purchased from Enzyme System Products (Dublin, CA). Soluble tropocollagen, bovine type I collagen was purchased from USB (Cleveland, OH) and C4-S was purchased from Sigma-Aldrich (St. Louis, MO).

3.4.2 Cathepsin K Titration

Enzyme concentrations were determined by active site titration using E-64 and the fluorogenic substrate, Z-FR-MCA, as previously described in project 1. Briefly, 5 µL diluted enzyme was mixed with 5 µL of various concentrations of E-64 solutions and incubated for 30 min on ice in 90 µL of activity buffer. 5µL of the mixture was used to determine the residual activity towards Z-FR-MCA. The excitation and emission wavelengths used for the assays were set at 380 nm and 460 nm, respectively, using the spectro-fluorometer Perkin-Elmer LB50.

3.4.3 Collagenase Assay

400 nM wild-type human cathepsin K was incubated with 0.6 mg/mL soluble bovine type I collagen and 200 nM C4-S. Activity buffer (100 mM sodium acetate, 2.5 mM dithiothreitol and EDTA, pH 5.5) was used to bring the reaction system volume to 50 µL. Wild-type human cathepsin K was expressed in *Pichia pastoris* and purified as previously described in 2.4.3 and 2.4.4. Dilutions of the inhibitors were made in solvents as listed in Table 1. 1 µL of various concentrations of the inhibitors was added to the reaction mixture. Soluble bovine type I collagen and collagen mixed with the same concentrations of wild-type cathepsin K and C4-S control group were used as controls. Both control groups were added 1µL of compounds' solvents to eliminate potential solvent effects on collagen degradation. All reaction mixtures were incubated at 28°C. After 4 hours, 1 µL of 100 µM E-64 was added to each reaction to inhibit residual cathepsin K activity. Reactions were kept at 4°C overnight and then ran on 8% SDS-PAGE gels. After Coomassie Blue staining, SDS-PAGE gels were scanned using the GeneSnap program (Syngene Inc. Frederick, MD) and density of

collagen's α1 bands were measured and compared to determine the inhibitory effect of potential inhibitors of the collagenase activity of cathepsin K.

3.4.4 Determination of IC50 Values of Inhibitors Based on Collagen Assay of Cathepsin K

SDS-PAGE gels containing the separated reaction mixtures were quantified by determining the α1 collagen band density using the GeneSnap program (Syngene Inc. Frederick, MD). Percentages of inhibition at different inhibitor concentrations compared to collagen controls were calculated and plotted using Prism software (Graphpad Company). In the collagen control group, the density of α1 band was set to 100%. Changes of α1 band densities relative to the control on the Coomassie Blue stained SDS-PAGE gel were determined. The X axis was log (inhibitor concentrations) and the Y axis was percentage of inhibition, which used the collagen control as 100%. Results were shown as mean± SD and were derived from three independent experiments

3.4.5 Gelatinase Assay

Gelatin was prepared by heating soluble bovine type I collagen at 95°C for 20 minutes and the concentration of it used for assay was 0.6 mg/mL. Gelatin was incubated with 5 nM human cathepsin K in activity buffer (100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM dithiothreitol, and EDTA). Various inhibitors were added to make the final concentration of 50 μM. The total reaction volume of the assays was 50 μL. Assay duration was half an hour at 28°C. 1 μL of 100 μM E-64 was added to each reaction to inhibit the remaining cathepsin K activity. After being stored at 4°C overnight, the reactions were analyzed using 8% SDS-PAGE gels.

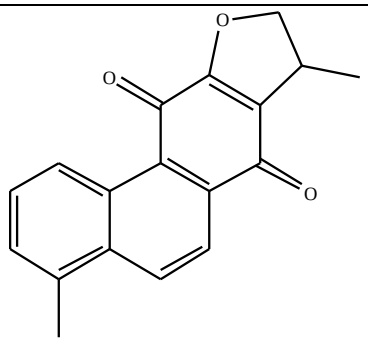
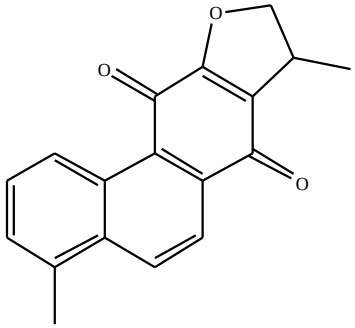
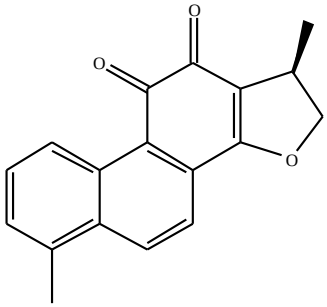
3.4.6 Z-FR-MCA Activity Measurement

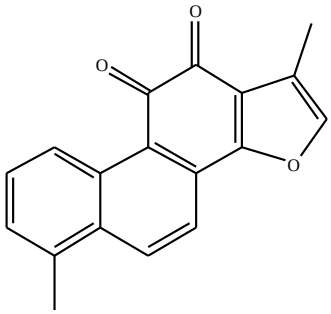
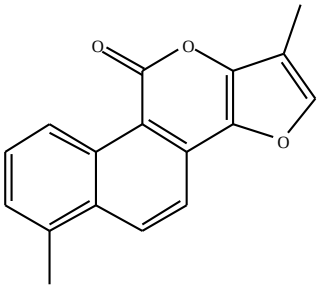
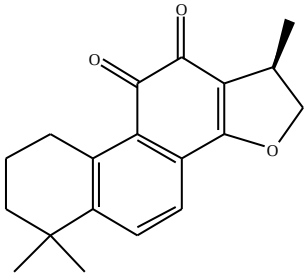
1.5 nM cathepsin K was used as the final concentration in assays. The control group was cathepsin K alone, incubated with the appropriate solvent used in the inhibitor experiments in a 1 mL activity buffer (100 mM sodium acetate buffer, pH 5.5, containing 2.5 mM dithiothreitol, and EDTA) for 30 seconds. 5 µL of 1 mM Z-FR-MCA was added into reaction system. The excitation and emission wavelengths used for the assays were set at 380 nm and 460 nm, respectively, using the spectro-fluorometer Perkin-Elmer LB50. Each reaction was monitored for 2 min. 12.5 µM of inhibitors were used as final drug concentration. The process was similar to that of control group as reactions were incubated for 30 seconds and 5 µL of 1 mM Z-FR-MCA was added and the reaction was followed for 2 min.

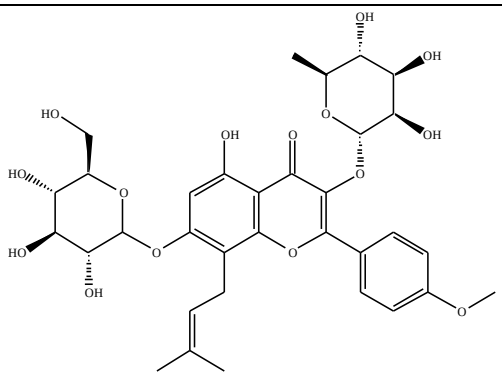
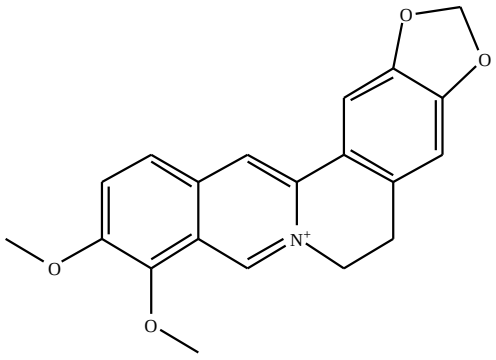
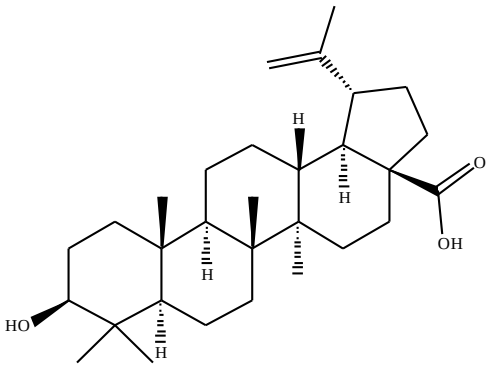
3.5 Results

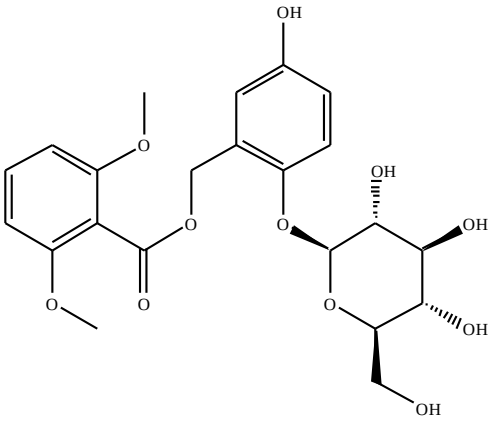
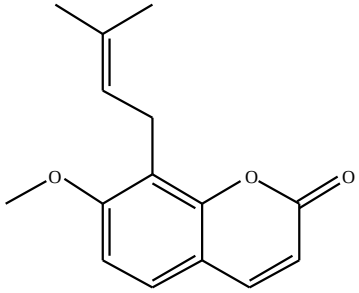
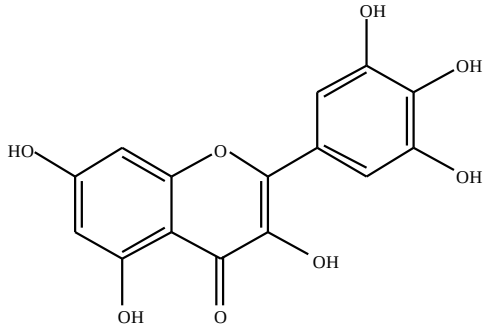
According to the functions, compounds tested in this project were divided into two groups: i) anti-resorptive compounds group and ii) compounds increasing bone formation (Table 10). In the anti-resorptive group, there are several compounds extracted from *Salvia miltiorrhiza* sharing similar structures. These compounds will be discussed together as tanshinone analogues. Other compounds in the anti-resorptive group do not have structural similarity. Active site inhibitors odanacatib and antipain were also evaluated for comparison.

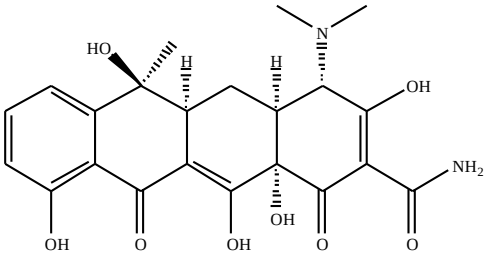
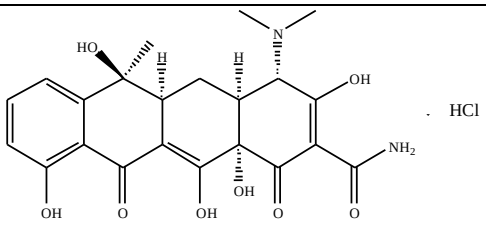
Table 10: List of Compound Groups Examined

Compound name	Structure	Chinese herb	Solvent and solubility
Anti-resorptive compounds Tanshinone analogues			
p-dihydrotanshinone (ChemFaces)		*Salvia miltiorrhiza*	DMSO **2.5 mM
p-dihydrotanshinone (MAYA)		*Salvia miltiorrhiza*	DMSO **2.5 mM
o-dihydrotanshinone		*Salvia miltiorrhiza*	Ethanol *3.6 mM

Compound name	Structure	Chinese herb	Solvent and solubility
Tanshinone I		*Salvia miltiorrhiza*	DMSO **5 mM
Tanshinlactone		*Salvia miltiorrhiza*	DMSO **10 mM
Cryptotanshinone		*Salvia miltiorrhiza*	DMSO *16.9 mM
Anti-resorptive compounds Other compounds			

Compound name	Structure	Chinese herb	Solvent and solubility
Icariin (Shanghai)		*Epimedium sagittatum*	DMSO *29.6 mM
Berberine		*Berberis aquifolium*	Water *3.0 mM
Compounds increasing bone formation			
Betulinic acid		*Betula pubescens*	DMSO *50 mM

Compound name	Structure	Chinese herb	Solvent and solubility
Curculigoside		*Curculigo orchioides*	DMSO **10 mM
Osthole		*Cnidium monnieri*	DMSO *200.6 mM
Myricetin		*Myrica cerifera*	DMSO **10 mM

Compound name	Structure	Chinese herb	Solvent and solubility
Tetracycline		NA	Ethanol **5 mM
Tetracycline. HCl		NA	Water *22.5 mM

Structures of the compounds were checked from PubChem.

** Information of solvents and solubility were taken from the product information sheets provided by the vendors.*

***Highest concentration used to prepare stock solutions.*

3.5.1 Determination of Potential Exosite Inhibitors with Collagenase Assay

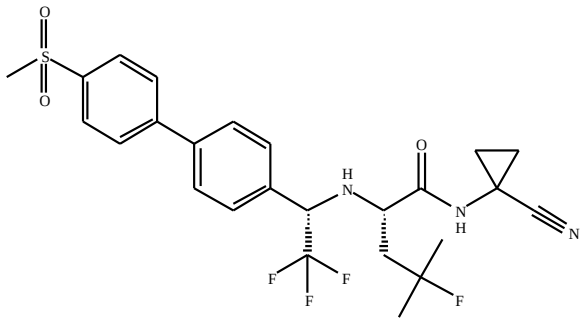
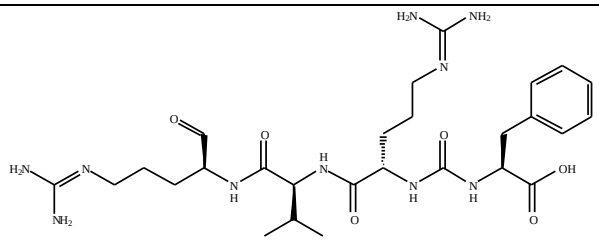
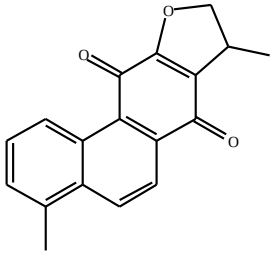
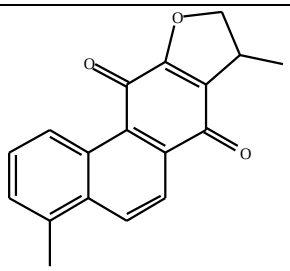
Compounds that could be considered as potential exosite inhibitors are expected to have two important characteristics. First, they must have an inhibitory effect on collagenase activity. Second, there should be no inhibition on the gelatinase activity and Z-FR-MCA hydrolysis activity. For compounds that have been chosen, the collagenase assay was performed first. SDS-PAGE gels of active site inhibitors are shown in Figure 16. Figure 17 shows SDS-PAGE gels of tanshinone analogues group. For other anti-resorptive inhibitors and

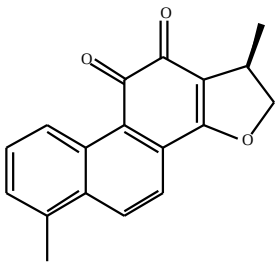
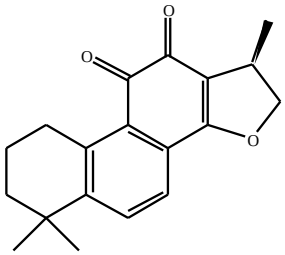
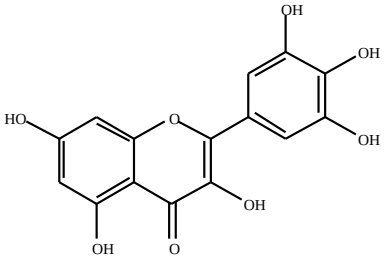
compounds increasing bone formation, the collagenase assay results are shown in Figures 18 and 19. Results are summarized for all effective inhibitors in Table 11.

10 out of the 17 compounds tested did not reveal any inhibition of the collagenolytic activity of cathepsin K at the maximal concentration tested (100 μM). These compounds included tanshinone I, tanshinlactone, betulinic acid, berberine, icariin, curculigoside, osthole, tetracycline, and tetracycline.HCl.

Among the effective compounds, *o*-dihydrotanshinone is capable of suppressing collagen degradation at a relatively low concentration with an IC_{50} of 10.1 ± 2.1 μM (Figure 17 and Table 11). *p*-dihydrotanshinone was tested for samples obtained from two vendors: The one from Faces has an IC_{50} of 11.1 ± 2.8 μM, while the one from Maya is 15.5 ± 2.0 μM. Cryptotanshinone, onother compound belonging to the tanshinone analogues group, has an IC_{50} of 22.8 ± 2.2 μM. Another effective compound is myricetin, whose IC_{50} is 37.3 ± 2.0 μM (Figure 19 and Table 11). Odanacatib, an active site inhibitor for cathepsin K developed by Merck Company, has an IC_{50} value of 0.4 ± 0.1 μM. As a known active site inhibitor, antipain revealed an IC_{50} value of 35.6 ± 2.3 μM (Figure 16 and Table 11).

Table 11: Effective Collagenase Exosite Inhibitors.

Compound (effective)	Structure	IC_{50} (μM)
Odanacatib		0.39±0.1
Antipain		35.6±2.3
p-Dihydrotanshinone (Chemfaces)		11.1±2.8
p-Dihydrotanshinone (Maya)		15.5 ±2.0

Compound (effective)	Structure	IC_{50} (μM)
o-Dihydrotanshinone		10.1±2.1
Cryptotanshinone		22.8±2.2
Myricetin		37.3±2.0

Structures of the compounds were taken from PubChem. IC_{50} values were determined by plotting the percentage of inhibition versus Log (concentration) of inhibitor. 400 nM human recombinant cathepsin K, 200 nM C4-S were added to the activity buffer in the presence of increasing concentartions of inhibitors. Three replicates were carried out for each inhibitor, and the IC_{50} values were determined using GraphPad Prism software.

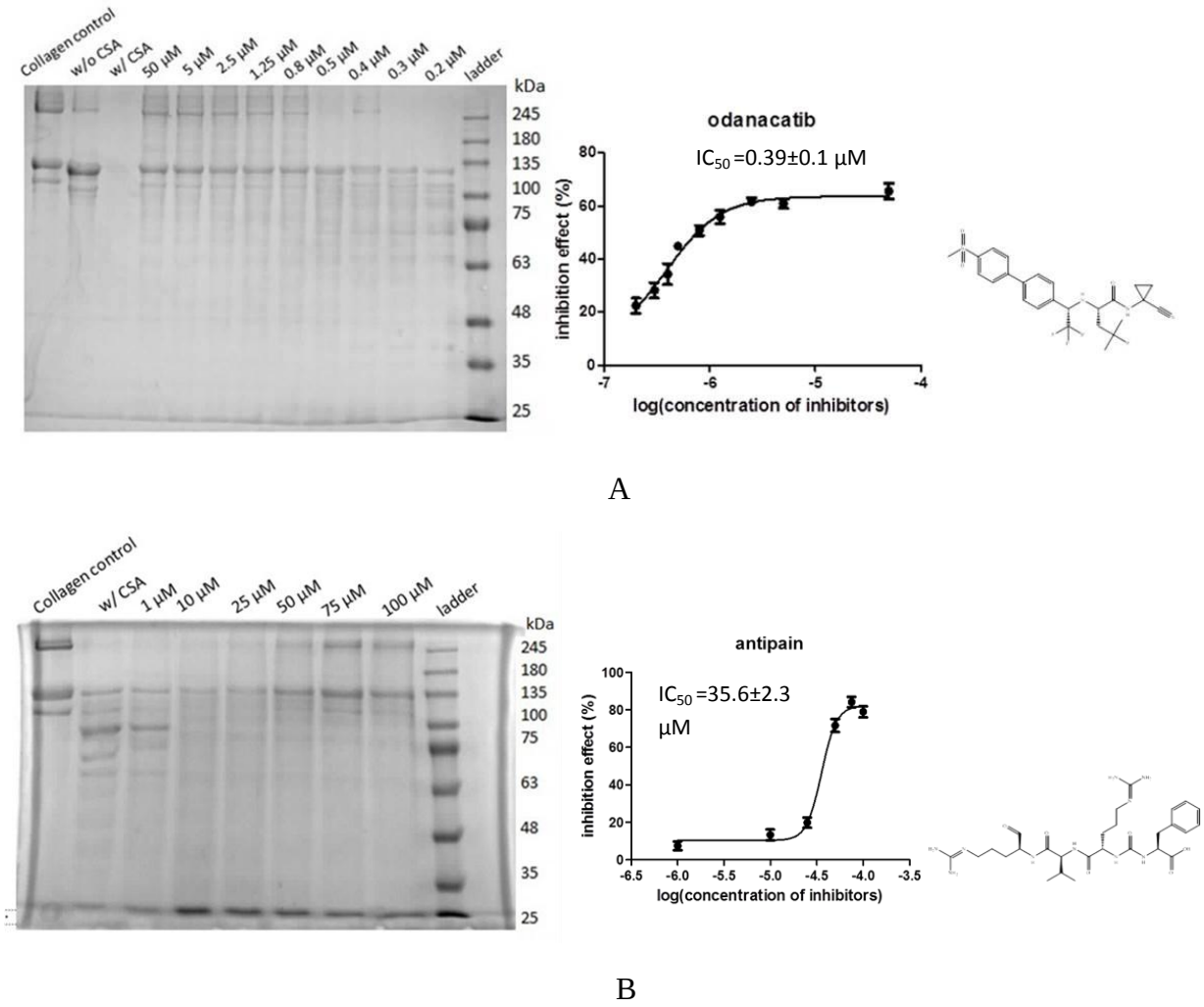
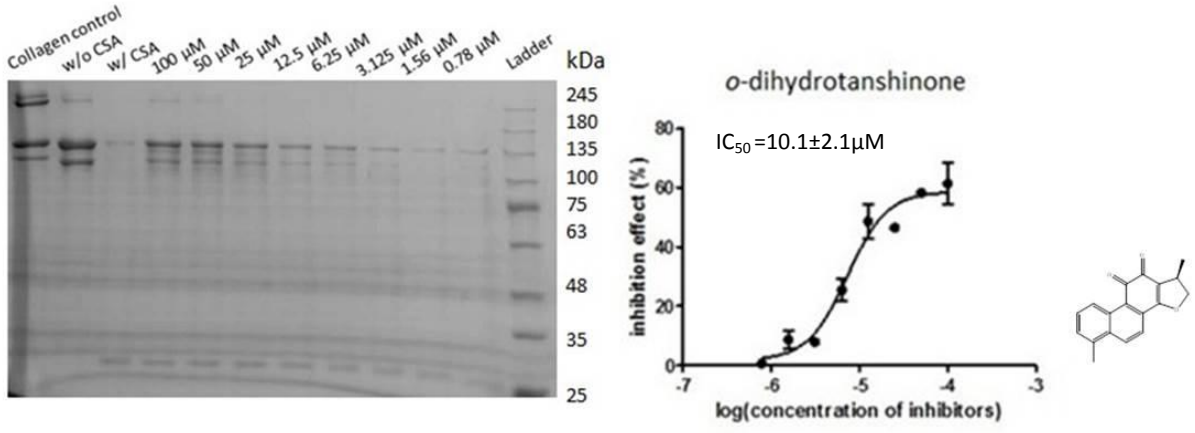
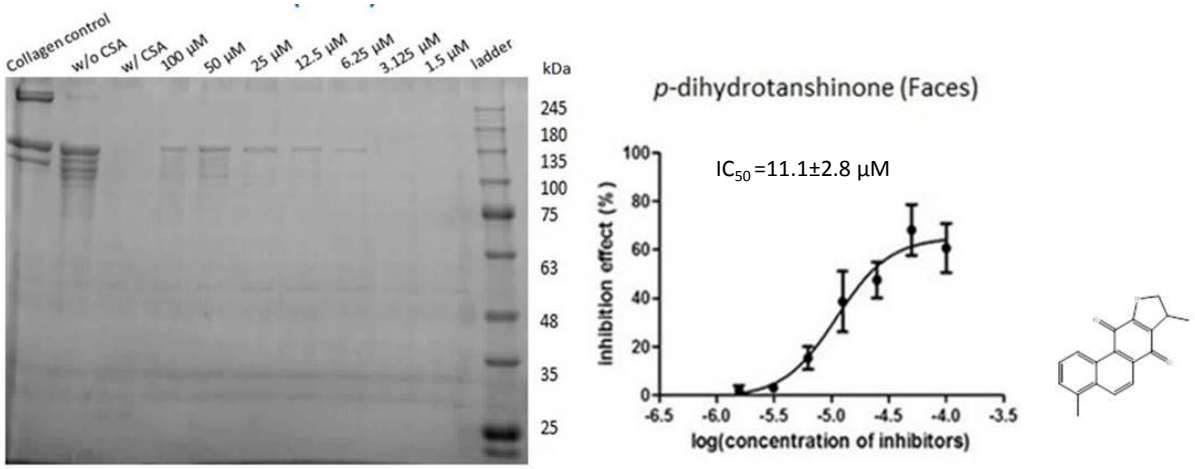


Figure 16. Cathepsin K Collagenase Assays in the Presence of Active Site Inhibitors.

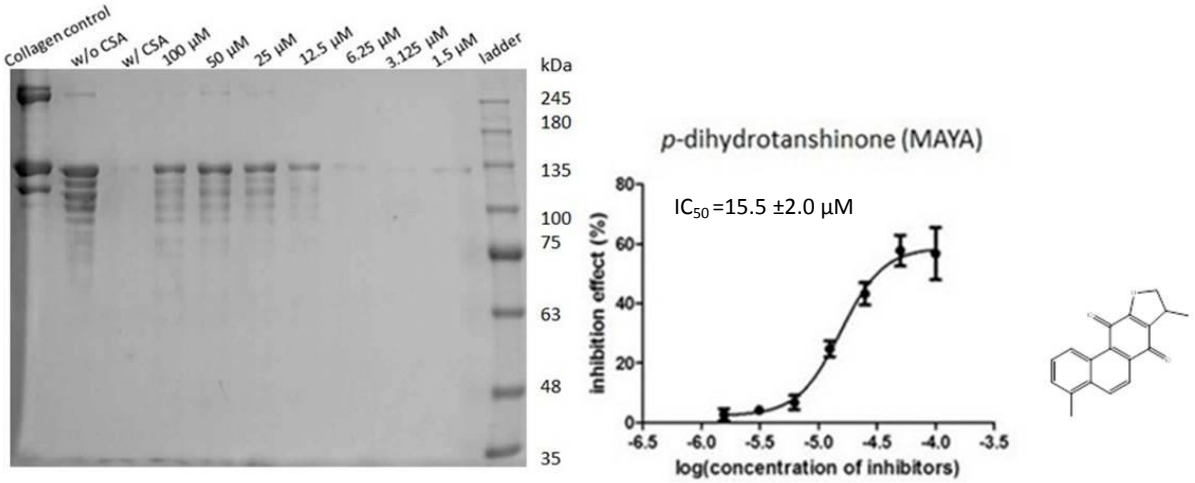
A. SDS-PAGE image, IC_{50} plots for type I collagen degradation and chemical structure of odanacatib. B. SDS-PAGE image, IC_{50} plots for type I collagen degradation and chemical structure of antipain. Inhibitors were diluted in their solvents. Gradients of inhibitors were used to study the effect of inhibition on collagenase degradation by cathepsin K. Assays were carried out with 400 nM cathepsin K and 200 nM C4-S. Solvent concentrations were constant in all assay conditions. The degradation of collagen was quantified as α-1 band density on the coomassie blue stained gel. Percentage of inhibition was obtained by dividing the α-1 band in the presence of inhibitor by the α-1 band in the absence of inhibitor. The percentage of inhibition was plotted against its corresponding inhibitor concentration to obtain the IC_{50} values. All assays were done in three independent experiments.


Collagen control
w/o CSA
w/ CSA
100 µM
50 µM
25 µM
12.5 µM
6.25 µM
3.125 µM
1.56 µM
0.78 µM
Ladder
kDa
245
180
135
100
75
63
48
35
25
o-dihydrotanshinone
IC50 =10.1±2.1µM
inhibition effect (%)
log(concentration of inhibitors)

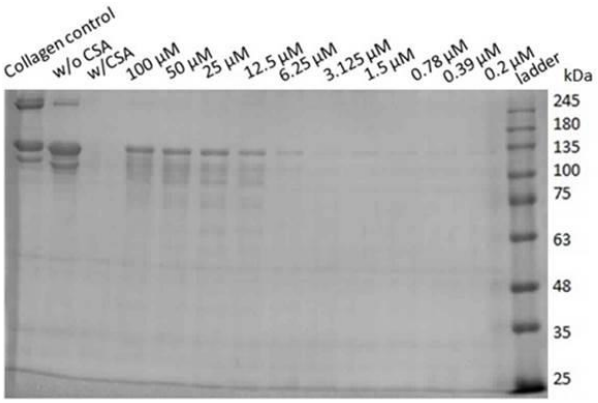
A

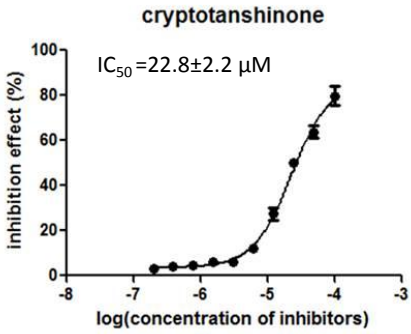

Collagen control
w/o CSA
w/ CSA
100 µM
50 µM
25 µM
12.5 µM
6.25 µM
3.125 µM
1.5 µM
ladder
kDa
245
180
135
100
75
63
48
35
25
p-dihydrotanshinone (Faces)
IC50 =11.1±2.8 µM
inhibition effect (%)
log(concentration of inhibitors)

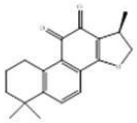
B


Collagen control
w/o CSA
w/ CSA
100 µM
50 µM
25 µM
12.5 µM
6.25 µM
3.125 µM
1.5 µM
ladder
kDa
245
180
135
100
75
63
48
35
p-dihydrotanshinone (MAYA)
IC50 =15.5 ±2.0 µM
inhibition effect (%)
log(concentration of inhibitors)

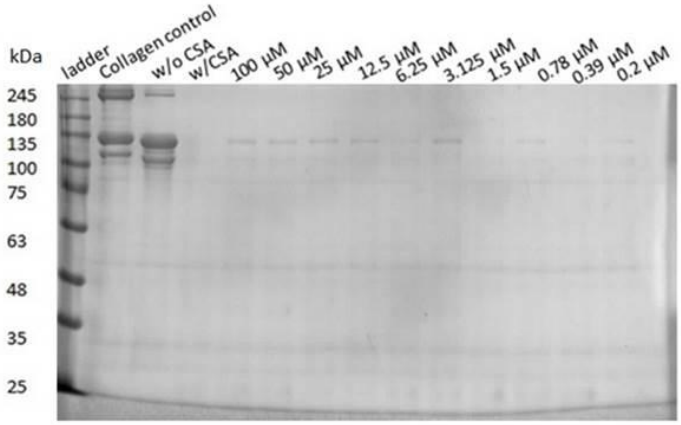
C





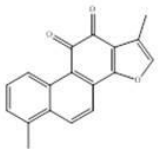


D

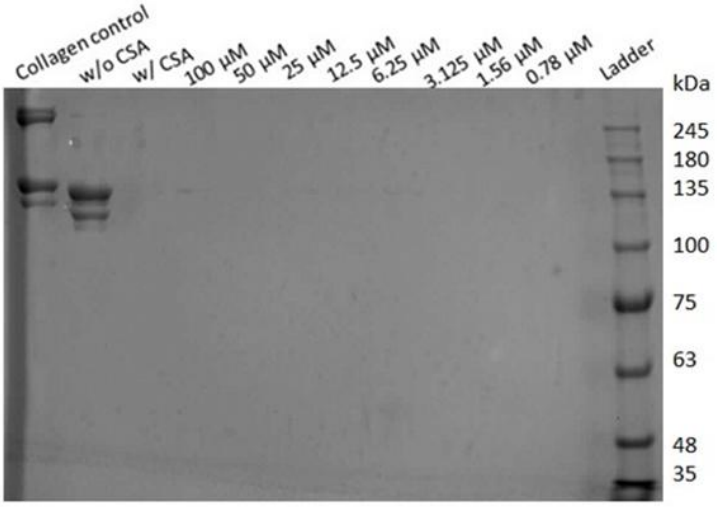


IC_{50} > 100 μM

Tanshinone I

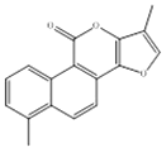


E



IC_{50} > 100 μM

Tanshinlactone



F

Figure 17. Cathepsin K Collagenase Assays in the Presence of Tanshinone Analogues Group Inhibitors.

A. SDS-PAGE image, IC_{50} plot for type I collagen degradation and chemical structure of *o*-dihydrotanshinone. B. SDS-PAGE image, IC_{50} plot for type I collagen degradation and chemical structure of *p*-dihydrotanshinone (Faces). C. SDS-PAGE image, IC_{50} plot for type I collagen degradation and chemical structure of *p*-dihydrotanshinone (MAYA). D. SDS-PAGE image, IC_{50} plot for type I collagen degradation and chemical structure of cryptotanshinone. E. SDS-PAGE image for type I collagen degradation and chemical structure of tanshinone I. F. SDS-PAGE image for type I collagen degradation and chemical structure of tanshinlactone. The details of this assay are described in Figure 16 and in the Methods section. All assays were done in three independent experiments.

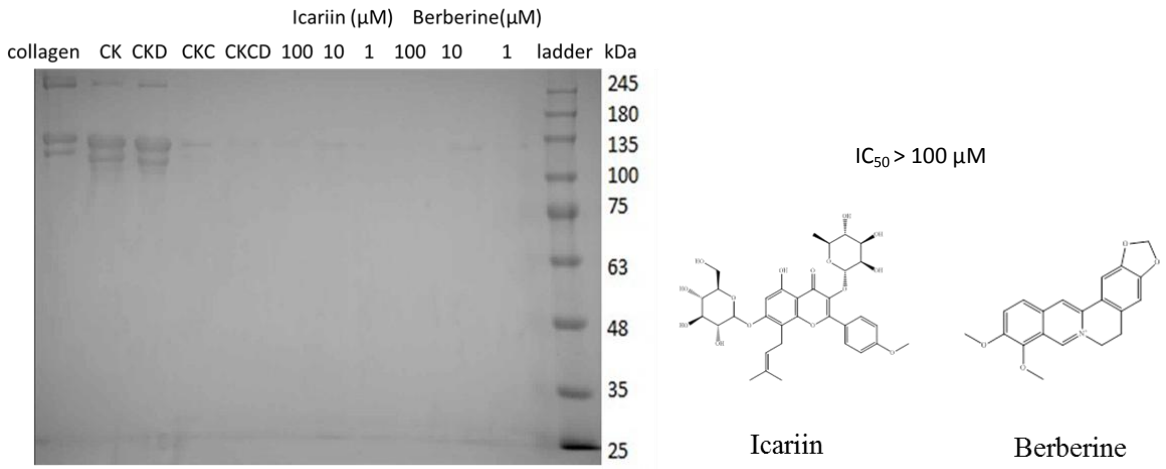
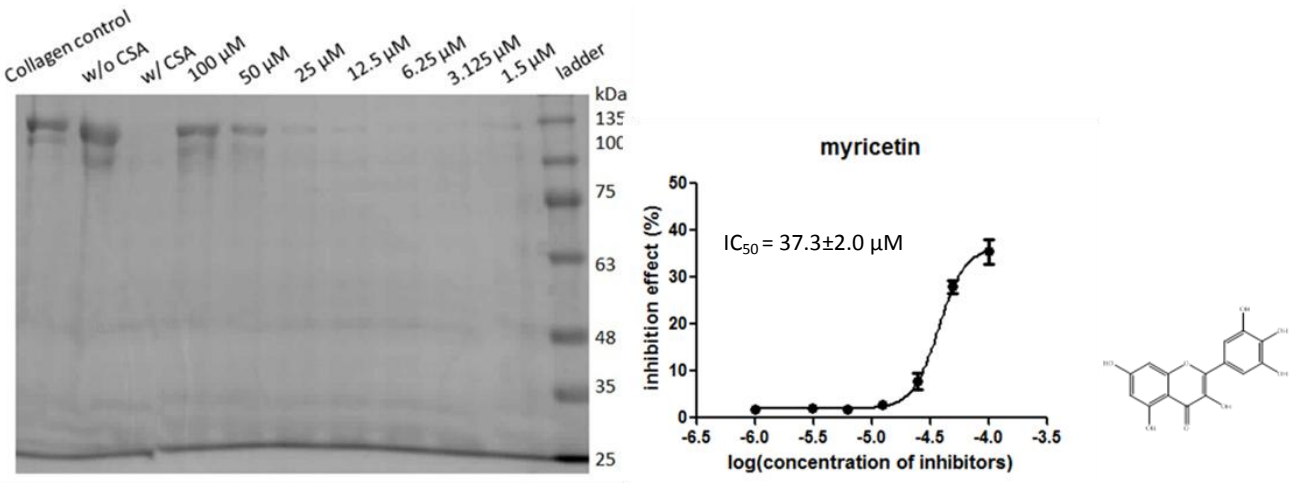
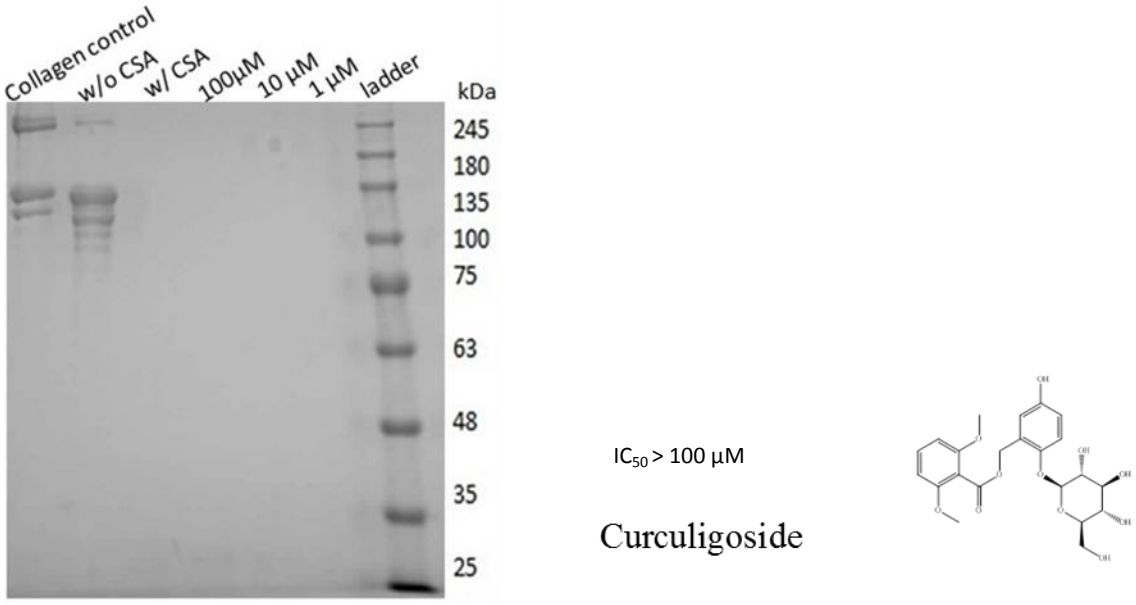


Figure 18. Cathepsin K Collagenase Assays in the Presence of Other Anti-resorptive Inhibitors

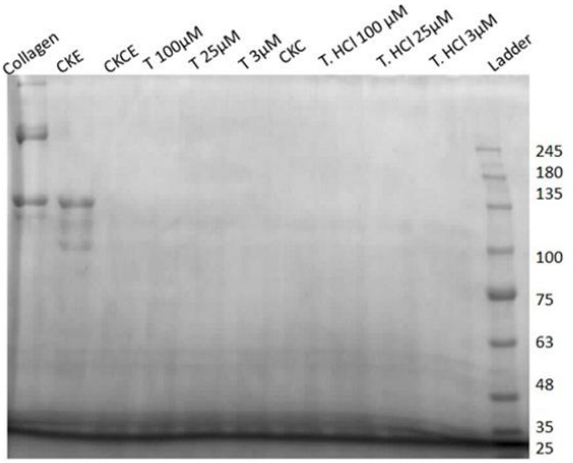
SDS-PAGE images for type I collagen degradation and chemical structures of icariin and berberine. The details of this assay are described in Figure 16 and the Methods section. CK refers to reaction of collagen and cathepsin K. CKD refers to reaction of collagen and cathepsin K in the presence of DMSO. CKC refers to reaction of collagen, cathepsin K, and CSA. CKCD refers to reaction of collagen, cathepsin K, CSA, and DMSO. All assays were done in three independent experiments.

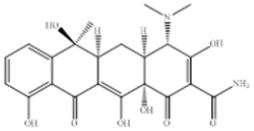

Collagen control
w/o CSA
w/ CSA
100 µM
50 µM
25 µM
12.5 µM
6.25 µM
3.125 µM
1.5 µM
ladder
kDa
135
100
75
63
48
35
25
myricetin
inhibition effect (%)
IC50 = 37.3±2.0 µM
log(concentration of inhibitors)

A

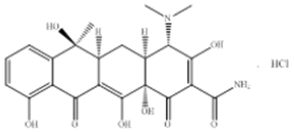

Collagen control
w/o CSA
w/ CSA
100µM
10 µM
1 µM
ladder
kDa
245
180
135
100
75
63
48
35
25
IC50 > 100 µM
Curculigoside

B





Tetracycline

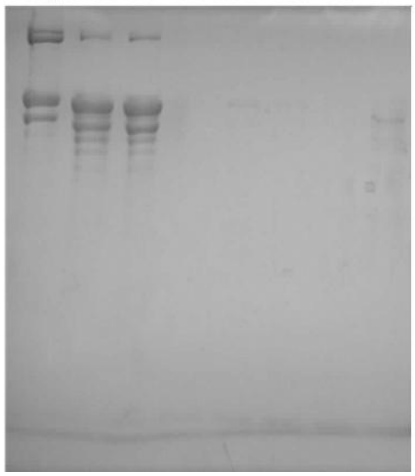


Tetracycline. HCl

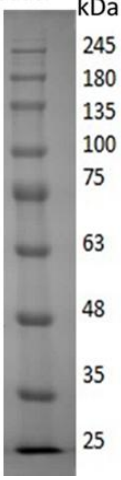
C

Betulinic acid(μM)

Collagen CK CKD CKC CKCD 100 10 1

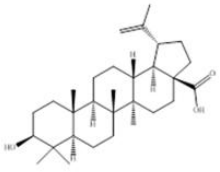


ladder

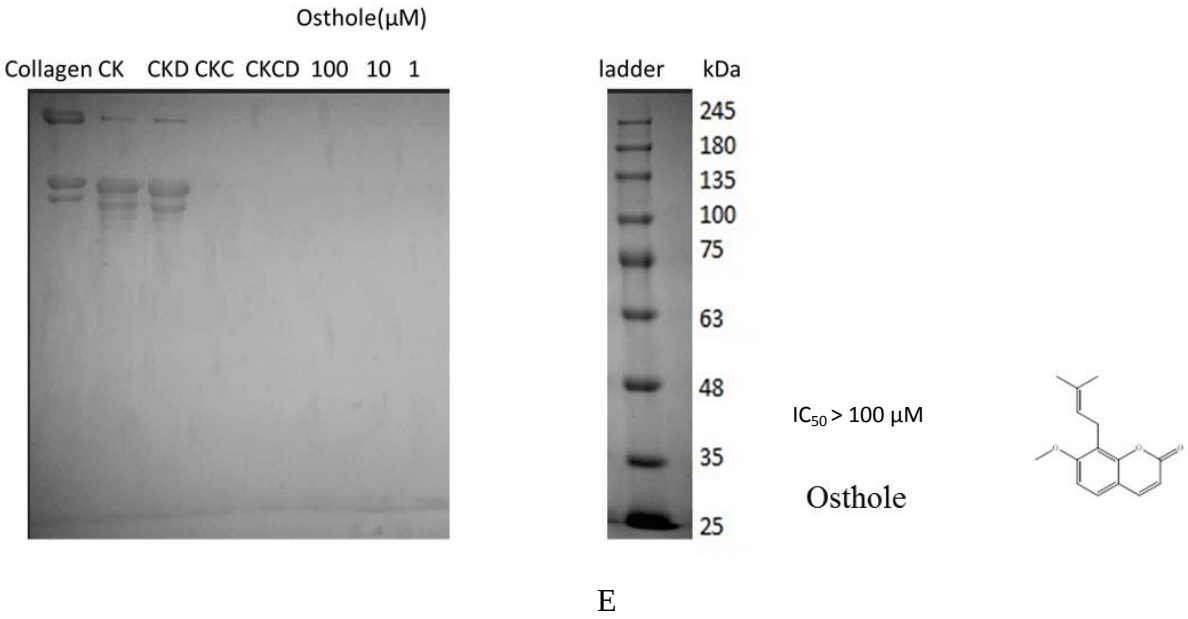


IC_{50} > 100 μM

Betulinic acid



D



E

Figure 19. Cathepsin K Collagenase Assays in the Presence of Compounds Increasing Bone Formation

A. SDS-PAGE image, IC_{50} plots for type I collagen degradation and chemical structure of myricetin. B. SDS-PAGE image for type I collagen degradation and chemical structure of curculigoside. C. SDS-PAGE images for type I collagen degradation and chemical structures of tetracycline and tetracycline. HCl. D. SDS-PAGE image for type I collagen degradation and chemical structure of betulinic acid. E. A. SDS-PAGE image for type I collagen degradation and chemical structure of osthole. The details of this assay are described in Figure 16. CK refers to reaction of collagen and cathepsin K. CKC refers to reaction of collagen, cathepsin K and CSA. CKE refers to reaction of collagen, cathepsin K and ethanol. CKCE refers to reaction of collagen, cathepsin K, CSA, in the presence of ethanol. CKD refers to reaction of collagen and cathepsin K in the presence of DMSO. CKC refers to reaction of collagen, cathepsin K, and CSA. CKCD refers to reaction of collagen, cathepsin K, CSA, and DMSO. All assays were done in three independent experiments.

3.5.2 Determination of the Effect of Inhibitors on Gelatinase Activity and Z-FR-MCA Activity

What makes exosite inhibitors different from active site inhibitors is that exosite inhibitors are expected to inhibit cathepsin K's collagenase activity without interfering with its other proteolytic activities. To further confirm whether potential compounds screened from collagenase assay will affect cathepsin K's other enzyme activites, the gelatinase and Z-FR-

MCA activity assays were performed. Odanacatib and antipain, two active site inhibitors, were used as controls for the gelatinase assay.

50 µM of all compounds were used in the gelatinase assay. As expected, odanacatib and antipain effectively blocked the gelatin degradation, as the inhibition percentages were 93.5 ±4.3% and 90.7 ± 3.1, respectively. The exosite inhibitor candidates did not reveal any inhibitory effects on gelatin degradation (Figure 20).

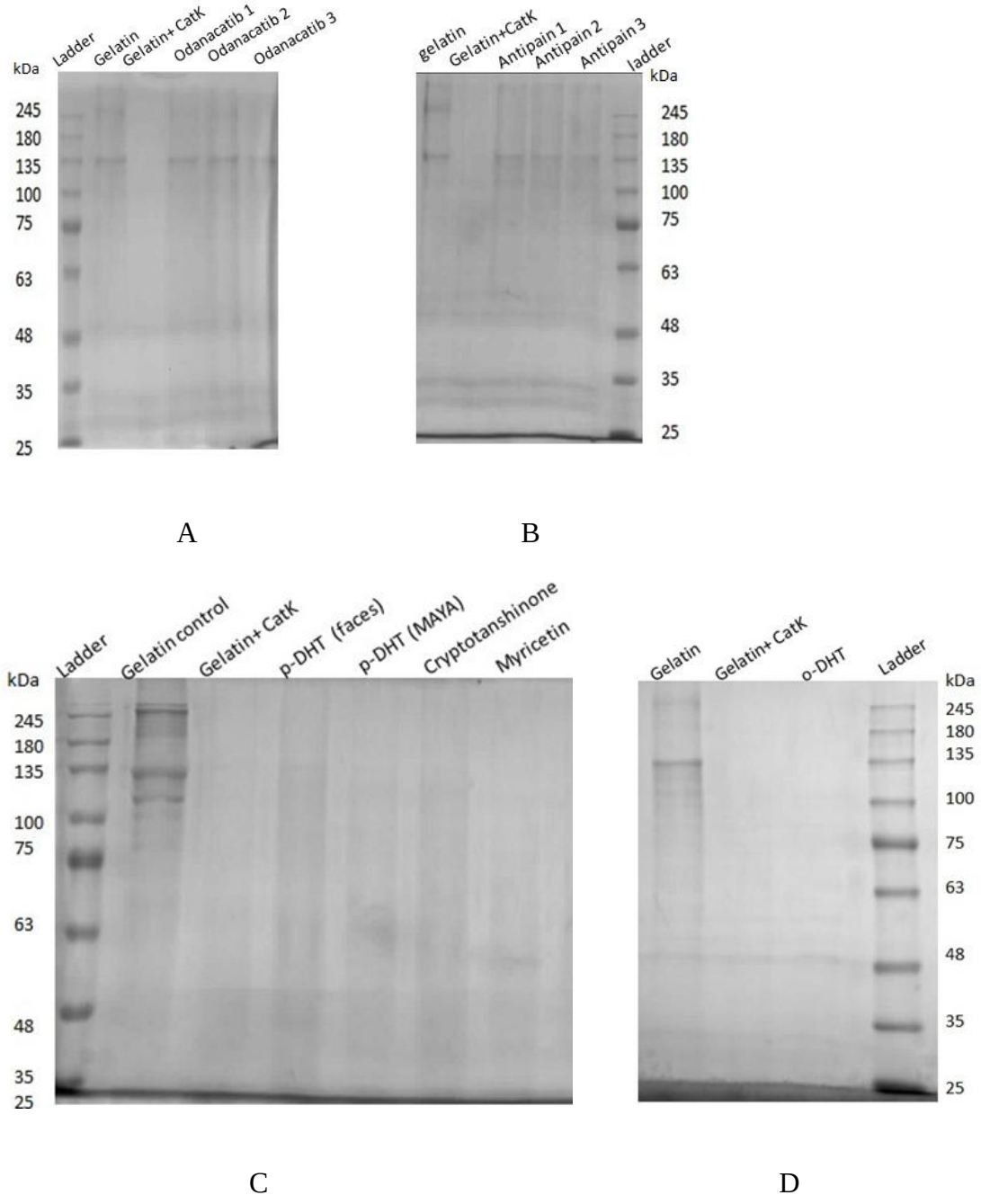


Figure 20. Gelatinase Assay of Cathepsin K with Various Exosite Inhibitors.

A. Gelatinase assay of odanacatib. B. Gelatinase assay of antipain. C. Gelatinase assay of p-dihydrotanshinone (two companies), cryptotanshinone and myricetin. D. Gelatinase assay of o-dihydrotanshinone. Assay was performed at 5 nM cathepsin K with 50 μM final concentration of inhibitors for 30 min at 28°C. Gelatin alone and gelatin with cathepsin K with corresponding solvent were used as control. Samples were run on an 8% SDS-PAGE gel. The density of α1 band was quantified to obtain the inhibition percentage. All assays were done in three independent experiments

In the Z-FR-MCA activity assay, odanacatib and antipain revealed an inhibition percentage of 98.0 ± 1% and 99 ± 0.6%, respectively. Myricetin and cryptotanshinone also revealed inhibitory effects, which were 65 ± 0% and 62 ±4 %, respectively. *o*-dihydrotanshinone, *p*-dihydrotanshinone (ChemFaces) and *p*-dihydrotanshinone (Maya) showed similar but much lower rates of inhibition: 28 ± 2%, 34 ± 2%, and 30 ± 2 % inhibition of Z-FR-MCA activity, rsspectively (Table 12).

Table 12: Effects of Exosite Inhibitors on the Active Site of Cathepsin K.

Compound	$[Inhibitor]_f$ (μM)	% Inhibition on substrates		[Inhibitor]/[E] Ratio	
		Z-FR-MCA	Gelatin	Z-FR-MCA	Gelatin
Odanacatib	12.5	98±1	94±4	8333	10000
Antipain	12.5	99±6	91 ± 3	8333	10000
Myricetin	12.5	65±0	0	8333	10000
p-Dihydrotanshinone (Chemfaces)	12.5	34±2	0	8333	10000
p-Dihydrotanshinone (Maya)	12.5	30±2	0	8333	10000
o-Dihydrotanshinone	12.5	28±2	0	8333	10000
Cryptotanshinone	12.5	62±4	0	8333	10000

Inhibition was determined using the Z-FR-MCA cleavage and gelatinase assays. For the gelatinase assay, the final inhibitor concentration was maintained at 50 μM, whereas the final concentration of cathepsin K was 5 nM. This translates into an enzyme: inhibitor ratio of 1: 10,000. The reaction time for the gelatinase assay was 30 min. For Z-FR-MCA activity assay, final inhibitor concentration was 12.5 μM and enzyme concentration was 1.5 nM. This translates into an enzyme: inhibitor ratio of 1: 8,333. The release of the fluorogenic MCA group was monitored for 2 min using the spectro-fluorometer Perkin-Elmer LB50. Z-FR-MCA activity assay was done in two independent experiments.

3.6 Discussion

Current active site inhibitors of cathepsin K are prone for side effects. The aim of this study was to develop exosite inhibitors for this protease. My project was focused on determining the effects of compounds used in bone disease treatment in TCM on cathepsin K activities.

3.6.1 Active Site Inhibitors

For active site inhibitors, not only collagen degradation, but also active site substrates cleavage could be suppressed. Gelatin and synthetic substrate, Z-FR-MCA, were used as active site substrates in this project. As well-known active site inhibitors, odanacatib and antipain (22) inhibited almost 100% of the gelatinolytic activity and the hydrolysis of Z-FR-MCA. Odanacatib, which is in clinical phase III trial for osteoporosis, is the most promising drug candidate, revealing an IC_{50} of 400 nM. It should be mentioned that the majority of inhibitors studied around the world have IC_{50} values in the range of pM when used in active-site directed assays. The IC_{50} of odanacatib for the collagenase assay, I obtained is much higher. Also antipain revealed an IC_{50} value of 35.6 ± 2.3 µM whereas its Ki-value in the Z-FR-MCA assay is 40 nM for cathepsin K (unpublished data, Brömme lab). To explain this, it should be noted that the IC_{50} values from collagenase assay do not present true binding constants and depend on the concentration of protease. In the assay used in this study, 400 nM cathepsin K is needed to observe the degradation of the f α1 bands on SDS-PAGE gel. When 400 nM enzyme is applied, the theoretical IC_{50} valuae that can be obtained is 200 nM. This illustrates that IC_{50} values obtained from the collagenase assay are substantially higher than the real binding constants for these inhibitors.

3.6.2 Tanshinone Analogues in Anti-resorptive Group

Tanshinone analogues group consists of various compounds purified from *Salvia miltiorrhiza*, the Chinese herbal medical plant, which has been widely used for skeletal, cardiovascular and cerebrovascular diseases in China (78).

Preliminary screening efforts identified dihydrotanshinone as potent anti-collagenase inhibitor of cathepsin K (unpublished data, Brömme lab). Dihydrotanshinone has two

conformations, which are para and ortho. For *o*-dihydrotanshinone, its IC_{50} is 10.1±2.1 μM. Although this value is 20 times higher than that of odanacatib, its inhibitory effect on collagenolytic degradation is still considerable. There was no inhibitory effect on the gelatinolytic activity of carthepsin K when the ratio of *o*-dihydrotanshinone and cathepsin K was 10,000:1, which illustrates that *o*-dihydrotanshinone does not interact with the active site. The inhibition of Z-FR-MCA hydrolysis was 28 ± 2%, at a ratio of 8333:1. Even though more inhibitors were applied to gelatinase assay compared to Z-FR-MCA activity assay, no inhibitory effect could be detected in gelatinase assay. It should be noted that the reaction time for gelatinase assay is 30 min, much longer than that for Z-FR-MCA assay, which is 2 min. It is possible that *o*-dihydrotanshinone has some low affinity reversible binding to the active site. When the reaction time is long enough, there is more chance for cathepsin K to depart from active site-bound *o*-dihydrotanshinone to allow the complete hydrolysis of gelatin (end-point determination).

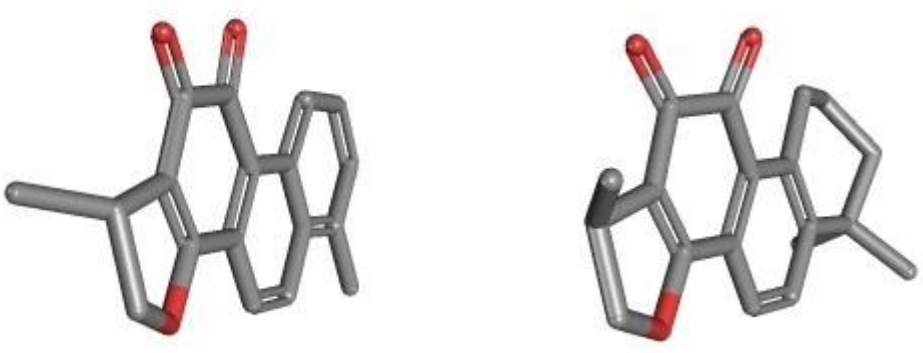
Two *p*-dihydrotanshinone produced from different companies displayed similar IC_{50} values to *o*-dihydrotanshinone, which render them as potential exosite inhibitors. Inhibition effects on gelatin and Z-FR-MCA of *p*-dihydrotanshinones were also comparable to the *o*-dihydrotanshinone derivative. This means that the structural difference between *p*-dihydrotanshinone and *o*-dihydrotanshinone is not critical for defining effective and ineffective inhibitors.

Another effective inhibitor derived from tanshinone analogues is cryptotanshinone, which has an IC_{50} of 22.8 ± 2.2 μM. Gelatinolytic degradation could not be suppressed by cryptotanshinone. Nevertheless, it had a 62 ± 4% inhibitory effect on Z-FR-MCA activity assay.

Tanshinone I and tanshinlactone were demonstrated to lack an anti-collagenolytic activity upto a concentration of 100 μM. Neither of them could inhibit cathepsin K degrading gelatin.

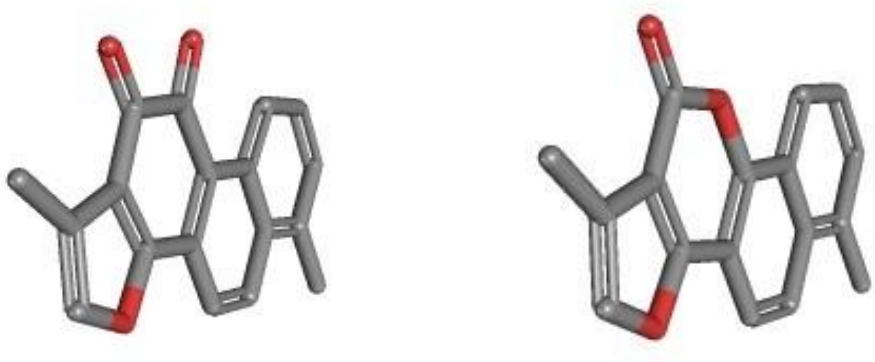
Even though *p*-dihydrotanshinone, *o*-dihydrotanshinone, and cryptotanshinone have considerable inhibitory effects on the collagenase activity of cathepsin K, it should be noticed these compounds have been identified by others to suppress osteoclastogenesis (64) (described in the introduction part). Osteoclastogenesis is the process of generating osteoclasts, which includes commitment, differentiation, multinucleation, and activation of precursor cells (79). Although it is effective to reduce bone resorption by inhibiting osteoclastogenesis, the communication of osteoclasts and osteoblasts could be affected, which may lead to the suppression of bone formation (81). Considering the double activities of dihydrotanshinones and cryptotanshinone, they may not represent perfect exosite inhibitors.

Taking the structures and functions of these tanshinone analogues into consideration, 2D structures are similar among them. What makes the inhibitory effect of cathepsin K on collagenase activity so different is an interesting question. I compared their 3D structures to better understand their overall structure and functioning relationships using pubchem website (Figure 21).



(a) (b)

A.



(c) (d)

B.

Figure 21. 3D Structures Analysis of Tanshinone Analogues.

Pubchem was used for 3D structures analysis. Grey is C and red is O. A.Tanshinone analogues that have inhibitory effects on collagenase activity of cathepsin K. (a) 3D structure of *o*-dihydrotanshinone (b) 3D structure of cryptotanshinone. B. Tanshinone analogues that are non-effective inhibitors of cathepsin K. (c) 3D structure of tanshinone I (d) 3D structure of tanshinlactone.

3D analysis illustrated that for effective inhibitors, *p*-dihydrotanshinone, *o*-dihydrotanshinone and cryptotanshinone have a major characteristic in common. They all have a five-membered ring, oxolane, which only has one double bond. In contrast, the ineffective inhibitors

tanshinone I and tanshinlactone, have the aromatic furan as five-membered ring, which has two double bonds. In the aromatic ring system the methyl substitute is planar to the ring whereas in the oxolane rings the methyl group bends to the side. It is likely that the planar orientation of the methyl ring hinders the binding of these tanshinone derivatives.

3.6.3 Other Anti-Resorptive Compounds

In addition to tanshinones, other anti-resorptive compounds analyzed in this project were icariin and berberine.

Icariin and berberine are the components of Er-Xian decoction, the formula for osteoporosis treatment used in TCM (68). Both of them were reported having inhibitory effects on osteoclastic bone resorption (68). However, none of the above compounds have inhibitory effects on the collagenolytic activity of cathepsin K. As previous studies showed (65-68), these two compounds are effective for bone resorption by repressing osteoclastogenesis, which is not relevant to the mechanism of collagen degradation by cathepsin K.

3.6.4 Compounds Increasing Bone Formation

Betulinic acid, curculigoside, tetracyclines, myricetin, and osthole have been shown to improve bone formation (70, 71, 72, 74, 75). I tested them with the collagenase assay to determine their effects on the collagenase activity of cathepsin K. Only myricetin showed inhibitory effect on collagenase activity of cathepsin K with an IC_{50} of 37.3 ± 2.0 μM. In contrast, myricetin did not inhibit the gelatinolytic activity, indicating the inhibition is specific for triple collagen degradation. However, myricetin revealed partial inhibition of Z-FR-MCA hydrolysis (65%). The potential reason for this discrepancy between the lack of

gelatinase inhibition and the partial Z-FR-MCA inhibition is the same as described above for the effective tanshinone.

In summary, potential exosite inhibitors were identified using collagenase, gelatinase, and Z-FR-MCA activity assays. Effective exosite inhbitors were: *o*-dihydrotanshinone, *p*-dihydrotanshinone, cryptotanshinone, and myricetin. *o*-dihydrotanshinone and *p*-dihydrotanshione were the most potent exosite inhibitors identified yet for cathepsin K. Cryptotanshinone also has a considerable IC_{50} value. It should be noted, that all these compounds have a demonstrated anti-osteoclastogenesis activity. This might result in limitations for drug development in the future as the aim should be to develop selective anticollagenase inhibitors for cathepsin K without affecting the viability and presence of osteoclasts. The mechanisms of effective inhibitors binding to cathepsin K are still unknown and this is a research area that requires further study. Structure-activity relationship studies (SAR) need to be applied to these compounds for their future improvement. An X-ray structure revealing the exact binding of one of the most potent tanshinones will further help to improve inhibitors by rational design.

Chapter 4: Conclusions and Suggestions for Future Work

4.1 Conclusions

Cathepsin K plays an important role not only in bone, but also in other tissues, like skin, lungs and brain. It should be noted that the active site is not the only critical site that affects the activity of a protease. Other sites that are distant from the active site, named exosites, also have a critical role in substrate recognition and turnover. These sites are referred to as protein-protein interaction sites and GAG binding sites for cathepsin K. To summarize my study, protein-GAG binding sites were first studied to understand the mechanism of the collagenase activity of cathepsin K. Then, considering the functions of exosites on the collagenase activity, exosite inhibitors were identified and characterized to selectively inhibit the collagenase activity without affecting other enzyme activities (Figure 22).

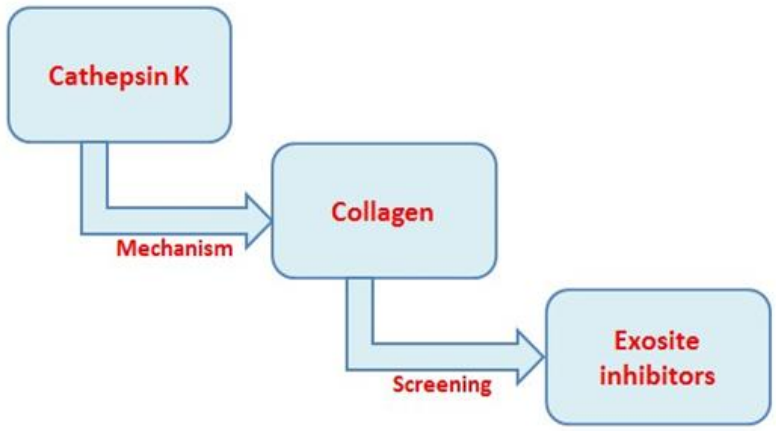


Figure 22. Overall Summary of My Projects

In project 1, the mechanism of collagen degradation by cathepsin K/GAG complexes was studied. The active cleft of cathepsin K is only 5Å, which is not sufficient to accommodate a

tropocollagen molecule with a diameter of 15Å. GAGs are needed to form a specific complex with cathepsin K to allow its collagenase activity. These complexes are likely needed to unravel the triple helix of collagen. GAG binding sites were studied based on tetramer and dimer cathepsin K models for their functions related to their collagenase activities. The tetramer and dimer models were derived from X-ray studies of cathepsin K/GAG complexes. Mutational analysis of protein-GAG interactions supported the existence and functions of the tetramer and dimer models in the degradation of soluble and insoluble collagens, respectively. However, the inhibitory effect on degrading tropocollagen by dimer GAG binding mutants was not expected. The cathepsin K hexamer model (advanced version of tetramer model) was employed to further explain the degradation of collagen. Considering the importance of GAG binding sites, exosite inhibitors could be exploited to interact with GAG binding sites so that the collagenolytic degradation could be selectively blocked.

In project 2, 15 compounds chosen from Traditional Chinese Medicine were examined to determine whether they could be potential exosite inhibitors. Compounds were divided into an anti-resorptive drug group and compounds increasing bone formation. The largest group of potential anti-resorptive compounds was presented by tanshinones, which are diterpenoids isolated from *Salvia miltirhozzia*. For the tanshinones, IC_{50} values for the inhibition of tropocollagen degradation were between 10-22 μM. Preliminary SAR studies illustrated that it might be critical for the inhibitory effect on the collagenase activity of cathepsin K that oxolane structure of ring D and the connected methyl group are not in the same plane. Other compounds that have anti-resorptive functions do not block the collagenase activity of cathepsin K. Interestingly, myricetin, a flavonoid known to increase bone formation, is an effective exosite inhibitor of cathepsin K.

More potential exosite inhibitors need to be identified to obtain more information for SAR analysis. Potent exosite inhibitors which avoid active site inhibition could be applied to drug design for a novel generation of osteoporosis inhibitors.

4.2 Suggestions for Future Work

This thesis has focused on the role of GAG binding sites to determine the mechanism of the collagenase activity in cathepsin K tetramer and dimer models and on the development of potential exosite inhibitors. The results obtained from these two projects will inspire future studies:

4.2.1 Chapter 2

1. Individual site directed mutagenesis to define the most critical residues involved in tetramer and dimer formation in the presence of GAGs.

2. Crystallization of inactive cathepsin K with GAGs and a synthetic triple helical peptide to determine the structure of a cathepsin K/GAG complex with collagen. This will offer more information about the fine mechanism of the collagenolytic activity of cathepsin K.

4.2.2 Chapter 3

1. The mechanism and binding sites of exosite inhibitors identified in this study are not clear. For further study, crystallizations of cathepsin K in the presence of inhibitors are needed to determine the exact binding sites of exosite inhibitors.

2. Scanning electron microscopy (SEM) and atomic force microscopy (AFM) analysis could be used to study the effects of exosite inhibitors on collagen fiber degradation and on the morphology of the collagenolytically active cathepsin K/GAG complexes.

3. Based on future 3D structures of tanshinone cathepsin K complexes, targeted SAR will be possible for the development of advanced inhibitors.

Additional screening for potential exosite inhibitors and in silico screening should be performed to identify more potent exosite inhibitors of different chemical scaffolds.

References

1. Lopez-Otin, C. and C. M. Overall (2002). "Protease degradomics: a new challenge for proteomics." Nat Rev Mol Cell Biol 3(7): 509-519.
2. Rawlings, N. D., et al. (2006). "MEROPS: the peptidase database." Nucleic Acids Res 34(Database issue): D270-272.
3. Turk, D. and G. Guncar (2003). "Lysosomal cysteine proteases (cathepsins): promising drug targets." Acta Crystallogr D Biol Crystallogr 59(Pt 2): 203-213.
4. Lecaille, F., et al. (2008). "Biochemical properties and regulation of cathepsin K activity." Biochimie 90(2): 208-226.
5. Isabel, E., et al. (2010). "The discovery of MK-0674, an orally bioavailable cathepsin K inhibitor." Bioorg Med Chem Lett 20(3): 887-892.
6. Costa, A. G., et al. (2011). "Cathepsin K: its skeletal actions and role as a therapeutic target in osteoporosis." Nat Rev Rheumatol 7(8): 447-456.
7. Li, Z., et al. (2002). "Collagenase activity of cathepsin K depends on complex formation with chondroitin sulfate." J Biol Chem 277(32): 28669-28676.
8. Brömme, D. and K. Okamoto (1995). "Human cathepsin O2, a novel cysteine protease highly expressed in osteoclastomas and ovary molecular cloning, sequencing and tissue distribution." Biol Chem Hoppe Seyler 376(6): 379-384.
9. Li, Z., et al. (2004). "Regulation of collagenase activities of human cathepsins by glycosaminoglycans." J Biol Chem 279(7): 5470-5479.
10. Gauthier, J. Y., et al. (2008). "The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K." Bioorg Med Chem Lett 18(3): 923-928.

11. Cherney, M. M., et al. (2011). "Structure-activity analysis of cathepsin K/chondroitin 4-sulfate interactions." J Biol Chem 286(11): 8988-8998.
12. Lecaille, F., et al. (2002). "Human and parasitic papain-like cysteine proteases: their role in physiology and pathology and recent developments in inhibitor design." Chem Rev 102(12): 4459-4488.
13. Schechter, I. and A. Berger (1967). "On the size of the active site in proteases. I. Papain." Biochem Biophys Res Commun 27(2): 157-162.
14. Scott, J. E. (1988). "Proteoglycan-fibrillar collagen interactions." Biochem J 252(2): 313-323.
15. Turk, B., et al. (2000). "Lysosomal cysteine proteases: more than scavengers." Biochim Biophys Acta 1477(1-2): 98-111.
16. Lecaille, F., et al. (2002). "Selective inhibition of the collagenolytic activity of human cathepsin K by altering its S2 subsite specificity." Biochemistry 41(26): 8447-8454.
17. Lecaille, F., et al. (2003). "Probing cathepsin K activity with a selective substrate spanning its active site." Biochem J 375(Pt 2): 307-312.
18. McGrath, M. E., et al. (1997). "Crystal structure of human cathepsin K complexed with a potent inhibitor." Nat Struct Biol 4(2): 105-109.
19. Matsumoto, K., et al. (1999). "Structural basis of inhibition of cysteine proteases by E-64 and its derivatives." Biopolymers 51(1): 99-107.
20. Brömme, D., et al. (1991). "Tight-binding inhibition of cathepsin S by cystatins." Biomed Biochim Acta 50(4-6): 631-635.

21. Bengtsson, E., et al. (2005). "Lack of the cysteine protease inhibitor cystatin C promotes atherosclerosis in apolipoprotein E-deficient mice." Arterioscler Thromb Vasc Biol 25(10): 2151-2156.
22. Brömme, D. and F. Lecaille (2009). "Cathepsin K inhibitors for osteoporosis and potential off-target effects." Expert Opin Investig Drugs 18(5): 585-600.
23. Katunuma, N. (2011). "Structure-based development of specific inhibitors for individual cathepsins and their medical applications." Proc Jpn Acad Ser B Phys Biol Sci 87(2): 29-39.
24. Ochi, Y., et al. (2011). "Effects of ONO-5334, a novel orally-active inhibitor of cathepsin K, on bone metabolism." Bone 49(6): 1351-1356.
25. Hou, W. S., et al. (2001). "Cathepsin k is a critical protease in synovial fibroblast-mediated collagen degradation." Am J Pathol 159(6): 2167-2177.
26. Tepel, C., et al. (2000). "Cathepsin K in thyroid epithelial cells: sequence, localization and possible function in extracellular proteolysis of thyroglobulin." J Cell Sci 113 Pt 24: 4487-4498.
27. Rivas, M. and J. R. Naranjo (2007). "Thyroid hormones, learning and memory." Genes Brain Behav 6 Suppl 1: 40-44.
28. Bhoola, K. D., et al. (1992). "Kinins--key mediators in inflammatory arthritis?" Br J Rheumatol 31(8): 509-518.
29. Monteiro, R. Q. (2005). "Targeting exosites on blood coagulation proteases." An Acad Bras Cienc 77(2): 275-280.

30. Overall, C. M. (2002). "Molecular determinants of metalloproteinase substrate specificity: matrix metalloproteinase substrate binding domains, modules, and exosites." Mol Biotechnol 22(1): 51-86.
31. Steffensen, B., et al. (1995). "Extracellular matrix binding properties of recombinant fibronectin type II-like modules of human 72-kDa gelatinase/type IV collagenase. High affinity binding to native type I collagen but not native type IV collagen." J Biol Chem 270(19): 11555-11566.
32. Clark, I. M. and T. E. Cawston (1989). "Fragments of human fibroblast collagenase. Purification and characterization." Biochem J 263(1): 201-206.
33. Somoza, J. R., et al. (2000). "Crystal structure of human cathepsin V." Biochemistry 39(41): 12543-12551.
34. Hardegger, L. A., et al. (2011). "Systematic investigation of halogen bonding in protein-ligand interactions." Angew Chem Int Ed Engl 50(1): 314-318.
35. Du, X., et al. (2013). "Elastin degradation by cathepsin V requires two exosites." J Biol Chem.
36. He, D., et al. (2007). "Comparative genomics of elastin: Sequence analysis of a highly repetitive protein." Matrix Biol 26(7): 524-540.
37. Bruckner, P. and M. van der Rest (1994). "Structure and function of cartilage collagens." Microsc Res Tech 28(5): 378-384.
38. Szpak, P. (2011). "Fish bone chemistry and ultrastructure: implications for taphonomy and stable isotope analysis." Journal of Archaeological Science 38(12): 3358-3372.
39. Shoulders, M. D. and R. T. Raines (2009). "Collagen structure and stability." Annu Rev Biochem 78: 929-958.

40. Hulmes, D. J. (1992). "The collagen superfamily--diverse structures and assemblies." Essays Biochem 27: 49-67.
41. Fields, G. B. (2013). "Interstitial collagen catabolism." J Biol Chem 288(13): 8785-8793.
42. Jarvelainen, H., et al. (2009). "Extracellular matrix molecules: potential targets in pharmacotherapy." Pharmacol Rev 61(2): 198-223.
43. Gelse, K., et al. (2003). "Collagens--structure, function, and biosynthesis." Adv Drug Deliv Rev 55(12): 1531-1546.
44. Lord, M. S. and J. M. Whitelock (2013). "Recombinant production of proteoglycans and their bioactive domains." FEBS J 280(10): 2490-2510.
45. Hochberg, M., et al. (2013). "Symptom and structure modification in osteoarthritis with pharmaceutical-grade chondroitin sulfate: what's the evidence?" Curr Med Res Opin 29(3): 259-267.
46. Velasco, C. R., et al. (2010). "Proteoglycans on bone tumor development." Drug Discov Today 15(13-14): 553-560.
47. Sasisekharan, R., et al. (2006). "Glycomics approach to structure-function relationships of glycosaminoglycans." Annu Rev Biomed Eng 8: 181-231.
48. Sobieszczanska, M., et al. (2013). "Osteoporosis: genetic determinants and relationship with cardiovascular disease." Adv Clin Exp Med 22(1): 119-124.
49. Ng, K. W. (2012). "Potential role of odanacatib in the treatment of osteoporosis." Clin Interv Aging 7: 235-247.
50. Warriner, A. H. and K. G. Saag (2013). "Osteoporosis diagnosis and medical treatment." Orthop Clin North Am 44(2): 125-135.

51. Barrett-Connor, E., et al. (2006). "Effects of raloxifene on cardiovascular events and breast cancer in postmenopausal women." N Engl J Med 355(2): 125-137.
52. Falgueyret, J. P., et al. (2005). "Lysosomotropism of basic cathepsin K inhibitors contributes to increased cellular potencies against off-target cathepsins and reduced functional selectivity." J Med Chem 48(24): 7535-7543.
53. Alves Pereira, D., et al. (2008). "Pycnodysostosis. A report of 3 clinical cases." Med Oral Patol Oral Cir Bucal 13(10): E633-635.
54. Xue, Y., et al. (2011). "Clinical and animal research findings in pycnodysostosis and gene mutations of cathepsin K from 1996 to 2011." Orphanet J Rare Dis 6: 20.
55. Li, Z., et al. (2008). "The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex." J Mol Biol 383(1): 78-91.
56. Nallaseth, F. S., et al. (2013). "The role of basic amino Acid surface clusters on the collagenase activity of cathepsin k." Biochemistry 52(44): 7742-7752.
57. Le Gall, C., et al. (2008). "Cathepsin K inhibitors as treatment of bone metastasis." Curr Opin Support Palliat Care 2(3): 218-222.
58. Yang, Y., et al. (2013). "The Role of Traditional Chinese Medicines in Osteogenesis and Angiogenesis." Phytother Res.
59. Di, M. Y. and J. L. Tang (2013). "Adaption and Application of the Four Phase Trials to Traditional Chinese Medicines." Evid Based Complement Alternat Med 2013: 128030.
60. Wang, Z., et al. (2013). "Molecular Imaging in Traditional Chinese Medicine Therapy for Neurological Diseases." Biomed Res Int 2013: 608430.

61. Lai, J. N., et al. (2013). "Observational Studies on Evaluating the Safety and Adverse Effects of Traditional Chinese Medicine." Evid Based Complement Alternat Med 2013: 697893.
62. Li, AR., etal. (1998). "Traditional Chinese medical treatment in Guzhi Shusong Zheng (Osteoporosis)." Anhui Scientific Technology Press: Hefei; 1–2.
63. Kim, H. H., et al. (2004). "Inhibition of osteoclast differentiation and bone resorption by tanshinone IIA isolated from Salvia miltiorrhiza Bunge." Biochem Pharmacol 67(9): 1647-1656.
64. Lee, S. Y., et al. (2005). "Inhibition of osteoclast differentiation by tanshinones from the root of Salvia miltiorrhiza bunge." Arch Pharm Res 28(8): 909-913.
65. Hu, Z., et al. (2011). "Berberine induces dendritic cell apoptosis and has therapeutic potential for rheumatoid arthritis." Arthritis Rheum 63(4): 949-959.
66. Li, H., et al. (1999). "The effect of kampo formulae on bone resorption in vitro and in vivo. II. Detailed study of berberine." Biol Pharm Bull 22(4): 391-396.
67. Xu, D., et al. (2010). "Preventive effects of berberine on glucocorticoid-induced osteoporosis in rats." Planta Med 76(16): 1809-1813.
68. Xue, L., et al. (2012). "Effects and interaction of icariin, curculigoside, and berberine in er-xian decoction, a traditional chinese medicinal formula, on osteoclastic bone resorption." Evid Based Complement Alternat Med 2012: 490843.
69. Qin, L. P., et al. (2003). "Total coumarins from fruits of Cnidium monnieri inhibit formation and differentiation of multinucleated osteoclasts of rats." Acta Pharmacol Sin 24(2): 181-186.

70. Tang, D. Z., et al. (2010). "Osthole stimulates osteoblast differentiation and bone formation by activation of beta-catenin-BMP signaling." J Bone Miner Res 25(6): 1234-1245.
71. Nie, Y., et al. (2013). "Medicinal plants of genus Curculigo: traditional uses and a phytochemical and ethnopharmacological review." J Ethnopharmacol 147(3): 547-563.
72. Lo, Y. C., et al. (2010). "Betulinic acid stimulates the differentiation and mineralization of osteoblastic MC3T3-E1 cells: involvement of BMP/Runx2 and beta-catenin signals." J Agric Food Chem 58(11): 6643-6649.
73. Martelanc, M., et al. (2007). "Determination of three major triterpenoids in epicuticular wax of cabbage (Brassica oleracea L.) by high-performance liquid chromatography with UV and mass spectrometric detection." J Chromatogr A 1164(1-2): 145-152.
74. Hsu, Y. L., et al. (2007). "Myricetin induces human osteoblast differentiation through bone morphogenetic protein-2/p38 mitogen-activated protein kinase pathway." Biochem Pharmacol 73(4): 504-514.
75. Payne, J. B. and L. M. Golub (2011). "Using tetracyclines to treat osteoporotic/osteopenic bone loss: from the basic science laboratory to the clinic." Pharmacol Res 63(2): 121-129.
76. Neale, J. R., et al. (2009). "Bone selective effect of an estradiol conjugate with a novel tetracycline-derived bone-targeting agent." Bioorg Med Chem Lett 19(3): 680-683.
77. Golub, L. M., et al. (1999). "A chemically modified nonantimicrobial tetracycline (CMT-8) inhibits gingival matrix metalloproteinases, periodontal breakdown, and extra-oral bone loss in ovariectomized rats." Ann N Y Acad Sci 878: 290-310.

78. Cui, Y., et al. (2011). "Characterization of Salvia miltiorrhiza ethanol extract as an anti-osteoporotic agent." BMC Complement Altern Med 11: 120.
79. Yavropoulou, M. P. and J. G. Yovos (2008). "Osteoclastogenesis--current knowledge and future perspectives." J Musculoskelet Neuronal Interact 8(3): 204-216.
80. Kavanagh, K. L., et al. (2006). "The molecular mechanism of nitrogen-containing bisphosphonates as antiosteoporosis drugs." Proc Natl Acad Sci U S A 103(20): 7829-7834.
81. Martin, T. J. and N. A. Sims (2005). "Osteoclast-derived activity in the coupling of bone formation to resorption." Trends Mol Med 11(2): 76-81.
82. Paul, A., et al. (2013). "Anticancer potential of myricanone, a major bioactive component of Myrica cerifera: novel signaling cascade for accomplishing apoptosis." J Acupunct Meridian Stud 6(4): 188-198.
83. Sun, P., et al. (2013). "An inhibitor of cathepsin K, icariin suppresses cartilage and bone degradation in mice of collagen-induced arthritis." Phytomedicine 20(11): 975-979.
84. Alves, M. F., et al. (2003). "S3 to S3' subsite specificity of recombinant human cathepsin K and development of selective internally quenched fluorescent substrates." Biochem J 373(Pt 3): 981-986.
85. Yasuda, Y., et al. (2004). "Cathepsin V, a novel and potent elastolytic activity expressed in activated macrophages." J Biol Chem 279(35): 36761-36770.
86. Chen, F., et al. (2013). "Estrogen inhibits RANKL-induced osteoclastic differentiation by increasing the expression of TRPV5 channel." J Cell Biochem.