5-2022

Systems Modeling of Collagenous Tissue Degradation: Protease and Mechanical Interactions

Amirreza Yeganegi

ayegane@clemson.edu

Follow this and additional works at: https://tigerprints.clemson.edu/all_dissertations

Recommended Citation

4-2022

Systems Modeling of Collagenous Tissue Degradation: Protease and Mechanical Interactions

Amirreza Yeganegi

SYSTEMS MODELING OF COLLAGENOUS TISSUE DEGRADATION:
PROTEASE AND MECHANICAL INTERACTIONS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Bioengineering

by
Amirreza Yeganegi
May 2022

Accepted by:
Dr. William Richardson, Committee Chair
Dr. Bruce Z. Gao
Dr. Robert Latour
Dr. Yongren Wu

ABSTRACT

Heart failure (HF) is a chronic, progressive condition defined as an abnormality of cardiac function with the inability of the heart muscle to pump enough blood to meet the body's requirements for metabolism. HF has various contributing pathologies, including hypertension (86 million Americans), myocardial infarction (MI, 800,000 Americans per year and 300,000 recurrent infarctions each year), both of which promote fibrosis. Myocardial fibrosis contributes to left ventricular (LV) dysfunction and is histologically defined by excessive deposition of fibrous tissue relative to the mass of cardiomyocytes within the myocardial tissue. Quantitatively, myocardial fibrosis is characterized by increased collagen volume fraction (CVF) or percentage of myocardial tissue with collagen fibers. Currently, there are no prescribed therapeutics for preventing cardiac fibrosis, and clinicians are unable to predict which patients at what time and to what extent are more likely to develop fibrosis. Collagen accumulation contributes to increased stiffness and loss of function in failing hearts, and cardiac fibrosis remains a significant barrier to the treatment and prevention of HF. Collagen remodeling is regulated by a complex network of extracellular interactions, including: (1) collagen secretion, (2) protease secretion, activation, and degradation of collagen (namely Matrix Metalloproteinase (MMP) and Cathepsins), and (3) tissue inhibitors of metalloproteinases (TIMP) secretion and inhibition of MMPs. Importantly, this network is sensitive to mechanical tension. Fibroblast expression of collagen, MMPs, and TIMPs all depend on tension, and it is known that an excessive amount of tension can damage matrix fibers. There is also evidence that protease degradation of collagen can depend on fiber tension. However, it is unknown how tension

affects collagen degradation by different proteases and protease mixes. The overarching objective of this dissertation is to develop a computational model of collagen turnover under combinatory chemo-mechano-conditions as a predictive tool for stratifying fibrotic risk for HF patients. Firstly, we tested the effect of tensile loading on collagenous tissue degradation by proteases. We picked four proteases and quantified the role of mechanical loading on the degradation of collagenous tissue by each protease. As matrix degradation leads to decaying force levels, sample degradation rate was quantified for different strain levels for each protease. Secondly, we developed a detailed biochemical network computational model of collagen I proteolysis capturing all interactions of type I collagen, four MMPs, and three TIMPs in a cell-free, well-stirred environment. We monitored the proteolytic activity of MMPs and inhibitory activity of TIMPs and then used the results from experimental data to fit five different hypothetical reaction topologies and determined kinetic rate constants for collagen degradation by MMPs, MMP inhibition by TIMPs, MMP and TIMP inactivation, MMP cannibalism, and MMP and TIMP distraction. We also used post-MI time courses of collagen, MMP, and TIMP levels in animal experiments from the literature to perform a parameter sensitivity analysis across the model reaction rates to identify which molecules or interactions are the essential regulators of ECM post-MI for both early and late time-periods. Lastly, we developed an ensemble classification algorithm for diagnosing HF patients with preserved ejection fraction (HFpEF) within a population of 459 individuals, including HFpEF patients and referent control patients. We concluded that machine learning algorithms could substantially improve the predictive value of circulating plasma biomarkers. Additionally, we built a mechanistic model to predict ECM

component degradation using a genetic algorithm to connect ECM remodeling to the plasma biomarkers to help us with HFpEF patients' classification. Our findings demonstrate that machine learning-based classification algorithms show promise as a non-invasive diagnostic tool for HFpEF patients' classification while also suggesting priority biomarkers for future mechanistic studies to elucidate more specific regulatory roles. Our work suggests that computational modeling can serve as a beneficial tool for HF prognosis and potentially developing novel therapeutics.

DEDICATION

This work is dedicated to my parents, Shiva and Hamid, who have sacrificed their life for me. You are the reason I wake up everyday and try to be a better human being. I am so grateful for having your love and support.

This work is also dedicated to my sister, Dr. Elaheh Yeganegi, who is the best sister anyone can ask for. Thank you for all of your support and love.

every moment of my life and showing me the meaning of love. To Elaheh—thank you for

being an awesome sister and emotional support through my PhD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

List of Figures (Continued)

CHAPTER 1

INTRODUCTION

1.1 Study Significance

An estimated 6.2 million American adults ≥20 years of age had heart failure (HF) between 2013 and 2016. HF has various contributing pathologies, including hypertension (86 million Americans), myocardial infarction (MI) and fibrosis. Myocyte hypertrophy resulting as a compensatory reaction to prior pathologies above enlarges the ventricle and reduces systolic and diastolic function, while collagen deposition stiffens the myocardial wall, further reducing the pump action[1–3].

Experimental studies show that collagen content is a critical determinant in left ventricular (LV) remodeling, and collagen content is sensitive to mechanical strain[4–6]. Several drugs used to control hypertension and post-MI HF were found to control myocardial remodeling and decrease interstitial fibrosis[7,8]. Direct stem cell transplantation into the healing infarct is already in use as an experimental therapy, and tissue-engineered replacement patches of the myocardium can also be a potential treatment[7]. However, collagen sensitivity to mechanical strain shows that many therapies targeting LV remodeling may have different results under different degrees of tension or different patient-specific biochemical levels. These therapies are developed mainly based on trial-and-error rather than from an understanding of the mechanical properties of the healing infarct and its coupling to the LV[7].

Collagen remodeling is a result of a network of extracellular interactions, including (1) collagen secretion, (2) protease secretion, activation, and degradation of collagen, and (3) tissue inhibitors of metalloproteinases (TIMP) secretion and inhibition of proteases[9–12]. The diversity of matrix-protease-TIMP interactions makes it very difficult to intuitively predict the effects of any individual matrix-, protease-, or TIMP-targeting therapy. It is even more challenging to predict therapeutic effects across patients' variabilities. There is great potential benefit in systematically predicting dynamic matrix turnover under various matrix, protease, and TIMP levels by constructing a computational network model of their interactions.

Several past studies have shown that mechanical loads can alter protease-mediated degradation of collagen, presumably due to altered molecular conformations of the collagen molecule's protease-binding sites. However, some groups report increased collagen turnover with increased loading, while other groups report decreased collagen turnover with loading[13–25]. Most of these groups used bacterial collagenase in their studies, a robust protease but physiologically less relevant than human matrix metalloproteinases (MMPs)[21]. It is also unknown which proteases are sensitive to strain and what are the levels of those sensitivities. It is very important to test the effect of strain on human protease isoform-specific degradation of collagenous tissue to better understand the sensitivity of each protease to strain, thereby elucidating the relative contributions of various protease isoforms to collagen turnover in the dynamically loaded heart.

In order to predict collagen remodeling under mechanical loading, our first step is to experimentally test the effect of mechanical strain on collagenous tissue degradation by

proteases. Then, we will computationally model collagen degradation by various mixes of proteases and TIMPs. Finally, we will test our modeling capability to predict the risk of HF within a clinical dataset.

1.2 Specific Aims

Aim 1: Mechanical Strain Modulates Extracellular Matrix Degradation Kinetics and Byproducts in Isoform-Specific Manner.

Decellularized porcine pericardium samples were treated with recombinant human MMP-1, MMP-8, MMP-9, Cathepsin K, or a protease-free control while subjected to different levels of mechanical strain (from ~5-40%). Isotropic displacement control was provided, and the degradation level of pericardium samples was measured using force decay data. The degradation products were also analyzed by mass spectrometry to assess how mechanical strain levels altered the degradome signatures.

Aim 2: Build and Analyze a Computational Model of Collagen Turnover Regulation by MMPs and TIMPs.

A detailed computational model of the biochemical network of collagen I proteolysis capturing all type I collagen interactions, four MMPs (MMP-1, -2, -8, and -9), and three TIMPs (TIMP-1, -2, and -4) in a cell-free, well-stirred environment were presented. Dye Quenched (DQ) collagen was used to monitor the proteolytic activity of MMPs and the inhibitory activity of TIMPs. The experimental results were then used to fit five different hypothetical reaction topologies using a system of ordinary differential equations (ODEs) to determine the kinetic rate constants of the collagen-MMP-TIMP

network. Post-MI time-courses of collagen, MMP, and TIMP levels in the animal heart were also used from the literature to perform a parameter sensitivity analysis across the model reaction rates to identify which molecules or interactions are the important regulators of ECM post-MI.

<u>Aim 3: Integrate Ensemble Machine Learning and Mechanistic Modeling to Identify HFpEF Patients from Matrix-Related Plasma Biomarkers</u>

We developed multiple advanced machine learning frameworks for the classification of HFpEF patients within a population of 459 individuals, including HFpEF patients and referent control patients. Additionally, we developed a mechanistic model for five different ECM component remodeling, including type I and type III collagen and three potential ECM candidates. A genetic fitting algorithm was then used to find the best-fit combination of rate parameters simultaneously in order to predict ECM components turnover for patient-specific data. The patient-specific results from the mechanistic model were then applied to all of the multiple machine learning models to improve the models' predictive capabilities.

CHAPTER 2

LITERATURE REVIEW

2.1 Heart Failure

Cardiovascular disease (CVD) is the leading cause of death worldwide and responsible for 31% of global mortality[26]. The estimated annual costs for CVD and strokes are $316.6 billion per year[26]. Heart failure (HF) is a chronic, progressive condition and an abnormality of cardiac function with the inability of the heart muscle to pump enough blood to meet the body's requirements for metabolism[27]. An estimated 6.2 million American adults ≥ 20 years of age had HF between 2013 and 2016, compared with an estimated 5.7 million between 2009 and 2012. The total percentage of the population experiencing HF was 2.42% in 2012. Projections show that HF will increase 46% from 2012 to 2030, resulting in more than 8 million people ≥18 years of age with HF, and the total percentage of the population with HF is predicted to increase from 2.42% in 2012 to 2.97% in 2030[26,28].

There are two types of left-sided HF. Systolic failure or HF with reduced ejection fraction (HFrEF) happens when the left ventricle (LV) loses its ability to contract normally. Diastolic failure or HF with preserved ejection fraction (HFpEF), also called diastolic dysfunction, is a condition resulting from the stiffness of heart muscle and causes the inability of the LV to relax normally. In this condition, the heart cannot appropriately fill with blood during the resting period between each beat[27,29].

HF has various contributing pathologies, including hypertension (86 million Americans), myocardial infarction (MI, 800,000 Americans per year[2] and 300,000 recurrent infarctions each year[7]), and fibrosis. MI is one of the major causes of death in the US and worldwide[30]. Progression to HF can occur in up to one-third of patients as a result of adverse remodeling of the collagenous scar. Excessive collagen accumulation can result in a stiff and non-compliant LV, and insufficient collagen in scar tissue can cause LV thinning and dilation[31]. Several studies show LV remodeling is the primary mechanism of death and disability after MI[31]. Myocyte hypertrophy resulting as a compensatory reaction to prior pathologies above enlarges the ventricle and reduces systolic and diastolic function, while collagen deposition stiffens the myocardial wall, further reducing the pump action.

Myocardial fibrosis contributes to LV dysfunction and is histologically defined by excessive deposition of fibrous tissue relative to the mass of cardiomyocytes within the myocardial tissue. Quantitatively, myocardial fibrosis is characterized by increased collagen volume fraction (CVF) or percentage of myocardial tissue with collagen fibers. Fibrosis stiffens the myocardial wall, thereby decreasing LV distensibility, contractility, and pump function.

2.1.1 Current standards of care for prevention and treatment

Several of the most prescribed drugs for HF include angiotensin-converting enzyme (ACE) inhibitors, angiotensin receptor blockers (ARBs), beta-blockers, and neprilysin inhibitors. Most treatments that show promising results are for HFrEF and are not effective against HFpEF. Treatments for chronic HFrEF are similar across the American and

European guidelines. Patients with symptomatic HFrEF should receive a combination of an (ACE)-I (or ARB if ACE-I is not tolerated), a beta-blocker, and a mineralocorticoid antagonist (MRA). Clinical trials have confirmed the beneficial effects of ACE-I in improving mortality in patients with HFrEF[32–34]. Several past HF trials show that ARBs are similarly effective in HFrEF when compared with ACE-I[34,35]. Currently, no specific drug therapy shows significantly improved mortality and morbidity in patients with HF and preserved ejection fraction[34,36].

While the use of therapeutics mentioned above has vastly improved patient care, it is important to note that many patients do not respond to these lines of treatment as reported in the randomized clinical trials. Also, many of these treatments target vasoconstricting processes and help with hypertension. But in many cases, residual hypertrophy or fibrosis still causes an overall reduction in heart function, leading to hospital admission, reduced quality of life, and a high risk of mortality for patients. Hypertension (i.e., highblood pressure) is associated with an increased risk of developing HF. A study on antihypertensive drug therapy showed a reduction in HF. They concluded that the preventive effect of antihypertensive therapy on HF stems from the blood pressure lowering effects rather than from drug classes[37].

It is important to develop new therapies that target cardiac fibrosis in addition to hypertrophy and high blood pressure. Furthermore, no FDA-approved drugs currently exist to target fibrosis, and clinicians cannot predict which patients at what time and to what extent are more likely to develop fibrosis. Therefore, understanding the mechanisms of cardiac fibrosis can directly impact patient outcomes by developing new therapeutics.

2.2 Scar formation

Myocardial fibrosis happens when an excessive amount of extracellular matrix (ECM) components accumulates in the myocardium. In this condition, the net ECM deposition is greater than ECM degradation by matrix proteases[38]. Various conditions can promote cardiac fibrosis, such as MI, hypertensive heart disease, diabetic hypertrophic cardiomyopathy, and idiopathic dilated cardiomyopathy[26,38–40]. A pathological remodeling of the ECM leads to fibrotic scars and the scar tissue in the heart causes several cardiac disfunction. Scar tissue causes myocardial matrix stiffening, and as a result, the ejection fraction reduces. It also impairs diastolic performance and can lead to death[26].

The human body has an impressive capacity to heal itself, especially after a minor injury; however, it cannot heal all defects, which is true for the heart. The most considerable fibrotic remodeling happens after acute cardiomyocyte death since mammalian myocardium has negligible regenerative capacity. The processes of ECM deposition and degradation are typically balanced in healthy myocardium. ECM deposition after an injury is a protective mechanism and can be helpful for wound healing and tissue regeneration. However, excessive ECM deposition can lead to impaired tissue function. Excessive collagen accumulation can result in a stiff and non-compliant LV, and an insufficient amount of collagen in scar tissue can cause LV thinning and dilation[31]. Several studies show LV remodeling is the principal mechanism of death and disability after MI[31].

Experimental studies show that collagen content is a critical determinant in LV remodeling, and collagen content is sensitive to mechanical strain[4–6]. Several drugs used to control hypertension and post-MI HF were found to control myocardial remodeling and

decrease interstitial fibrosis[7,8]. Direct stem cell transplantation into the healing infarct is already in use as an experimental therapy, and tissue-engineered replacement patches of the myocardium can also be a potential treatment[7]. However, collagen sensitivity to mechanical strain shows that many therapies targeting LV remodeling may have different results under different degrees of tension or different patient-specific biochemical levels. These therapies are developed mainly based on trial-and-error rather than an understanding of the mechanical properties of the healing infarct and its coupling to the LV[7]. Collagen remodeling is a result of a network of extracellular interactions, including (1) collagen secretion, (2) protease secretion, activation, and degradation of collagen, and (3) tissue inhibitors of metalloproteinases (TIMP) secretion and inhibition of proteases[10–12].

2.2.1 Collagen

The ECM surrounds cells and forms a connective structure critical for overall tissue function. ECM also creates the cellular environments required during development and morphogenesis[41–43]. ECM comprises many substances, but only collagen fibrils and proteoglycans are present in all connective tissues; however, they are in different forms and shapes[44].

Collagen is the most abundant protein in the human body and comprises one-third of the body's total protein[45]. Collagen synthesis, formation alteration, and degradation are essential processes in many physiological processes such as development and many diseases. There are more than 25 types of collagens that vary in the nature, length of helix,

9

and size of the non-helical portions. The most common collagen type is fibrillar types I and III, which are composed of three polypeptide subunits that exist in a triple helix form[45,46].

The distinctive feature of collagen is its triple helix structure in which three parallel polypeptide strands in a left-handed, helical conformation coil about each other. This structure mandates every third residue to be glycine (Gly), and the final formation can be a repeating sequence of Xaa Yaa Gly, where Xaa and Yaa can be any amino acids[45].

Type I collagen is the most abundant form of collagen in human tissue and is found in skin, tendon, and bone in large amounts. The triple helix structure of the collagen I molecule is comprised of two identical "a1" chains and one "a2" chain. The diameter of the triple-helical molecule is 1.5 nm, and its length is 300 nm. Form and structure of ECM depend mainly on collagen, and stiff collagen fibrils self-assemble into large-scale structures such as fibers and sheets. They can make large and long parallel arrays found in tendons and ligaments or wide regular sheets found in cornea[44,46–50]. The primary role of fibril forming collagens (collagen I, II, III, V, and XI) is bearing and transmitting mechanical loads along their main axis[13].

2.2.2 Proteases

Fibroblasts in vertebrate animals secrete matrix metalloproteinases (MMPs) and cathepsins to modulate the ECM and degrade the collagen protein. Both MMP and cathepsin are involved in matrix formation, remodeling, and homeostasis. MMPs are zinc-dependent proteinases that can degrade the native collagen triple helix as well as other ECM components relevant to LV remodeling. In normal physiological conditions, the

activities of MMPs are regulated at the level of transcription, activation of the precursor zymogens, interaction with specific ECM components, and inhibition by inhibitors[5,41,51]. Currently, the MMP family is composed of 25 proteinases that can be categorized into five groups based on the substrate they tend to degrade. The collagenases (MMP-1, -8, and -13) can cleave fibrillar-type collagens at a particular site located between Gly775 and Ile776. The gelatinases (MMP-2 and -9) can degrade gelatins. The stromelysins (MMP-3 and -10) and matrilysins (MMP-7 and -26) are broad-spectrum proteinases, and the membrane-type MMPs (MT-MMPs) are anchored to the plasma membrane[5,41,51–55]. Some MMPs can degrade multiple substrates, and these categorizations can sometimes overlap. MMP-9 was first thought to be only gelatinase, but recent studies showed its ability to degrade full-length interstitial collagens[51].

Cathepsins are the superfamily of cysteine proteases that comprises 11 members. Some of its members participate in ECM remodeling and can proteolyze ECM. Cathepsins K and S are involved in elastin degradation in cardiovascular diseases. Cathepsin K is the most potent mammalian collagenase, which is able to cleave type I and type II collagen in the native triple helix as well as in the telopeptide regions. Cathepsin S can degrade elastin and maintain its proteolytic activity at neutral pH, making it unique among the cathepsins, which generally prefer acidic environments[56–59].

Bacterial collagenases are metalloproteinases involved in the degradation of the ECM of animal cells. Until today, bacterial collagenases do not have a proper and well-defined classification, and there is great disagreement regarding the correct identification

11

of bacterial collagenases. Clostridial collagenases were the first bacterial collagenase identified which are the reference for newly discovered collagenolytic enzymes[60].

2.2.3 Tissue inhibitors of metalloproteinases

The TIMP family currently include four different members (TIMPs-1 to 4)[61]. MMPs in active and inactive forms can be inhibited by their physiological TIMPs. Typically, the degenerative potential of the MMPs is mainly balanced by TIMPs. Disruption of this MMP-TIMP balance can result in disorders such as rheumatoid and osteoarthritis, atherosclerosis, tumor growth, metastasis, and fibrosis[61].

2.3 Effect of strain on collagen degradation by proteases

In the past few decades, several studies have shown that mechanical loads can alter protease-mediated degradation of collagen, presumably due to altered molecular conformations of the collagen molecule's protease-binding sites (Table 1). These studies worked on different scales of collagen, from a single molecule to tissue. The primary goal of these experiments was to understand whether mechanical strain increases or decreases the degradation rate of collagen. The first of these studies was conducted by Huang and Yannas in 1977 and found that mechanically loading reconstituted collagen I tapes reduced their degradation by bacterial collagenase[14]. Since then, more reports have agreed with this result, but others are contradictory; some groups report increased collagen turnover with increased loading, while other groups report decreased collagen turnover with loading[13–25]. Most of these groups used bacterial collagenase in their studies which is robust but

physiologically less relevant than MMPs[21]. It is also unknown which proteases are sensitive to strain and what are the level of these sensitivities.

It is important to mention that collagenous structures are well designed to carry tensile loads by forming different structures: individual molecules assemble into microfibrils, which can form larger fibers, which can organize into interconnected networks which can be found in macro-scale gels and tissues[23].

Table 1: Literature report of effect of loading on collagen degradation

Reference	Collagen structure	Collagenase type	Degradation effect	
Camp (2011)	Single molecule	Bacterial collagenase	Decrease	
Adhikari (2011)	Homotrimeric peptide	MMP-1	Increase	
Adhikari (2012)	Single molecule	MMP-1	Increase	
Bhole (2009)	Fibril Network	Bacterial collagenase	Decrease	
Flynn (2010)	Fibril Network	MMP-8	Decrease	
Flynn (2013)	Single fibril	Bacterial collagenase	Decrease	
Huang (1977)	Reconstituted tape	Bacterial collagenase	Decrease	
Nabeshima (1996)	Tendon-tibia units	Bacterial collagenase	Decrease	
Ellsmere (1999)	Pericardial tissue	Bacterial collagenase	Increase	
Ruberti (2005)	Corneal tissue	Bacterial collagenase	Decrease	
Wyatt (2009)	Rat-tail tissue	Bacterial collagenase	Decrease	
Zareian (2010)	Corneal tissue	Bacterial collagenase	Decrease	
Yi (2016)	Lung tissue	Bacterial collagenase	Increase	Decrease
Ghazanfari (2016)	Pericardial tissue	Bacterial collagenase	Decrease	

2.3.1 Molecular Scale

Starting at the individual collagen molecule, Camp et al. and Adhikari et al. both worked on the individual collagen molecule and reported that loading could affect collagen monomer degradation, but each reported an opposite result[16,17,19].

In 2011 Camp et al. showed that tensile loads higher than 3 pN dramatically reduced (10×) the enzymatic degradation rate of recombinant human type I collagen monomers by bacterial collagenase compared to unloaded controls. In order to investigate whether the change of degradation rate stems from molecular mechanics, they used a parallel, single-molecule, mechanochemical reaction assay. They modulated the force on the collagen tethers by changing the height of the magnet above the glass surface to which the superparamagnetic (SPM) beads are connected by the collagen link. A maximum force of about 12 pN could be applied to tethered beads. They divided their experimental series into three categories: "zero force" (Brownian tether forces ~0.06 pN), "low force" (averaging 3.6 ± 1.1 pN), and "high force" (averaging 9.4 ± 1.3 pN). The forces were achieved by changing the magnet stack heights to ∞, 2.6 mm or 1.1 mm above the surface of the glass. Their loaded and unloaded collagen-tethered beads were exposed to 5.56 µM enzyme (Clostridium histolyticum) during the experiment. They reported that the stiffness of the collagen encounters a rapid increase from the low to high force. The data obtained from the 0 pN experimental series, where the beads were collagen-tethered but unloaded, showed a 10-fold increase in the rate of enzymatic digestion relative to the low and high force series[17].

In contrast to Camp et al., Adhikari et al. reported that mechanical load could increase protease degradation of collagen. In 2011, Adhikari et al. showed that the application of $\sim$10 pN in extensional force causes a $\sim$100-fold increase in the degradation rate of collagen by MMP-1. They used a single molecule magnetic tweezers assay to study the effect of force on collagen degradation by MMP-1. Because of the conflicting result of the effects of load on collagen degradation, they provided a quantitative, single-molecule assay on a homogeneous substrate instead of whole tissue or reconstituted collagen. Proteolysis of a collagen trimer results in bead detachment from the coverslip. They measured bead detachment as a function of time and MMP-1 concentration.

$$f(t) = ae^{-kt} + c \qquad \text{(Eq. 1)}$$

Where f(t) is the fraction of beads still attached at time t, k is the detachment rate, and c likely reflects nonspecifically attached beads. They postulate that the apparent differences between previous experiments with their results is probably because of the structural differences between isolated collagen trimers and collagen fibrils, which contain hundreds of trimers. In order to confirm that MMP-1 cleaved the model peptide at the expected recognition site, they used matrix-assisted laser desorption/ionization mass spectrometry.

An important distinction between Adhikari's results versus Camp et al.'s is that the model collagen peptide used by Adhikari et al. was a homotrimeric molecule. In contrast, the collagen molecules of Camp et al. were likely to be primarily heterotrimeric[19].

To test the role of this distinction, Adhikari et al. followed up their first report with a second similar study with the same magnetic bead-based setup (single-molecule magnetic tweezers assay) to test the degradation of heterotrimeric collagen I instead of a

homotrimeric molecule by both MMP-1 and bacterial collagenase (Clostridium histolyticum). They showed that applying 16 pN of force caused an 8-fold increase in collagen proteolysis rates by MMP-1 but did not affect cleavage rates by bacterial collagenase. They confirmed that the bead detachment resulted from collagen degradation. They investigated the effect of both MMP-1 and bacterial collagenase on both the collagen antibody and streptavidin and found neither was degraded by either protease. The experiments were performed using 3 μM MMP-1, without applying force and with forces between 0.2 pN and 16.5 pN. Their results indicate that an applied force of 16.5 pN leads to an 8-fold increase in the observed proteolysis rate. But loading did not affect cleavage rates by the bacterial collagenase (contrasting Camp et al., who observed a significant decrease in degradation with bacterial collagenase).

Adhikari's observations suggest that Clostridium collagenase degrades collagen independent of the unwinding process, which is essential for degradation by MMP-1. Since different proteases attack and cleave collagen molecules in various sites, the protease type and collagen molecular assembly may be the cause of the contradictory degradation rate in response to strain. Clearly, there is a difference between the degradation of hetero- vs. homotrimer, as well as a difference between mammalian and bacterial collagenases. Still, these differences do not fully explain the opposite responses observed by the two groups[16].

In an effort to explore these experimental discrepancies of degradation under strain, Teng and Hwang employed computational simulations of collagen molecular dynamics and protease binding[62]. Adhikari et al.'s setup linked individual collagen molecules between a magnetic bead and glass surface in a way that allowed the bead to rotate during

16

loading freely. One key distinction in Teng et al.'s simulations was allowing the end of the collagen molecule to rotate which facilitated molecular unwinding under load, which could enhance protease cleavage. Camp et al.'s setup, however, exposed beads to a volume of collagen I molecules; Teng et al. hypothesized that the setup details could result in multiple collagen molecules bound together between the relatively large beads and the glass surface, thereby restricting the conformational motion of the individual molecules and the beads. It is not known, though, whether this was indeed the case. Additional simulations also predicted that molecular unwinding of the triple helix could result in either stabilization or destabilization depending on specific peptide sequences (e.g., imino-rich vs. imino-poor domains). Clearly, there are a variety of very subtle factors that can dramatically affect the mechano-sensitivity of collagen molecule degradation by proteases.

2.3.2 Fibril and Fibril Network Scale

As mentioned above, collagenous structures are not simply a group of individual molecules but rather organized hierarchies of many collagen monomers polymerizing into fibrils and fibril networks. As such, several studies have also tested how mechanically loading these collagen polymer structures affects protease-mediated degradation. These studies have collectively agreed that at the fibril scale, load tends to reduce the degradation rate of collagen by a variety of collagenases. These reports suggest that the procedures that cause a single collagen molecule to respond differently to strain do not apply to collagen fibril and micro-networks.

In 2009 Bhole et al. showed that networks of collagen fibrils exposed to collagenase (Clostridium histolyticum) persist longer under strain compared with unstrained, free collagen fibrils. They used a pair of micropipettes to generate micro-network strain. They also investigated the effect of bacterial collagenase on unstrained collagen gels. They used DIC microscopy in order to investigate the loss of single collagen molecule as a result of degradation even when the fibril diameters become very small (Figure 1). They measured the time required for strained and unstrained fibrils exposed to bacterial collagenase to degrade and found that the unstrained collagen fibrils were degraded significantly faster. In addition to investigating the mechanochemistry of collagen/bacterial collagenase, they performed experiments using MMP-8 as the catabolic enzyme. They found that mechanical strain causes a decrease in the degradation of collagen fibril by a physiologically relevant collagenase[22].

In 2010 Flynn et al. investigated the effect of mechanical strain on the degradation rate of reconstituted collagen fibrils by MMP-8. They strained reconstituted type I collagen micronetworks between micropipettes while collagen was exposed to active MMP-8. Relative degradation rates for loaded and unloaded fibrils were tracked using DIC imaging. They found that mechanical loading significantly increased the degradation time of loaded fibrils. They discovered that strained, reconstituted collagen fibrils persist substantially longer in the presence of MMP-8 than paired, unstrained control fibrils. They concluded that increased resistance of the collagen under loading is the result of a strain-induced reduction in enzymatic cleavage rate[21].

In 2013 Flynn et al. examined the effect of tensile strain on the individual bovine collagen fibrils degradation by bacterial collagenase. They held collagen fibrils at three levels of tension. Fibrils held at zero-load failed rapidly and consistently (20 min), while fibrils at 1.8 pN/monomer failed more slowly (35-55 min), and fibrils at 23.9 pN/monomer did not exhibit detectable degradation. They stretched each fibril to either zero-load, low-load (2 pN/monomer), or high-load (24 pN/monomer) and exposed to Clostridium histolyticum bacterial collagenase. Fibril load was held constant (load-control). The enzymes were examined in 300s intervals to quantify enzymatic degradation rate via calculated reduction in fibril stiffness. Figure 1B and C demonstrate fibrils with no force applied to them failed during the initial 1200 s, low-load fibrils failed during 2100-3300 s, and high-load fibrils did not fail at 14400 s, suggesting that applying tensile strain decreases the collagen degradation rate by bacterial collagenase[18].
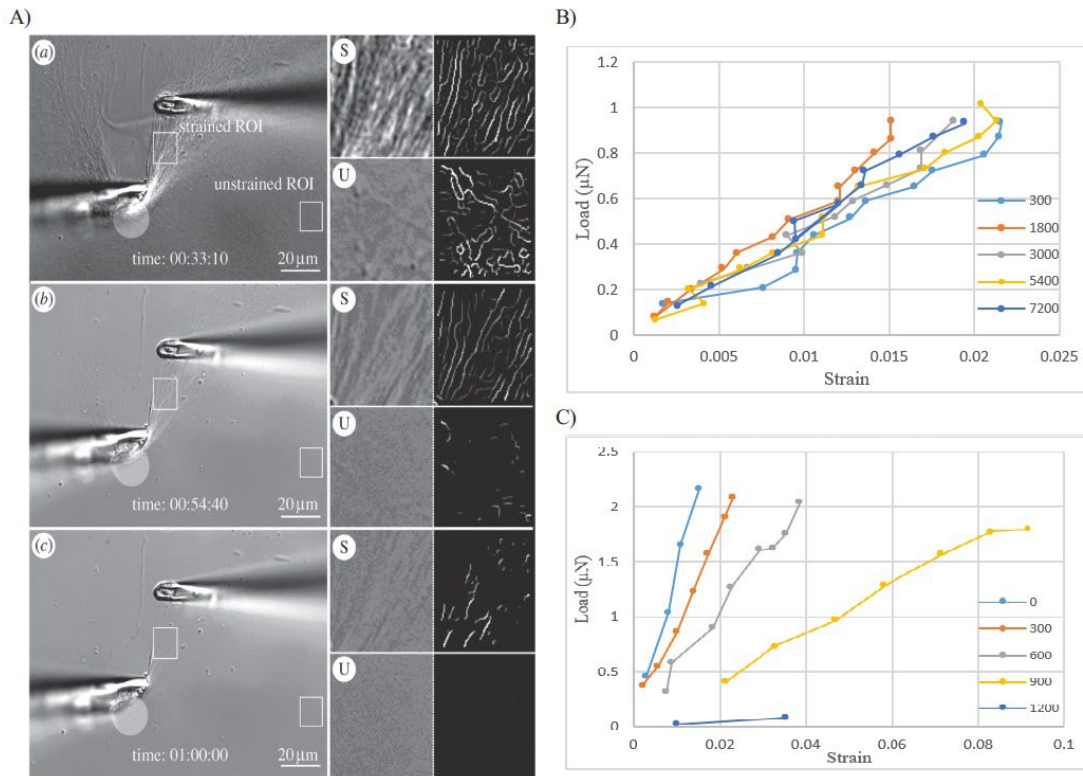
Figure 1: Optical tracking of collagen networks in vitro has demonstrated slower degradation rates of fibrils stretched between micropipettes compared to unstretched fibrils in the same image (A). This optical observation is consistent with changes in collagen fibril mechanical properties measured at different time points during protease-mediated degradation of fibrils either loaded or unloaded (B/C). Specifically, the slope of load vs. strain plots diminished only slightly in loaded fibrils (B), while this slope for unloaded fibrils diminished very quickly, indicating a greater loss in mechanical integrity (C).

2.3.3 Gel or Tissue Scale

Motivated by physiologic relevance and/or experimental feasibility, several groups have investigated the mechano-dependency of collagen degradation at the macro scale using reconstituted gels or explanted tissues. These studies have reported a variety of responses, including increased, decreased, and even a V-shaped degradation rate versus strain, suggesting that degradation rate has a minimum in a specific strain. One of the main reasons for these differences is the collagenase they used. It is believed that the difference

between bacterial collagenases and MMP lies in the specificity of the enzyme. Bacterial collagenase is capable of degrading both native and denatured collagen at multiple sites; however, mammalian collagenases such as MMP-1 are much more restrictive and only degrade native collagen at a single site. Moreover, the initial orientation and alignment of collagen molecules are important. Collagen molecules in the lung and pericardium are not initially aligned; however, corneal tissue consists of layers of aligned collagen. Furthermore, in a macro environment, other neighbor collagen molecules may limit the motion of the cleavage domain under mechanical strain, thereby protecting it from cleavage[63,64].

To our knowledge, the first report characterizing this behavior was by Huang and Yannas, who reconstituted bovine collagen into thin tapes and then loaded each tape to a range of strains between 1%–7% through a hook and suture connected to a load-cell[14]. They stretched a length of the collagen tape rapidly to a fixed extension and then monitored the relaxation of the force exerted by the tape. When the fibers were immersed in a solution containing bacterial collagenase, the force relaxed continuously until the specimen failed. They have found that up to the failure point, the force (F), could be represented by a single negative exponential term:

$$F = F_0 e^{-\frac{t}{\tau}} \qquad \text{(Eq. 2)}$$

Treating the tapes with bacterial collagenase led to matrix degradation as assessed by measuring the relaxation rate of the force required to maintain the collagen tape at a given strain. Interestingly the relationship between loading and degradation was biphasic, with a minimum degradation rate of around 4% strain and higher degradation at strains

below and above 4% (Figure 2A). The minimum degradation rate appears to occur at the strain level of 4%, where the suggested uncrimping of the fibers is completed and the slope of the stress-strain curve shows an abrupt rise.

One explanation for the decrease in degradation rate at low strain levels (1-4%) can be a drop in the enzyme flux rate into the substrate due to the uncrimping process. One possible mechanism that can explain the increase in degradation rate with strain at higher strain levels (4-7%) is collagen opening the new site of the enzymatic attack.

In agreement with Huang and Yannas, Nabeshima et al. observed strain-induced decreases in collagen degradation for the first time in explanted tissues[24]. Specifically, they subjected rabbit patella-patellar tendon-tibia units to bacterial collagenase with or without 4% strain, then measured tissue stiffness and maximum failure force. In the unstrained tissues, collagenase treatments induced substantial decreases in both stiffness and failure force, but these reductions were significantly inhibited in tendons subjected to mechanical strain.

Other investigators have also measured strain-induced protection of collagen to proteases across a variety of tissue samples, including the Ruberti group, who found that uniaxially loading bovine corneal tissues while treating with bacterial collagenase resulted in degradation of collagen fibers perpendicular to the uniaxial load direction (i.e., unstrained fibers) but much less degradation of collagen fibers parallel to the load direction (i.e., strained fibers)[13].

Wyatt and colleagues degraded rat tail tendon fibers with bacterial collagenase while subjecting the tendon to variable strains between 1%–10%[65]. The stress-relaxation

behavior was measured while holding each fiber at a given strain level, and the relaxation rate was used as an estimate of degradation. Interestingly, they found no effect of loading on degradation at low strains, then a precipitous drop in degradation around 3% strain that held constant across higher strains.

Ghazanfari et al. subjected decellularized porcine pericardial samples to uniaxial or biaxial strains in combination with bacterial collagenase, while monitoring the stress-relaxation response during degradation[20]. Across a strain range of 0–40%, degradation was minimum (in fact, near-zero) around 20% strain, with increasing degradation rates at lower and higher strain levels (Figure 2B). Further, the group imaged the resulting matrix structure after degradation and found that the tissue matrix (though initially unaligned) could become highly aligned after stretching + degradation. This study highlighted two significant findings: (1) the biphasic, V-shaped relationship between degradation and strain strongly supported the original findings of Huang and Yannas from 40 years prior, and (2) the results supported Ruberti's previous findings that strain-dependent degradation of fibers can alter matrix alignment in tissues by preferentially degrading fibers of a particular orientation relative to the principal strain directions.
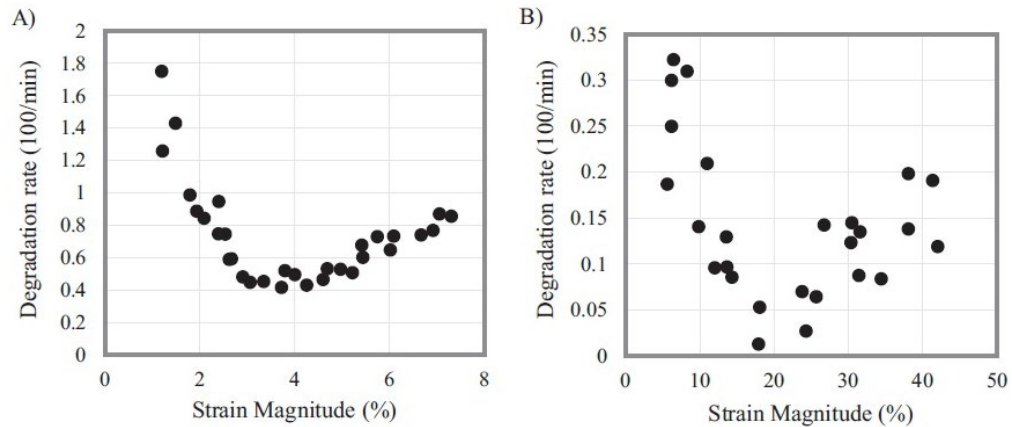
Figure 2: Protease-mediated degradation of reconstituted collagen tapes (A) and decellularized pericardial tissue (B) is reduced in the presence of mechanical strain. Interestingly, this strain-protection of collagen seems to follow a biphasic relationship with minimal degradation at an intermediate strain level

In partial agreement with Ghazanfari et al., Yi et al. subjected mouse lung tissue strips to static uniaxial strains of 0%, 20%, 40%, or 80% while treating with bacterial collagenase[25]. Stress-strain curves from before and after degradation treatments indicated the highest stiffness (i.e., presumably lowest degradation) in tissues subjected to 20% strain, with slightly lower stiffness in the unloaded group. However, the 40% and 80% strain groups indicated much lower stiffnesses (i.e., higher degradation) than the unloaded group.

Ellsmere et al. also measured increased degradation of intact tissue in the presence of mechanical load[15]. They subjected bovine pericardium strips to a constant uniaxial force until 30% tissue extension was reached (defined as the tissue failure point). In the presence of bacterial collagenase, they found that increasing the static load on tissue samples decreased their time to failure, that is, increased the degradation rate.

In sum, studies of the mechano-regulation of collagen degradation at the macro tissue scale have highlighted a nuanced relationship between loading and degradation.

There are several factors to consider when investigating this relationship at the tissue scale. First, it seems that different levels of strain can produce highly variable effects on degradation, which is possibly related to the transmission of loads from the macro sample (where most strains/loads are applied and monitored) down to the micro individual molecule (where protease binding and cleavage occur). Since collagen structures are organized via hierarchical packing and crimping, the global strain will engage individual molecules nonlinearly (e.g., low strains may simply uncrimp fibers without actually straining molecules). It is also important to note that whole-tissue preparations will include not just collagen but a host of other matrix proteins, glycoproteins, proteoglycans, and so on that can have significant effects on collagen organization, load transmission, and resulting conformational changes of individual molecules, and the varied results in mechano-dependent degradation seen across studies of individual collagen molecules. Lastly, the mechanical loading of tissues can potentially alter matrix porosity and the resulting flux of proteases, which could clearly confound degradation measurement results. It is important to note these limitations and continue investigating the mechano-regulation of degradation across all collagen length scales.

2.4 Systems Modeling of collagen-protease interactions

Collagen turnover depends on a highly regulated balance between collagens, MMPs, and TIMPs. The diversity of matrix-MMP-TIMP interactions makes it very difficult to intuitively predict the effects of any individual matrix-, MMP-, or TIMP-targeting therapy. A computational model of collagen-MMP-TIMP could help us to better

understand these interactions, enable us to systematically predict dynamic matrix turnover under diverse matrix, MMP, and TIMP levels, and eventually help us to use the model as a screening tool across patients with potential heart diseases. Very few computational models have been previously published for matrix-MMP-TIMP interactions, and all of them are focused on a small number of MMPs and TIMPs.

Karaggiannis and Popel's computational model was limited to collagen I, MMP-2, MMP-14, and TIMP-2. They included all known interactions between collagen I and two MMPs and one TIMP and represented network reactions as mass-action ordinary differential equations (ODEs) based on Michaelis-Menten kinetics in their model and previously reported reaction rates for rate parameters. Their model could quantitatively describe the activation of the MMP-2 proenzyme (pro-MMP2), the ectodomain shedding of MT1-MMP, and the collagenolysis caused by both enzymes. Their results indicated that pro-MMP-2 activation reaches its maximum at intermediate inhibitor levels and is suppressed at high TIMP levels. They also introduced and described quantitively the proteolytic synergism of MMP-2 and MT1-MMP[66].

Karaggiannis and Popel continued their work with the same computational model limited to collagen I, MMP-2, MT1-MMP, and TIMP-2 to investigate the effect of proteolytic potential on endothelial cell migration. The reaction network describes the interactions of the proteins, assuming that ability of the cell to carry out proteolysis is the rate-limiting step of the migration. They showed that at high collagen content, proteolysis was carried out primarily by MT1-MMP, whereas at lower concentrations, MT1-MMP and

MMP-2 worked together in the proteolysis process. They indicate that TIMP-2 is a regulator of the proteolysis process[67].

Vempati, Karaggiannis, and Popel used a computational model to quantify the MMP-9 activation and inhibition rates. They determine kinetic rate constants for MMP-9 activation by MMP-3, MMP-10, MMP-13, and trypsin; inhibition by the TIMP-1 and TIMP-2; and MMP-9 deactivation. They validated their model using existing biochemical experimental data and reported that inhibition due to a single binding step could not describe MMP-9 inhibition by TIMP-1. They also theoretically characterized the MMP-3/TIMP-2/pro-MMP-9 and MMP-3/TIMP-1/pro-MMP-9 systems revealing that these systems differ significantly in their time scales of activation and inhibition such that MMP-9 is able to temporarily overshoot its final equilibrium value in the latter[68].

These studies highlighted the ability of computational modeling to elucidate the possible interactions between substrate, MMPs and TIMPs which can help us understand the kinetics of these interactions. These findings also show that there are possible unknown interactions between MMPs and TIMPs which can make complex molecules that can be investigated using computational modeling.

Barry and Platt used a computational model and introduced cathepsin cannibalism, a novel mechanism by which cathepsins degrade each other in addition to the substrate. They investigated the proteolytic activity of cathepsin S and cathepsin K. They indicated that a reduction in total hydrolysis of elastin and type I collagen happened compared with computationally predicted values derived from individual cathepsin assays. Furthermore, they showed that cathepsin K activity was prevented, and collagen was preserved from

degradation by cathepsin K when a 10-fold ratio of cathepsin S cocultured with highly collagenolytic cathepsin K. Computational modeling helped them better understand combined proteolytic activities of cathepsins toward substrates and each other[57].

The Platt group expanded their work on the cathepsin proteolytic network and substrate degradation by developing a mechanistic model consist of a system of ODEs characterizing the cathepsin K, L, and S proteolytic network with elastin and gelatin as the substrates. They characterized the kinetic rates of individual cathepsins on both elastin and gelatin, including cathepsin-on-cathepsin binding and catalytic activity, and obtained interaction rates for each pair of cathepsin K, L, and S. At last, they integrated all three cathepsin interactions and substrate degradation activities. Their simulations indicated improved predictions of substrate degradation in a multiple protease network after including reaction terms of autodigestion, inactivation, cannibalism, and distraction. With this network model, they simulated the effects of changes to this proteolytic network for additional substrates[56].

2.5 Introduction to dissertation aims

The following chapters of this dissertation will focus on mechano-chemo interactions across a collagen-MMP-TIMP network through experimental and computational studies. In chapter 3, we tested the effect of tensile loading on collagenous tissue degradation by proteases. We picked four proteases and quantified the role of mechanical loading on the degradation of collagenous tissue by each protease. As matrix degradation leads to decaying force levels, the sample degradation rate was quantified for

different strain levels for each protease. In chapter 4, we presented a detailed computational model of a biochemical network of collagen I proteolysis capturing all interactions of type I collagen, four MMPs, and three TIMPs in a cell-free, well-stirred environment. We monitored the proteolytic activity of MMPs and inhibitory activity of TIMPs and then used the results from experimental data to fit five different hypothetical reaction topologies and determined kinetic rate constants for collagen degradation by MMPs, MMP inhibition by TIMPs, MMP and TIMP inactivation, MMP cannibalism, and MMP and TIMP distraction. We also used post-MI time courses of collagen, MMP, and TIMP levels in animal experiments from the literature to perform a parameter sensitivity analysis across the model reaction rates to identify which molecules or interactions are the important regulators of ECM post-MI for both early and late time-periods. In chapter 5, we developed an ensemble classification algorithm for diagnosing HF patients with preserved ejection fraction (HFpEF) within a population of 459 individuals, including HFpEF patients and referent control patients. We concluded that machine learning algorithms could substantially improve the predictive value of circulating plasma biomarkers. Additionally, we built a mechanistic model to predict ECM component degradation using a genetic algorithm to connect ECM remodeling to the plasma biomarkers to help us with HFpEF patients' classification. Our findings demonstrate that machine learning-based classification algorithms show promise as a non-invasive diagnostic tool for HFpEF patients' classification, while also suggesting priority biomarkers for future mechanistic studies to elucidate more specific regulatory roles.

CHAPTER 3

AIM 1: MECHANICAL STRAIN MODULATES EXTRACELLULAR MATRIX
DEGRADATION KINETICS AND BYPRODUCTS IN ISOFORM-SPECIFIC
MANNER

3.1 Introduction

The extracellular matrix (ECM) forms connective tissue around cells and serves as

a major regulator of tissue structure and function across development and disease[41–43].

ECM is comprised of many underlying molecules, most notably collagen fibrils and

proteoglycans that are present in all connective tissues in various isoforms and

assemblies[44]. Collagen is the most abundant protein in the human body and comprises one-

third of the total protein[45]. There are more than 25 types of collagens, with the most

common being fibrillar types I and III, composed of three polypeptide subunits that exist

in a triple helix form[45,46]. The form and structure of ECM depend mainly on collagen, and

stiff collagen fibrils self-assemble into large-scale structures such as fibers and sheets[46].

The primary role of fibril forming collagens (I, II, III, V, and XI) is bearing and transmitting

mechanical loads along their main axis[7].

Collagen turnover is centrally involved in many diseases, including cardiac fibrosis,

pulmonary fibrosis, wound healing, cancer metastasis, and myocardial infarction (MI),

where collagen accumulates in the infarct zone to form collagenous scar tissue[18,31].

Progression to heart failure (HF) can occur in up to one-third of patients as a result of

adverse remodeling of the collagenous scar[31]. The degradation of collagen is mediated by

various proteases, primarily different isoforms of matrix metalloproteinases (MMPs) and cathepsins[18]. MMPs are zinc-dependent proteinases that can degrade the native collagen triple helix as well as other ECM components relevant to LV remodeling. The MMP family consists of many different members that can be categorized into different groups based on the substrate they prefer to degrade. The collagenases (MMP-1, MMP-8, and MMP-13) can cleave fibrillar-type collagens, and the gelatinases (MMP-2 and MMP-9) can degrade gelatins[21,41,51]. MMP-9 was first thought to be only gelatinase, but recent studies showed that it is also able to degrade full-length interstitial collagens[51]. Cathepsins are a superfamily of cysteine proteases, and some members can proteolyze ECM. Cathepsin K is the most potent mammalian collagenase, and it can cleave type I collagen in the native triple helix and the telopeptide regions[58,59].

Notably, previous studies have established that mechanical strain can modulate the enzymatic degradation of collagen fibers by altering the 'mechanochemistry' of collagen-protease binding[18,20] presumably due to altered molecular conformations of the collagen molecule's protease-binding sites (Table 1). However, some groups report increased collagen turnover with increased loading, while other groups report decreased collagen turnover with loading[4,5,13,15–23,25,69,70]. In this current study, we sought to test the hypothesis that different types of proteases might exhibit different sensitivities to mechanical loading. Specifically, we measured the degradation of porcine pericardium samples treated with MMP-1, MMP-8, MMP-9, or cathepsin K while subjected to different levels of mechanical loading. We also measured the proteomic signatures of degradation byproducts to assess

how mechanical loading alters the particular types of products released after the degradation of porcine pericardium samples by each protease.

3.2 Materials and Methods

In order to assess the effect of tensile loading on collagenous tissue degradation, we subjected porcine pericardium samples to different levels of equibiaxial tensile displacement with or without different proteases. Porcine pericardium samples were collected from fresh pig hearts at a local slaughterhouse. Pericardium samples were decellularized in standard decellularization protocol[71]. Upon receipt, tissue samples were placed in beakers of ddH2O with ice. Tissue samples were then placed in tube containers at 4° C for 24 hours for cell lysis. Samples were then placed in decellularization solution containing 50 mM Tris, 0.15% (v/v) Triton x-100, 0.25% deoxycholic acid-sodium salt, 0.1% EDTA and 0.02% sodium azide and gently shaken at room temperature for three days. The decellularization solution was changed with a new decellularization solution on day three, and samples were shaken for three more days. Samples were then rinsed two times with ddH2O for 10 minutes at room temperature, then two times with 70% ethanol for 10 minutes at room temperature, and finally another two times with ddH2O for 10 minutes at room temperature. Samples were then placed in 2x concentration of DNase/RNase solution containing DNase (stock = 3480 U/mg), RNase (stock = 97.1 U/mg), 5 mmol MgCl and DPBS in 37° C for 24 hours with shaking. Samples were then washed twice with ddH2O and kept sterile in FBS at 4°C.

Pericardium samples were subjected to different mechanical strains (from ~5-40%) and different proteases in order to identify the effect of strain on collagenous tissue degradation by each protease. Recombinant human proMMP-1, proMMP-8, proMMP-9, and cathepsin K solutions were purchased commercially (Enzo Life Sciences, Farmingdale, New York), and proMMPs were activated based on provided activation protocol. ProMMP-1 was activated by 8μl of 0.5 mg/ml trypsin added to 100 μl of proMMP-1 and incubated for 30 minutes at 37°C. ProMMP-8 was activated by 2μl of 0.5 mg/ml trypsin added to 100 μl of proMMP-1 and incubated for 30 minutes at 37°C. ProMMP-9 was activated by 20μl of 0.5 mg/ml TCPK-trypsin added to 100 μl of proMMP-1 and incubated for 30 minutes at 37°C. After activation, 300 μl of ddH2O was added to all MMP samples to reach the final volume of 400 μl and the final concentration of 2.5 μg/ml.

Thin, square pericardium samples (7 mm × 7 mm × 0.2 mm) were biaxially stretched using a commercially available planar biaxial testing system (BioTester; CellScale, Waterloo, Canada) for 30 seconds to different levels of static strain in PBS at 37°C (Figure 3A). Biaxial tensile displacement control was used in order to load all fibers within the tissue regardless of their orientation. A custom rig was designed to hold the protease solution to allow us to achieve the desired high protease concentration within a temperature-controlled bath. Samples were placed in the rig and kept under protease solution during the whole process (Figure 3B). After the loading step, samples were held at constant extension under displacement control for one hour to allow for viscoelastic stress relaxation, after which the PBS bath was replaced with either a protease solution or

33

protease-free control, and forces were monitored for two additional hours (Figure 3C). Matrix degradation led to decaying force levels, and sample stress degradation decay rate was quantified as the decay constant (b) of an exponential fit to the force versus time curve for every strain level for each protease as done previously[20]

$$F = F_0 e^{-bt} \qquad \text{(Eq. 3)}$$

An F-stat, non-linear, quadratic regression analysis of degradation decay rate on strain level was performed in order to test the accuracy of our predictions. The quadratic model contains constant, linear, and squared terms.

Degradation fragments were potentially released into the solution in the form of either large degradation products or small degradation products while the degradation process was happening. The solution containing either a protease solution or protease-free control was preserved after a 3-hour relaxation and degradation process. The solutions were then analyzed using silver stain analysis to confirm degradation, study the difference between protease treated samples and controls, and compare the level of degradation between proteases (Figure 4).

In addition, samples (i.e., solutions containing the degradation products) were prepared for mass spectrometry (MS) analysis first by dissolving the protein lysate in 8M Urea/1M NH4HCO3 buffer, followed by 1h reduction at 37°C with 120 mM Tris(2-carboxyethyl)phosphine (TCEP). Protein alkylation was then performed with 160 mM iodoacetamide (IAA) for 30 min at room temperature with shaking. The samples were then diluted 8-fold with water to reduce the urea content, pH was adjusted to 8.0, and trypsin was added at a ratio of 1:25. Digestion occurred at 37°C overnight with shaking. Trypsin

was deactivated by acidifying samples to pH <3.0 using formic acid. Samples were desalted and purified using 1cc C18 cartridge columns, and peptides were recovered in 0.1% formic acid. Samples were subjected to nano-LC-MS/MS analysis using an UltiMate 3000 RSLCnano system (ThermoFisher) coupled to a Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (ThermoFisher) via a nanoelectrospray ionization source. For each injection, 4µL (~1 µg) of the sample was first trapped on an Acclaim PepMap 100 20 mm × 0.075 mm trapping column (ThermoFisher Cat# 164,535; 5 µL/min at 98/2 v/v water/acetonitrile with 0.1% formic acid). Analytical separation was then performed over a 95 min gradient (flow rate of 250 nL/min) of 4–25% acetonitrile using a 2 µm EASY-Spray PepMap RSLC C18 75 µm × 250 mm column (ThermoFisher Cat# ES802A) with a column temperature of 45ºC. MS1 was performed at 70,000 resolution, with automatic gain control (AGC) target of 3×106 ions and a maximum injection time (IT) of 100 ms. MS2 spectra were collected by the data-dependent acquisition of the top 15 most abundant precursor ions with a charge greater than 1 per MS1 scan, with dynamic exclusion enabled for 20 s. Precursor ions isolation window was 1.5 m/z, and normalized collision energy was 27. MS2 scans were performed at 17,500 resolution, maximum IT of 50 ms, and AGC target of 1×105 ions. Proteome Discoverer 2.5 was used for raw data analysis, with default search parameters including oxidation (15.995 Da on M) as a variable modification and carbamidomethyl (57.021 Da on C) as a fixed modification. Data were searched against the NCBI Sus scrofa reference proteome (Taxonomy ID 9823). Peptide-spectrum matches were filtered to a 1% false discovery rate (FDR) and grouped into unique peptides while

maintaining a 1% FDR at the peptide level. Peptides were grouped into proteins using the rules of strict parsimony, and proteins were filtered to 1% FDR.
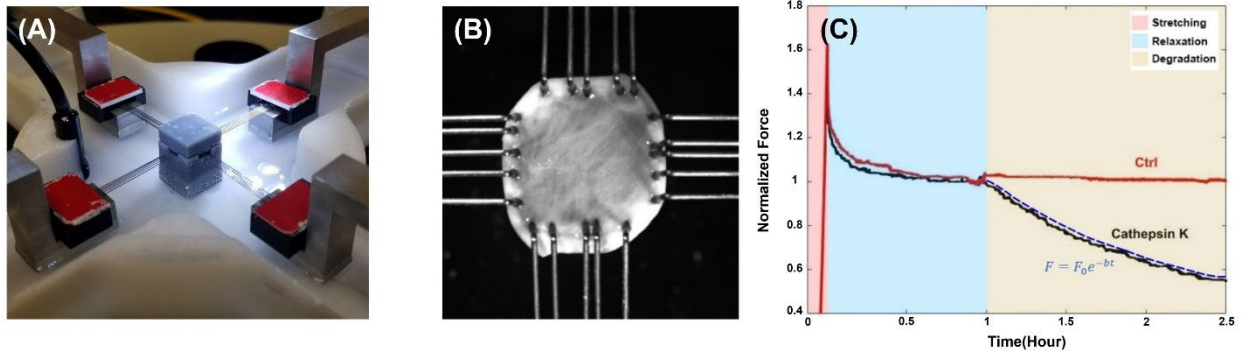
Figure 3: A) Decellularized pericardium stretched biaxially with CellScale and the custom rig used for holding high concentration protease solution within a temperature-controlled water bath. B) Decellularized pericardium sample under biaxial stretch. C) Example force-time curve of pericardium under constant extension subjected to protease or protease-free control.

3.3 Results

It has been shown in previous studies that tensile loading can both increase and decrease collagen degradation by bacterial collagenase, MMP-1, and MMP-8. We stretched collagenous samples to different levels of static strain and treated them with the same concentrations of MMP-1, MMP-8, MMP-9, cathepsin K or protease-free buffer control. After two hours, we analyzed the stress-strain data, found the force experienced by the samples, and observed that collagenous tissue degradation depends on protease type and level of mechanical strain. Across different strain levels, pericardium samples treated with protease-free buffer control maintained steady forces during the 2-hour treatment period, leading to near-zero degradation products by silver stain (Figure 4A, C) and force decay constants (Figure 5E). In contrast, samples treated with any of the four proteases

36

exhibited drops in force levels during the treatment period indicating substantial degradation of the matrix. Specifically, we found that strain can increase and decrease collagenous tissue degradation based on protease type and strain level (Figure 5).

Repeating the degradation tests across multiple strain levels demonstrated that the particular rate of degradation was affected by the sample strain (as previous studies have shown), but the specific influence of strain depended upon the type of protease. Specifically, degradation by cathepsin K was increased with increasing strains, and degradation by MMP-1, MMP-8, and MMP-9 first decreased and then increased with the strain magnitude forming a V-shaped curve. The stress degradation decay rate reaches a minimum at a strain level of approximately $\varepsilon_{min} = 20\%$ for MMP-1 and MMP-8 and approximately $\varepsilon_{min} = 25\%$ for MMP-9. Across nearly all strain levels, MMP-1 demonstrated the highest degree of force decay while the other three proteases swapped degradation rate rankings based on the particular level of strain. F-stat statistical analysis shows that the increase in degradation decay for cathepsin K and the V-shaped curve for MMP-1, MMP-8, and MMP-9 are statistically significant (p-value < 0.05)

Silver stain analysis revealed that the degradation products were released to the solution and the level of degradation was significantly higher for protease-treated samples compared to control samples. A one-way ANOVA test was performed on both large and small degradation products. Analysis of small degradation products shows that MMP-1 and MMP-9 treated samples produce significantly more of the smaller degradation products. Table 2 shows every protein detected in proteomic analysis in most of the samples in each protease category after removing all proteins detected in the control

samples. Proteomics analysis results show that MMP-9 releases more protein types into the solutions. These results reveal that most of the proteins released into the solution by each protease are unique to the protease type; however, there are some overlaps between the proteases (Figure 6B). Excitingly, the degraded protein signature released into the solution depends on strain magnitude. Figure 6C and D show that there are some proteins in each protease type that are only released into the solution in either low stretch or high stretch. To our knowledge, this finding has never been shown before with any kind of protease or tissue.
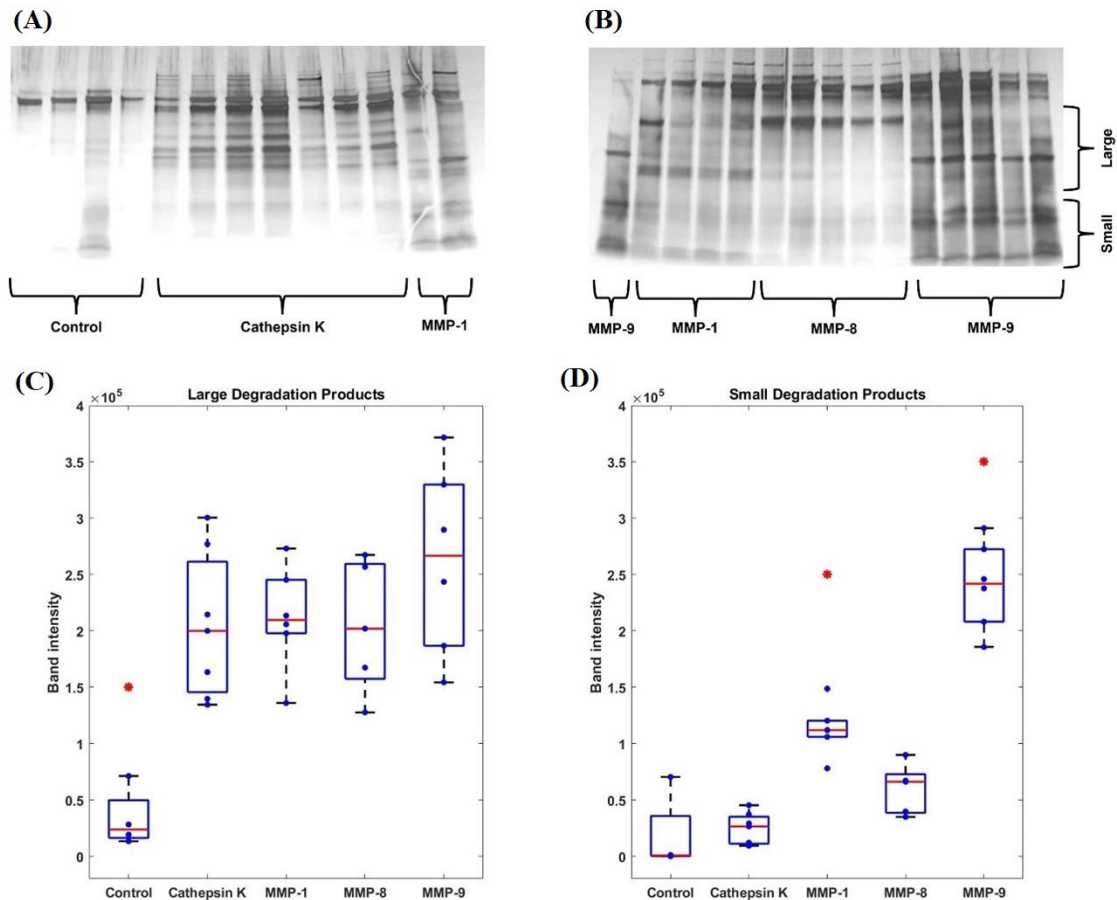
Figure 4: Silver stain analysis of control samples compared with protease treated samples. A and B) shows the different degradation products released to the solution by each protease sample and protease free control. C) Analysis of large

degradation products, all proteases degrade collagenous tissue more that protease free control. D) Small degradation products, MMP-9 is able to degrade collagen fragment into smaller parts.
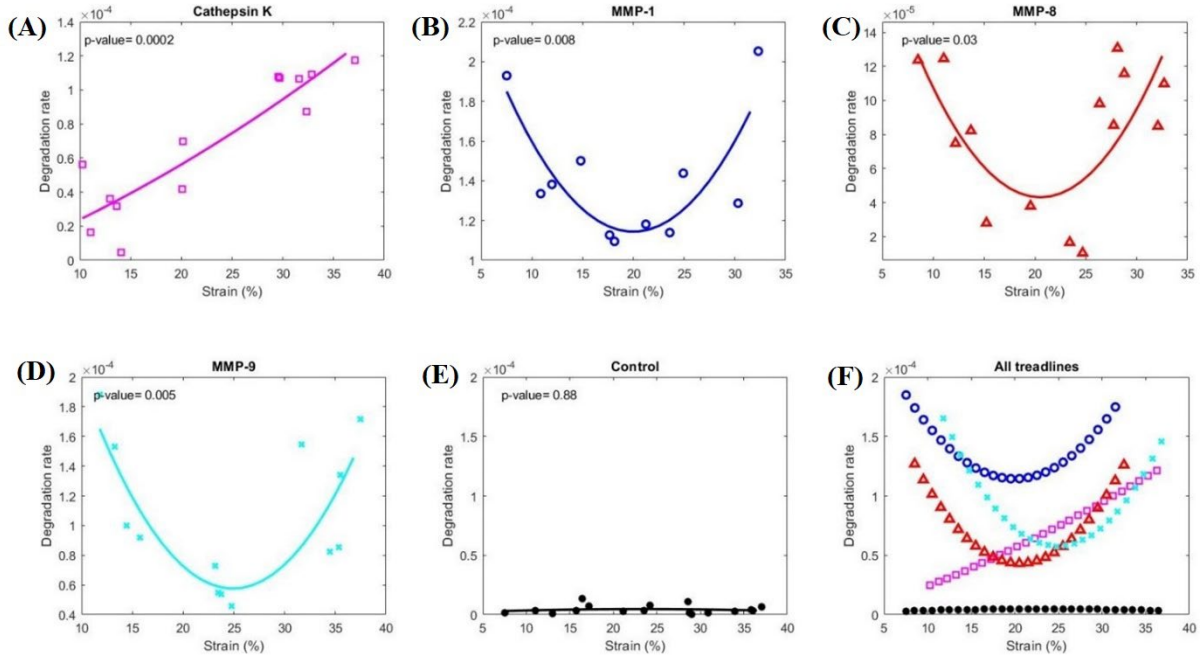
Figure 5: Stress degradation decay rates of biaxially loaded pericardium samples subjected to proteases across various strains. Stress degradation decay rate depends on sample strain as well as protease type. A) Degradation by cathepsin K was increased with increasing strains. B, C, and D) degradation by MMP-1, MMP-8 and MMP-9 first decreased and then increased with the strain magnitude forming a V-shaped curve. E) No degradation detected in protease free control samples. E) degradation decay regime for all proteases.

Table 2: Complete list of proteins detected in proteomic analysis after removing proteins detected in the control samples

CatK	MMP1		MMP8		MMP9			
C1R	ACAN	FBN2	ACAN	LAMA1	A1BG	DEFA1	IGHG4	PON3
COL1A2	ADAM23	FGFR4	ADAMTSL4	LYG1	A2M	DNAJC3	IL17B	PSMD7
COL5A3	ADNP	GDI2	AHNAK	NCAM1	ACTN4	DOCK2	IL4R	PZP
COL6A1	APOLD1	GPC1	ALDOC	NINL	ACTR10	DST	JUP	QSOX1
COL6A3	C1orf54	HDLBP	APOB	NIT2	ADAM28	EEF2	KERA	RAB11FIP3
DHRS7C	CD274	HSPA8	APOLD1	ODAD4	ADM2	EMILIN2	KIF20B	RELN
ERFE	CEP131	HYOU1	CD274	PDCD6IP	AFM	EPHB4	LAMA1	RETN
FMOD	CEP290	IL17B	CEP290	PIP	ARSG	F9	LAMA2	ROCK1
GDF9	CHADL	LAMB1	CHADL	PLG	CALML5	FAN1	LAMC3	SDCBP
H4C1	COL1A2	LTBP1	COL1A2	SELP	CAT	FAP	LIPH	SEMA3F
ICOS	COL8A1	LYG1	CYRIB	SEMA3F	CD200R1	FERMT3	LOX	SEMA6B
ITIH2	CYRIB	NINL	FBN2	SERPINB13	CD274	FGFR4	LRIG3	SERPINB13
ITIH3	DOCK2	NIT2	GALNT1	SERPINC1	CEP131	FIBIN	LRRTM2	SERPINC1
LAMA4	DST	NUDCD2	HSP90AB1	SLC12A1	CEP290	FRAS1	LYG1	SLC12A1
SPATA20	F10	HSP90AB1	IL17B	WNT10A	CES4A	GALNT1	MAPT	SORL1
WNT6	FAN1	PDCD6IP	KIF20B		CFAP43	GARS1	MVP	SPINT1
	PKD1	SPAG17			CFI	GDI2	NCAPH2	SST
	PLG	UNC13D			CHADL	GSDMD	NINL	TSPEAR
	QSOX1				CLCA3P	HMGN2	NIT2	UNC13D
	RBP3	SEMA6D			COL28A1	HSP90AA1	ODAD4	VWF
	REN	SERPINB13			COL4A3	HSP90AB1	PLA2G3	WNT10A
	SLC12A1	SERPINB6			CYRIB	HUWE1	PLAA	PON3
	TUBB4B	WNT10A						
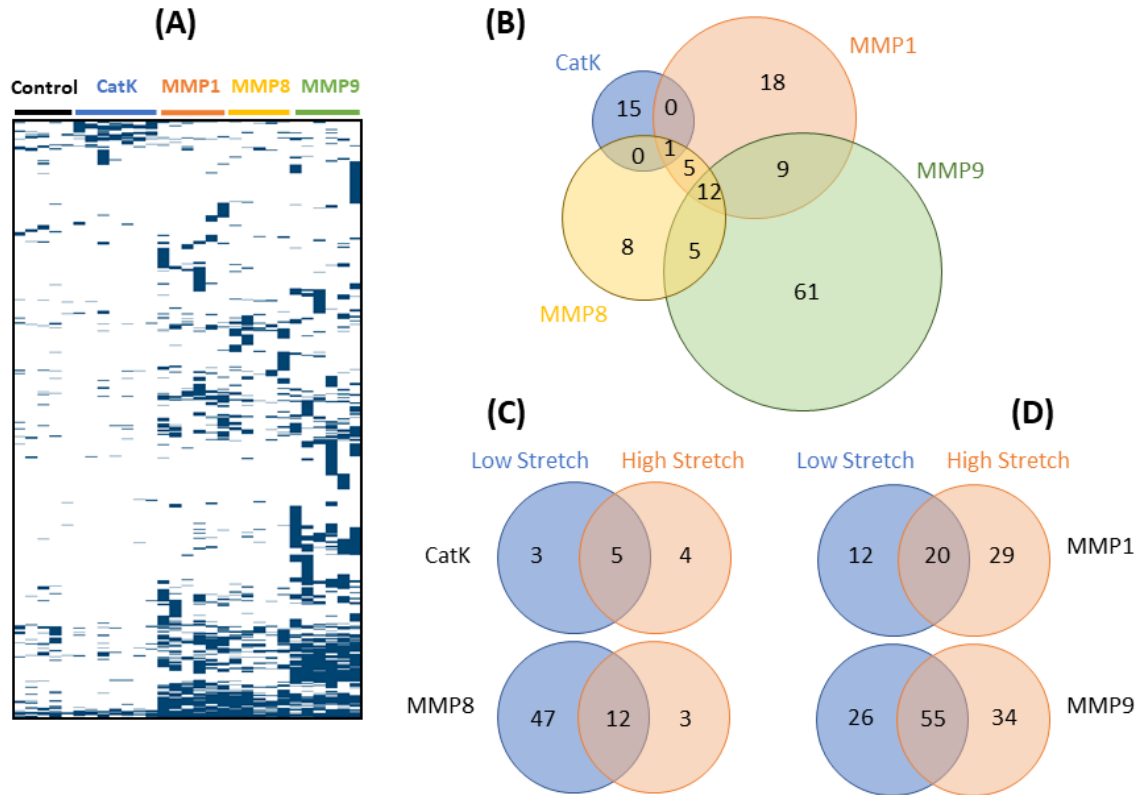
Figure 6: Proteomic results for every protease. A) Heat map of all proteins released to the solution in each protease B) MMP-9 release more variation of protein types into the solutions. Most of the proteins released into the solution by each protease is unique to the protease type, however there are some overlaps between the proteases. C and D) Proteins released to the solution depends on strain rate. Some proteins in each protease type are only released into the solution in either low stretch or high stretch magnitude.

3.4 Discussion

In the current study, we aimed to study the effect of strain on collagenous tissue degradation by different proteases by applying biaxial displacement to square pericardium samples in the presence of protease. Our results demonstrate that the degree of mechanical strain can significantly alter protease-mediated stress degradation decay rates of collagenous tissues like pericardium, but the particular effect depends on the particular type of protease. We showed that cathepsin K-mediated degradation was enhanced with

increasing strain, and MMP-1, MMP-8, and MMP-9-mediated degradation was first decreased and then increased by strain, forming a V-shaped curve with a minimum at strain level of approximately $\varepsilon_{min} = 20 - 25\%$. These findings are consistent with past studies by Huang (1977) and Ghazanfari (2016), who found that collagen degradation by bacterial collagenase was influenced by strain and discovered bacterial collagenase-mediated degradation rate also formed a V-shaped curve[14,20]. Huang found that the lowest degradation rate in a reconstituted collagen occurred at 4% strain. However, Ghazanfari reported the minimum degradation value was around 20% strain in pericardial tissue. The collagen in the reconstituted construct is comprised of relatively straight fibers compared to the pericardium, and the shift in the curve might be a result of the larger strain needed to unfold the fibers. Overall, these results highlight that the relative contributions of different proteases to tissue turnover depend not only on protease concentrations but also on the local level of tissue's mechanical conditions. These results reveal that (1) high protease isoform expression doesn't mean that protease dominates degradation – it depends on localized tissue mechanics (note this could be dynamically or spatially modulated during dynamic/spatial changes in strains like post-MI or tendon healing). (2) categorization of proteases as collagenase vs. gelatinase may be modulated by tissue mechanics.

Silver stain analysis confirms that the force decay is the result of the degradation of the load-bearing component of the porcine pericardium samples. It also shows that MMP-1 and MMP-9 are able to degrade the components of the solution to even smaller fragments. One-way ANOVA analysis on large degradation products confirms that all of the proteases

degrade collagenous tissue more than protease-free control. These findings align with previous findings suggesting that MMP-9 is able to degrade collagen fragments into smaller parts. One-way ANOVA analysis also confirms that MMP-9 degrades the collagenous tissue into smaller degradation products significantly more than other proteases[31]. Even further, proteomic analysis revealed distinctly different protein signatures of the degradation profiles generated by each protease type as well as the profiles generated between low strain vs. high strain tests. It is important to mention that change in ECM structure is critical in cellular response regulation and tissue remodeling. Degradation of ECM proteins generates various ECM proteins that interact with cell surface receptors and alters inflammatory, fibrogenic, angiogenic, and reparative cascades[72]. These ECM fragments modulate fibroblast function either directly (changing fibroblast signal transduction) or indirectly (changing fibroblast activation process), indicating their role in tissue remodeling and fibrosis[73].

Our approach comes with a few notable limitations. First, the collagen fibers in pericardium samples are sometimes aligned in variable directions. This could cause some variation in the stress experience in different tissue samples with the same strain. Using aligned, pure collagen gels can potentially eliminate these variations. Furthermore, we recognize that the pericardium is not only comprised of collagen I, and the cardiac protease is not solely MMP-1, 8, and 9. We selected this subset of enzymes as representative of different types of MMPs, such as collagenase and gelatinase, and the goal of Aim 1 was to investigate the behavior of different protease types on collagen degradation under mechanical loading. Expanding this study across more proteases and adding TIMPs,

different protease and TIMP cocktails, and a variety of different combinations of mechanical strain, including static and dynamic strains, can expand our knowledge of the effect of tensile loading on collagenous degradation by proteases. We also acknowledge that the stress concentration around the holes made by the tissue holders might cause some decay in the force calculated by the load cells. In spite of these limitations, our collective results highlight that therapeutic strategies for modulating collagen turnover might benefit from tailoring such strategies to the particular mechanical context in-vivo.

CHAPTER 4

AIM 2: BUILD AND ANALYSE A COMPUTATIONAL MODEL OF COLLAGEN
TURNOVER REGULATION BY MMPS AND TIMPS.

4.1 Introduction

Collagen is the most abundant structural protein in the human extracellular matrix

(ECM), comprising one-third of the total protein in the human body[45]. Matrix

metalloproteinases (MMPs) are zinc-dependent enzymes that participate in ECM

degradation, and they cleave all structural components of the ECM. There are different

groups of MMPs based on their in-vitro substrate preferences, including, for example, the

collagenases (MMP-1, -8, and -13) that cleave fibrillar type collagens at a highly conserved

site and the gelatinases (MMP-2 and -9) that degrade gelatins[5,41,74]. Tissue inhibitors of

metalloproteinases (TIMPs) are specific inhibitors of MMPs which bind to activated

MMPs with 1:1 stoichiometry and control the activity of MMPs in tissues to help refine

the balance between deposition and destruction of collagen[5,41,75].

Collagen turnover is centrally involved in many diseases, including tissue fibrosis,

cancer metastasis, wound healing, and myocardial infarction (MI)[3]. Collagen turnover

depends on a highly regulated balance between collagens, MMPs, and TIMPs, but the

diversity of matrix-MMP-TIMP interactions makes it very difficult to intuitively predict

the effects of any individual matrix-, MMP-, or TIMP-targeting therapy. A computational

model of collagen-MMP-TIMP could help us to better understand these interactions, enable us to systematically predict dynamic matrix turnover under diverse matrix, MMP, and TIMP levels, and eventually help us to use the model as a screening tool across patient-specific disease conditions[76–78].

Very few computational models have been previously published for matrix-MMP-TIMP interactions[66–68], and all of them are focused on a small number of MMPs and TIMPs. For example, Karaggiannis and Popel's computational model was limited to collagen I, MMP-2, MMP-14, and TIMP-2[66]. They included all known interactions between collagen I and two MMPs and one TIMP and represented each reaction as mass-action ordinary differential equations (ODEs) based on Michaelis-Menten kinetics in their model and previously reported reaction rates for rate parameters. Vempati, Karaggiannis, and Popel also used ODE to capture two MMPs and two TIMPs interactions[68]. These studies highlighted the ability of computational modeling to elucidate the possible interactions between substrate and MMPs and TIMPs and help us understand the kinetics of these interactions. These findings also show that there are possible unknown MMP-MMP and MMP-TIMP interactions which can make complex molecules that can be investigated using computational modeling[56].

We use a systematic approach to 1) characterize the kinetics of individual MMPs on the substrate and their substrate degradation potential; 2) characterize the inhibitory kinetics of individual TIMP on MMP, and 3) include all MMP and TIMP interactions in the system. With this network model, we predicted substrate degradation with four MMPs and three TIMPs.

4.2 Materials and Methods

In order to experimentally test the effect of different MMP-TIMP combinations on collagen degradation, we prepared a 96 well plate with a range of different mixtures of Dye Quenched (DQ), MMPs, and TIMPs (Figure 7). We used DQ collagen to monitor the proteolytic activity of different MMPs and MMP-TIMP mixtures and used a large-scale system of nonlinear ODEs capturing the interactions of collagen I with four MMPs and three TIMPs to record reaction rates between these species. The DQ collagen labeled with fluorescein so heavily that the fluorescence signal is quenched in intact DQ collagen. This substrate can be digested by most collagenases and gelatinases to yield highly fluorescent peptides. The increase in fluorescence signal is proportional to proteolytic activity and can be monitored using a fluorescence microplate reader. DQ collagen was purchased commercially (Thermo Fisher Scientific, Waltham, Massachusetts) and reconstituted to 1 mg/ml by adding 1 mL of deionized water (ddH2O) to the DQ substrate. The solution was then heated to 50°C and agitated in ultrasonic water bath for 5 minutes to facilitate dissolving.

Recombinant human proMMP-1, proMMP-2, proMMP-8, and proMMP-9, solutions were purchased commercially (Enzo Life Sciences, Farmingdale, New York). ProMMPs were activated using provided activation protocol. ProMMP-1 was activated by 8µl of 0.5 mg/ml trypsin added to 100 µl of proMMP-1 and incubated for 30 minutes at 37°C. ProMMP-2 was activated by 2mM APMA (final concentration) for two hours at 37°C. ProMMP-8 was activated by 2µl of 0.5 mg/ml trypsin added to 100 µl of proMMP-1 and incubated for 30 minutes at 37°C. ProMMP-9 was activated by 20µl of 0.5 mg/ml

TCPK-trypsin added to 100 µl of proMMP-9 and incubated for 30 minutes at 37°C. Recombinant protein TIMP-1, TIMP-2, and TIMP-4 in lyophilized form were purchased commercially (ProSci, Poway, California) and reconstituted in ddH2O. The starting concentration of DQ collagen was 25 µg/ml for all conditions and the final volume of collagen-MMP-TIMP in each sample was 200 µl (5 µg of DQ collagen in every sample). Two different MMP concentration was used $\frac{m_{MMP}}{m_{Collagen}} = \frac{1}{50}$ (0.1 µg MMP in each well) and $\frac{m_{MMP}}{m_{Collagen}} = \frac{1}{25}$ (0.2 µg MMP in each well). $\frac{m_{TIMP}}{m_{MMP}}$ was either 1/2 or 1/4 by keeping $m_{TIMP} = 0.05 \, \mu g$ in all samples.

The samples with lower concentrations of MMPs contain 0.1 µg of either single or each MMP in the mix and the samples with higher concentration of MMPs contain 0.2 µg of either single or each MMP in the mix. Each sample contains 0.05 µg of either single TIMP or each TIMP in the mix. We ran the experiment in a fluorescent microplate reader in 37°C letting the proteases, inhibitors and substrate interact with each other. MMPs degrade DQ collagen and release fluorescein attached to the collagen and the product (degraded collagen) can be measured using fluorescent microplate reader (BioTek, Synergy 4) with absorption maxima at 495 nm and fluorescence emission maxima at 515 nm.

We then used the results from experimental data to fit the ODE model and we used a genetic fitting algorithm to find an optimal (best-fit) combination of rate parameters, including MMP degradation rates and TIMP inhibition rates, etc[79,80]. Mass action kinetics was the basis for constructing our ODE models of MMP proteolytic and TIMP inhibition

activity. A system of ordinary differential equations (based on Michaelis–Menten kinetics) was used to mechanistically describe the MMP-substrate, MMP-TIMP, and MMP-MMP interactions.

M1 T1	M1 T1	M1, M2, M8 T1	M1, M2, M8 T1	M1, M2, M8 T1, T2	M1, M2, M8 T1, T2	M1, M2, M8 T1, T2, T3	M1, M2, M8, M9 T1
M1 T2	M1 T2	M1, M2, M8 T2	M1, M2, M8 T2	M1, M2, M8 T1, T4	M1, M2, M8 T1, T4	M1, M2, M9 T1, T2, T3	M1, M2, M8, M9 T2
M1 T4	M1 T4	M1, M2, M8 T4	M1, M2, M8 T4	M1, M2, M8 T2, T4	M1, M2, M8 T2, T4	M1, M8, M9 T1, T2, T3	M1, M2, M8, M9 T4
M2 T1	M2 T1	M1, M2, M9 T1	M1, M2, M9 T1	M1, M2, M9 T1, T2	M1, M2, M9 T1, T2	M2, M8, M9 T1, T2, T3	M1, M2, M8, M9 T1
M2 T2	M2 T2	M1, M2, M9 T2	M1, M2, M9 T2	M1, M2, M9 T1, T4	M1, M2, M9 T1, T4	M1, M2, M8 T1, T2, T3	M1, M2, M8, M9 T2
M2 T4	M2 T4	M1, M2, M9 T4	M1, M2, M9 T4	M1, M2, M9 T2, T4	M1, M2, M9 T2, T4	M1, M2, M9 T1, T2, T3	M1, M2, M8, M9 T4
M8 T1	M8 T1	M1, M8, M9 T1	M1, M8, M9 T1	M1, M8, M9 T1, T2	M1, M8, M9 T1, T2	M1, M8, M9 T1, T2, T3	M1, M2, M8, M9 T1, T2
M8 T2	M8 T2	M1, M8, M9 T2	M1, M8, M9 T2	M1, M8, M9 T1, T4	M1, M8, M9 T1, T4	M2, M8, M9 T1, T2, T3	M1, M2, M8, M9 T1, T4
M8 T4	M8 T4	M1, M8, M9 T4	M1, M8, M9 T4	M1, M8, M9 T2, T4	M1, M8, M9 T2, T4	M1, M2, M8, M9 T1, T2, T3	M1, M2, M8, M9 T2, T4
M9 T1	M9 T1	M2, M8, M9 T1	M2, M8, M9 T1	M2, M8, M9 T1, T2	M2, M8, M9 T1, T2	M1, M2, M8, M9 T1, T2, T3	M1, M2, M8, M9 T1, T2
M9 T2	M9 T2	M2, M8, M9 T2	M2, M8, M9 T2	M2, M8, M9 T1, T4	M2, M8, M9 T1, T4		M1, M2, M8, M9 T1, T4
M9 T4	M9 T4	M2, M8, M9 T4	M2, M8, M9 T4	M2, M8, M9 T2, T4	M2, M8, M9 T2, T4		M1, M2, M8, M9 T2, T4

Figure 7: 96 well plate layout showing all of different conditions and included MMPs and TIMPs. All wells contain 5 µg of DQ collagen. The yellow wells contain 0.1 µg of mentioned MMPs, and blue wells contain 0.2 µg of mention MMPs. Each well contain 0.05 µg of mentioned TIMPs.

Computational model ODE solutions of collagen-MMP-TIMP interactions were approximated using the built-in ode45 function in MATLAB (MathWorks). These models were constructed using mass action kinetics describing MMPs binding and hydrolysis of substrates and TIMPs inhibiting the MMPs. Our baseline model assumed one MMP binds and catalyzes collagen with associated k_{deg} rates (degradation) and one TIMP binds to one

MMP with associated k_{inh} rates (inhibition) (Figure 8). Along with this baseline model, we tested the predictive accuracies of four additional model reaction topologies based on various assumptions (Figure 8). The second model assumed one MMP binds to one substrate with associated k_{on} and k_{off} rates and then cleaves the collagen to form degraded collagen and free enzyme with k_{deg} rate. TIMPs binding to MMPs also have "on" and "off" reaction rates instead of only one binding reaction rate. In the third model, MMP and TIMP inactivation terms were added (k_{inact}). In the fourth model MMP cannibalism terms were added which means that MMPs can interact with the same MMP in active and inactive form. In the final model, inactive MMPs and inactive TIMPs can be distracted and react with other TIMPs and MMPs as previously demonstrated in protease network interactions[56]. Similar interaction terms between different MMPs were also added.

The initial model had 16 parameters and final model had 163 parameters, indicated in supplementary equations. The models were fit to experimental data using the genetic algorithm in MATLAB by minimizing the difference between experimental and model predicted degraded collagen formation. The objective function that the genetic algorithm tries to minimize by finding the best combination of rate parameters was defined by adding sum of squared errors of experiment and simulation for all conditions. Experimental data were collected every 10 minutes for 6 hours, and thus, the simulations were also performed for 6 hours with time steps of 10 minutes generating same number of data point as the experimental data. The sum of squared error for each condition was calculated by adding the square of the difference of experimental data points and simulation results for each data point.

The genetic algorithm consisted of 50 generations, each with a population size of 100 sets of parameters. The model tries to generate combination of parameters that minimizes the objective function. We used an ensemble method to minimize the possibility of the model being trapped in local minimum. The same genetic algorithm consisted of 50 generations, each with a population size of 100, was repeated 10 times that uses different initial value for parameters sets and tries to find the minimum objective function simultaneously. The top 20% of best-fitting parameter sets were used as "parents" for the next round of simulations; these sets were re-used in the next series of simulations; remaining parameter sets were generated by crossing the values contained in each set[81].

The initial constraints for the upper and lower bounds of kinetic parameters were set to 0 and 20 ($M \times hr^{-1}$ for bimolecular reaction coefficients; hr^{-1} for unimolecular reaction coefficients) in an effort to avoid constraining the algorithm too tightly. The initial concentration of DQ collagen and all MMPs and TIMPs are known. For each iteration of the genetic algorithm, random sets of rate parameters were generated and used to predict degraded collagen as the output using the given initial protein and enzyme levels for that condition. For each set of parameters, predicted changes in output levels for each condition and difference between experimental data points and model output for each data point was calculated. At the end of the simulations, we chose the optimal combination of rate parameters which minimized the error.
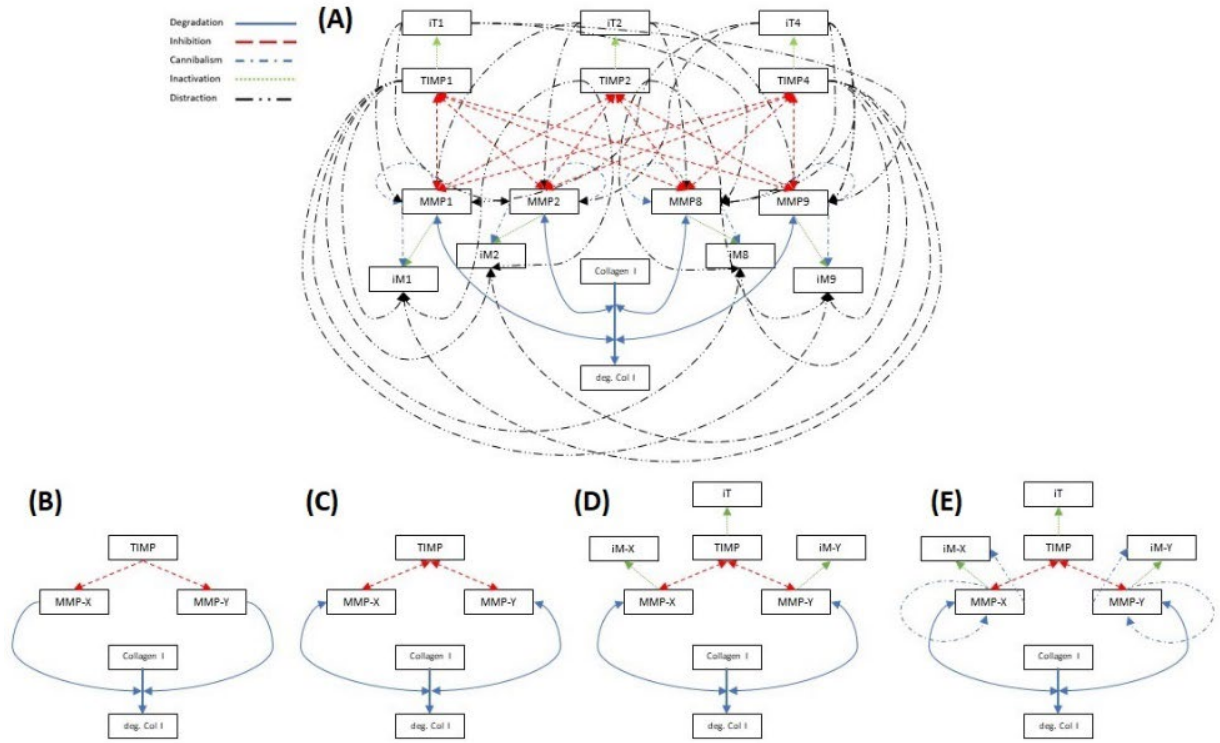
Figure 8: Schematic of all different model topologies showing all included reactions. Blue solid line shows degradation of collagen by MMPs. Red, dash line shows inhibition of MMPs by TIMPs. Blue, long dash dot line shows cannibalism of MMPs. Green dot line shows inactivation of MMPs and TIMPs. Black dash double dot line shows distraction of TIMP with inactive MMPs or distraction of MMPs by inactive TIMPs. A) The final model, containing MMP degradation, inhibition, and cannibalism, MMP and TIMP inactivation and MMP and TIMP distraction terms. B) Baseline model with one degradation and one inhibition term. C) Second model with "on" and "off" terms for MMP degradation and TIMP inhibition terms D) MMP and TIMP inactivation terms were added to the second model. E) MMP cannibalism terms were added. MMPs interact with the same MMP in active and inactive form.

A simplified ODE system based on the fifth model network is shown below. Equations 4-11 show a simplified interactions of collagen I with one MMP, one TIMP. The complete interaction can be found in the supplementary materials.

$$\frac{d(C1)}{dt} = (k_{prod_{C1}} \times PINP) - (k_{on_{M1.C1}} \times MMP1 \times C1) + (k_{off_{M1.C1}} \times M1.C1) \quad \text{(Eq. 4)}$$

$$\frac{d(dC1)}{dt} = (2 \times k_{deg_{M1.C1}} \times M1.C1) - (k_{on_{M1.dC1}} \times MMP1 \times dC1) + (k_{off_{M1.dC1}} \times M1.dC1) \quad \text{(Eq. 5)}$$

$$\frac{d(ddC1)}{dt} = (2 \times k_{deg_{M1.dC1}} \times M1.dC1) \qquad \text{(Eq. 6)}$$

$$\frac{d(MMP1)}{dt} = -(k_{on_{M1.C1}} \times MMP1 \times C1) + (k_{off_{M1.C1}} \times M1.C1) + (k_{deg_{M1.C1}} \times$$
$$M1.C1) - (k_{on_{M1.dC1}} \times MMP1 \times dC1) + (k_{off_{M1.dC1}} \times M1.dC1) + (k_{deg_{M1.dC1}} \times$$
$$M1.dC1) - (2 \times k_{on_{M1.M1}} \times MMP1 \times MMP1) + (2 \times k_{off_{M1.M1}} \times M1.M1) +$$
$$(k_{deg_{M1.M1}} \times M1.M1) - (k_{dist_{on_{M1}}} \times MMP1 \times iM1) + (k_{dist_{off_{M1}}} \times M1.iM1) +$$
$$(k_{dist_{deg_{M1}}} \times M1.iM1) - (k_{inh_{on_{M1.T1}}} \times MMP1 \times TIMP1) + (k_{inh_{off_{M1.T1}}} \times$$
$$M1.T1) + (k_{TIMP2_{T1}} \times M1.T1) - (k_{inactive_{M1}} \times MMP1) \qquad \text{(Eq. 7)}$$

$$\frac{d(dM1)}{dt} = (k_{deg_{M1.M1}} \times M1.M1) + (k_{dist\_deg_{M1}} \times M1.iM1) + (k_{off_{M1.C1}} \times M1.C1)$$

$$\text{(Eq. 8)}$$

$$\frac{d(iM1)}{dt} = (k_{inactive_{M1}} \times MMP1) - (k_{dist_{on_{M1}}} \times MMP1 \times iM1) + (k_{dist_{off_{M1}}} \times$$
$$M1.iM1) + (k_{TIMP1_{T1}} \times M1.T1) - (k_{dist_{on_{T1}}} \times TIMP1 \times iM1) + (k_{dist_{off_{T1}}} \times$$
$$T1.iM1) \qquad \text{(Eq. 9)}$$

$$\frac{d(TIMP1)}{dt} = -(k_{inh_{on_{M1.T1}}} \times MMP1 \times TIMP1) + (k_{inh_{off_{M1.T1}}} \times M1.T1) +$$
$$(k_{TIMP1_{T1}} \times M1.T1) - (k_{dist_{on_{T1}}} \times TIMP1 \times iM1) + (k_{dist_{off_{T1}}} \times T1.iM1) -$$
$$(k_{inactive_{T1}} \times TIMP1) \qquad \text{(Eq. 10)}$$

$$\frac{d(iT1)}{dt} = +(k_{TIMP2_{T1}} \times M1.T1) + (k_{inactive_{T1}} \times TIMP1) \qquad \text{(Eq.11)}$$

We also mined the literature to map time-courses of collagen, MMP, and TIMP levels in the heart measured post-MI in animal experiments[82,83](Figure 9). The concentration of collagen, MMPs and TIMPs varies at different time points after MI. Thus, we used these concentrations in each time point as initial condition and performed a parameter sensitivity analysis across the model reaction rates wherein we modulate each reaction rates in the network from 0.5-2x times their baseline values, one-at-a-time, and calculated the resulting changes in the primary simulation output which is degraded collagen concentration (e.g., increasing k_{on} for MMP-1 represents increasing MMP-1-collagen I binding and increasing k_{deg} represents increasing collagen I degradation by MMP-1).
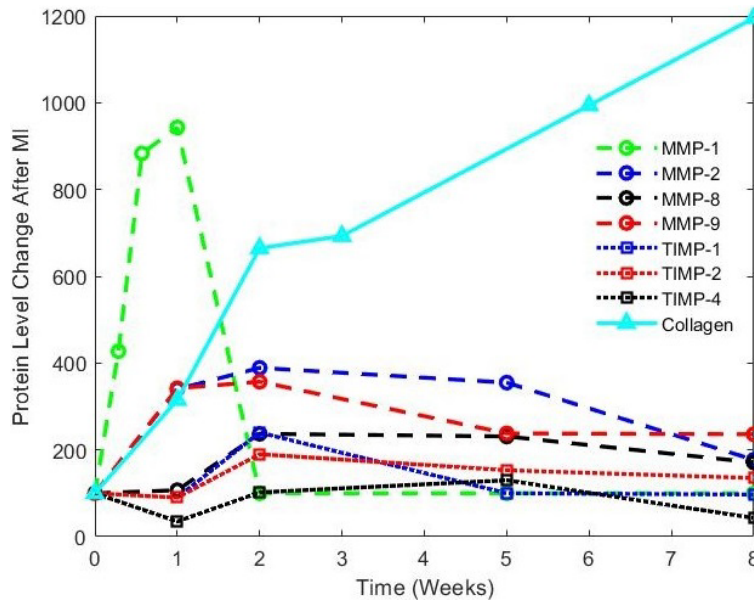
Figure 9: Levels of collagen, MMP, and TIMP in the heart measured following MI in mice shows that different MMPs are dominant in different time points.

This sensitivity analysis was repeated using inputs that match stimulation levels at 1-week, 2-weeks, 5-weeks, and 8-weeks post-MI in order to identify which species are the

important regulators of ECM post-MI for early and late time periods. We used an ensemble method to calculate parameter sensitivity. For each of the five models, we performed 10 different simulations resulting in 10 different sets of parameter rates. We changed each reaction rate in the network from 0.5 - 2x times their baseline values and fit a linear function to 9 different collagen remaining values (0.5x 0.8x 0.9x 0.95x 1 1.05x 1.1x 1.2x 2x). The mean of the slope of the 10 resulting lines for each kinetic rate parameter calculated from 10 different simulations for each model was reported as the sensitivity of the model to that specific parameter for each time point after MI. We performed the same method for calculating the sensitivity of the model to the initial conditions[84].

4.3. Results

We incubated four different MMP isoforms with two different $\frac{MMP}{Collagen}$ concentration ratios and three different TIMP types in different combinations with DQ collagen. We performed the experiments to record the degraded collagen concentrations at different time points for different concentrations and cocktails of MMP and TIMP. We then fit five different model topologies separately to the experimental data. Figure 10 shows the results of the final model with all possible interactions, fit to experimental data for single MMPs and different TIMPs at the same time in low concentrations and high concentrations. We used different simulations for low concentrations and high concentrations in an effort to accurately fit the model to the experimental data. These results show our model's capability to capture different regimes. Overall high concentration of MMPs degrades collagen faster and more efficiently. Surprisingly, MMP-2, which is a

gelatinase, shows more degradation than MMP-1 and MMP-8 (which are both collagenases), especially in low concentrations. We can also highlight that different TIMPs are dominant in each MMP solution. For example, TIMP-4 inhibits MMP-1 more effectively than other TIMPs, but TIMP-1 inhibits MMP-2 more effectively.

In order to find the best model describing the mass-action kinetics, we constructed five different models and compared their final errors with each other for low concentration and high concentration samples of single MMPs with different TIMPs. We took advantage of an ensemble method wherein each model has 10 different outputs since every simulation was repeated 10 times independently. Figure 11 shows the sum of the squared error of all conditions for five different topologies. This comparison indicates that adding "on" and "off" terms to the collagen degradation kinetics from the first model to the second model substantially improved the objective function. However, the most significant improvement is evident when comparing the second and third models. It is also revealed that although the most complex model, which captures all possible interactions between species, showed the best results (especially in low concentration cases), the third model containing only degradation, inhibition, and inactivation terms and without any cannibalism or distraction terms was able to minimize the sum of squared errors to a comparable level to model 4 and model 5. Additionally, our models generally performed better predicting kinetics of DQ collagen with a low concentration of MMPs.
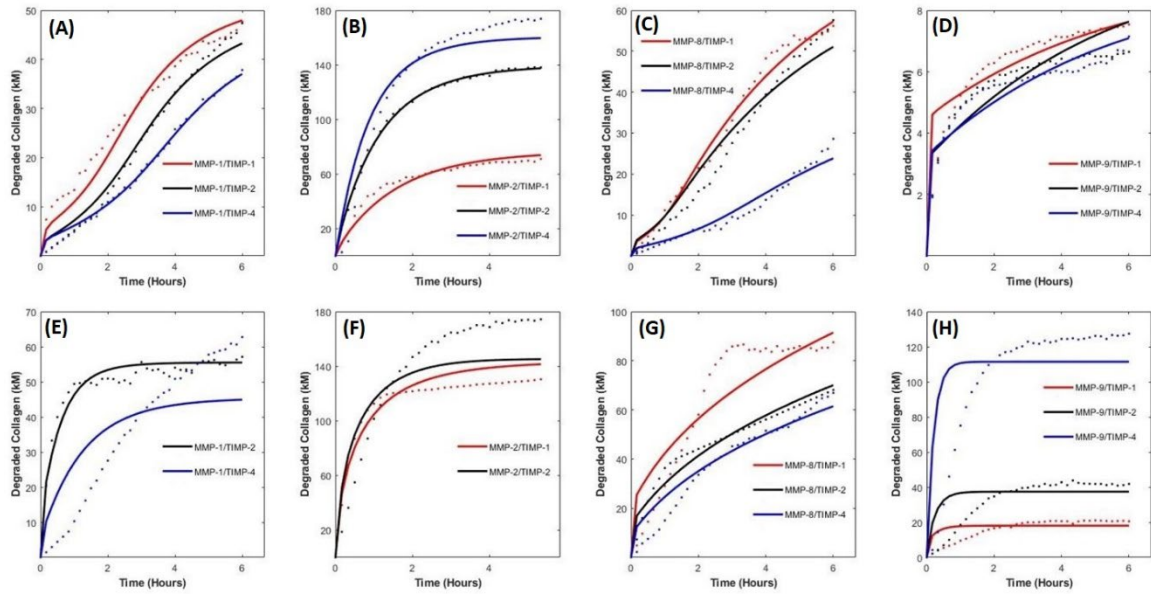
Figure 10: The final model results(solid lines) fitted to the experimental data (dots) with all possible interactions for single MMPs and different TIMPs fitted at the same time for low concentrations or high concentrations A) Low concentration MMP-1 and all TIMPs. B) Low concentration MMP-2 and all TIMPs. C) Low concentration MMP-8 and all TIMPs. D) Low concentration MMP-9 and all TIMPs. E) High concentration MMP-1 and all TIMPs. F) High concentration MMP-2 and all TIMPs. G) High concentration MMP-8 and all TIMPs. H) High concentration MMP-9 and all TIMPs.
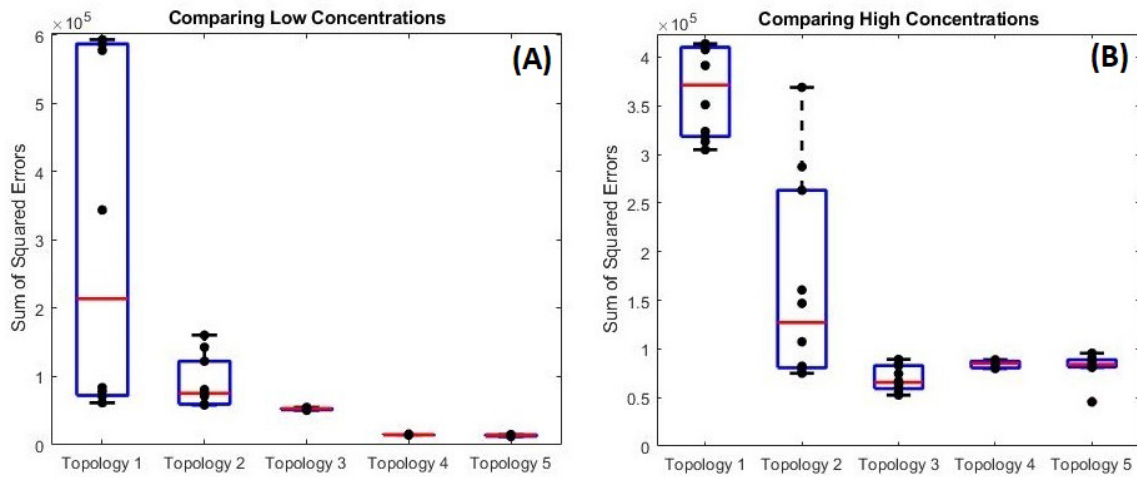
Figure 11: Simulation results (sum of squared error of all conditions) across five different topologies. Each dot is an objective function at the end of one of 10 ensemble simulation for each topology A) Low concentration of MMPs and B) High concentration of MMPs

To investigate the difference between the model performances across all samples, we tried to fit each model to all conditions at the same time, separately for low and high concentrations. Figure 12 shows the heatmaps of relative error ($\frac{Experimental - Simulation}{Experimental}$) for each condition across five models and indicates the predictive capability of each model. These results show how each model performed for any specific condition, which revealed general improvement in predictions with increasing model complexity.
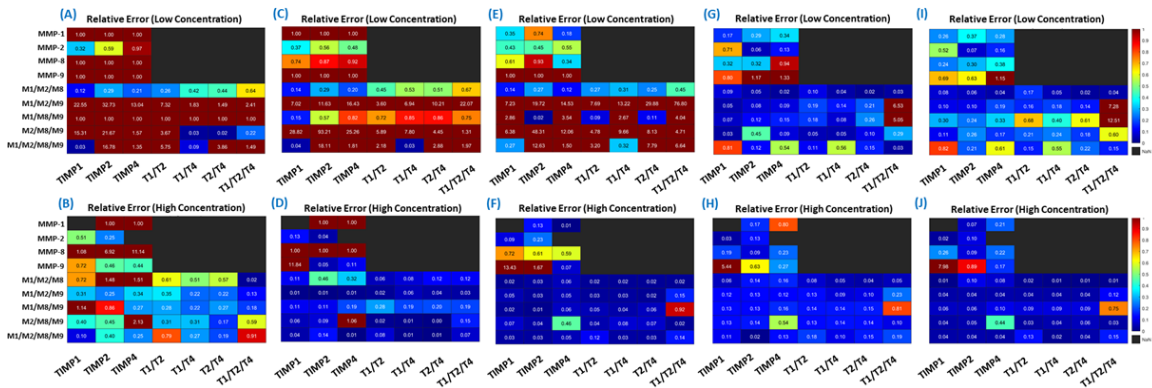
Figure 12: Heatmaps of relative error for each condition across five models' showing improvement from baseline model to the final model.

We also performed a parameter sensitivity analysis across the model reaction rates at 1-week, 2-weeks, 5-weeks, and 8-weeks post-MI in order to identify which species are the most important regulators of ECM post-MI for both early and late time periods. We modulated each reaction rate in the network from 0.5 - 2x times their baseline values, one at a time, ran the simulation for 6 hours, and calculated the resulting changes in collagen remaining concentration. The sensitivity was then reported as the slope of the line fitted to the collagen concentration result from each change in the corresponding parameter. This sensitivity analysis was repeated to identify the most important species of ECM post-MI for both early and late time periods and investigate which parameter change will change the primary simulation output more drastically. Generally, the results from sensitivity analysis showed that the concentration of collagen is more sensitive to the parameters at later time points. It is revealed that the collagen concentration is not greatly sensitive to TIMPs inhibition terms or TIMP inactivation terms. It is also important to mention that collagen concentration is more sensitive to MMP-2 degradation and inactivation terms in

58

every time point compared to other MMPs. Predictably, collagen concentration changes drastically when we modulated the initial values for each species, and it is most sensitive to MMP-2 and MMP-9 initial concentration after collagen concentration.
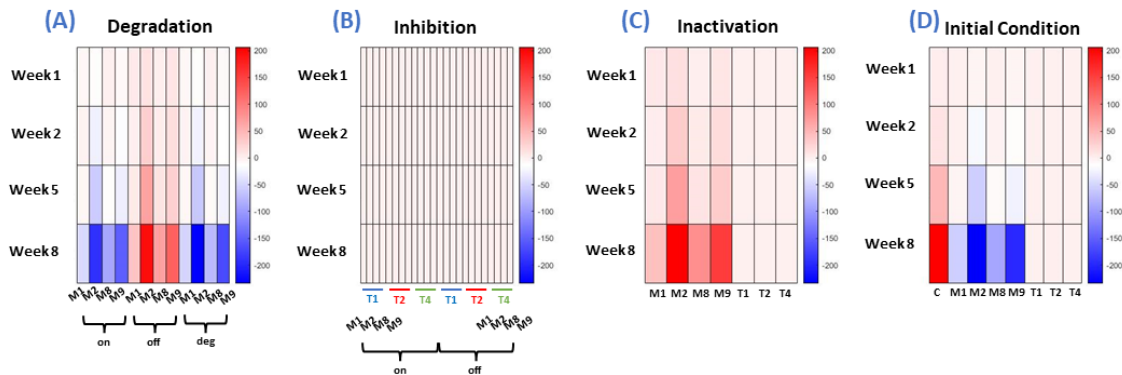
Figure 13: Sensitivity analysis for the rate parameters and initial conditions at 1-week, 2-weeks, 5-weeks and 8 weeks post-MI showing that the concentration of collagen is more sensitive to the parameters at later time points. A) Degradation terms B) Inhibition terms C) Inactivation terms D) Sensitivity of the model to the initial conditions.

4.4 Discussion

The network of collagen-MMP-TIMP interactions is complicated by multifaceted crosstalk and nonlinear relationships that make it unclear which molecule or reaction is the most important. We built a computational model that can integrate all multivariate interactions and capture the reaction parameters to investigate the kinetics of collagen-MMP-TIMP interactions. We fit the model to experimental data to better understand possible interactions and identify the key regulators of the system behavior. We were able to predict collagen degradation regime to a good extent for single MMP cases for both low and high concentrations. In-vitro experimental data helped us understand which MMP is

59

capable of degrading collagen faster, but only by comparing these models we were able to identify which MMP's kinetic terms are essential for better predicting collagen-MMP-TIMP interaction. Notably, MMP inactivation terms are more important than cannibalism and distraction terms which was impossible to predict without using the models.

These models can additionally be helpful when used for predicting patients' in-vivo data, especially after MI. Following MI, the concentration of ECM components, MMPs, and TIMPs change drastically, and the model can be helpful to predict collagen deposition to collagen degradation ratio as a mean to monitor LV remodeling. Predicting collagen concentration after MI can help physicians develop potential treatments specifically modulated for each patient or for optimizing drug time-course strategies.

It is important to note a few limitations of this work. First, the models performed well when used against single MMPs, but the performance decreased when used to predict all conditions simultaneously. We tried to capture all the interactions possible between species, but there is not enough evidence to prove that all of these interactions are happening. Also, there is not enough insight into the different possible interactions between species. For example, there is evidence showing that the activation process of MMPs is not a simple one-step process and happens in multiple steps forming complex molecules. We restricted the upper band and the lower band for each parameter in the simulations, but there is not enough evidence showing the relevance of the calculated rate parameters to physiological values since most of the kinetic parameters are still unknown in the literature. DQ collagen was used because it enabled us to quantitatively trace the degradation process; however, it would be beneficial to monitor collagen degradation within 3-dimensional

tissues. It is also possible that the fluorescein released from the DQ collagen to the solution interfere with the collagen molecule binding site and the MMP degradation process. Our experimental work can be expanded using more conditions and different concentrations of MMPs and TIMPs, potentially allowing the model to improve its performance.

CHAPTER 5

AIM 3: INTEGRATE ENSEMBLE MACHINE LEARNING AND MECHANISTIC
MODELING TO IDENTIFY HFPEF PATIENTS FROM MATRIX-RELATED
PLASMA BIOMARKERS.

5.1 Introduction

Heart failure (HF) is a chronic, progressive condition and an abnormality of cardiac function with the inability of the heart muscle to pump enough blood to meet the body's requirements for metabolism[27]. Diastolic failure or HF with preserved ejection fraction (HFpEF) is a condition resulting from the stiffness of heart muscle that causes the inability of the left ventricle (LV) to relax and fill normally. Currently, no treatment has yet been proven to reduce morbidity and mortality of HFpEF patients[29]. An estimated 6.2 million American adults ≥ 20 years of age had HF between 2013 and 2016, compared with an estimated 5.7 million between 2009 and 2012[2,26]. HF has various contributing pathologies, including hypertension (86 million Americans), myocardial infarction (MI, 800,000 Americans per year[2], and 300,000 recurrent infarction each year[7]), and fibrosis. Myocardial fibrosis contributes to left LV dysfunction and is histologically defined by excessive deposition of fibrous tissue relative to the mass of cardiomyocytes within the myocardial tissue. The predominant contributor to cardiac fibrosis is fibrillar collagen concentration (type I and III) which is regulated by changes in the relative ratio between collagen production and degradation[85–91,92]. Collagen progression toward remodeling can be tracked using certain byproducts. Specifically, the cleavage of the N-terminal propeptide of

procollagen type I and type III (PINP and PIIINP) produces fragments which is an indicator of collagen production, and the carboxyl-terminal telopeptide of collagen type I (CITP) which is a product of collagen degradation[93].

Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs), which can inhibit the activity of MMPs, are key regulators of collagen degradation processes. MMPs are zinc-dependent proteinases that can degrade the native collagen triple helix as well as other ECM components relevant to LV remodeling. Currently, the MMP family is composed of 25 proteinases that can be categorized into five groups based on the substrate they prefer to degrade. The MMPs focused herein are MMP-1 (collagenase), MMP-2 (gelatinase), MMP-3 (stromelysin), MMP-7 (stromelysin), MMP-8 (collagenase), and MMP-9 (gelatinase). The TIMP family currently includes four different members (TIMPs-1 to 4). MMPs in active and inactive forms can be inhibited by their physiological TIMPs[5,41,51].

Previous studies attempting the prediction or risk stratification for HFpEF have each generally focused on information from just a single domain - for example, either demographics/clinical history, imaging-based approaches, or the use of blood-based measurements (i.e., biomarkers). However, strategies that use a multidomain approach to identify key variables in a predictive model remain to be established. Accordingly, we sought to apply machine learning approaches, which provide a comprehensive and unbiased approach for evaluating variables from multiple domains in patients with patients with HFpEF in order to develop refined prediction models. Recent studies indicate that

machine learning techniques can be helpful in the classification of different pathologies and have shown predictive capabilities for many different data types and prediction tasks[94–98]. In our present study, we look to expand on previous work that investigates the use of remodeling-related biomarkers for the classification of diastolic HF.[99] To improve the predictive power of these biomarkers for classification, we utilized multiple advanced machine learning frameworks.

Additionally, we developed a mechanistic model for five different ECM component remodeling, including type I and type III collagen and three potential ECM candidates. A genetic fitting algorithm was then used to find the best-fit combination of rate parameters simultaneously in order to predict ECM components turnover for patient-specific data. The patient-specific results from the mechanistic model were then applied to all of the multiple machine learning models to improve the models' predictive capabilities.

5.2 Methods

Patient data were previously collected for 480 individuals as described by Zile and colleagues[99]. The original study recruited volunteers between 2004-2006 from health fairs, physician referral, and echocardiographic studies. All patients provided written informed consent and the research protocol was reviewed and approved by the institutional review board at the Medical University of South Carolina. Patients were excluded if they had evidence of a clinical condition that might modulate circulating plasma profiles including pulmonary disease, end stage renal disease, rheumatological disease, poorly controlled

diabetes, a recent major surgical procedure, MI, active infection, or other fibrotic and/or inflammatory conditions.

The patient data included basic information as well as mechanical biomarkers derived from echocardiogram data: sex, age, height, weight, body surface area (BSA), heart rate after a brisk, six minute hall walk, systolic blood pressure, diastolic blood pressure, pulse pressure, MMP-1, MMP-2, MMP-3, MMP-7, MMP-8, MMP-9, TIMP-1, TIMP-2, TIMP-3, TIMP-4, PINP, PIIINP, CITP, cardiotrophin (CT-1), N-terminal propeptide of brain natriuretic peptide (NT-proBNP), soluble receptor for advanced glycation end products (sRAGE), osteopontin, LV internal diameter end diastole (LVIDd), left ventricular internal diameter end systole (LVIDs), left ventricular end diastolic volume and index (EDV and EDVi), left ventricular end systolic volume and index (ESV and ESVi), stroke volume (SV), ejection fraction (EF), peak systolic stress (PSS), end systolic stress (ESS), and end diastolic stress (EDS). The list of biomarkers was chosen to include molecules that were (1) established to have some mechanistic link to tissue remodeling from previous studies, and (2) previously validated to be detectable in the patient plasma samples.

The patients were divided into two categories: referent control patients (n = 400) and patients with HFpEF (n = 59). The definitions for these categories were based on previous studies and Lahey Clinic and the HF and Echocardiography Associations of the European Society of Cardiology[99]. Briefly, HFpEF diagnosis required (1) clinical signs of

HF using Framingham Criteria, Boston Criteria, exercise testing, or quality of life survey, (2) preserved EF of >50%, (3) Normal LVEDVi <90 mL/m^2, and (4) evidence of diastolic LV dysfunction.
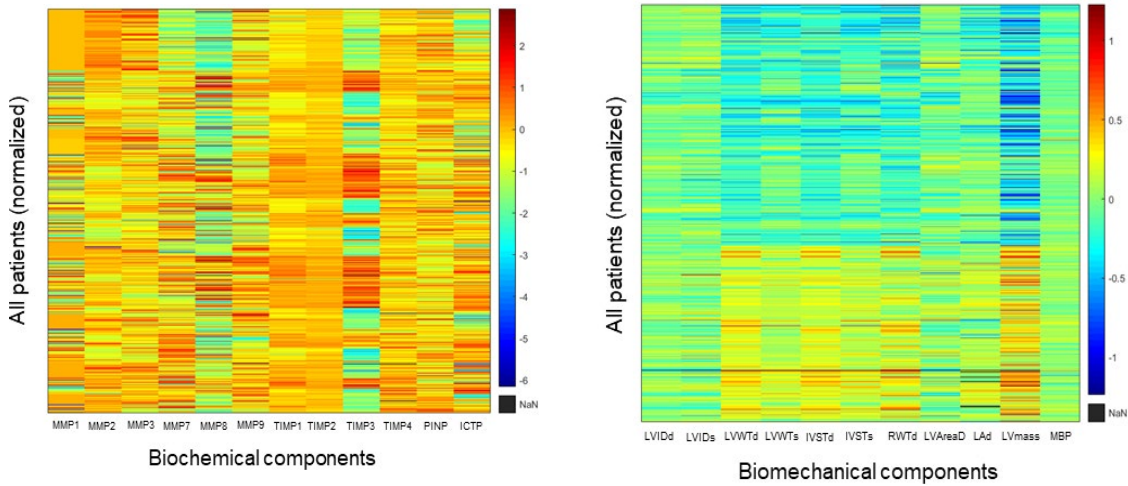
Figure 14: Biochemical and biomechanical component of all patients showing variability between features and patients

Across the full dataset of 38 features × 480 patients, 1.3% of data points were missing. We excluded patients that had 10 or more of mentioned 38 features missing which in result reduces patient number from 480 to 459. For the rest of the patients with missing data, we normalized the data using a z-score approach to normalize the data with respect to the sample means and standard deviations. We then found the five nearest-neighbors for each patient that had minimal mean squared errors across all features, and we imputed each individual missing value with the average of that variable across the five nearest-neighbors for each patient. Feature averages and variability for each group are shown in Table 3.

Table 3: Summary statistics for Controls vs HFpEF patients

Feature	Control		HFpEF	
	Average	Std. Dev.	Average	Std. Dev.
Patient Number	400		60	
Female (%)	65.0		61.0	
Age (y)	57.51	13.15	66.30	11.81
Height (cm)	168.02	11.41	168.67	9.55
Weight (kg)	82.44	21.08	92.95	24.07
Systolic BP (mm Hg)	130.02	12.54	139.63	18.37
Diastolic BP (mm Hg)	76.88	7.83	76.88	9.32
Pulse Pressure (mm Hg)	53.14	10.78	62.75	16.19
Heart Rate	68.24	10.26	68.80	10.71
MMP1 (ng/mL)	0.79	0.64	0.89	0.73
MMP2 (ng/mL)	340.09	144.50	418.32	170.64
MMP3 (ng/mL)	9.88	5.88	11.38	5.86
MMP7 (ng/mL)	1.66	1.19	2.13	1.38
MMP8 (ng/mL)	2.84	3.70	1.95	1.52
MMP9 (ng/mL)	110.13	88.17	125.08	67.60
TIMP1 (ng/mL)	74.64	24.61	84.44	24.89
TIMP2 (ng/mL)	79.59	14.63	81.83	12.95
TIMP3 (ng/mL)	8.03	8.54	5.62	6.67
TIMP4 (ng/mL)	1.47	0.64	1.86	0.76
PINP (ng/mL)	36.56	21.15	40.58	24.82
PIIINP (ng/mL)	7.33	1.94	9.14	3.19
CITP (ng/mL)	3.11	1.90	3.80	3.19
CT1 (ng/mL 10^{-3})	0.05	0.10	0.02	0.04
pro-NT BNP (pg/m)	87.89	93.47	210.67	253.91
sRAGE (ng/mL)	3.27	2.57	2.86	1.97
OSTEO (ng/mL)	78.25	42.41	93.15	42.15
BSA (m^2)	1.95	0.27	2.07	0.29
LVIDd (cm)	4.71	0.47	4.80	0.54
LVIDs (cm)	2.87	0.41	2.93	0.45
EDV (mL)	102.14	23.50	110.03	29.00
EDVi (mL/m^2)	52.70	11.02	53.41	13.46
ESV (mL)	32.75	11.26	35.02	13.11
ESVi (mL/m^2)	16.82	5.29	16.86	6.31
SV (mm Hg)	69.39	16.61	75.02	20.52
PSS (g/cm^2)	50.76	14.07	44.67	12.48
ESS (g/cm^2)	36.96	10.24	31.40	8.98
EDS (g/cm^2)	14.46	4.04	16.44	6.27
EF (%)	68	7	68	7

5.2.1 Mechanistic Model

We built a comprehensive mechanistic model in order to predict ECM remodeling using a large-scale system of nonlinear ODEs capturing the interactions of ECM components contributing to fibrosis, including type I collagen, type III collagen, three potential ECM candidates, procollagen type I and type III (PINP and PIIINP) and a network consisting of MMP-1, MMP-2, MMP-3, MMP-7, MMP-8, MMP-9, and their tissue inhibitors, TIMP-1, TIMP-2, TIMP-3, and TIMP-4. Figure 10 displays the simple mechanistic model used for one substrate molecule (for example, collagen type I) turnover, where solid blue lines show collagen degradation by MMPs. Red, dash lines show inhibition of MMPs by TIMPs. Blue, long dash-dot lines show cannibalism of MMPs. Green dot lines show the inactivation of MMPs and TIMPs. Black dash double dot lines show distraction of TIMP with inactive MMPs or distraction of MMPs by inactive TIMPs. Equations 4-11 show a simplified interactions of collagen I with one MMP, one TIMP. The complete interaction can be found in the supplementary materials.
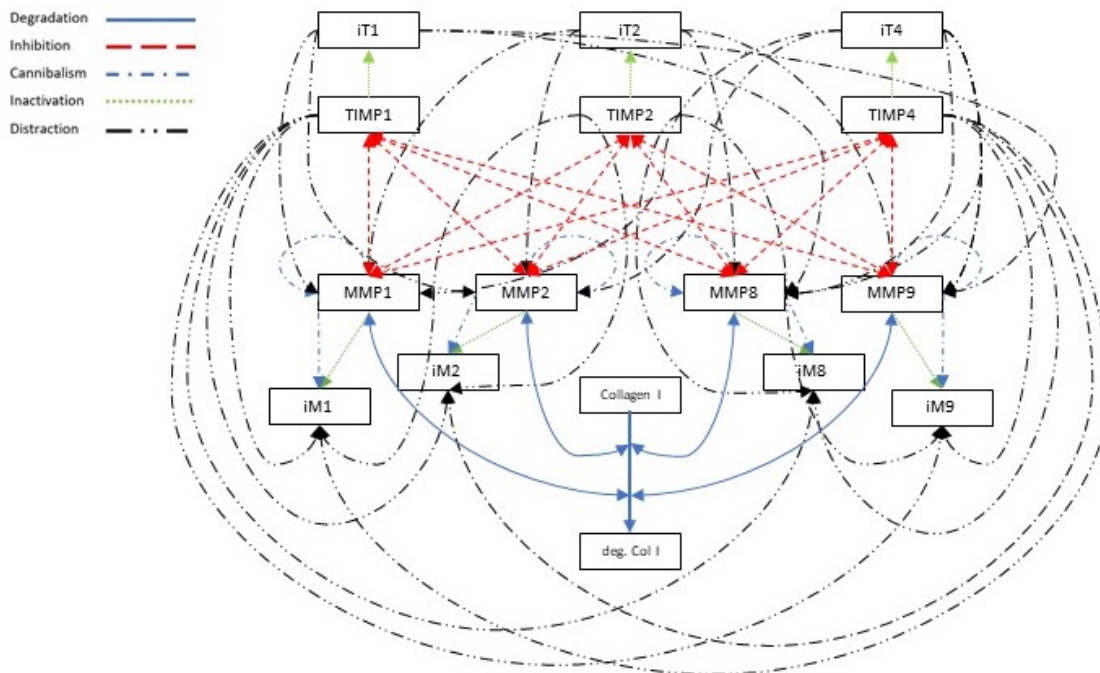
Figure 15: Schematic of the model topology showing all included reactions. The model containing MMP degradation, inhibition, and cannibalism, MMP and TIMP inactivation and MMP and TIMP distraction terms

The sum of all five ECM components remaining at the end of the simulation for each patient will be used to distinguish HF patients from control patients. Due to the lack of physiologically relevant studies regarding kinetic coefficients of the various reactions involving the proteases and their inhibitors, a genetic fitting algorithm was used to find the best-fit combination of rate parameters simultaneously in order to predict ECM turnover for patient-specific data. The genetic algorithm consisted of 30 generations, each with a population size of 50 sets of parameters. This approach generates a combination of parameters that can distinguish HF patients from control patients by ranking them based on the remaining ECM components as the parameter we want to monitor that causes fibrosis.

The model predicts substrate degradation and ranks the patients at the end of the simulations based on the ECM degradation product for each patient using measured values for MMPs, TIMPs, PINP, and PIIINP of that patient. The initial constraints for the upper and lower bounds of kinetic parameters were set to 0 and 15 ($M \times hr^{-1}$ for biomolecular reaction coefficients; hr^{-1} for unimolecular reaction coefficients) in an effort to avoid constraining the algorithm too tightly. Since we have shown that matrix degradation by proteases often follows a biphasic relationship with mechanical stimulation, substrate degradation parameters were set to a quadradic function of pulse pressure for each patient ($k_1 \times \lambda^2 + k_2 \times \lambda + k_3$) in order to provide a mechanical strain component to the mechanistic model.

The model will go through 80% of patient data as the training data set and uses each patient's MMP value as the initial value for MMPs, each TIMP value as the initial value for TIMPs, and PINP and PIIINP as collagen I and collagen III production rate. The initial value for each ECM component was set to the product of the average of all MMP values and a constant. The initial values for ECM components were approximated since the dataset did not include markers that could be used as the initial value for our model. We generated five different random training data sets in an effort to minimize the possibility of errors stemming from data selection.

The model then uses a genetic fitting algorithm to find an optimal (best-fit) combination of rate parameters, including collagen I and collagen III production rate, MMP

degradation rates, TIMP inhibition rates, etc[79,80]. For each iteration of the algorithm, random sets of rate parameters will be generated and used to predict normalized output levels for each patient, given initial protein and enzyme levels for that patient. For each set of parameters, predicted changes in output levels will be calculated for each patient, and top 13% of patients with the highest ECM accumulation (as the indicator of fibrosis) were chosen as HF patients since 13% in the original data experienced HF. The model compares the results from the simulation to the clinical diagnosis and calculates true positive (TP), false positive (FP), true negative (TN), false negative (FN), sensitivity ($\frac{TP}{TP+FN}$), and specificity ($\frac{TN}{TN+FP}$). In order to minimize the model bias on either sensitivity or specificity, $1 - \frac{sensitivity+specificity}{2}$ was used as a metric for the best-fitting parameters. The top 20% of best-fitting parameter sets were then used as "parents" for the next round of simulations, remaining parameter sets were generated by crossing the values contained in each set, and this new pool of parameter sets were tested against the clinical diagnosis as before[81].

This sequence was repeated for 30 generation to increase sensitivity and specificity of the model predictions for the training patient dataset. The new variables (five ECM component content remaining at the end of the simulations) were used as five new features within the same training data set in the machine learning framework in order to provide a mechanistic component to the classification.

5.2.2 Machine Learning Algorithms for Classification

The complete patient matrix was imported into the MATLAB Classification Learner within the Statistics and Machine Learning Toolbox, R2021b. We used seven

different machine learning algorithms for patient classification: logistic regression, discriminate analysis, naïve Bayes, support vector machines, k-nearest neighbors, a built-in ensemble algorithm, neural network, and also our own ensemble algorithm.

These supervised algorithms are often used for binary classification problems, and they can be used for both probabilistic and deterministic models. The logistic regression approach uses a logistic function to model a binary dependent variable, and the classifier models the class probabilities as a function of the linear combination of predictors[100]. [MP 20]. The discriminant analysis assumes that different classes generate data based on different Gaussian distributions[101].The naive Bayes algorithm leverages the Bayes theorem and assumes that predictors are conditionally independent, given the class[102]. A support vector machine model for classification seeks to find a hyperplane within the subspace that can best separate data from each class[103]. The k-nearest neighbors' approach finds an object's k-nearest neighbors through a distance metric and uses the neighbors' classes for the object's own classification[104]. Ensemble approaches use multiple machine learning algorithms in order to improve the performance of the model over the individual components of the ensemble. Here we used either bagging or boosting approaches. The bagging approach uses a training dataset to generate new sample data and then trains the model. Boosting method uses a weak learning model to start and then re-weight the model in each iteration by adding the weak model to stronger classifiers[105,106]. The neural network model is a feedforward, fully connected model for classification. There is a connection between the first fully connected layer of the neural network and the input data, and each subsequent layer has a connection from the previous layer. Each fully connected layer

multiplies the input by a weight matrix and then adds a bias vector. The final fully connected layer produces the network's output. We also used a voting ensemble that averaged the classification probabilities from each of the seven individual algorithms and then classified each patient according to the average.

5.2.3 Model optimization

For fitting the algorithms, patients were randomly split into a training subset (80% of the patient data) and a testing subset (20% of the patient data); the hyperparameters were then optimized using the training subset only, and the performance metrics (accuracy, receiver-operator-characteristic curves, etc.) were calculated on the testing subset. All reported performance indicators, therefore, represent validation performance with data that were not used for model fitting. In order to ensure that our conclusions were not an artifact of a single, randomized 80-20 split, we repeated the process five times (i.e., five different testing-training splits) and calculated the same performance metrics each time using different training sets. Each of the approaches, except the logistic regression approach, was optimized using a Bayesian hyperparameter optimization approach over 100 iterations[107]. This method uses a Gaussian process model of an objective function, a Bayesian update procedure for modifying the model, and an acquisition function that needs to be maximized[108].

The optimizable hyperparameters for each machine learning approach are as follows: the discriminate type, including linear, diagonal linear, quadratic, and linear quadratic for the discriminate analysis approach; the distribution (Gaussian or Kernel) as

well as the kernel type for the naïve Bayes approach; the kernel function (Gaussian, linear, quadratic, and cubic), the box constraint level, and the kernel scale for the support vector machine approach; the number of neighbors, the distance metric (Euclidean, city block, Chebyshev, cubic, Mahalanobis, cosine, correlation, Spearman, Hamming, and Jaccard), and the distance weight (equal, inverse, squared inverse) for the k-nearest neighbor approach, the method (AdaBoost, RUSBoost, LogitBoost, GentleBoost, Bag), the maximum number of splits, the number of learners, the learning rate, and the number of sampling predictors fir the ensemble approach, number of fully connected layers, activation (ReLU, Tanh, None, Sigmoid), regularization strength (Lambda), Standardize (Yes or No) for the neural network approach[107].

5.2.4 Model testing and Feature selection

After training simulations for every training data sets, the accuracies of the testing groups were calculated, with true positive (TP), true negative (TN), false positive (FP), and false negative (FN) using the formula for accuracy of a binary classification model:

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

Additionally, precision, F1 scores and Matthews Correlation Coefficients were calculated as follows:

$$F1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

The number of samples in each category is largely different from each other, and accuracy is heavily weighted towards the category with the larger number of samples; therefore, accuracy is not a complete measure of a classifier's ability. In order to minimize this bias, these additional statistical evaluations of the classifiers were used[109]. The F1 score is frequently used to evaluate the predictive power of a classifier, but it doesn't account for the true negative group. We employed the Matthews Correlation Coefficient (MCC), which is an indicator of the model's capability of correct prediction of both the positive and negative classes.[110]

We used a Minimum Redundancy Maximum Relevancy (MRMR) algorithm to help us understand the importance of each feature[111]. MRMR algorithm uses either mutual information of the features or the amount of uncertainty of a feature that can be minimized by the use of another feature to find a set of features from the training set that can effectively represent the response and minimize the redundancy of the feature MRMR algorithm ranks the feature and the higher weighting features are more confident for the model in the effort to predict the responses.

5.3 Results

5.3.1 Mechanistic Model

We performed a genetic algorithm to find the best combination of rate parameters and the five ECM components remaining at the end of simulations. Figure 16 shows the result of the best-performing simulation curves for each of the five ECM components over time for all patients. The new variables (five ECM component content remaining at the end of the simulations) were used as five new features within the same training data set in the machine learning framework in order to provide a mechanistic component to the classification.
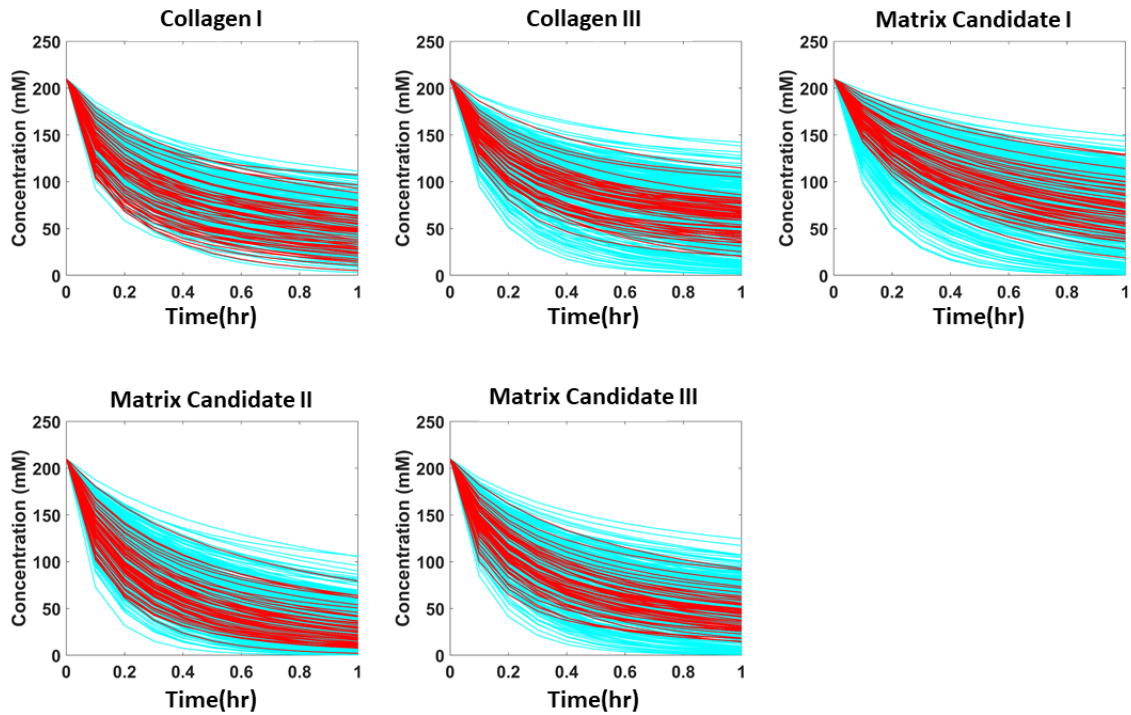
Figure 16: Concentration change of the five ECM component content for every patient in the dataset. Blue are control patients and red are HFpEF patients. We used the remaining ECM component at the end of the simulations as a new feature adding five new features to the patient's data

5.3.2 Machine Learning Algorithms for Classification

After the simulations for classification were completed, receiver-operating-characteristic (ROC) curves were recorded to assess predictive performance (Figure 19).
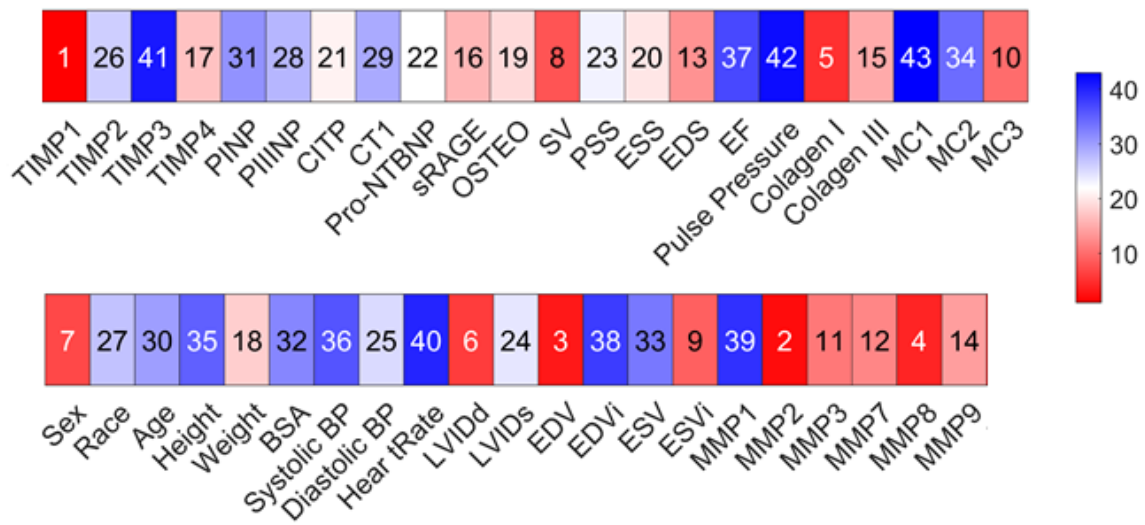
Figure 17: Feature selection by MRMR identified a ranked list of each feature's importance to HFpEF classification. Note the red colors denote higher performance rankings while blue colors denote lower performance ranking.

Most machine learning algorithms overperformed the previously reported, HFpEF

classification by Zile et al[99]. (Figure 18)
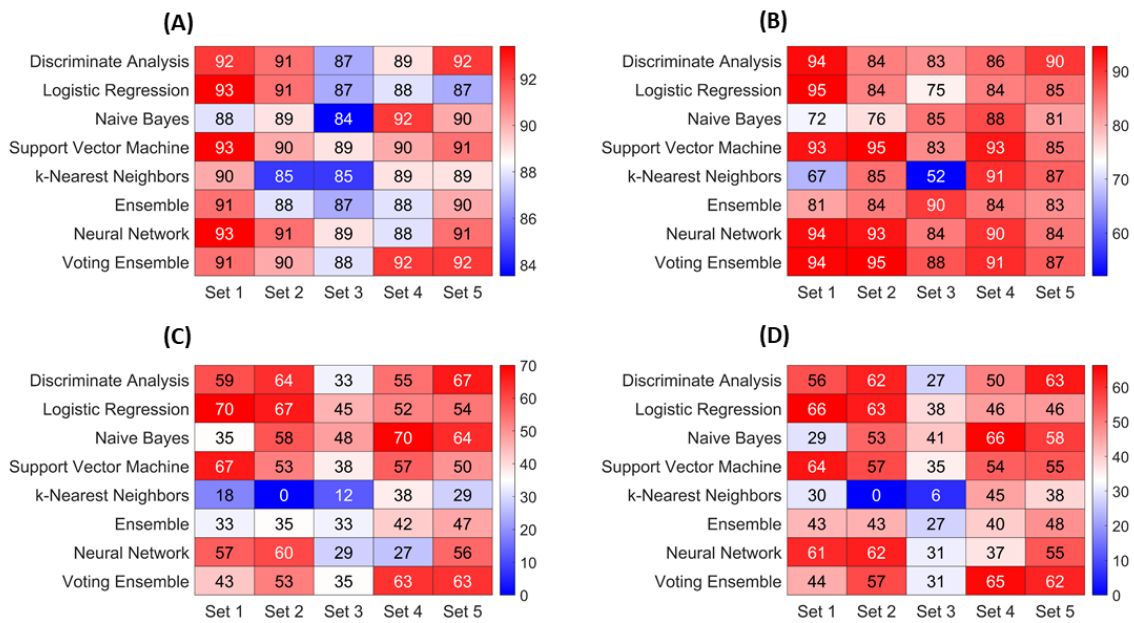
Figure 18: Performance metrics (in %) for five different testing set for each machine learning algorithm included (A) Accuracy, (B) ROC area-under-the-curve (AUC), (C) F1 scores, (D)Matthews Correlation Coefficients (MCC), classification of HFpEF patients.

Overall performance of the various algorithms was compared across F1 scores, MCCs, AUCs, and accuracy levels, with various algorithms reaching as high as 95% AUC and 93% accuracy (Figure 18). Across all classification tasks, our voting ensemble algorithm showed the most consistently high performance, reaching as high as 95% AUC.
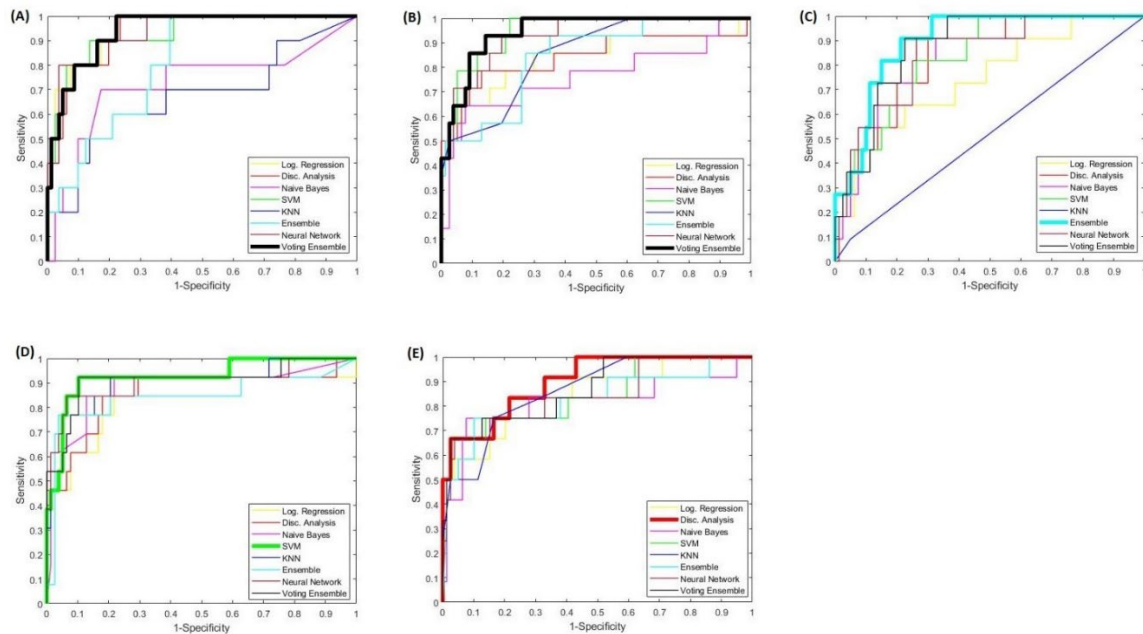
Figure 19: Receiver operating characteristic (ROC) curves for classification of HFpEF patients for all of the five different testing datasets. The voting ensemble approach had the most consistent performance. Figures A through E shows the results of five different training testing splits.

These results were largely consistent across five different randomized training-testing splits of the dataset, which showed only minor variability in performance (standard deviations for the ensemble AUC was 3.6%).

In order to better interpret our modeling predictions, we performed an MRMR feature selection approach for generating a ranked list of feature importance (Figure 17). The most heavily weighted features for the HFpEF classification included some basic features like sex, as well as some plasma biomarkers like TIMP1, MMP2, and MMP8 and collagen I coming from mechanistic modeling.

In general, the non-MMP and non-TIMP plasma biomarkers ranked very low. Interestingly, apart from a few volume measures, other echo-based variables ranked relatively poorly in the feature importance analysis.

5.4 Discussion

The progression of hypertension-induced HFpEF is complex and utilizes various mechanistic pathways. Machine learning approaches offer predictive power without the need to fully understand these mechanistic complexities. We used advanced machine learning techniques to deal with the need for more detailed kinetic studies involving specific roles played by key factors of matrix turnover. After utilizing multiple machine learning approaches, the voting ensemble method had consistently high predictive power based on accuracy and the AUC. This is not surprising given the ensemble approach's ability to combine the high-performing predictions from the individual algorithms while discounting the low-performing predictions.

The MRMR algorithm is used as the feature selection method since it is repeatedly used in biological applications, and it can be used for different machine learning methods. It uses the relationship of each variable in the training dataset to the response variable in order to rank the feature and help with the feature selection process[111–113]. In this study, features were strongly related to each other (such as the proteases and their tissue inhibitors or various echo-based features calculated from other echo features). The MRMR algorithm helps us investigate the importance of each feature compared to the rest of the features and ensures that the variables selected are not similar to one another. After feature analysis was conducted with the MRMR algorithm, MMPs 2, 8, and TIMP-1 were found as important features for HFpEF classification. These findings corroborate the previous study by Zile et al. of HFpEF classification through logistic regression[99]. Previous work has found that MMP-3 and MMP-9 were significant predictive biomarkers of HFpEF. TIMP-2 and TIMP-

4 plasma levels are also increased in HFpEF patients, and it has been shown that MMP-2 and MMP-7 are increased while MMP-8 levels are decreased in HFpEF patients compared to control patients[90,99]. NT-proBNP, which is categorized as one of the most important biomarkers in hypertrophic disease, ranked relatively low in the feature rankings for the classification tasks[114]. Notably, feature selection revealed that only a few important features came from echocardiographic data relating to the heart's structure. For the HFpEF classification, ventricular volumes emerged as highly rated features, which is supported by previous evidence of ventricular mechanics contributing to pathologies that impair ventricular function, such as cardiomyopathy[115,116].

It is important to note several key limitations in the current work that limit the model's predictive power. First, the patient dataset consisted of 459 patients. Many of these patients were previously associated with cardiology clinics and studies. A larger, prospective study with more diverse patients would be needed in order to further validate the results and expand these classification models with additional patients. We believe that extracellular remodeling is a key factor in the HF process. A limitation of this study is that the set of biomarkers that participate in extracellular remodeling is limited. Our machine learning methods can benefit from more arrays of biomarkers, specifically ECM remodeling biomarkers.

The mechanistic model was utilized to consider the role of fibrosis in HF and consider the interaction between the ECM component and MMP and TIMPs. Although we restricted the upper band and the lower band for each parameter in the simulation, there is not enough evidence showing the relevance of the calculated rate parameters to

physiological values since most of the kinetic parameters are still unknown in the literature. Also, we tried to capture all of the interactions possible between species, but there is not enough evidence to prove that all of these interactions are happening. We built a mechanistic model that includes the interaction of three potential ECM components in addition to fibrillar collagen type I and types III since the biomarkers measured in the patient dataset have multiple functions in addition to fibrotic turnover. However, each of these biomarkers might have a specific role in heart function. Current studies for the classification of HFpEF patients show the promise of MR-proANP, MR-proADM, troponins, sST2, GDF-15, and galectin-3 as additional biomarkers[117].

Our work shows that advanced machine learning techniques could be used as a promising tool for the classification of HFpEF patients. These techniques can benefit from larger and more diverse patient data with additional physiologically related biomarkers to prove the clinical application of our ensemble algorithm.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE STUDIES

6.1 Conclusions

Cardiac fibrosis is critical in a variety of HF patients, including post-MI as well as non-ischemic diseases like hypertension, and there are no currently prescribed therapeutics for fibrotic scar regulation. The field is also unable to predict fibrosis due to the inability to assess which patients at what time and to what extent are more likely to develop fibrosis. Confounding variabilities include 1) mechanical variabilities, 2) temporal variabilities, and 3) patient-to-patient variabilities. Computational modeling can help capture these variabilities in order to help us with novel treatments. Our investigation aimed to identify key regulators of cardiac fibrosis and understand the ECM turnover mechanism and ECM interaction with matrix-related proteins through a combination of in-vitro experiments and computational simulations exploring collagen-MMP-TIMP chemo- and mechano-interactions.

In the first aim of this investigation, in an effort to study the mechanical variabilities, we tested the effect of tensile loading on collagenous tissue degradation by proteases. Decellularized porcine pericardium samples were treated with MMP1, MMP8, MMP9, Cathepsin K, or a protease-free control while subjected to different levels of mechanical strain (from ~5-40%). Isotropic displacement control was provided, and the degradation level of pericardium samples was measured using force decay data. The stress degradation decay vs. strain rate curves indicate that particular rate of degradation was

affected by the sample strain, but the particular influence of strain depended upon the type of protease. Specifically, degradation by cathepsin K was increased with increasing strains, and degradation by MMP-1, MMP-8, and MMP-9 first decreased and then increased with the strain magnitude forming a V-shaped curve. The degradation products were also analyzed by mass spectrometry to assess how mechanical strain levels altered the degradome signatures, which revealed unique shifts in degradation products for all four protease types. Overall, these results highlight that the relative contributions of different proteases to tissue turnover depend not only on protease concentrations but also on the local level of tissue's mechanical conditions.

In the second aim of this investigation, in an effort to study the temporal variabilities, we presented a detailed computational model of the biochemical network of collagen I proteolysis capturing all interactions of type I collagen, four MMPs (MMP-1, -2, -8, and -9), and three TIMPs (TIMP-1, -2, and -4) in a cell-free, well-stirred environment. We used DQ collagen to monitor the proteolytic activity of MMPs and the inhibitory activity of TIMPs. We then used the results from experimental data to fit five different hypothetical reaction topologies in order to investigate their respective accuracies. We determined kinetic rate constants for collagen degradation by MMPs, MMP inhibition by TIMPs, MMP and TIMP inactivation rates, MMP cannibalism rates, as well as MMP and TIMP distraction rates. We also used post-MI time-courses of collagen, MMP, and TIMP levels in the animal heart from the literature to perform a parameter sensitivity analysis across the model reaction rates to identify which molecules or interactions are the important regulators of ECM post-MI for both early and late time-periods.

In vitro experimental data helped us understand which MMP is capable of degrading collagen faster and more effectively, but only by comparing these models we were able to identify which MMP's kinetic terms are essential for better predicting collage-MMP-TIMP interaction. Notably, MMP-2 is able to degrade collagen faster and more, and MMP inactivation terms are more important than cannibalism and distraction terms which were impossible to predict without using the models. Generally, the results from sensitivity analysis show that the concentration of collagen is more sensitive to the parameters at later time points (8 weeks after MI) compared with earlier time points (1 week after MI). It is revealed that the collagen concentration is not greatly sensitive to TIMPs inhibition terms or TIMP inactivation terms. It is also worth mentioning that collagen concentration is more sensitive to MMP-2 degradation and inactivation terms in every time point. Predictably, collagen concentration changes drastically when we modulated the initial values for each species, and it is most sensitive to MMP-2 and MMP-9 initial concentration after collagen concentration.

In the third aim of this investigation, in an effort to study the patient-to-patient variabilities, we hypothesized that machine learning algorithms could substantially improve the predictive value of circulating plasma biomarkers by leveraging more sophisticated statistical approaches. We first developed an ensemble classification algorithm for diagnosing HFpEF within a population of 459 individuals, including HFpEF patients and referent control patients. Algorithms showed strong diagnostic performance with receiver-operating-characteristic curve (ROC) areas of 0.95 for identifying HFpEF patients using demographic information, plasma biomarkers related to ECM remodeling,

86

echocardiogram data, and ECM component prediction resulting from our mechanistic model. We built a mechanistic model to predict ECM components using a genetic algorithm to connect ECM remodeling to the plasma biomarkers and help us with HFpEF patients' classification. Our findings demonstrate that machine learning-based classification algorithms show promise as a non-invasive diagnostic tool for HFpEF while also suggesting priority biomarkers for future mechanistic studies to elucidate more specific regulatory roles.

Although no treatment has yet been proven to reduce morbidity and mortality of HFpEF patients, and there is still a gap for targeted, anti-fibrotic therapies, computational modeling serves as an important tool to investigate the complex matrix-MMP-TIMP mechano-chemo-interaction network and can help testing and identifying novel treatments in order to address mechanical, temporal, and patient-to-patient variabilities.

6.2 Recommendations for future studies

6.2.1 The effect of tensile loading on collagen degradation by proteases

The collagen fibers in pericardium samples are randomly aligned. This could cause some variation in the stress experience in different tissue samples with the same strain. Using aligned, pure collagen gels can potentially eliminate these variations. Furthermore, we recognize that the pericardium is not only comprised of collagen I, and the cardiac protease is not solely MMP-1, 8, and 9. We selected this subset of enzymes as representative of different types of MMPs, such as collagenase and gelatinase, and the goal of Aim 1 was to investigate the behavior of different protease types on collagenous tissue

degradation under mechanical loading. Expanding this study across more proteases and adding TIMPs, different protease and TIMP concentration and protease cocktails, and a variety of different combinations of mechanical strain, including static and dynamic strains, can expand our knowledge of the effect of tensile loading on collagenous tissue degradation by proteases.

In addition, we acknowledge that MMPs are activated by MMPs and other enzymes, and these interactions are critical to overall MMP activity in vivo. The reason we did not test those interactions in our experimental setup is that we are purposefully trying to reduce each condition to one protease-matrix interaction per pericardium sample in order to quantify the isolated effect of mechanical load on each of those specific interactions before we then re-combine all of those interactions (including activation steps) in the computational model. Moreover, we studied the effect of tensile loading on collagenous tissue with one protease in vitro; however, each of these proteases might have a specific role in heart function. Expanding this study to in-vivo applications can reveal some interesting facts about ECM remodeling.

6.2.2 Computational Modeling of Collagen Turnover by Protease and Inhibitor Combinations

Many studies tried to calculate Michaelis-Menten reaction rates which could be used as our model parameters; however, these parameters could potentially be chosen incorrectly since these studies have been done in different animals and different experimental environments. Thus, we built our own model and tried to come up with the

best set of reaction parameters that can explain collagen-MMP-TIMP interactions. The in-vitro experiments can be expanded across multiple MMP-TIMP combinations (a variety of MMP and TIMP types and concentrations). The study can be expanded with more data points in order to help the models find better fits to the experimental data. Additionally, there could be hidden reactions between MMPs or TIMPs affecting activation, degradation, and inhibition rates that are not previously studied and can change the network structure. Replacing DQ collagen with collagen gels and quantifying collagen degradation using real-time imaging enables the possibility of stretching the samples and opens new doors to investigate the time-dependent effects of combinations of MMPs and TIMPs on collagen degradation. The model then can be expanded by adding terms that can explain the effects of mechanical strain.

6.2.3 Machine Learning Models Identifies HFpEF Patients from Matrix-Related Plasma Biomarkers

The machine learning approach shows promising results for the classification of HFpEF patients. A more extensive, prospective study with more diverse patients would further validate the results and expand these classification models with additional patients. Additionally, our machine learning methods can benefit from more arrays of biomarkers, specifically ECM remodeling biomarkers. Our mechanistic model of collagen degradation using patient-specific protein data allowed us to come up with a set of parameters that, in addition to plasma biomarkers, helped us accurately diagnose HF patients. We used serum-based protein measurements for our model inputs and outputs. We acknowledge serum levels are an aggregate of collagen turnover across many organs, not just the heart; thus,

89

future studies could seek to employ micro-dialysis technology currently being used by our collaborator to sample interstitial fluid directly from myocardial tissue in these patients. Furthermore, the mechanistic model used in this study could be integrated with the machine learning models. The models go through every patient, uses a set of arbitrary rate parameters for all of the reaction rates, calculates the ECM component remaining at the end of the simulations and add that value for each patient to the list of the variables used for the classification models. The models then calculate TP, TN, FP, FN, sensitivity and specificity and the genetic algorithm tries to make the best sets of parameters to minimize objective functions. Future models can benefit from this method in an effort to improve the accuracy of the models.

APPENDICES

Appendix A

Aim 2 Supplemental Equations

The code for all of interactions in topology 1:

```
function dydt = DQ_ODE_function1(t,y, k_degM1, k_degM2, k_degM8, k_degM9,...
                                 k_inhT1M1, k_inhT1M2, k_inhT1M8,
                                 k_inhT1M9,...
                                 k_inhT2M1, k_inhT2M2, k_inhT2M8,
                                 k_inhT2M9,...
                                 k_inhT4M1, k_inhT4M2, k_inhT4M8,
                                 k_inhT4M9)


C1 = 1;
dC1 = 2;

M1 = 3;
M2 = 4;
M8 = 5;
M9 = 6;

T1 = 7;
T2 = 8;
T4 = 9;

M1T1 = 10;
M1T2 = 11;
M1T4 = 12;

M2T1 = 13;
M2T2 = 14;
M2T4 = 15;


M8T1 = 16;
M8T2 = 17;
M8T4 = 18;

M9T1 = 19;
M9T2 = 20;
M9T4 = 21;


%%%%%%%%%%%%%%%%%%%%%%%%%%%% REACTIONS %%%%%%%%$%%%%%%%%%%%%%%%%%%%%%%%%
dydt = zeros(21,1);


dydt(C1) = - k_degM1*y(M1)*y(C1) - k_degM2*y(M2)*y(C1)  - k_degM8*y(M8)*y(C1)
- k_degM9*y(M9)*y(C1);
```

92

```
dydt(dC1) = k_degM1*y(M1)*y(C1) + k_degM2*y(M2)*y(C1) + k_degM8*y(M8)*y(C1) +
k_degM9*y(M9)*y(C1);



dydt(M1) =  - k_inhT1M1*y(M1)*y(T1) - k_inhT2M1*y(M1)*y(T2) -
k_inhT4M1*y(M1)*y(T4);

dydt(M2) =  - k_inhT1M2*y(M2)*y(T1) - k_inhT2M2*y(M2)*y(T2) -
k_inhT4M2*y(M2)*y(T4);

dydt(M8) =  - k_inhT1M8*y(M8)*y(T1) - k_inhT2M8*y(M8)*y(T2) -
k_inhT4M8*y(M8)*y(T4);

dydt(M9) =  - k_inhT1M9*y(M9)*y(T1) - k_inhT2M9*y(M9)*y(T2) -
k_inhT4M9*y(M9)*y(T4);



dydt(T1) = -k_inhT1M1*y(M1)*y(T1) - k_inhT1M2*y(M2)*y(T1) -
k_inhT1M8*y(M8)*y(T1) - k_inhT1M9*y(M9)*y(T1);

dydt(T2) = -k_inhT2M1*y(M1)*y(T2) - k_inhT2M2*y(M2)*y(T2) -
k_inhT2M8*y(M8)*y(T2) - k_inhT2M9*y(M9)*y(T2);

dydt(T4) = -k_inhT4M1*y(M1)*y(T4) - k_inhT4M2*y(M2)*y(T4) -
k_inhT4M8*y(M8)*y(T4) - k_inhT4M9*y(M9)*y(T4);



dydt(M1T1) = k_inhT1M1*y(M1)*y(T1);
dydt(M1T2) = k_inhT2M1*y(M1)*y(T2);
dydt(M1T4) = k_inhT4M1*y(M1)*y(T4);


dydt(M2T1) = k_inhT1M2*y(M2)*y(T1);
dydt(M2T2) = k_inhT2M2*y(M2)*y(T2);
dydt(M2T4) = k_inhT4M2*y(M2)*y(T4);


dydt(M8T1) = k_inhT1M8*y(M8)*y(T1);
dydt(M8T2) = k_inhT2M8*y(M8)*y(T2);
dydt(M8T4) = k_inhT4M8*y(M8)*y(T4);


dydt(M9T1) = k_inhT1M9*y(M9)*y(T1);
dydt(M9T2) = k_inhT2M9*y(M9)*y(T2);
dydt(M9T4) = k_inhT4M9*y(M9)*y(T4);


end
```

The code for all of interactions in topology 2:

```
function dydt = DQ_ODE_function2(t,y, k_on_M1C1,k_off_M1C1, k_deg_M1C1,...
                                 k_inh_on_T1M1, k_inh_on_T2M1,
                                 k_inh_on_T4M1,...
                                 k_inh_off_T1M1, k_inh_off_T2M1,
                                 k_inh_off_T4M1,...
                                 k_on_M2C1,k_off_M2C1, k_deg_M2C1,...

                                 k_inh_on_T1M2, k_inh_on_T2M2,
                                 k_inh_on_T4M2,...
                                 k_inh_off_T1M2, k_inh_off_T2M2,
                                 k_inh_off_T4M2,...
                                 k_on_M8C1,k_off_M8C1, k_deg_M8C1,...

                                     k_inh_on_T1M8, k_inh_on_T2M8,
                                     k_inh_on_T4M8,...
                                 k_inh_off_T1M8, k_inh_off_T2M8,
                                 k_inh_off_T4M8,...
                                 k_on_M9C1,k_off_M9C1, k_deg_M9C1,...

                             k_inh_on_T1M9, k_inh_on_T2M9,
                             k_inh_on_T4M9,...
                              k_inh_off_T1M9, k_inh_off_T2M9,
                             k_inh_off_T4M9)

        C1 = 1;
        dC1 = 2;

        M1 = 3;
        M2 = 4;
        M8 = 5;
        M9 = 6;

        M1C1 = 7;
        M2C1 = 8;
        M8C1 = 9;
        M9C1 = 10;


        T1 = 11;
        T2 = 12;
        T4 = 13;

        M1T1 = 14;
        M1T2 = 15;
        M1T4 = 16;

        M2T1 = 17;
        M2T2 = 18;
        M2T4 = 19;
```

```
M8T1 = 20;
M8T2 = 21;
M8T4 = 22;


M9T1 = 23;
M9T2 = 24;
M9T4 = 25;


%%%%%%%%%%%%%%%%%%%%%%%%%%% REACTIONS %%%%%%%$%%%%%%%%%%%%%%%%%%%%%%%
dydt = zeros(25,1);


dydt(C1) = - k_on_M1C1*y(M1)*y(C1) + k_off_M1C1*y(M1C1) -
k_on_M2C1*y(M2)*y(C1) + k_off_M2C1*y(M2C1) - k_on_M8C1*y(M8)*y(C1) +
k_off_M8C1*y(M8C1) - k_on_M9C1*y(M9)*y(C1) + k_off_M9C1*y(M9C1);

dydt(dC1) = k_deg_M1C1*y(M1C1)  + k_deg_M2C1*y(M2C1)  + k_deg_M8C1*y(M8C1)  +
k_deg_M9C1*y(M9C1);



dydt(M1) =   - k_on_M1C1*y(M1)*y(C1)  + k_off_M1C1*y(M1C1) +
k_deg_M1C1*y(M1C1)+...
            - k_inh_on_T1M1*y(M1)*y(T1) - k_inh_on_T2M1*y(M1)*y(T2) -
k_inh_on_T4M1*y(M1)*y(T4) +  k_inh_off_T1M1*y(M1T1) + k_inh_off_T2M1*y(M1T2) +
k_inh_off_T4M1*y(M1T4);

dydt(M2) =   - k_on_M2C1*y(M2)*y(C1)  + k_off_M2C1*y(M2C1) +
k_deg_M2C1*y(M2C1)+...
            - k_inh_on_T1M2*y(M2)*y(T1) - k_inh_on_T2M2*y(M2)*y(T2) -
k_inh_on_T4M2*y(M2)*y(T4) +  k_inh_off_T1M2*y(M2T1) + k_inh_off_T2M2*y(M2T2) +
k_inh_off_T4M2*y(M2T4);

dydt(M8) =   - k_on_M8C1*y(M8)*y(C1)  + k_off_M8C1*y(M8C1) +
k_deg_M8C1*y(M8C1)+...
            - k_inh_on_T1M8*y(M8)*y(T1) - k_inh_on_T2M8*y(M8)*y(T2) -
k_inh_on_T4M8*y(M8)*y(T4) +  k_inh_off_T1M8*y(M8T1) + k_inh_off_T2M8*y(M8T2) +
k_inh_off_T4M8*y(M8T4);

dydt(M9) =   - k_on_M9C1*y(M9)*y(C1)  + k_off_M9C1*y(M9C1) +
k_deg_M9C1*y(M9C1)+...
            - k_inh_on_T1M9*y(M9)*y(T1) - k_inh_on_T2M9*y(M9)*y(T2) -
k_inh_on_T4M9*y(M9)*y(T4) +  k_inh_off_T1M9*y(M9T1) + k_inh_off_T2M9*y(M9T2) +
k_inh_off_T4M9*y(M9T4);



dydt(M1C1) =  k_on_M1C1*y(M1)*y(C1) - k_off_M1C1*y(M1C1) - k_deg_M1C1*y(M1C1);

dydt(M2C1) =  k_on_M2C1*y(M2)*y(C1) - k_off_M2C1*y(M2C1) - k_deg_M2C1*y(M2C1);
```

```
dydt(M8C1) =  k_on_M8C1*y(M8)*y(C1) - k_off_M8C1*y(M8C1) - k_deg_M8C1*y(M8C1);

dydt(M9C1) =  k_on_M9C1*y(M9)*y(C1) - k_off_M9C1*y(M9C1) - k_deg_M9C1*y(M9C1);



dydt(T1) =  k_inh_off_T1M1*y(M1T1) - k_inh_on_T1M1*y(M1)*y(T1)+...
            k_inh_off_T1M2*y(M2T1) - k_inh_on_T1M2*y(M2)*y(T1)+...
            k_inh_off_T1M8*y(M8T1) - k_inh_on_T1M8*y(M8)*y(T1)+...
            k_inh_off_T1M9*y(M9T1) - k_inh_on_T1M9*y(M9)*y(T1);

dydt(T2) =  k_inh_off_T2M1*y(M1T2) - k_inh_on_T2M1*y(M1)*y(T2)+...
            k_inh_off_T2M2*y(M2T2) - k_inh_on_T2M2*y(M2)*y(T2)+...
            k_inh_off_T2M8*y(M8T2) - k_inh_on_T2M8*y(M8)*y(T2)+...
            k_inh_off_T2M9*y(M9T2) - k_inh_on_T2M9*y(M9)*y(T2);

dydt(T4) =  k_inh_off_T4M1*y(M1T4) - k_inh_on_T4M1*y(M1)*y(T4)+...
            k_inh_off_T4M2*y(M2T4) - k_inh_on_T4M2*y(M2)*y(T4)+...
            k_inh_off_T4M8*y(M8T4) - k_inh_on_T4M8*y(M8)*y(T4)+...
            k_inh_off_T4M9*y(M9T4) - k_inh_on_T4M9*y(M9)*y(T4);


dydt(M1T1) = k_inh_on_T1M1*y(M1)*y(T1) - k_inh_off_T1M1*y(M1T1);
dydt(M1T2) = k_inh_on_T2M1*y(M1)*y(T2) - k_inh_off_T2M1*y(M1T2);
dydt(M1T4) = k_inh_on_T4M1*y(M1)*y(T4) - k_inh_off_T4M1*y(M1T4);

dydt(M2T1) = k_inh_on_T1M2*y(M2)*y(T1) - k_inh_off_T1M2*y(M2T1);
dydt(M2T2) = k_inh_on_T2M2*y(M2)*y(T2) - k_inh_off_T2M2*y(M2T2);
dydt(M2T4) = k_inh_on_T4M2*y(M2)*y(T4) - k_inh_off_T4M2*y(M2T4);

dydt(M8T1) = k_inh_on_T1M8*y(M8)*y(T1) - k_inh_off_T1M8*y(M8T1);
dydt(M8T2) = k_inh_on_T2M8*y(M8)*y(T2) - k_inh_off_T2M8*y(M8T2);
dydt(M8T4) = k_inh_on_T4M8*y(M8)*y(T4) - k_inh_off_T4M8*y(M8T4);

dydt(M9T1) = k_inh_on_T1M9*y(M9)*y(T1) - k_inh_off_T1M9*y(M9T1);
dydt(M9T2) = k_inh_on_T2M9*y(M9)*y(T2) - k_inh_off_T2M9*y(M9T2);
dydt(M9T4) = k_inh_on_T4M9*y(M9)*y(T4) - k_inh_off_T4M9*y(M9T4);


end
```

The code for all of interactions in topology 3:

```
function dydt = DQ_ODE_function3(t,y, k_on_M1C1,k_off_M1C1, k_deg_M1C1,...
                                 k_inh_on_T1M1, k_inh_on_T2M1,
                                 k_inh_on_T4M1,...
                                 k_inh_off_T1M1, k_inh_off_T2M1,
                                 k_inh_off_T4M1,...
                                 k_on_M2C1,k_off_M2C1, k_deg_M2C1,...

                                     k_inh_on_T1M2, k_inh_on_T2M2,
                                     k_inh_on_T4M2,...
                                 k_inh_off_T1M2, k_inh_off_T2M2,
                                 k_inh_off_T4M2,...
                                 k_on_M8C1,k_off_M8C1, k_deg_M8C1,...

                                     k_inh_on_T1M8, k_inh_on_T2M8,
                                     k_inh_on_T4M8,...
                                 k_inh_off_T1M8, k_inh_off_T2M8,
                                 k_inh_off_T4M8,...
                                 k_on_M9C1,k_off_M9C1, k_deg_M9C1,...

                                     k_inh_on_T1M9, k_inh_on_T2M9,
                                     k_inh_on_T4M9,...
                                 k_inh_off_T1M9, k_inh_off_T2M9,
                                 k_inh_off_T4M9,...

                                 k_inact_M1,k_inact_M2,k_inact_M8,k_in
                                 act_M9,...
                                 k_inact_T1,k_inact_T2,k_inact_T4)

C1 = 1;
dC1 = 2;

M1 = 3;
M2 = 4;
M8 = 5;
M9 = 6;

M1C1 = 7;
M2C1 = 8;
M8C1 = 9;
M9C1 = 10;


T1 = 11;
T2 = 12;
T4 = 13;

M1T1 = 14;
M1T2 = 15;
M1T4 = 16;
```

```
M2T1 = 17;
M2T2 = 18;
M2T4 = 19;


M8T1 = 20;
M8T2 = 21;
M8T4 = 22;

M9T1 = 23;
M9T2 = 24;
M9T4 = 25;

iM1 = 26;
iM2 = 27;
iM8 = 28;
iM9 = 29;

iT1 = 30;
iT2 = 31;
iT4 = 32;


%%%%%%%%%%%%%%%%%%%%%%%%%% REACTIONS %%%%%%%$%%%%%%%%%%%%%%%%%%%%%%%%
dydt = zeros(32,1);


dydt(C1) = - k_on_M1C1*y(M1)*y(C1) + k_off_M1C1*y(M1C1) -
k_on_M2C1*y(M2)*y(C1) + k_off_M2C1*y(M2C1) - k_on_M8C1*y(M8)*y(C1) +
k_off_M8C1*y(M8C1) - k_on_M9C1*y(M9)*y(C1) + k_off_M9C1*y(M9C1);

dydt(dC1) = k_deg_M1C1*y(M1C1)  + k_deg_M2C1*y(M2C1)  + k_deg_M8C1*y(M8C1)  +
k_deg_M9C1*y(M9C1);



dydt(M1) =   - k_on_M1C1*y(M1)*y(C1)  + k_off_M1C1*y(M1C1) +
k_deg_M1C1*y(M1C1)+...
            - k_inh_on_T1M1*y(M1)*y(T1) - k_inh_on_T2M1*y(M1)*y(T2) -
k_inh_on_T4M1*y(M1)*y(T4) +  k_inh_off_T1M1*y(M1T1) + k_inh_off_T2M1*y(M1T2) +
k_inh_off_T4M1*y(M1T4) - k_inact_M1*y(M1);

dydt(M2) =   - k_on_M2C1*y(M2)*y(C1)  + k_off_M2C1*y(M2C1) +
k_deg_M2C1*y(M2C1)+...
            - k_inh_on_T1M2*y(M2)*y(T1) - k_inh_on_T2M2*y(M2)*y(T2) -
k_inh_on_T4M2*y(M2)*y(T4) +  k_inh_off_T1M2*y(M2T1) + k_inh_off_T2M2*y(M2T2) +
k_inh_off_T4M2*y(M2T4) - k_inact_M2*y(M2);

dydt(M8) =   - k_on_M8C1*y(M8)*y(C1)  + k_off_M8C1*y(M8C1) +
k_deg_M8C1*y(M8C1)+...
```

```
            - k_inh_on_T1M8*y(M8)*y(T1) - k_inh_on_T2M8*y(M8)*y(T2) -
k_inh_on_T4M8*y(M8)*y(T4) +  k_inh_off_T1M8*y(M8T1) + k_inh_off_T2M8*y(M8T2) +
k_inh_off_T4M8*y(M8T4) - k_inact_M8*y(M8);

dydt(M9) =    - k_on_M9C1*y(M9)*y(C1)  + k_off_M9C1*y(M9C1) +
k_deg_M9C1*y(M9C1)+...
            - k_inh_on_T1M9*y(M9)*y(T1) - k_inh_on_T2M9*y(M9)*y(T2) -
k_inh_on_T4M9*y(M9)*y(T4) +  k_inh_off_T1M9*y(M9T1) + k_inh_off_T2M9*y(M9T2) +
k_inh_off_T4M9*y(M9T4) - k_inact_M9*y(M9);


dydt(M1C1) =  k_on_M1C1*y(M1)*y(C1) - k_off_M1C1*y(M1C1) - k_deg_M1C1*y(M1C1);

dydt(M2C1) =  k_on_M2C1*y(M2)*y(C1) - k_off_M2C1*y(M2C1) - k_deg_M2C1*y(M2C1);

dydt(M8C1) =  k_on_M8C1*y(M8)*y(C1) - k_off_M8C1*y(M8C1) - k_deg_M8C1*y(M8C1);

dydt(M9C1) =  k_on_M9C1*y(M9)*y(C1) - k_off_M9C1*y(M9C1) - k_deg_M9C1*y(M9C1);


dydt(iM1) =  k_inact_M1*y(M1);
dydt(iM2) =  k_inact_M2*y(M2);
dydt(iM8) =  k_inact_M8*y(M8);
dydt(iM9) =  k_inact_M9*y(M9);

dydt(T1) =  k_inh_off_T1M1*y(M1T1) - k_inh_on_T1M1*y(M1)*y(T1)+...
            k_inh_off_T1M2*y(M2T1) - k_inh_on_T1M2*y(M2)*y(T1)+...
            k_inh_off_T1M8*y(M8T1) - k_inh_on_T1M8*y(M8)*y(T1)+...
            k_inh_off_T1M9*y(M9T1) - k_inh_on_T1M9*y(M9)*y(T1) -
k_inact_T1*y(T1);

dydt(T2) =  k_inh_off_T2M1*y(M1T2) - k_inh_on_T2M1*y(M1)*y(T2)+...
            k_inh_off_T2M2*y(M2T2) - k_inh_on_T2M2*y(M2)*y(T2)+...
            k_inh_off_T2M8*y(M8T2) - k_inh_on_T2M8*y(M8)*y(T2)+...
            k_inh_off_T2M9*y(M9T2) - k_inh_on_T2M9*y(M9)*y(T2) -
k_inact_T2*y(T2);

dydt(T4) =  k_inh_off_T4M1*y(M1T4) - k_inh_on_T4M1*y(M1)*y(T4)+...
            k_inh_off_T4M2*y(M2T4) - k_inh_on_T4M2*y(M2)*y(T4)+...
            k_inh_off_T4M8*y(M8T4) - k_inh_on_T4M8*y(M8)*y(T4)+...
            k_inh_off_T4M9*y(M9T4) - k_inh_on_T4M9*y(M9)*y(T4) -
k_inact_T4*y(T4);


dydt(M1T1) = k_inh_on_T1M1*y(M1)*y(T1) - k_inh_off_T1M1*y(M1T1);
dydt(M1T2) = k_inh_on_T2M1*y(M1)*y(T2) - k_inh_off_T2M1*y(M1T2);
dydt(M1T4) = k_inh_on_T4M1*y(M1)*y(T4) - k_inh_off_T4M1*y(M1T4);

dydt(M2T1) = k_inh_on_T1M2*y(M2)*y(T1) - k_inh_off_T1M2*y(M2T1);
dydt(M2T2) = k_inh_on_T2M2*y(M2)*y(T2) - k_inh_off_T2M2*y(M2T2);
dydt(M2T4) = k_inh_on_T4M2*y(M2)*y(T4) - k_inh_off_T4M2*y(M2T4);
```

```
dydt(M8T1) = k_inh_on_T1M8*y(M8)*y(T1) - k_inh_off_T1M8*y(M8T1);
dydt(M8T2) = k_inh_on_T2M8*y(M8)*y(T2) - k_inh_off_T2M8*y(M8T2);
dydt(M8T4) = k_inh_on_T4M8*y(M8)*y(T4) - k_inh_off_T4M8*y(M8T4);


dydt(M9T1) = k_inh_on_T1M9*y(M9)*y(T1) - k_inh_off_T1M9*y(M9T1);
dydt(M9T2) = k_inh_on_T2M9*y(M9)*y(T2) - k_inh_off_T2M9*y(M9T2);
dydt(M9T4) = k_inh_on_T4M9*y(M9)*y(T4) - k_inh_off_T4M9*y(M9T4);



dydt(iT1) = k_inact_T1*y(T1);
dydt(iT2) = k_inact_T2*y(T2);
dydt(iT4) = k_inact_T4*y(T4);


end
```

The code for all of interactions in topology 4:

```
function dydt = DQ_ODE_function5(t,y, k_on_M1C1,k_off_M1C1,
k_deg_M1C1,k_on_M1dC1, k_off_M1dC1, k_deg_M1dC1,...
                                        k_inh_on_T1M1, k_inh_on_T2M1,
k_inh_on_T4M1,...
                                        k_inh_off_T1M1, k_inh_off_T2M1,
k_inh_off_T4M1,...
                                        k_on_M2C1,k_off_M2C1,
k_deg_M2C1,k_on_M2dC1, k_off_M2dC1, k_deg_M2dC1,...
                                        k_inh_on_T1M2, k_inh_on_T2M2,
k_inh_on_T4M2,...
                                        k_inh_off_T1M2, k_inh_off_T2M2,
k_inh_off_T4M2,...
                                        k_on_M8C1,k_off_M8C1,
k_deg_M8C1,k_on_M8dC1, k_off_M8dC1, k_deg_M8dC1,...
                                        k_inh_on_T1M8, k_inh_on_T2M8,
k_inh_on_T4M8,...
                                        k_inh_off_T1M8, k_inh_off_T2M8,
k_inh_off_T4M8,...
                                        k_on_M9C1,k_off_M9C1,
k_deg_M9C1,k_on_M9dC1, k_off_M9dC1, k_deg_M9dC1,...
                                        k_inh_on_T1M9, k_inh_on_T2M9,
k_inh_on_T4M9,...
                                        k_inh_off_T1M9, k_inh_off_T2M9,
k_inh_off_T4M9,...

k_inact_M1,k_inact_M2,k_inact_M8,k_inact_M9,...
                                        k_inact_T1,k_inact_T2,k_inact_T4,...
                                        k_on_M1M1, k_off_M1M1, k_deg_M1M1,
k_dist_on_M1, k_dist_off_M1, k_dist_deg_M1,...
                                        k_on_M2M2, k_off_M2M2, k_deg_M2M2,
k_dist_on_M2, k_dist_off_M2, k_dist_deg_M2,...
                                        k_on_M8M8, k_off_M8M8, k_deg_M8M8,
k_dist_on_M8, k_dist_off_M8, k_dist_deg_M8,...
                                        k_on_M9M9, k_off_M9M9, k_deg_M9M9,
k_dist_on_M9, k_dist_off_M9, k_dist_deg_M9)

C1 = 1;
dC1 = 2;

M1 = 3;
M2 = 4;
M8 = 5;
M9 = 6;

M1C1 = 7;
M2C1 = 8;
M8C1 = 9;
M9C1 = 10;
```

```
T1 = 11;
T2 = 12;
T4 = 13;

M1T1 = 14;
M1T2 = 15;
M1T4 = 16;

M2T1 = 17;
M2T2 = 18;
M2T4 = 19;


M8T1 = 20;
M8T2 = 21;
M8T4 = 22;

M9T1 = 23;
M9T2 = 24;
M9T4 = 25;

iM1 = 26;
iM2 = 27;
iM8 = 28;
iM9 = 29;

iT1 = 30;
iT2 = 31;
iT4 = 32;

M1dC1 = 33;
M2dC1 = 34;
M8dC1 = 35;
M9dC1 = 36;

ddC1 = 37;

M1M1 = 38;
dM1 = 39;
M1iM1 = 40;


M2M2 = 41;
dM2 = 42;
M2iM2 = 43;

M8M8 = 44;
dM8 = 45;
M8iM8 = 46;


M9M9 = 47;
```

```
dM9 = 48;
M9iM9 = 49;
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%% REACTIONS %%%%%%%$%%%%%%%%%%%%%%%%%%%%%%%%%%
dydt = zeros(49,1);


dydt(C1) = - k_on_M1C1*y(M1)*y(C1) + k_off_M1C1*y(M1C1) -
k_on_M2C1*y(M2)*y(C1) + k_off_M2C1*y(M2C1) - k_on_M8C1*y(M8)*y(C1) +
k_off_M8C1*y(M8C1) - k_on_M9C1*y(M9)*y(C1) + k_off_M9C1*y(M9C1);

dydt(dC1) = 2*k_deg_M1C1*y(M1C1) - k_on_M1dC1*y(M1)*y(dC1) +
k_off_M1dC1*y(M1dC1) + 2*k_deg_M2C1*y(M2C1) - k_on_M2dC1*y(M2)*y(dC1) +
k_off_M2dC1*y(M2dC1) + 2*k_deg_M8C1*y(M8C1) - k_on_M8dC1*y(M8)*y(dC1) +
k_off_M8dC1*y(M8dC1) + 2*k_deg_M9C1*y(M9C1) - k_on_M9dC1*y(M9)*y(dC1) +
k_off_M9dC1*y(M9dC1);



dydt(M1) =   - k_on_M1C1*y(M1)*y(C1)  + k_off_M1C1*y(M1C1) +
k_deg_M1C1*y(M1C1) - k_on_M1dC1*y(M1)*y(dC1)  + k_off_M1dC1*y(M1dC1) +
k_deg_M1dC1*y(M1dC1)+...
          - 2*k_on_M1M1*y(M1)*y(M1)  + 2*k_off_M1M1*y(M1M1) +
k_deg_M1M1*y(M1M1) - k_dist_on_M1*y(M1)*y(iM1)  + k_dist_off_M1*y(M1iM1) +
k_dist_deg_M1*y(M1iM1)+...
          - k_inh_on_T1M1*y(M1)*y(T1) - k_inh_on_T2M1*y(M1)*y(T2) -
k_inh_on_T4M1*y(M1)*y(T4) +  k_inh_off_T1M1*y(M1T1) + k_inh_off_T2M1*y(M1T2) +
k_inh_off_T4M1*y(M1T4) - k_inact_M1*y(M1);

dydt(M2) =   - k_on_M2C1*y(M2)*y(C1)  + k_off_M2C1*y(M2C1) +
k_deg_M2C1*y(M2C1) - k_on_M2dC1*y(M2)*y(dC1)  + k_off_M2dC1*y(M2dC1) +
k_deg_M2dC1*y(M2dC1)+...
          - 2*k_on_M2M2*y(M2)*y(M2)  + 2*k_off_M2M2*y(M2M2) +
k_deg_M2M2*y(M2M2) - k_dist_on_M2*y(M2)*y(iM2)  + k_dist_off_M2*y(M2iM2) +
k_dist_deg_M2*y(M2iM2)+...
          - k_inh_on_T1M2*y(M2)*y(T1) - k_inh_on_T2M2*y(M2)*y(T2) -
k_inh_on_T4M2*y(M2)*y(T4) +  k_inh_off_T1M2*y(M2T1) + k_inh_off_T2M2*y(M2T2) +
k_inh_off_T4M2*y(M2T4) - k_inact_M2*y(M2);

dydt(M8) =   - k_on_M8C1*y(M8)*y(C1)  + k_off_M8C1*y(M8C1) +
k_deg_M8C1*y(M8C1) - k_on_M8dC1*y(M8)*y(dC1)  + k_off_M8dC1*y(M8dC1) +
k_deg_M8dC1*y(M8dC1)+...
          - 2*k_on_M8M8*y(M8)*y(M8)  + 2*k_off_M8M8*y(M8M8) +
k_deg_M8M8*y(M8M8) - k_dist_on_M8*y(M8)*y(iM8)  + k_dist_off_M8*y(M8iM8) +
k_dist_deg_M8*y(M8iM8)+...
          - k_inh_on_T1M8*y(M8)*y(T1) - k_inh_on_T2M8*y(M8)*y(T2) -
k_inh_on_T4M8*y(M8)*y(T4) +  k_inh_off_T1M8*y(M8T1) + k_inh_off_T2M8*y(M8T2) +
k_inh_off_T4M8*y(M8T4) - k_inact_M8*y(M8);
```

```
dydt(M9) =    - k_on_M9C1*y(M9)*y(C1)  + k_off_M9C1*y(M9C1) +
k_deg_M9C1*y(M9C1) - k_on_M9dC1*y(M9)*y(dC1)  + k_off_M9dC1*y(M9dC1) +
k_deg_M9dC1*y(M9dC1)+...
              - 2*k_on_M9M9*y(M9)*y(M9)  + 2*k_off_M9M9*y(M9M9) +
k_deg_M9M9*y(M9M9) - k_dist_on_M9*y(M9)*y(iM9)  + k_dist_off_M9*y(M9iM9) +
k_dist_deg_M9*y(M9iM9)+...
              - k_inh_on_T1M9*y(M9)*y(T1) - k_inh_on_T2M9*y(M9)*y(T2) -
k_inh_on_T4M9*y(M9)*y(T4) +  k_inh_off_T1M9*y(M9T1) + k_inh_off_T2M9*y(M9T2) +
k_inh_off_T4M9*y(M9T4) - k_inact_M9*y(M9);


dydt(M1C1) =  k_on_M1C1*y(M1)*y(C1) - k_off_M1C1*y(M1C1) - k_deg_M1C1*y(M1C1);

dydt(M2C1) =  k_on_M2C1*y(M2)*y(C1) - k_off_M2C1*y(M2C1) - k_deg_M2C1*y(M2C1);

dydt(M8C1) =  k_on_M8C1*y(M8)*y(C1) - k_off_M8C1*y(M8C1) - k_deg_M8C1*y(M8C1);

dydt(M9C1) =  k_on_M9C1*y(M9)*y(C1) - k_off_M9C1*y(M9C1) - k_deg_M9C1*y(M9C1);


dydt(M1dC1) =  k_on_M1dC1*y(M1)*y(dC1) - k_off_M1dC1*y(M1dC1) -
k_deg_M1dC1*y(M1dC1);

dydt(M2dC1) =  k_on_M2dC1*y(M2)*y(dC1) - k_off_M2dC1*y(M2dC1) -
k_deg_M2dC1*y(M2dC1);

dydt(M8dC1) =  k_on_M8dC1*y(M8)*y(dC1) - k_off_M8dC1*y(M8dC1) -
k_deg_M8dC1*y(M8dC1);

dydt(M9dC1) =  k_on_M9dC1*y(M9)*y(dC1) - k_off_M9dC1*y(M9dC1) -
k_deg_M9dC1*y(M9dC1);


dydt(ddC1) = 2*k_deg_M1dC1*y(M1dC1) + 2*k_deg_M2dC1*y(M2dC1) +
2*k_deg_M8dC1*y(M8dC1) + 2*k_deg_M9dC1*y(M9dC1);


dydt(M1M1) = k_on_M1M1*y(M1)*y(M1)  - k_off_M1M1*y(M1M1) - k_deg_M1M1*y(M1M1);
dydt(dM1) =  k_deg_M1M1*y(M1M1) + k_dist_deg_M1*y(M1iM1);
dydt(M1iM1) = k_dist_on_M1*y(M1)*y(iM1) - k_dist_off_M1*y(M1iM1) -
k_dist_deg_M1*y(M1iM1);


dydt(M2M2) = k_on_M2M2*y(M2)*y(M2)  - k_off_M2M2*y(M2M2) - k_deg_M2M2*y(M2M2);
dydt(dM2) =  k_deg_M2M2*y(M2M2) + k_dist_deg_M2*y(M2iM2);
dydt(M2iM2) = k_dist_on_M2*y(M2)*y(iM2) - k_dist_off_M2*y(M2iM2) -
k_dist_deg_M2*y(M2iM2);


dydt(M8M8) = k_on_M8M8*y(M8)*y(M8)  - k_off_M8M8*y(M8M8) - k_deg_M8M8*y(M8M8);
dydt(dM8) =  k_deg_M8M8*y(M8M8) + k_dist_deg_M8*y(M8iM8);
```

```
dydt(M8iM8) = k_dist_on_M8*y(M8)*y(iM8) - k_dist_off_M8*y(M8iM8) -
k_dist_deg_M8*y(M8iM8);


dydt(M9M9) = k_on_M9M9*y(M9)*y(M9)  - k_off_M9M9*y(M9M9) - k_deg_M9M9*y(M9M9);
dydt(dM9) =  k_deg_M9M9*y(M9M9) + k_dist_deg_M9*y(M9iM9);
dydt(M9iM9) = k_dist_on_M9*y(M9)*y(iM9) - k_dist_off_M9*y(M9iM9) -
k_dist_deg_M9*y(M9iM9);


dydt(iM1) =  k_inact_M1*y(M1) - k_dist_on_M1*y(M1)*y(iM1)  +
k_dist_off_M1*y(M1iM1);
dydt(iM2) =  k_inact_M2*y(M2) - k_dist_on_M2*y(M2)*y(iM2)  +
k_dist_off_M2*y(M2iM2);
dydt(iM8) =  k_inact_M8*y(M8) - k_dist_on_M8*y(M8)*y(iM8)  +
k_dist_off_M8*y(M8iM8);
dydt(iM9) =  k_inact_M9*y(M9) - k_dist_on_M9*y(M9)*y(iM9)  +
k_dist_off_M9*y(M9iM9);

dydt(T1) =  k_inh_off_T1M1*y(M1T1) - k_inh_on_T1M1*y(M1)*y(T1)+...
            k_inh_off_T1M2*y(M2T1) - k_inh_on_T1M2*y(M2)*y(T1)+...
            k_inh_off_T1M8*y(M8T1) - k_inh_on_T1M8*y(M8)*y(T1)+...
            k_inh_off_T1M9*y(M9T1) - k_inh_on_T1M9*y(M9)*y(T1) -
k_inact_T1*y(T1);

dydt(T2) =  k_inh_off_T2M1*y(M1T2) - k_inh_on_T2M1*y(M1)*y(T2)+...
            k_inh_off_T2M2*y(M2T2) - k_inh_on_T2M2*y(M2)*y(T2)+...
            k_inh_off_T2M8*y(M8T2) - k_inh_on_T2M8*y(M8)*y(T2)+...
            k_inh_off_T2M9*y(M9T2) - k_inh_on_T2M9*y(M9)*y(T2) -
k_inact_T2*y(T2);

dydt(T4) =  k_inh_off_T4M1*y(M1T4) - k_inh_on_T4M1*y(M1)*y(T4)+...
            k_inh_off_T4M2*y(M2T4) - k_inh_on_T4M2*y(M2)*y(T4)+...
            k_inh_off_T4M8*y(M8T4) - k_inh_on_T4M8*y(M8)*y(T4)+...
            k_inh_off_T4M9*y(M9T4) - k_inh_on_T4M9*y(M9)*y(T4) -
k_inact_T4*y(T4);


dydt(M1T1) = k_inh_on_T1M1*y(M1)*y(T1) - k_inh_off_T1M1*y(M1T1);
dydt(M1T2) = k_inh_on_T2M1*y(M1)*y(T2) - k_inh_off_T2M1*y(M1T2);
dydt(M1T4) = k_inh_on_T4M1*y(M1)*y(T4) - k_inh_off_T4M1*y(M1T4);

dydt(M2T1) = k_inh_on_T1M2*y(M2)*y(T1) - k_inh_off_T1M2*y(M2T1);
dydt(M2T2) = k_inh_on_T2M2*y(M2)*y(T2) - k_inh_off_T2M2*y(M2T2);
dydt(M2T4) = k_inh_on_T4M2*y(M2)*y(T4) - k_inh_off_T4M2*y(M2T4);

dydt(M8T1) = k_inh_on_T1M8*y(M8)*y(T1) - k_inh_off_T1M8*y(M8T1);
dydt(M8T2) = k_inh_on_T2M8*y(M8)*y(T2) - k_inh_off_T2M8*y(M8T2);
dydt(M8T4) = k_inh_on_T4M8*y(M8)*y(T4) - k_inh_off_T4M8*y(M8T4);

dydt(M9T1) = k_inh_on_T1M9*y(M9)*y(T1) - k_inh_off_T1M9*y(M9T1);
dydt(M9T2) = k_inh_on_T2M9*y(M9)*y(T2) - k_inh_off_T2M9*y(M9T2);
```

```
dydt(M9T4) = k_inh_on_T4M9*y(M9)*y(T4) - k_inh_off_T4M9*y(M9T4);


dydt(iT1) = k_inact_T1*y(T1);
dydt(iT2) = k_inact_T2*y(T2);
dydt(iT4) = k_inact_T4*y(T4);


end
```

The code for all of interactions in topology 5:

```
function dydt = DQ_ODE_function7(t,y, k_on_M1C1,k_off_M1C1,
k_deg_M1C1,k_on_M1dC1, k_off_M1dC1, k_deg_M1dC1,...
                                         k_inh_on_T1M1, k_inh_on_T2M1,
k_inh_on_T4M1,...
                                         k_inh_off_T1M1, k_inh_off_T2M1,
k_inh_off_T4M1,...
                                         k_on_M2C1,k_off_M2C1,
k_deg_M2C1,k_on_M2dC1, k_off_M2dC1, k_deg_M2dC1,...
                                         k_inh_on_T1M2, k_inh_on_T2M2,
k_inh_on_T4M2,...
                                         k_inh_off_T1M2, k_inh_off_T2M2,
k_inh_off_T4M2,...
                                         k_on_M8C1,k_off_M8C1,
k_deg_M8C1,k_on_M8dC1, k_off_M8dC1, k_deg_M8dC1,...
                                         k_inh_on_T1M8, k_inh_on_T2M8,
k_inh_on_T4M8,...
                                         k_inh_off_T1M8, k_inh_off_T2M8,
k_inh_off_T4M8,...
                                         k_on_M9C1,k_off_M9C1,
k_deg_M9C1,k_on_M9dC1, k_off_M9dC1, k_deg_M9dC1,...
                                         k_inh_on_T1M9, k_inh_on_T2M9,
k_inh_on_T4M9,...
                                         k_inh_off_T1M9, k_inh_off_T2M9,
k_inh_off_T4M9,...

k_inact_M1,k_inact_M2,k_inact_M8,k_inact_M9,...
                                         k_inact_T1,k_inact_T2,k_inact_T4,...
                                         k_on_M1M1, k_off_M1M1, k_deg_M1M1,
k_dist_on_M1, k_dist_off_M1, k_dist_deg_M1,...
                                         k_on_M2M2, k_off_M2M2, k_deg_M2M2,
k_dist_on_M2, k_dist_off_M2, k_dist_deg_M2,...
                                         k_on_M8M8, k_off_M8M8, k_deg_M8M8,
k_dist_on_M8, k_dist_off_M8, k_dist_deg_M8,...
                                         k_on_M9M9, k_off_M9M9, k_deg_M9M9,
k_dist_on_M9, k_dist_off_M9, k_dist_deg_M9,...
                                         k_on_M1M2, k_off_M1M2, k_deg_M1M2,...
                                         k_on_M1M8, k_off_M1M8, k_deg_M1M8,...
                                         k_on_M1M9, k_off_M1M9, k_deg_M1M9,...
                                         k_on_M2M1, k_off_M2M1, k_deg_M2M1,...
                                         k_on_M2M8, k_off_M2M8, k_deg_M2M8,...
                                         k_on_M2M9, k_off_M2M9, k_deg_M2M9,...
                                         k_on_M8M1, k_off_M8M1, k_deg_M8M1,...
                                         k_on_M8M2, k_off_M8M2, k_deg_M8M2,...
                                         k_on_M8M9, k_off_M8M9, k_deg_M8M9,...
                                         k_on_M9M1, k_off_M9M1, k_deg_M9M1,...
                                         k_on_M9M2, k_off_M9M2, k_deg_M9M2,...
                                         k_on_M9M8, k_off_M9M8, k_deg_M9M8,...
```

```
k_dist_deg_M1iM2,...                    k_dist_on_M1iM2, k_dist_off_M1iM2,

k_dist_deg_M1iM8,...                     k_dist_on_M1iM8, k_dist_off_M1iM8,

k_dist_deg_M1iM9,...                     k_dist_on_M1iM9, k_dist_off_M1iM9,

k_dist_deg_M2iM1,...                     k_dist_on_M2iM1, k_dist_off_M2iM1,

k_dist_deg_M2iM8,...                     k_dist_on_M2iM8, k_dist_off_M2iM8,

k_dist_deg_M2iM9,...                     k_dist_on_M2iM9, k_dist_off_M2iM9,

k_dist_deg_M8iM1,...                     k_dist_on_M8iM1, k_dist_off_M8iM1,

k_dist_deg_M8iM2,...                     k_dist_on_M8iM2, k_dist_off_M8iM2,

k_dist_deg_M8iM9,...                     k_dist_on_M8iM9, k_dist_off_M8iM9,

k_dist_deg_M9iM1,...                     k_dist_on_M9iM1, k_dist_off_M9iM1,

k_dist_deg_M9iM2,...                     k_dist_on_M9iM2, k_dist_off_M9iM2,

k_dist_deg_M9iM8,...                     k_dist_on_M9iM8, k_dist_off_M9iM8,

k_TIMP1_T1, k_TIMP1_T2, k_TIMP1_T4,...   k_TIMP2_T1, k_TIMP2_T2, k_TIMP2_T4,

k_dist_on_T4,...                         k_dist_on_T1, k_dist_on_T2,

k_dist_off_T4)                           k_dist_off_T1, k_dist_off_T2,

C1 = 1;
dC1 = 2;

M1 = 3;
M2 = 4;
M8 = 5;
M9 = 6;

M1C1 = 7;
M2C1 = 8;
M8C1 = 9;
M9C1 = 10;


T1 = 11;
T2 = 12;
T4 = 13;

M1T1 = 14;
M1T2 = 15;
M1T4 = 16;
```

```
M2T1 = 17;
M2T2 = 18;
M2T4 = 19;


M8T1 = 20;
M8T2 = 21;
M8T4 = 22;

M9T1 = 23;
M9T2 = 24;
M9T4 = 25;

iM1 = 26;
iM2 = 27;
iM8 = 28;
iM9 = 29;

iT1 = 30;
iT2 = 31;
iT4 = 32;

M1dC1 = 33;
M2dC1 = 34;
M8dC1 = 35;
M9dC1 = 36;

ddC1 = 37;

M1M1 = 38;
dM1 = 39;
M1iM1 = 40;


M2M2 = 41;
dM2 = 42;
M2iM2 = 43;

M8M8 = 44;
dM8 = 45;
M8iM8 = 46;

M9M9 = 47;
dM9 = 48;
M9iM9 = 49;

M1M2 = 50;
M1M8 = 51;
M1M9 = 52;
M2M8 = 53;
M2M9 = 54;
```

```
M8M9 = 55;


M1iM2 = 56;
M1iM8 = 57;
M1iM9 = 58;

M2iM1 = 59;
M2iM8 = 60;
M2iM9 = 61;

M8iM1 = 62;
M8iM2 = 63;
M8iM9 = 64;

M9iM1 = 65;
M9iM2 = 66;
M9iM8 = 67;


T1iM1 = 68;
T2iM1 = 69;
T4iM1 = 70;

T1iM2 = 71;
T2iM2 = 72;
T4iM2 = 73;

T1iM8 = 74;
T2iM8 = 75;
T4iM8 = 76;

T1iM9 = 77;
T2iM9 = 78;
T4iM9 = 79;


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% REACTIONS %%%%%%%%$%%%%%%%%%%%%%%%%%%%%%%%%%
dydt = zeros(79,1);


dydt(C1) = - k_on_M1C1*y(M1)*y(C1) + k_off_M1C1*y(M1C1) -
k_on_M2C1*y(M2)*y(C1) + k_off_M2C1*y(M2C1) - k_on_M8C1*y(M8)*y(C1) +
k_off_M8C1*y(M8C1) - k_on_M9C1*y(M9)*y(C1) + k_off_M9C1*y(M9C1);

dydt(dC1) = 2*k_deg_M1C1*y(M1C1) - k_on_M1dC1*y(M1)*y(dC1) +
k_off_M1dC1*y(M1dC1) + 2*k_deg_M2C1*y(M2C1) - k_on_M2dC1*y(M2)*y(dC1) +
k_off_M2dC1*y(M2dC1) + 2*k_deg_M8C1*y(M8C1) - k_on_M8dC1*y(M8)*y(dC1) +
k_off_M8dC1*y(M8dC1) + 2*k_deg_M9C1*y(M9C1) - k_on_M9dC1*y(M9)*y(dC1) +
k_off_M9dC1*y(M9dC1);
```

dydt(M1) =   - k_on_M1C1*y(M1)*y(C1) + k_off_M1C1*y(M1C1) + k_deg_M1C1*y(M1C1) - k_on_M1dC1*y(M1)*y(dC1) + k_off_M1dC1*y(M1dC1) + k_deg_M1dC1*y(M1dC1)+...
          - 2*k_on_M1M1*y(M1)*y(M1) + 2*k_off_M1M1*y(M1M1) + k_deg_M1M1*y(M1M1) - k_inact_M1*y(M1) - k_dist_on_M1*y(M1)*y(iM1) + k_dist_off_M1*y(M1iM1) + k_dist_deg_M1*y(M1iM1)+...
          - k_inh_on_T1M1*y(M1)*y(T1) - k_inh_on_T2M1*y(M1)*y(T2) - k_inh_on_T4M1*y(M1)*y(T4) + k_inh_off_T1M1*y(M1T1) + k_inh_off_T2M1*y(M1T2) + k_inh_off_T4M1*y(M1T4)+...
          k_TIMP2_T1*y(M1T1) + k_TIMP2_T2*y(M1T2) + k_TIMP2_T4*y(M1T4)+...
          - k_on_M1M2*y(M1)*y(M2) + k_off_M1M2*y(M1M2) + k_deg_M1M2*y(M1M2) - k_on_M1M8*y(M1)*y(M8) + k_off_M1M8*y(M1M8) + k_deg_M1M8*y(M1M8) - k_on_M1M9*y(M1)*y(M9) + k_off_M1M9*y(M1M9) + k_deg_M1M9*y(M1M9)+...
          - k_on_M2M1*y(M1)*y(M2) + k_off_M2M1*y(M1M2) - k_on_M8M1*y(M1)*y(M8) + k_off_M8M1*y(M1M8) - k_on_M9M1*y(M1)*y(M9) + k_off_M9M1*y(M1M9)+...
          - k_dist_on_M1iM2*y(M1)*y(iM2) + k_dist_off_M1iM2*y(M1iM2) + k_dist_deg_M1iM2*y(M1iM2) - k_dist_on_M1iM8*y(M1)*y(iM8) + k_dist_off_M1iM8*y(M1iM8) + k_dist_deg_M1iM8*y(M1iM8) - k_dist_on_M1iM9*y(M1)*y(iM9) + k_dist_off_M1iM9*y(M1iM9) + k_dist_deg_M1iM9*y(M1iM9);

dydt(M2) =   - k_on_M2C1*y(M2)*y(C1) + k_off_M2C1*y(M2C1) + k_deg_M2C1*y(M2C1) - k_on_M2dC1*y(M2)*y(dC1) + k_off_M2dC1*y(M2dC1) + k_deg_M2dC1*y(M2dC1)+...
          - 2*k_on_M2M2*y(M2)*y(M2) + 2*k_off_M2M2*y(M2M2) + k_deg_M2M2*y(M2M2) - k_inact_M2*y(M2) - k_dist_on_M2*y(M2)*y(iM2) + k_dist_off_M2*y(M2iM2) + k_dist_deg_M2*y(M2iM2)+...
          - k_inh_on_T1M2*y(M2)*y(T1) - k_inh_on_T2M2*y(M2)*y(T2) - k_inh_on_T4M2*y(M2)*y(T4) + k_inh_off_T1M2*y(M2T1) + k_inh_off_T2M2*y(M2T2) + k_inh_off_T4M2*y(M2T4)+...
          k_TIMP2_T1*y(M2T1) + k_TIMP2_T2*y(M2T2) + k_TIMP2_T4*y(M2T4)+...
          - k_on_M2M1*y(M1)*y(M2) + k_off_M2M1*y(M1M2) + k_deg_M2M1*y(M1M2) - k_on_M2M8*y(M2)*y(M8) + k_off_M2M8*y(M2M8) + k_deg_M2M8*y(M2M8) - k_on_M2M9*y(M2)*y(M9) + k_off_M2M9*y(M2M9) + k_deg_M2M9*y(M2M9)+...
          - k_on_M1M2*y(M1)*y(M2) + k_off_M1M2*y(M1M2) - k_on_M8M2*y(M2)*y(M8) + k_off_M8M2*y(M2M8) - k_on_M9M2*y(M2)*y(M9) + k_off_M9M2*y(M2M9)+...
          - k_dist_on_M2iM1*y(M2)*y(iM1) + k_dist_off_M2iM1*y(M2iM1) + k_dist_deg_M2iM1*y(M2iM1) - k_dist_on_M2iM8*y(M2)*y(iM8) + k_dist_off_M2iM8*y(M2iM8) + k_dist_deg_M2iM8*y(M2iM8) - k_dist_on_M2iM9*y(M2)*y(iM9) + k_dist_off_M2iM9*y(M2iM9) + k_dist_deg_M2iM9*y(M2iM9);

dydt(M8) =   - k_on_M8C1*y(M8)*y(C1) + k_off_M8C1*y(M8C1) + k_deg_M8C1*y(M8C1) - k_on_M8dC1*y(M8)*y(dC1) + k_off_M8dC1*y(M8dC1) + k_deg_M8dC1*y(M8dC1)+...

```
              - 2*k_on_M8M8*y(M8)*y(M8)  + 2*k_off_M8M8*y(M8M8) +
k_deg_M8M8*y(M8M8) - k_inact_M8*y(M8) - k_dist_on_M8*y(M8)*y(iM8)  +
k_dist_off_M8*y(M8iM8) + k_dist_deg_M8*y(M8iM8)+...
              - k_inh_on_T1M8*y(M8)*y(T1) - k_inh_on_T2M8*y(M8)*y(T2) -
k_inh_on_T4M8*y(M8)*y(T4) +  k_inh_off_T1M8*y(M8T1) + k_inh_off_T2M8*y(M8T2) +
k_inh_off_T4M8*y(M8T4)+...
              k_TIMP2_T1*y(M8T1) + k_TIMP2_T2*y(M8T2) +
k_TIMP2_T4*y(M8T4)+...
              - k_on_M8M1*y(M1)*y(M8) + k_off_M8M1*y(M1M8) + k_deg_M8M1*y(M1M8)
- k_on_M8M2*y(M2)*y(M8) + k_off_M8M2*y(M2M8) + k_deg_M8M2*y(M2M8) -
k_on_M8M9*y(M8)*y(M9) + k_off_M8M9*y(M8M9) + k_deg_M8M9*y(M8M9)+...
              - k_on_M1M8*y(M1)*y(M8) + k_off_M1M8*y(M1M8) -
k_on_M2M8*y(M2)*y(M8) + k_off_M2M8*y(M2M8) - k_on_M9M8*y(M8)*y(M9) +
k_off_M9M8*y(M8M9)+...
              - k_dist_on_M8iM1*y(M8)*y(iM1)  + k_dist_off_M8iM1*y(M8iM1) +
k_dist_deg_M8iM1*y(M8iM1) - k_dist_on_M8iM2*y(M8)*y(iM2)   +
k_dist_off_M8iM2*y(M8iM2) + k_dist_deg_M8iM2*y(M8iM2) -
k_dist_on_M8iM9*y(M8)*y(iM9)  + k_dist_off_M8iM9*y(M8iM9) +
k_dist_deg_M8iM9*y(M8iM9);


dydt(M9) =   - k_on_M9C1*y(M9)*y(C1)  + k_off_M9C1*y(M9C1) +
k_deg_M9C1*y(M9C1) - k_on_M9dC1*y(M9)*y(dC1)  + k_off_M9dC1*y(M9dC1) +
k_deg_M9dC1*y(M9dC1)+...
              - 2*k_on_M9M9*y(M9)*y(M9)  + 2*k_off_M9M9*y(M9M9) +
k_deg_M9M9*y(M9M9) - k_inact_M9*y(M9) - k_dist_on_M9*y(M9)*y(iM9)  +
k_dist_off_M9*y(M9iM9) + k_dist_deg_M9*y(M9iM9)+...
              - k_inh_on_T1M9*y(M9)*y(T1) - k_inh_on_T2M9*y(M9)*y(T2) -
k_inh_on_T4M9*y(M9)*y(T4) +  k_inh_off_T1M9*y(M9T1) + k_inh_off_T2M9*y(M9T2) +
k_inh_off_T4M9*y(M9T4)+...
              k_TIMP2_T1*y(M9T1) + k_TIMP2_T2*y(M9T2) +
k_TIMP2_T4*y(M9T4)+...
              - k_on_M9M1*y(M1)*y(M9) + k_off_M9M1*y(M1M9) + k_deg_M9M1*y(M1M9)
- k_on_M9M2*y(M2)*y(M9) + k_off_M9M2*y(M2M9) + k_deg_M9M2*y(M2M9) -
k_on_M9M8*y(M8)*y(M9) + k_off_M9M8*y(M8M9) + k_deg_M9M8*y(M8M9)+...
              - k_on_M1M9*y(M1)*y(M9) + k_off_M1M9*y(M1M9) -
k_on_M2M9*y(M2)*y(M9) + k_off_M2M9*y(M2M9) - k_on_M8M9*y(M8)*y(M9) +
k_off_M8M9*y(M8M9)+...
              - k_dist_on_M9iM1*y(M9)*y(iM1)  + k_dist_off_M9iM1*y(M9iM1) +
k_dist_deg_M9iM1*y(M9iM1) - k_dist_on_M9iM2*y(M9)*y(iM2)   +
k_dist_off_M9iM2*y(M9iM2) + k_dist_deg_M9iM2*y(M9iM2) -
k_dist_on_M9iM8*y(M9)*y(iM8)  + k_dist_off_M9iM8*y(M9iM8) +
k_dist_deg_M9iM8*y(M9iM8);



dydt(M1C1) =  k_on_M1C1*y(M1)*y(C1) - k_off_M1C1*y(M1C1) - k_deg_M1C1*y(M1C1);

dydt(M2C1) =  k_on_M2C1*y(M2)*y(C1) - k_off_M2C1*y(M2C1) - k_deg_M2C1*y(M2C1);

dydt(M8C1) =  k_on_M8C1*y(M8)*y(C1) - k_off_M8C1*y(M8C1) - k_deg_M8C1*y(M8C1);
```

```
dydt(M9C1) =  k_on_M9C1*y(M9)*y(C1) - k_off_M9C1*y(M9C1) - k_deg_M9C1*y(M9C1);



dydt(M1dC1) =  k_on_M1dC1*y(M1)*y(dC1) - k_off_M1dC1*y(M1dC1) -
k_deg_M1dC1*y(M1dC1);

dydt(M2dC1) =  k_on_M2dC1*y(M2)*y(dC1) - k_off_M2dC1*y(M2dC1) -
k_deg_M2dC1*y(M2dC1);

dydt(M8dC1) =  k_on_M8dC1*y(M8)*y(dC1) - k_off_M8dC1*y(M8dC1) -
k_deg_M8dC1*y(M8dC1);

dydt(M9dC1) =  k_on_M9dC1*y(M9)*y(dC1) - k_off_M9dC1*y(M9dC1) -
k_deg_M9dC1*y(M9dC1);



dydt(ddC1) = 2*k_deg_M1dC1*y(M1dC1) + 2*k_deg_M2dC1*y(M2dC1) +
2*k_deg_M8dC1*y(M8dC1) + 2*k_deg_M9dC1*y(M9dC1);



dydt(M1M1) = k_on_M1M1*y(M1)*y(M1)  - k_off_M1M1*y(M1M1) - k_deg_M1M1*y(M1M1);

dydt(dM1) =  k_deg_M1M1*y(M1M1) + k_dist_deg_M1*y(M1iM1) + k_deg_M2M1*y(M1M2)
+ k_deg_M8M1*y(M1M8) + k_deg_M9M1*y(M1M9) + k_dist_deg_M2iM1*y(M2iM1) +
k_dist_deg_M8iM1*y(M8iM1) + k_dist_deg_M9iM1*y(M9iM1);

dydt(iM1) =  k_inact_M1*y(M1) - k_dist_on_M1*y(M1)*y(iM1)  +
k_dist_off_M1*y(M1iM1) + k_TIMP1_T1*y(M1T1) + k_TIMP1_T2*y(M1T2) +
k_TIMP1_T4*y(M1T4)+...
              - k_dist_on_T1*y(T1)*y(iM1) - k_dist_on_T2*y(T2)*y(iM1) -
k_dist_on_T4*y(T4)*y(iM1)+...
              + k_dist_off_T1*y(T1iM1) + k_dist_off_T2*y(T2iM1) +
k_dist_off_T4*y(T4iM1)+...
              - k_dist_on_M2iM1*y(M2)*y(iM1) + k_dist_off_M2iM1*y(M2iM1) -
k_dist_on_M8iM1*y(M8)*y(iM1) + k_dist_off_M8iM1*y(M8iM1) -
k_dist_on_M9iM1*y(M9)*y(iM1) + k_dist_off_M9iM1*y(M9iM1);

dydt(M1iM1) = k_dist_on_M1*y(M1)*y(iM1) - k_dist_off_M1*y(M1iM1) -
k_dist_deg_M1*y(M1iM1);


dydt(M1M2) = k_on_M1M2*y(M1)*y(M2)  - k_off_M1M2*y(M1M2) - k_deg_M1M2*y(M1M2)
+ k_on_M2M1*y(M1)*y(M2)  - k_off_M2M1*y(M1M2) - k_deg_M2M1*y(M1M2);
dydt(M1M8) = k_on_M1M8*y(M1)*y(M8)  - k_off_M1M8*y(M1M8) - k_deg_M1M8*y(M1M8)
+ k_on_M8M1*y(M1)*y(M8)  - k_off_M8M1*y(M1M8) - k_deg_M8M1*y(M1M8);
dydt(M1M9) = k_on_M1M9*y(M1)*y(M9)  - k_off_M1M9*y(M1M9) - k_deg_M1M9*y(M1M9)
+ k_on_M9M1*y(M1)*y(M9)  - k_off_M9M1*y(M1M9) - k_deg_M9M1*y(M1M9);
dydt(M2M8) = k_on_M2M8*y(M2)*y(M8)  - k_off_M2M8*y(M2M8) - k_deg_M2M8*y(M2M8)
+ k_on_M8M2*y(M2)*y(M8)  - k_off_M8M2*y(M2M8) - k_deg_M8M2*y(M2M8);
```

```
dydt(M2M9) = k_on_M2M9*y(M2)*y(M9)  - k_off_M2M9*y(M2M9) - k_deg_M2M9*y(M2M9)
+ k_on_M9M2*y(M2)*y(M9)  - k_off_M9M2*y(M2M9) - k_deg_M9M2*y(M2M9);
dydt(M8M9) = k_on_M8M9*y(M8)*y(M9)  - k_off_M8M9*y(M8M9) - k_deg_M8M9*y(M8M9)
+ k_on_M9M8*y(M8)*y(M9)  - k_off_M9M8*y(M8M9) - k_deg_M9M8*y(M8M9);




dydt(M1iM2) = k_dist_on_M1iM2*y(M1)*y(iM2) - k_dist_off_M1iM2*y(M1iM2) -
k_dist_deg_M1iM2*y(M1iM2);
dydt(M1iM8) = k_dist_on_M1iM8*y(M1)*y(iM8) - k_dist_off_M1iM8*y(M1iM8) -
k_dist_deg_M1iM8*y(M1iM8);
dydt(M1iM9) = k_dist_on_M1iM9*y(M1)*y(iM9) - k_dist_off_M1iM9*y(M1iM9) -
k_dist_deg_M1iM9*y(M1iM9);




dydt(M2M2) = k_on_M2M2*y(M2)*y(M2)  - k_off_M2M2*y(M2M2) - k_deg_M2M2*y(M2M2);

dydt(dM2) =  k_deg_M2M2*y(M2M2) + k_dist_deg_M2*y(M2iM2) + k_deg_M1M2*y(M1M2)
+ k_deg_M8M2*y(M2M8) + k_deg_M9M2*y(M2M9) + k_dist_deg_M1iM2*y(M1iM2) +
k_dist_deg_M8iM2*y(M8iM2) + k_dist_deg_M9iM2*y(M9iM2);

dydt(iM2) =  k_inact_M2*y(M2) - k_dist_on_M2*y(M2)*y(iM2)  +
k_dist_off_M2*y(M2iM2) + k_TIMP1_T1*y(M2T1) + k_TIMP1_T2*y(M2T2) +
k_TIMP1_T4*y(M2T4)+...
            - k_dist_on_T1*y(T1)*y(iM2) - k_dist_on_T2*y(T2)*y(iM2) -
k_dist_on_T4*y(T4)*y(iM2)+...
            + k_dist_off_T1*y(T1iM2) + k_dist_off_T2*y(T2iM2) +
k_dist_off_T4*y(T4iM2)+...
            - k_dist_on_M1iM2*y(M1)*y(iM2) + k_dist_off_M1iM2*y(M1iM2) -
k_dist_on_M8iM2*y(M8)*y(iM2) + k_dist_off_M8iM2*y(M8iM2) -
k_dist_on_M9iM2*y(M9)*y(iM2) + k_dist_off_M9iM2*y(M9iM2);

dydt(M2iM2) = k_dist_on_M2*y(M2)*y(iM2) - k_dist_off_M2*y(M2iM2) -
k_dist_deg_M2*y(M2iM2);

dydt(M2iM1) = k_dist_on_M2iM1*y(M2)*y(iM1) - k_dist_off_M2iM1*y(M2iM1) -
k_dist_deg_M2iM1*y(M2iM1);
dydt(M2iM8) = k_dist_on_M2iM8*y(M2)*y(iM8) - k_dist_off_M2iM8*y(M2iM8) -
k_dist_deg_M2iM8*y(M2iM8);
dydt(M2iM9) = k_dist_on_M2iM9*y(M2)*y(iM9) - k_dist_off_M2iM9*y(M2iM9) -
k_dist_deg_M2iM9*y(M2iM9);




dydt(M8M8) = k_on_M8M8*y(M8)*y(M8)  - k_off_M8M8*y(M8M8) - k_deg_M8M8*y(M8M8);

dydt(dM8) =  k_deg_M8M8*y(M8M8) + k_dist_deg_M8*y(M8iM8) + k_deg_M1M8*y(M1M8)
+ k_deg_M2M8*y(M2M8) + k_deg_M9M8*y(M8M9) + k_dist_deg_M1iM8*y(M1iM8) +
k_dist_deg_M2iM8*y(M2iM8) + k_dist_deg_M9iM8*y(M9iM8);
```

```
dydt(iM8) =  k_inact_M8*y(M8) - k_dist_on_M8*y(M8)*y(iM8)  +
k_dist_off_M8*y(M8iM8) + k_TIMP1_T1*y(M8T1) + k_TIMP1_T2*y(M8T2) +
k_TIMP1_T4*y(M8T4)+...
                - k_dist_on_T1*y(T1)*y(iM8) - k_dist_on_T2*y(T2)*y(iM8) -
k_dist_on_T4*y(T4)*y(iM8)+...
                + k_dist_off_T1*y(T1iM8) + k_dist_off_T2*y(T2iM8) +
k_dist_off_T4*y(T4iM8)+...
                - k_dist_on_M1iM8*y(M1)*y(iM8) + k_dist_off_M1iM8*y(M1iM8) -
k_dist_on_M2iM8*y(M2)*y(iM8) + k_dist_off_M2iM8*y(M2iM8) -
k_dist_on_M9iM8*y(M9)*y(iM8) + k_dist_off_M9iM8*y(M9iM8);

dydt(M8iM8) = k_dist_on_M8*y(M8)*y(iM8) - k_dist_off_M8*y(M8iM8) -
k_dist_deg_M8*y(M8iM8);

dydt(M8iM1) = k_dist_on_M8iM1*y(M8)*y(iM1) - k_dist_off_M8iM1*y(M8iM1) -
k_dist_deg_M8iM1*y(M8iM1);
dydt(M8iM2) = k_dist_on_M8iM2*y(M8)*y(iM2) - k_dist_off_M8iM2*y(M8iM2) -
k_dist_deg_M8iM2*y(M8iM2);
dydt(M8iM9) = k_dist_on_M8iM9*y(M8)*y(iM9) - k_dist_off_M8iM9*y(M8iM9) -
k_dist_deg_M8iM9*y(M8iM9);

dydt(M9M9) = k_on_M9M9*y(M9)*y(M9)  - k_off_M9M9*y(M9M9) - k_deg_M9M9*y(M9M9);

dydt(dM9) =  k_deg_M9M9*y(M9M9) + k_dist_deg_M9*y(M9iM9) + k_deg_M1M9*y(M1M9)
+ k_deg_M2M9*y(M2M9) + k_deg_M8M9*y(M8M9) + k_dist_deg_M1iM9*y(M1iM9) +
k_dist_deg_M2iM9*y(M2iM9) + k_dist_deg_M8iM9*y(M8iM9);

dydt(iM9) =  k_inact_M9*y(M9) - k_dist_on_M9*y(M9)*y(iM9)  +
k_dist_off_M9*y(M9iM9) + k_TIMP1_T1*y(M9T1) + k_TIMP1_T2*y(M9T2) +
k_TIMP1_T4*y(M9T4)+...
                - k_dist_on_T1*y(T1)*y(iM9) - k_dist_on_T2*y(T2)*y(iM9) -
k_dist_on_T4*y(T4)*y(iM9)+...
                + k_dist_off_T1*y(T1iM9) + k_dist_off_T2*y(T2iM9) +
k_dist_off_T4*y(T4iM9)+...
                - k_dist_on_M1iM9*y(M1)*y(iM9) + k_dist_off_M1iM9*y(M1iM9) -
k_dist_on_M2iM9*y(M2)*y(iM9) + k_dist_off_M2iM9*y(M2iM9) -
k_dist_on_M8iM9*y(M8)*y(iM9) + k_dist_off_M8iM9*y(M8iM9);


dydt(M9iM9) = k_dist_on_M9*y(M9)*y(iM9) - k_dist_off_M9*y(M9iM9) -
k_dist_deg_M9*y(M9iM9);

dydt(M9iM1) = k_dist_on_M9iM1*y(M9)*y(iM1) - k_dist_off_M9iM1*y(M9iM1) -
k_dist_deg_M9iM1*y(M9iM1);
dydt(M9iM2) = k_dist_on_M9iM2*y(M9)*y(iM2) - k_dist_off_M9iM2*y(M9iM2) -
k_dist_deg_M9iM2*y(M9iM2);
dydt(M9iM8) = k_dist_on_M9iM8*y(M9)*y(iM8) - k_dist_off_M9iM8*y(M9iM8) -
k_dist_deg_M9iM8*y(M9iM8);

dydt(T1) =  k_inh_off_T1M1*y(M1T1) - k_inh_on_T1M1*y(M1)*y(T1) +
k_TIMP1_T1*y(M1T1) - k_dist_on_T1*y(T1)*y(iM1) + k_dist_off_T1*y(T1iM1)+...
```

```
            k_inh_off_T1M2*y(M2T1) - k_inh_on_T1M2*y(M2)*y(T1) +
k_TIMP1_T1*y(M2T1) - k_dist_on_T1*y(T1)*y(iM2) + k_dist_off_T1*y(T1iM2)+...
            k_inh_off_T1M8*y(M8T1) - k_inh_on_T1M8*y(M8)*y(T1) +
k_TIMP1_T1*y(M8T1) - k_dist_on_T1*y(T1)*y(iM8) + k_dist_off_T1*y(T1iM8)+...
            k_inh_off_T1M9*y(M9T1) - k_inh_on_T1M9*y(M9)*y(T1) +
k_TIMP1_T1*y(M9T1) - k_dist_on_T1*y(T1)*y(iM9) + k_dist_off_T1*y(T1iM9) -
k_inact_T1*y(T1);


dydt(T2) =  k_inh_off_T2M1*y(M1T2) - k_inh_on_T2M1*y(M1)*y(T2) +
k_TIMP1_T2*y(M1T2) - k_dist_on_T2*y(T2)*y(iM1) + k_dist_off_T2*y(T2iM1)+...
            k_inh_off_T2M2*y(M2T2) - k_inh_on_T2M2*y(M2)*y(T2) +
k_TIMP1_T2*y(M2T2) - k_dist_on_T2*y(T2)*y(iM2) + k_dist_off_T2*y(T2iM2)+...
            k_inh_off_T2M8*y(M8T2) - k_inh_on_T2M8*y(M8)*y(T2) +
k_TIMP1_T2*y(M8T2) - k_dist_on_T2*y(T2)*y(iM8) + k_dist_off_T2*y(T2iM8)+...
            k_inh_off_T2M9*y(M9T2) - k_inh_on_T2M9*y(M9)*y(T2) +
k_TIMP1_T2*y(M9T2) - k_dist_on_T2*y(T2)*y(iM9) + k_dist_off_T2*y(T2iM9) -
k_inact_T2*y(T2);

dydt(T4) =  k_inh_off_T4M1*y(M1T4) - k_inh_on_T4M1*y(M1)*y(T4) +
k_TIMP1_T4*y(M1T4) - k_dist_on_T4*y(T4)*y(iM1) + k_dist_off_T4*y(T4iM1)+...
            k_inh_off_T4M2*y(M2T4) - k_inh_on_T4M2*y(M2)*y(T4) +
k_TIMP1_T4*y(M2T4) - k_dist_on_T4*y(T4)*y(iM2) + k_dist_off_T4*y(T4iM2)+...
            k_inh_off_T4M8*y(M8T4) - k_inh_on_T4M8*y(M8)*y(T4) +
k_TIMP1_T4*y(M8T4) - k_dist_on_T4*y(T4)*y(iM8) + k_dist_off_T4*y(T4iM8)+...
            k_inh_off_T4M9*y(M9T4) - k_inh_on_T4M9*y(M9)*y(T4) +
k_TIMP1_T4*y(M9T4) - k_dist_on_T4*y(T4)*y(iM9) + k_dist_off_T4*y(T4iM9) -
k_inact_T4*y(T4);


dydt(M1T1) = k_inh_on_T1M1*y(M1)*y(T1) - k_inh_off_T1M1*y(M1T1) -
k_TIMP1_T1*y(M1T1) - k_TIMP2_T1*y(M1T1);
dydt(M1T2) = k_inh_on_T2M1*y(M1)*y(T2) - k_inh_off_T2M1*y(M1T2) -
k_TIMP1_T2*y(M1T2) - k_TIMP2_T2*y(M1T2);
dydt(M1T4) = k_inh_on_T4M1*y(M1)*y(T4) - k_inh_off_T4M1*y(M1T4) -
k_TIMP1_T4*y(M1T4) - k_TIMP2_T4*y(M1T4);

dydt(M2T1) = k_inh_on_T1M2*y(M2)*y(T1) - k_inh_off_T1M2*y(M2T1) -
k_TIMP1_T1*y(M2T1) - k_TIMP2_T1*y(M2T1);
dydt(M2T2) = k_inh_on_T2M2*y(M2)*y(T2) - k_inh_off_T2M2*y(M2T2) -
k_TIMP1_T2*y(M2T2) - k_TIMP2_T2*y(M2T2);
dydt(M2T4) = k_inh_on_T4M2*y(M2)*y(T4) - k_inh_off_T4M2*y(M2T4) -
k_TIMP1_T4*y(M2T4) - k_TIMP2_T4*y(M2T4);

dydt(M8T1) = k_inh_on_T1M8*y(M8)*y(T1) - k_inh_off_T1M8*y(M8T1) -
k_TIMP1_T1*y(M8T1) - k_TIMP2_T1*y(M8T1);
dydt(M8T2) = k_inh_on_T2M8*y(M8)*y(T2) - k_inh_off_T2M8*y(M8T2) -
k_TIMP1_T2*y(M8T2) - k_TIMP2_T2*y(M8T2);
dydt(M8T4) = k_inh_on_T4M8*y(M8)*y(T4) - k_inh_off_T4M8*y(M8T4) -
k_TIMP1_T4*y(M8T4) - k_TIMP2_T4*y(M8T4);
```

```
dydt(M9T1) = k_inh_on_T1M9*y(M9)*y(T1) - k_inh_off_T1M9*y(M9T1) -
k_TIMP1_T1*y(M9T1) - k_TIMP2_T1*y(M9T1);
dydt(M9T2) = k_inh_on_T2M9*y(M9)*y(T2) - k_inh_off_T2M9*y(M9T2) -
k_TIMP1_T2*y(M9T2) - k_TIMP2_T2*y(M9T2);
dydt(M9T4) = k_inh_on_T4M9*y(M9)*y(T4) - k_inh_off_T4M9*y(M9T4) -
k_TIMP1_T4*y(M9T4) - k_TIMP2_T4*y(M9T4);


dydt(iT1) =  k_TIMP2_T1*y(M1T1) + k_inact_T1*y(T1) + k_TIMP2_T1*y(M2T1) +
k_TIMP2_T1*y(M8T1) + k_TIMP2_T1*y(M9T1);
dydt(iT2) =  k_TIMP2_T2*y(M1T2) + k_inact_T2*y(T2) + k_TIMP2_T2*y(M2T2) +
k_TIMP2_T2*y(M8T2) + k_TIMP2_T2*y(M9T2);
dydt(iT4) =  k_TIMP2_T4*y(M1T4) + k_inact_T4*y(T4) + k_TIMP2_T4*y(M2T4) +
k_TIMP2_T4*y(M8T4) + k_TIMP2_T4*y(M9T4);

dydt(T1iM1) = k_dist_on_T1*y(T1)*y(iM1) - k_dist_off_T1*y(T1iM1);
dydt(T2iM1) = k_dist_on_T2*y(T2)*y(iM1) - k_dist_off_T2*y(T2iM1);
dydt(T4iM1) = k_dist_on_T4*y(T4)*y(iM1) - k_dist_off_T4*y(T4iM1);

dydt(T1iM2) = k_dist_on_T1*y(T1)*y(iM2) - k_dist_off_T1*y(T1iM2);
dydt(T2iM2) = k_dist_on_T2*y(T2)*y(iM2) - k_dist_off_T2*y(T2iM2);
dydt(T4iM2) = k_dist_on_T4*y(T4)*y(iM2) - k_dist_off_T4*y(T4iM2);

dydt(T1iM8) = k_dist_on_T1*y(T1)*y(iM8) - k_dist_off_T1*y(T1iM8);
dydt(T2iM8) = k_dist_on_T2*y(T2)*y(iM8) - k_dist_off_T2*y(T2iM8);
dydt(T4iM8) = k_dist_on_T4*y(T4)*y(iM8) - k_dist_off_T4*y(T4iM8);


dydt(T1iM9) = k_dist_on_T1*y(T1)*y(iM9) - k_dist_off_T1*y(T1iM9);
dydt(T2iM9) = k_dist_on_T2*y(T2)*y(iM9) - k_dist_off_T2*y(T2iM9);
dydt(T4iM9) = k_dist_on_T4*y(T4)*y(iM9) - k_dist_off_T4*y(T4iM9);

end
```

Figure 20: Time of course of all of the conditions, separated by each sample, showing increase in signal with time for all of the conditions indicating increase in collagen degradation. Red showing the low concentrations and blue showing high concentration
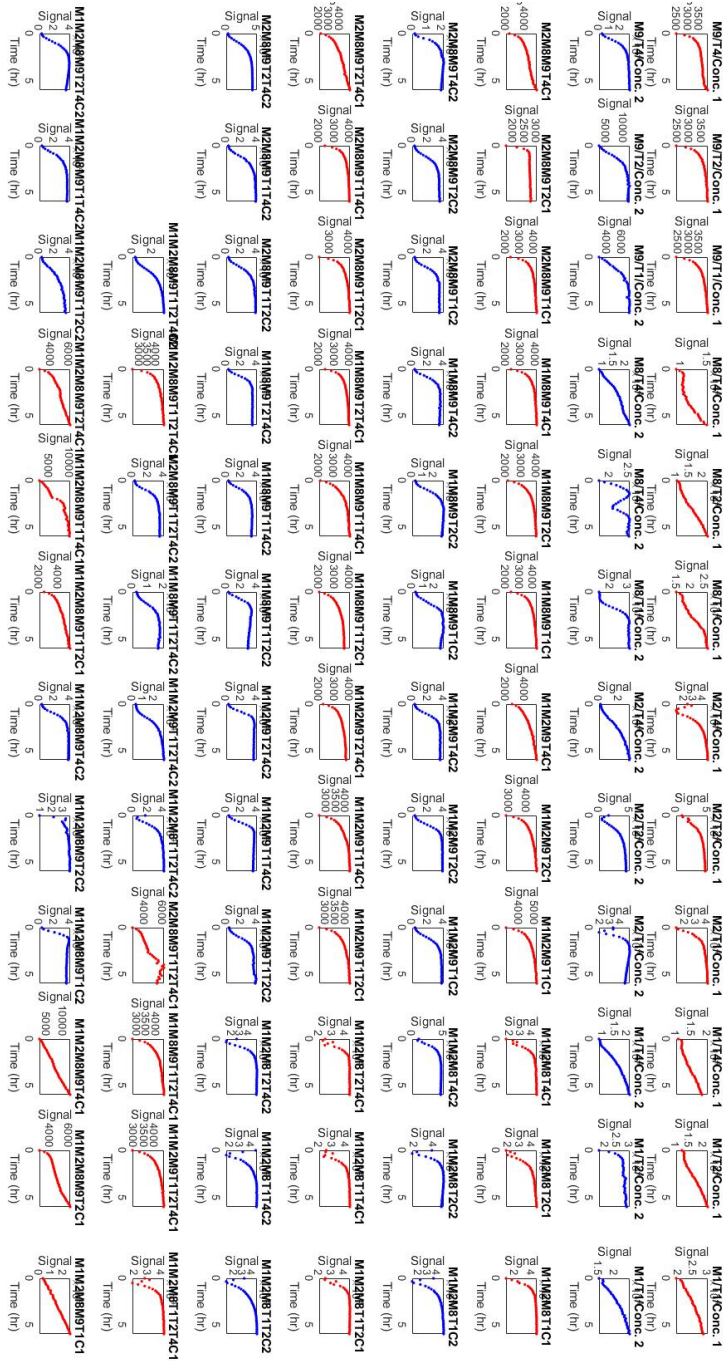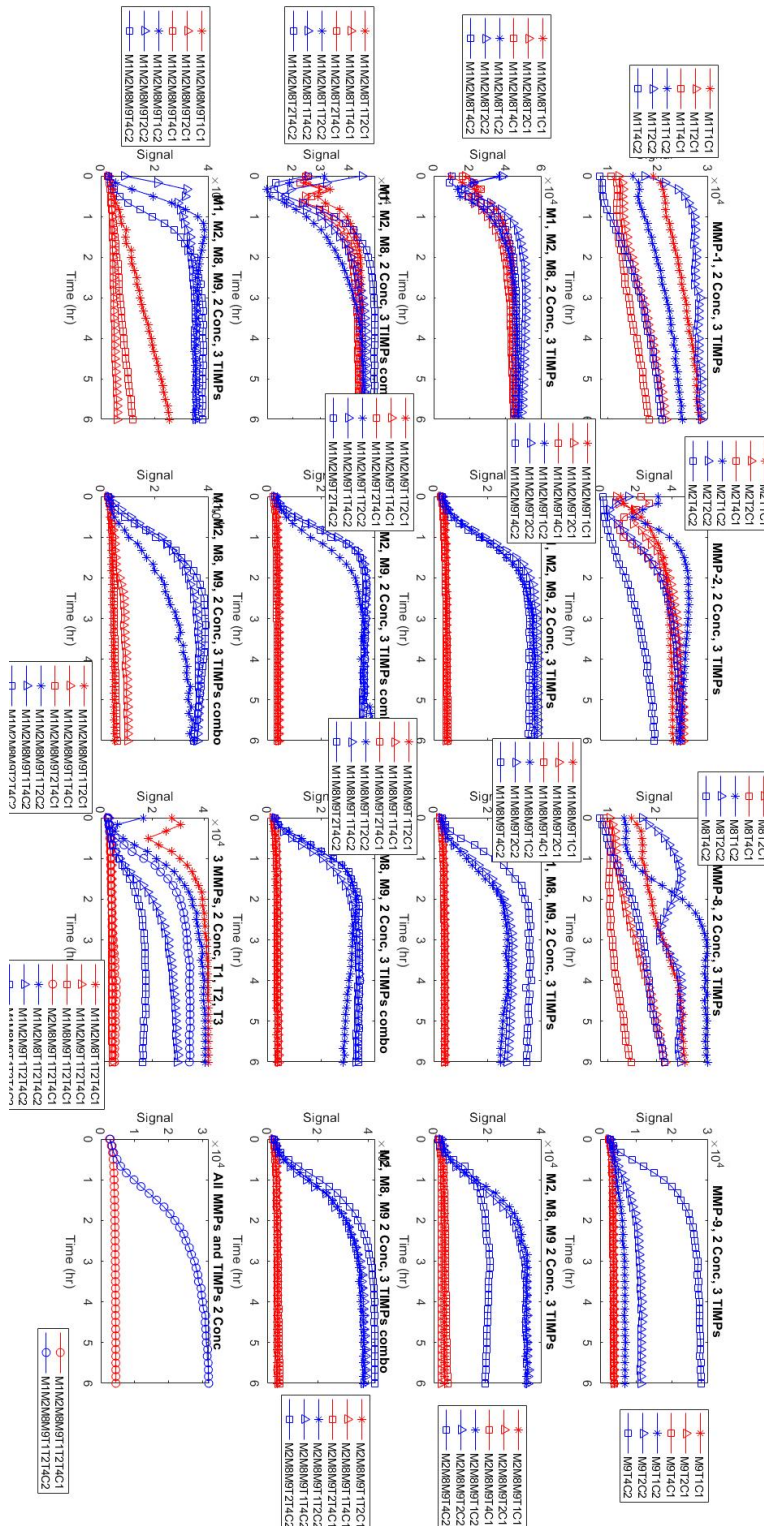
Figure 21: Time of course of all of the conditions showing increase in signal with time for all of the conditions indicating increase in collagen degradation. Red showing the low concentrations and blue showing high concentration

REFERENCES

1.   Richardson, W. J., Clarke, S. A., Quinn, T. A. & Holmes, J. W. Physiological Implications of Myocardial Scar Structure. in *Comprehensive Physiology* 1877–1909 (John Wiley & Sons, Inc., 2015). doi:10.1002/cphy.c140067.

2.   Virani, S. S. *et al. Heart disease and stroke statistics—2020 update: A report from the American Heart Association. Circulation* (2020). doi:10.1161/CIR.0000000000000757.

3.   Jugdutt, B. I. Ventricular remodeling after infarction and the extracellular collagen matrix: When is enough enough? *Circulation* **108**, 1395–1403 (2003).

4.   Bhole, A. P. *et al.* Mechanical strain enhances survivability of collagen micronetworks in the presence of collagenase: implications for load-bearing matrix growth and stability. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **367**, 3339–3362 (2009).

5.   Flynn, B. P. *et al.* Mechanical strain stabilizes reconstituted collagen fibrils against enzymatic degradation by mammalian collagenase matrix metalloproteinase 8 (MMP-8). *PLoS One* **5**, 21–23 (2010).

6.   Breen, E. C. Mechanical strain increases type I collagen expression in pulmonary fibroblasts in vitro. *J. Appl. Physiol.* **88**, 203–209 (2000).

7.   Holmes, J. W., Borg, T. K. & Covell, J. W. Structure and mechanics of healing myocardial infarcts. *Annu. Rev. Biomed. Eng.* **7**, 223–253 (2005).

8.   Jugdutt, B. I. Remodeling of the myocardium and potential targets in the collagen degradation and synthesis pathways. *Curr. Drug Targets-Cardiovascular Hematol. Disord.* **3**, 1–30 (2003).

9.   Hinz, B. The extracellular matrix and transforming growth factor-β1: Tale of a strained relationship. *Matrix Biol.* **47**, 54–65 (2015).

10.   Davis, G. E., Bayless, K. J., Davis, M. J. & Meininger, G. A. Regulation of tissue injury responses by the exposure of matricryptic sites within extracellular matrix molecules. *Am. J. Pathol.* **156**, 1489–1498 (2000).

11.   Maquart, F. X., Bellon, G., Pasco, S. & Monboisse, J. C. Matrikines in the regulation of extracellular matrix degradation. *Biochimie* **87**, 353–360 (2005).

12.   Weathington, N. M. *et al.* A novel peptide CXCR ligand derived from extracellular matrix degradation during airway inflammation. *Nat. Med.* **12**, 317–323 (2006).

13. Ruberti, J. W. & Hallab, N. J. Strain-controlled enzymatic cleavage of collagen in loaded matrix. *Biochem. Biophys. Res. Commun.* **336**, 483–489 (2005).

14. Huang, C. & Yannas, I. V. Mechanochemical studies of enzymatic degradation of insoluble collagen fibers. *J. Biomed. Mater. Res.* **11**, 137–154 (1977).

15. Ellsmere, J. C., Khanna, R. A. & Lee, J. M. Mechanical loading of bovine pericardium accelerates enzymatic degradation. *Biomaterials* **20**, 1143–1150 (1999).

16. Adhikari, A. S., Glassey, E. & Dunn, A. R. Conformational dynamics accompanying the proteolytic degradation of trimeric collagen I by collagenases. *J. Am. Chem. Soc.* **134**, 13259–13265 (2012).

17. Camp, R. J. *et al.* Molecular mechanochemistry: low force switch slows enzymatic cleavage of human type I collagen monomer. *J. Am. Chem. Soc.* **133**, 4073–4078 (2011).

18. Flynn, B. P., Tilburey, G. E. & Ruberti, J. W. Highly sensitive single-fibril erosion assay demonstrates mechanochemical switch in native collagen fibrils. *Biomech. Model. Mechanobiol.* **12**, 291–300 (2013).

19. Adhikari, A. S., Chai, J. & Dunn, A. R. Mechanical load induces a 100-fold increase in the rate of collagen proteolysis by MMP-1. *J. Am. Chem. Soc.* **133**, 1686–1689 (2011).

20. Ghazanfari, S., Driessen-Mol, A., Bouten, C. V. C. & Baaijens, F. P. T. Modulation of collagen fiber orientation by strain-controlled enzymatic degradation. *Acta Biomater.* **35**, 118–126 (2016).

21. Flynn, B. P. *et al.* Mechanical strain stabilizes reconstituted collagen fibrils against enzymatic degradation by mammalian collagenase matrix metalloproteinase 8 (MMP-8). *PLoS One* **5**, e12337 (2010).

22. Bhole, A. P. *et al.* Mechanical strain enhances survivability of collagen micronetworks in the presence of collagenase: implications for load-bearing matrix growth and stability. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **367**, 3339–3362 (2009).

23. Zareian, R. *et al.* Probing collagen/enzyme mechanochemistry in native tissue with dynamic, enzyme-induced creep. *Langmuir* **26**, 9917–9926 (2010).

24. Nabeshima, Y., Grood, E. S., Sakurai, A. & Herman, J. H. Uniaxial tension inhibits tendon collagen degradation by collagenase in vitro. *J. Orthop. Res.* **14**,

123–130 (1996).

25.   Yi, E. *et al.* Mechanical forces accelerate collagen digestion by bacterial collagenase in lung tissue strips. *Front. Physiol.* **7**, 287 (2016).

26.   Hinderer, S. & Schenke-Layland, K. Cardiac fibrosis–A short review of causes and therapeutic strategies. *Adv. Drug Deliv. Rev.* **146**, 77–82 (2019).

27.   Denolin, H., Kuhn, H., Krayenbuehl, H., Loogen, F. & Reale, A. The defintion of heart failure. *Eur. Heart J.* **4**, 445–448 (1983).

28.   Tsao, C. W. *et al.* Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association. *Circulation* **145**, e153–e639 (2022).

29.   Shah, S. J. *et al.* Prevalence and correlates of coronary microvascular dysfunction in heart failure with preserved ejection fraction: PROMIS-HFpEF. *Eur. Heart J.* **39**, 3439–3450 (2018).

30.   Wang, J. *et al.* MiR-101a loaded extracellular nanovesicles as bioactive carriers for cardiac repair. *Nanomedicine Nanotechnology, Biol. Med.* **27**, 102201 (2020).

31.   Lindsey, M. L. *et al.* A novel collagen matricryptin reduces left ventricular dilation post-myocardial infarction by promoting scar formation and angiogenesis. *J. Am. Coll. Cardiol.* **66**, 1364–1374 (2015).

32.   Pfeffer, M. A. *et al.* Effect of captopril on mortality and morbidity in patients with left ventricular dysfunction after myocardial infarction: results of the Survival and Ventricular Enlargement Trial. *N. Engl. J. Med.* **327**, 669–677 (1992).

33.   Investigators*, S. Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. *N. Engl. J. Med.* **327**, 685–691 (1992).

34.   Muiesan, M. L. *et al.* Current pharmacological therapies in heart failure patients. *High Blood Press. Cardiovasc. Prev.* **24**, 107–114 (2017).

35.   Pfeffer, M. A. *et al.* Valsartan, captopril, or both in myocardial infarction complicated by heart failure, left ventricular dysfunction, or both. *N. Engl. J. Med.* **349**, 1893–1906 (2003).

36.   McMurray, J. J. V *et al.* Effects of candesartan in patients with chronic heart failure and reduced left-ventricular systolic function taking angiotensin-converting-enzyme inhibitors: the CHARM-Added trial. *Lancet* **362**, 767–771 (2003).

37. Thomopoulos, C., Parati, G. & Zanchetti, A. Effects of blood pressure-lowering treatment. 6. Prevention of heart failure and new-onset heart failure–meta-analyses of randomized trials. *J. Hypertens.* **34**, 373–384 (2016).

38. Kong, P., Christia, P. & Frangogiannis, N. G. The pathogenesis of cardiac fibrosis. *Cell. Mol. life Sci.* **71**, 549–574 (2014).

39. Disertori, M., Masè, M. & Ravelli, F. Myocardial fibrosis predicts ventricular tachyarrhythmias. *Trends Cardiovasc. Med.* **27**, 363–372 (2017).

40. Jellis, C., Martin, J., Narula, J. & Marwick, T. H. Assessment of nonischemic myocardial fibrosis. *J. Am. Coll. Cardiol.* **56**, 89–97 (2010).

41. Visse, R. & Nagase, H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ. Res.* **92**, 827–839 (2003).

42. Hay, E. D. *Cell biology of extracellular matrix*. (Springer Science & Business Media, 1991).

43. Humphrey, J. D., Dufresne, E. R. & Schwartz, M. A. Mechanotransduction and extracellular matrix homeostasis. *Nat. Rev. Mol. cell Biol.* **15**, 802–812 (2014).

44. Ottani, V., Martini, D., Franchi, M., Ruggeri, A. & Raspanti, M. Hierarchical structures in fibrillar collagens. *Micron* **33**, 587–596 (2002).

45. Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **78**, 929–958 (2009).

46. Guo, C. & Kaufman, L. J. Flow and magnetic field induced collagen alignment. *Biomaterials* **28**, 1105–1114 (2007).

47. Tonge, T. K., Ruberti, J. W. & Nguyen, T. D. Micromechanical modeling study of mechanical inhibition of enzymatic degradation of collagen tissues. *Biophys. J.* **109**, 2689–2700 (2015).

48. Prockop, D. J. & Fertala, A. The collagen fibril: the almost crystalline structure. *J. Struct. Biol.* **122**, 111–118 (1998).

49. Parkinson, J., Kadler, K. E. & Brass, A. Simple physical model of collagen fibrillogenesis based on diffusion limited aggregation. *J. Mol. Biol.* **247**, 823–831 (1995).

50. Williams, B. Collagen fibril formation-optimal in vitro conditions and preliminary

kinetic results. *J Biol Chem* **253**, 6578–6585 (1978).

51. Iyer, R. P., Patterson, N. L., Fields, G. B. & Lindsey, M. L. The history of matrix metalloproteinases: milestones, myths, and misperceptions. *Am. J. Physiol. Circ. Physiol.* **303**, H919–H930 (2012).

52. Bannikov, G. A. *et al.* Serum haptoglobin–matrix metalloproteinase 9 (Hp–MMP 9) complex as a biomarker of systemic inflammation in cattle. *Vet. Immunol. Immunopathol.* **139**, 41–49 (2011).

53. Ashcroft, G. S. *et al.* Age-related differences in the temporal and spatial regulation of matrix metalloproteinases (MMPs) in normal skin and acute cutaneous wounds of healthy humans. *Cell Tissue Res.* **290**, 581–591 (1997).

54. Ardi, V. C., Kupriyanova, T. A., Deryugina, E. I. & Quigley, J. P. Human neutrophils uniquely release TIMP-free MMP-9 to provide a potent catalytic stimulator of angiogenesis. *Proc. Natl. Acad. Sci.* **104**, 20262–20267 (2007).

55. Apple, F. S., Smith, S. W., Pearce, L. A. & Murakami, M. M. Assessment of the multiple-biomarker approach for diagnosis of myocardial infarction in patients presenting with symptoms suggestive of acute coronary syndrome. *Clin. Chem.* **55**, 93–100 (2009).

56. Ferrall-Fairbanks, M. C., Kieslich, C. A. & Platt, M. O. Reassessing enzyme kinetics: Considering protease-as-substrate interactions in proteolytic networks. *Proc. Natl. Acad. Sci.* **117**, 3307–3318 (2020).

57. Barry, Z. T. & Platt, M. O. Cathepsin S cannibalism of cathepsin K as a mechanism to reduce type I collagen degradation. *J. Biol. Chem.* **287**, 27723–27730 (2012).

58. Li, W. A. *et al.* Detection of femtomole quantities of mature cathepsin K with zymography. *Anal. Biochem.* **401**, 91–98 (2010).

59. Chen, B. & Platt, M. O. Multiplex zymography captures stage-specific activity profiles of cathepsins K, L, and S in human breast, lung, and cervical cancer. *J. Transl. Med.* **9**, 1–13 (2011).

60. Duarte, A. S., Correia, A. & Esteves, A. C. Bacterial collagenases–a review. *Crit. Rev. Microbiol.* **42**, 106–126 (2016).

61. Bode, W. *et al.* Insights into MMP-TIMP interactions. *Ann. N. Y. Acad. Sci.* **878**, 73–91 (1999).

62. Teng, X. & Hwang, W. Chain registry and load-dependent conformational dynamics of collagen. *Biomacromolecules* **15**, 3019–3029 (2014).

63. Manka, S. W. *et al.* Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. *Proc. Natl. Acad. Sci.* **109**, 12461–12466 (2012).

64. French, M. F., Mookhtiar, K. A. & Van Wart, H. E. Limited proteolysis of type I collagen at hyperreactive sites by class I and II Clostridium histolyticum collagenases: complementary digestion patterns. *Biochemistry* **26**, 681–687 (1987).

65. Wyatt, K. E.-K., Bourne, J. W. & Torzilli, P. A. Deformation-dependent enzyme mechanokinetic cleavage of type I collagen. (2009).

66. Karagiannis, E. D. & Popel, A. S. A Theoretical Model of Type I Collagen Proteolysis by Matrix Metalloproteinase (MMP) 2 and Membrane Type 1 MMP in the Presence of Tissue Inhibitor of Metalloproteinase 2*[boxs]. *J. Biol. Chem.* **279**, 39105–39114 (2004).

67. Karagiannis, E. D. & Popel, A. S. Distinct modes of collagen type I proteolysis by matrix metalloproteinase (MMP) 2 and membrane type I MMP during the migration of a tip endothelial cell: insights from a computational model. *J. Theor. Biol.* **238**, 124–145 (2006).

68. Vempati, P., Karagiannis, E. D. & Popel, A. S. A biochemical model of matrix metalloproteinase 9 activation and inhibition. *J. Biol. Chem.* **282**, 37585–37596 (2007).

69. Daulagala, A. C. *et al.* A simple method to test mechanical strain on epithelial cell monolayers using a 3D-printed stretcher. in *Permeability Barrier* 235–247 (Springer, 2020).

70. Rogers, J. D., Yeganegi, A. & Richardson, W. J. Mechano-regulation of fibrillar collagen turnover by fibroblasts. in *Mechanobiology Handbook* 511–530 (CRC Press, 2018).

71. McGuire, R., Borem, R. & Mercuri, J. The fabrication and characterization of a multi-laminate, angle-ply collagen patch for annulus fibrosus repair. *J. Tissue Eng. Regen. Med.* **11**, 3488–3493 (2017).

72. Frangogiannis, N. G. The extracellular matrix in ischemic and nonischemic heart failure. *Circ. Res.* **125**, 117–146 (2019).

73. de Castro Brás, L. E. & Frangogiannis, N. G. Extracellular matrix-derived peptides in tissue remodeling and fibrosis. *Matrix Biol.* **91**, 176–187 (2020).

74. Yabluchanskiy, A., Ma, Y., Iyer, R. P., Hall, M. E. & Lindsey, M. L. Matrix metalloproteinase-9: Many shades of function in cardiovascular disease. *Physiology* **28**, 391–403 (2013).

75. Apte, S. S., Olsen, B. R. & Murphy, G. The Gene Structure of Tissue Inhibitor of Metalloproteinases (TIMP)-3 and Its Inhibitory Activities Define the Distinct TIMP Gene Family∗. *J. Biol. Chem.* **270**, 14313–14318 (1995).

76. Rogers, J. D., Aguado, B. A., Watts, K. M., Anseth, K. S. & Richardson, W. J. Network modeling predicts personalized gene expression and drug responses in valve myofibroblasts cultured with patient sera. *Proc. Natl. Acad. Sci.* **119**, e2117323119 (2022).

77. Watts, K. & Richardson, W. J. Effects of Sex and 17 β-Estradiol on Cardiac Fibroblast Morphology and Signaling Activities In Vitro. *Cells* **10**, 2564 (2021).

78. Potter, M. J. & Richardson, W. J. Fabrication and characterization methods for investigating cell-matrix interactions in environments possessing spatial orientation heterogeneity. *Acta Biomater.* **136**, 420–428 (2021).

79. Kiew, P. L., Ahmad, Z. & Don, M. M. A hybrid of back propagation neural network and genetic algorithm for optimization of collagen extraction from Malaysian cultured catfish (Hybrid Clarias sp.). *Biotechnol. bioprocess Eng.* **18**, 257–265 (2013).

80. Chawla, A., Mukherjee, S. & Karthikeyan, B. Characterization of human passive muscles for impact loads using genetic algorithm and inverse finite element methods. *Biomech. Model. Mechanobiol.* **8**, 67–76 (2009).

81. Mathworks. How the genetic algorithm works. mathworks.com Web site. https://www.mathworks.com/help/gads/how-the-genetic-algorithm-works.html. Updated 2020.

82. Peterson, J. T., Li, H., Dillon, L. & Bryant, J. W. Evolution of matrix metalloprotease and tissue inhibitor expression during heart failure progression in the infarcted rat. *Cardiovasc. Res.* **46**, 307–315 (2000).

83. Cleutjens, J. P. M., Kandala, J. C., Guarda, E., Guntaka, R. V & Weber, K. T. Regulation of collagen degradation in the rat myocardium after infarction. *J. Mol. Cell. Cardiol.* **27**, 1281–1292 (1995).

84. Hamby, D. M. A review of techniques for parameter sensitivity analysis of environmental models. *Environ. Monit. Assess.* **32**, 135–154 (1994).

85.   Nagase, H. Matrix metalloproteinases. *Zinc Met. Heal. Dis.* 173–224 (1996).

86.   Woessner, J. F., Woessner, J. F. & Nagase, H. *Matrix metalloproteinases and TIMPs*. (Oxford University Press, 2000).

87.   Spinale, F. G. Myocardial matrix remodeling and the matrix metalloproteinases: influence on cardiac form and function. *Physiol. Rev.* **87**, 1285–1342 (2007).

88.   González, A. *et al.* Biochemical markers of myocardial remodelling in hypertensive heart disease. *Cardiovasc. Res.* **81**, 509–518 (2009).

89.   Essa, E. M. *et al.* Changes in plasma profiles of matrix metalloproteinases (MMPs) and tissue inhibitors of MMPs in stress-induced cardiomyopathy. *J. Card. Fail.* **18**, 487–492 (2012).

90.   Ahmed, S. H. *et al.* Matrix metalloproteinases/tissue inhibitors of metalloproteinases: relationship between changes in proteolytic determinants of matrix composition and structural, functional, and clinical manifestations of hypertensive heart disease. *Circulation* **113**, 2089–2096 (2006).

91.   Ward, M. *et al.* Ensemble Machine Learning Model Identifies HFpEF Patients from Matrix-Related Plasma Biomarkers. *Am. J. Physiol. Circ. Physiol.* (2022).

92.   Hasankhani, F. & Khademi, A. Efficient and fair heart allocation policies for transplantation. *MDM policy Pract.* **2**, 2381468317709475 (2017).

93.   Mak, G. J. *et al.* Natural history of markers of collagen turnover in patients with early diastolic dysfunction and impact of eplerenone. *J. Am. Coll. Cardiol.* **54**, 1674–1682 (2009).

94.   Hanna, A., Shinde, A. V & Frangogiannis, N. G. Validation of diagnostic criteria and histopathological characterization of cardiac rupture in the mouse model of nonreperfused myocardial infarction. *Am. J. Physiol. Circ. Physiol.* **319**, H948–H964 (2020).

95.   Sanchez de la Nava, A. M., Atienza, F., Bermejo, J. & Fernández-Avilés, F. Artificial intelligence for a personalized diagnosis and treatment of atrial fibrillation. *Am. J. Physiol. Circ. Physiol.* **320**, H1337–H1347 (2021).

96.   Liem, D. A. *et al.* Phrase mining of textual data to analyze extracellular matrix protein patterns across cardiovascular disease. *Am. J. Physiol. Circ. Physiol.* **315**, H910–H924 (2018).

97.   Jin, T., Nguyen, N. D., Talos, F. & Wang, D. ECMarker: interpretable machine

learning model identifies gene expression biomarkers predicting clinical outcomes and reveals molecular mechanisms of human disease in early stages. *Bioinformatics* **37**, 1115–1124 (2021).

98.   Jagga, Z. & Gupta, D. Machine learning for biomarker identification in cancer research–developments toward its clinical application. *Per. Med.* **12**, 371–387 (2015).

99.   Zile, M. R. *et al.* Plasma biomarkers that reflect determinants of matrix composition identify the presence of left ventricular hypertrophy and diastolic heart failure. *Circ. Hear. Fail.* **4**, 246–256 (2011).

100. Dobson, A. J. & Barnett, A. G. *An introduction to generalized linear models*. (Chapman and Hall/CRC, 2018).

101. Fisher, R. A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **7**, 179–188 (1936).

102. Devroye, L., Györfi, L. & Lugosi, G. *A probabilistic theory of pattern recognition*. vol. 31 (Springer Science & Business Media, 2013).

103. Dietrich, R., Opper, M. & Sompolinsky, H. Statistical mechanics of support vector networks. *Phys. Rev. Lett.* **82**, 2975 (1999).

104. Cover, T. & Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. theory* **13**, 21–27 (1967).

105. Breiman, L. Bagging predictors. *Mach. Learn.* **24**, 123–140 (1996).

106. Freund, Y. & Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **55**, 119–139 (1997).

107. athworks. Hyperparameter Optimization in Classification Learner App. (2020).

108. Snoek, J., Larochelle, H. & Adams, R. P. Practical bayesian optimization of machine learning algorithms. *Adv. Neural Inf. Process. Syst.* **25**, (2012).

109. Sokolova, M., Japkowicz, N. & Szpakowicz, S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. in *Australasian joint conference on artificial intelligence* 1015–1021 (Springer, 2006).

110. Chicco, D. & Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC*

Genomics **21**, 1–13 (2020).

111. Ding, C. & Peng, H. Minimum redundancy feature selection from microarray gene expression data. *J. Bioinform. Comput. Biol.* **3**, 185–205 (2005).

112. Li, B.-Q., Feng, K.-Y., Chen, L., Huang, T. & Cai, Y.-D. Prediction of protein-protein interaction sites by random forest algorithm with mRMR and IFS. (2012).

113. Ma, C., Dong, X., Li, R. & Liu, L. A computational study identifies HIV progression-related genes using mRMR and shortest path tracing. *PLoS One* **8**, e78057 (2013).

114. Grewal, J. *et al.* Usefulness of N-terminal pro–brain natriuretic peptide and brain natriuretic peptide to predict cardiovascular outcomes in patients with heart failure and preserved left ventricular ejection fraction. *Am. J. Cardiol.* **102**, 733–737 (2008).

115. Hara, Y., Hamada, M. & Hiwada, K. Left ventricular end-systolic wall stress is a potent prognostic variable in patients with dilated cardiomyopathy. *Jpn. Circ. J.* **63**, 196–200 (1999).

116. Carter-Storch, R. *et al.* End-systolic wall stress in aortic stenosis: comparing symptomatic and asymptomatic patients. *Open Hear.* **6**, e001021 (2019).

117. Gaggin, H. K. & Januzzi Jr, J. L. Biomarkers and diagnostics in heart failure. *Biochim. Biophys. Acta (BBA)-Molecular Basis Dis.* **1832**, 2442–2450 (2013).