UNIVERSITY OF CALIFORNIA,
IRVINE

Modification of Recombinant Human Collagen with Matrix Metalloproteinase Cleavage Sites

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Biomedical Engineering

by

Dale Russell Crakes

Thesis Committee:
Associate Professor Szu-Wen Wang, Co-Chair
Professor Nancy Da Silva, Co-Chair
Assistant Professor Wendy Liu

2015

© 2015 Dale Russell Crakes

DEDICATION

To

my parents, friends, colleagues, and mentors

in recognition of their patience and support

"It is not the critic who counts; not the man who points out how the strong man stumbles, or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs, who comes short again and again, because there is no effort without error and shortcoming; but who does actually strive to do the deeds; who knows great enthusiasms, the great devotions; who spends himself in a worthy cause; who at the best knows in the end the triumph of high achievement, and who at the worst, if he fails, at least fails while daring greatly, so that his place shall never be with those cold and timid souls who neither know victory nor defeat."

-Theodore Roosevelt

"Citizenship In A Republic"

TABLE OF CONTENTS

Page

LIST OF FIGURES v

LIST OF TABLES vi

ACKNOWLEDGMENTS vii

ABSTRACT OF THE THESIS viii

CHAPTER 1: Introduction
- Background 1
  - Tissue Engineering Scaffolds 1
  - Collagen 2
  - Matrix Metalloproteinases 4
  - Recombinant Collagen Expression in Yeast 6
  - Modular Collagen Gene Design 8
- Motivation and Objectives 10

CHAPTER 2: Collagen Mutant Creation and Degradation
- Introduction 12
- Materials and Methods 13
  - Mutagenesis of Primary Fragments for MMP-Site Mutants 13
  - Secondary Fragments 15
  - Full Length Collagen Genes 15
  - Protein Expression and Purification 16
  - Circular Dichroism 20
  - Anti-Collagen Western Blot 20
  - MMP Digestion Assay 21
  - Microrheology Glass Slide Silanization 21
  - Polystyrene Microparticle Preparation 22
  - Hydrogel Formation in Microrheology Slides 23
  - Microfluidic Device Design and Photolithography 24
- Results and Discussion 26
  - Construction Strategy of MMP Variants of Recombinant Collagen 26
  - Collagen Characterization 28
  - Failure of rCol-1M-PF7 to be Successfully Purified 31
  - Digestion of Collagen Variants with MMP-1 39
  - Time-Point MMP Digestion of Collagen Variants 39
  - Collagen Hydrogel Formation 40
- Conclusions 44
- Future Directions 45

CHAPTER 3: Stabilization of Recombinant Collagen on Solid Surfaces for Cell Adhesion
Introduction 47
Materials and Methods 49
Recombinant Collagen Design, Expression and Purification 49
HT-1080 Cell Culture 49
Calcein AM Cell Adhesion Assay 50
Results and Discussion 51
Quantitative Cell Adhesion Analysis 51
Qualitative Cell Adhesion Analysis 54
Conclusions 58

References 58

Supplementary Information
Possible Reasons for Failure of rCol-1M-PF7 to Purify 69
SDS-PAGE Staining to Detect Digested Collagen Fragments 70
Cell Adhesion Assay Execution Discussion 71
Modular Collagen Gene PF, SF, and Full Gene Sequences 72

LIST OF FIGURES

		Page
Figure 2-1	Map of MMP Site Placement on Collagen III Gene	18
Figure 2-2	Microfluidic Device Photoresist Template	24
Figure 2-3	Final PDMS Microfluidic Device	25
Figure 2-4	Collagen CD Spectra	29
Figure 2-5	Collagen Temperature Scans	30
Figure 2-6	Western Blot of rCol-1M-PF7 After Expression	32
Figure 2-7	Western Blot of rCol-1M-PF7 With Protease Inhibitor	34
Figure 2-8	MMP Assay	37
Figure 2-9	Time Point MMP Assay	39
Figure 2-10	Hydrogel Fluorescent Tracer Particles	41
Figure 3-1	Cell Adhesion Fluorescence Data	52
Figure 3-2	Cell Adhesion Cell Images	55

LIST OF TABLES

		Page
Table 2-1	Mutagenesis Primers For MMP Mutant Primary Fragments	14
Table 2-2	Collagen Variant Apparent Melting Temperatures	30
Table 2-3	MMP-Variant Expected Digestion Sizes	35

ACKNOWLEDGMENTS

I would like to express the deepest appreciation to my committee chair, Associate Professor Szu-Wen Wang, for recruiting me to her lab and giving me a chance to grow under her guidance and leadership.

Similarly, my heartfelt thanks go to Professor Nancy Da Silva for her patience, immense scientific knowledge and experience, and for all her help when my research didn't go as planned.

I would like to thank my committee member Assistant Professor Wendy Liu, for her continued support of our lab and tissue engineering knowledge.

Finally, I would also like to thank the friends I have made during my time at UCI. Without them I would be lost.

This work was supported by the National Science Foundation (CBET-1034566).

ABSTRACT OF THE THESIS

Modification of Recombinant Human Collagen with Matrix Metalloproteinase Cleavage Sites

By

Dale Russell Crakes

Master of Science in Biomedical Engineering

University of California, Irvine, 2015

Associate Professor Szu-Wen Wang, Co-Chair

Professor Nancy Da Silva, Co-Chair

Collagen is the most abundant component of the extracellular matrix, and recombinant collagen is commonly used in tissue engineering applications. Matrix metalloproteinases (MMPs) are a specific family of enzymes that can cleave and degrade collagen. Control of the degradability of a tissue engineered matrix is an important attribute, and the modulation of MMP-cleavable sites in collagen could lead to better control over the digestibility and ease of cell infiltration in a collagen matrix. Using a computer-optimized, modular collagen III gene, recombinant collagens were designed in which native MMP-cleavable sites were deleted, added, and moved along the length of the collagen molecule, and then expressed in *Saccharomyces cerevisiae*. MMP-1 was used to digest each of the collagen variants, and more complete digestion was seen in collagen variants featuring more MMP sites. Hydrogel formation featuring recombinant collagen was explored, but future effort is still needed before consistent hydrogel formation can be obtained. EDC / NHS chemistry was used to covalently stabilize our recombinant collagen, which has slightly decreased thermostability compared to human collagen. Cell adhesion

assays were used to confirm the increased stability of the collagen and allow cell-based experiments to be performed at 37 °C.

Chapter 1: Introduction

Background

The extracellular matrix (ECM) is a complex network of proteins, proteoglycans and cell signals that interact dynamically with cells in tissues [1]. Not only does the ECM provide a structural scaffolding to support cells, but it also provides physical and biochemical cues to its constituent cells. Cell differentiation [2], morphogenesis [3], homeostasis [4], migration [5], and gene expression levels [6] are all influenced by the ECM. Because of the importance of the ECM, it is a common inspiration and basis for tissue engineering efforts, which seek to recreate the ECM with an artificial scaffold to ultimately create new tissues [7]. Artificial scaffolds typically take two forms: biomimetic, based on natural proteins and polymers, and synthetic, which are based on polymers and other materials not found naturally. Each type of scaffold has its advantages and disadvantages, with many different viable options within each category.

Tissue Engineering Scaffolds

Synthetic scaffolds are made from biologically compatible, biodegradable polymers and can be tuned and controlled to have desired mechanical and structural properties. Synthetic scaffolds commonly used in tissue engineering include poly(ethylene glycol) (PEG) [8], poly(lactic acid-co-glycolic acid) (PLGA) [9], and poly(caprolactone) [10] among others [11]. These polymers all offer biocompatibility and resistance to inflammatory responses from cells they encapsulate, as well as the ability to easily tune mechanical and degradation properties. However, a major drawback of synthetic scaffolds is that they lack possible sites for cell recognition. Although groups have circumvented this by

incorporating the cell-adhesion sequence arginine-glycine-aspartic acid (RGD) into their synthetic scaffolds [12], the need to incorporate cell-recognizable motifs into the matrix is a liability because all potential cell interactions must be considered when designing the scaffold. These interactions include biodegradability, appropriate mechanical properties, and ability to support biochemical signals such as cell growth factors [13].

Biological scaffolds, on the other hand, already have mechanical and biochemical properties similar to the natural tissues they hope to mimic [14]. Collagen [15], fibronectin [16], hyaluronic acid [17], elastin [18], chitosan [19] and alginate [20] are all among naturally-derived extracellular matrices used as scaffolds in tissue engineering. With the exception of alginate, all of these biopolymers contain natural cell adhesion sites, which allow the cells to easily interact with the matrix. Although the specific matrix requirements will differ for each application and / or desired tissue to be mimicked, there are some general properties that artificial scaffolds must have. Biocompatibility, mechanical strength, cell adhesion ability, cell degradability, porosity, and matrix reproducibility are all central requirements of a good tissue engineering platform [13]. Because of its role as the major component of the ECM, collagen makes for an attractive biopolymer to use in a wide variety of applications [21], and is the matrix of choice for our research.

Collagen

Collagen makes up ~30% of all proteins in humans and the majority of the ECM [22, 23]. At least 28 different collagens have been reported in vertebrates, although not all collagens are found in the ECM [24]. The collagen family mainly imparts structural integrity, specifically tensile strength, on the ECM [25, 26], but also dynamically interacts

with cells around it and is often remade and remodeled by them [27]. The characteristic structure of collagen is its triple helix, which is a result of the three amino acid repeating sequence Gly-X-Y, where X is typically proline and Y is typically hydroxyproline. Glycine, the smallest amino acid, sits in the middle of the triple helix. Each triple helix is made up of 3 left-handed α strands, which form together to make a right-handed triple helix coil [28]. Although the triple helices of some collagens exist as homotrimers, with the triple helix being made of 3 of the same protein (such as collagen types II, III, VII, VIII, and X), that is not a requisite. Heterotrimers of 2 or more different α chains can assemble to form the triple helix, such as in collagen types I, IV, V, VI, IV and XI [29]. The triple helix is a very stable structure, which comes mainly from hydroxylated prolines often found in the "Y" position of the Gly-X-Y repeating sequence [30-32]. This post-translational hydroxylation is performed by the enzyme prolyl-4-hydroxylase [33], and is a necessary component to add when engineering recombinant collagen systems like yeast [21, 34-36]. Other post-translational modifications happen to the collagen molecule, such as hydroxylation of lysines [37] and glycosylation [38]. The collagen triple helix makes it very resistant to proteases, with the exception of a family of enzymes known as matrix metalloproteinases (MMPs).

Collagen III, specifically, is a fibrillar collagen found in skin, muscle, blood vessels, lungs, and other tissues, most commonly found with collagen I [29, 39]. It is a homotrimeric protein consisting of roughly 1000 amino acids per chain, 300 nm long and 1 nm wide [40]. The fact that it is a homotrimer makes it an attractive target for recombinant expression, because the gene for only one collagen chain needs to be added to whatever expression system is being used. This also reduces the number of genes that must be modified if

changes are to be made to the collagen sequence, as shown in previous research by our laboratory [36, 41-43].

The prevalence of collagen III in the ECM makes it an excellent candidate for use in tissue engineering applications. To tune the collagen for specific desired traits, our lab genetically modifies the amino acid sequence to incur new, non-native characteristics, such as the modulation of cell adhesion sites [43] or the addition of non-native cysteines [42]. Because the ability of cells to infiltrate and remodel a matrix is extremely important [44], this work shows MMP-cleavable sites that have been introduced to modify collagen III to offer increased control over the ability of the cells to degrade a collagen matrix.

Matrix Metalloproteinases

MMPs are a family of 23 proteolytic enzymes, at least one of which can cleave every component of the ECM [45] [46]. MMPs have been implicated in many physiological processes, such as ECM remodeling during embryonic development [47], bone ossification [48], angiogenesis [49-54], wound healing [55, 56], cancer development and metastasis [57, 58], and innate immunity [59-61]. Because of this broad role of MMPs, sites that are susceptible to MMP cleavage have been used in the creation of multiple synthetic and biomimetic scaffolds for tissue engineering purposes [8, 44, 62-67]. MMP-1, specifically, is an interstitial collagenase, and was originally known by the names collagenase A [29] and collagenase-1 [68, 69]. It is expressed by fibroblasts [63], macrophages [70], and many cancer cell types [71], as well as during embryonic development [72].

MMP-1 consists of two domains: an N-terminal catalytic domain, which is linked to a C-terminal hemopexin domain [73]. The hemopexin domain consists of a four-bladed "β-

propeller" which acts in concert with the catalytic domain to cleave through collagen [74], and also helps to mediate interactions between the enzyme and the protein it is cleaving [75]. The catalytic domain contains a zinc atom in the active site of the enzyme [76]. The catalytic pro-domain covers this zinc atom and prevents enzymatic activity, until the MMP-1 is activated and this pro-domain is removed. Collagenases (MMP-1, MMP-2, MMP-8, MMP-13, and MMP-14) [62] digest interstitial, fibrillar collagens such as collagen III in a single locus along the collagen molecule [77]. The hemopexin and catalytic domains of MMP-1 work together to locally unwind the collagen triple helix at this single location before hydrolyzing the peptide bonds in each collagen monomer [78] [79].

The MMP-cleavable site is remarkable for its low triple helical stability, which helps the enzyme to unwind the triple helix and cleave peptide bonds [80, 81]. However, despite the need for lowered triple helix stability, a triple helix is still advantageous for MMP-1 to cleave the collagen, as it has decreased cleavage abilities through denatured collagens such as gelatin [77]. In interstitial collagens (types I, II and III), collagenase MMPs can cleave through the triple helix at one location roughly ¾ of the length of the molecule from the N-terminus [69]. The minimal collagen III sequence that is cleavable by MMP-1 is *GPLGIAGITGARGLA* [62]. MMP-1 cleaves this sequence between the glycine-isoleucine residues in this sequence [81]. Although MMP-1 can still cleave this sequence after slight modifications, changes that result in the tightening of the collagen triple helix confer resistance to proteolytic degradation [81]. MMP-1 can cleave different sequences occurring in other collagens, which shows some substrate redundancy that is characteristic of the family of MMPs overall [82].

Recombinant Collagen Expression in Yeast

For this work, the yeast *Saccharomyces cerevisiae* was used to express recombinant human collagen III and variants thereof. Recombinant collagens offer advantages over tissue-derived collagens and ECM proteins. Recombinant collagen offers no risk of animal virus contamination [83], very high batch-to-batch reproducibility with little variation [34], and consists of the same amino acid sequence and structure as native collagen, including a higher purity of full trimeric collagen compared to "off-the-shelf" collagens isolated from animal sources [34]. Many different systems have been used to express recombinant human collagen or collagen mimetics, including yeast [32, 36, 41-43, 84, 85], plants [86-88], and bacteria [89-91] (although it should be noted that bacterial systems do not contain the molecular machinery needed to hydroxylate collagen, and can only form non-hydroxylated gelatins or short hydroxylated collagen fragments, but no full length collagens) [92]. Yeast are advantageous for recombinant collagen expression because they can be grown to high cell densities and grow in simple media [93], which allows for the production of more recombinant protein.

One of the most important factors for the ability of a system to faithfully recreate triple helical collagen is the capacity to hydroxylate the collagen prolines. In yeast, this means that in addition to the desired collagen gene, the genes for the α and β subunits of prolyl-4-hydroxylase must also be transfected into the cell [32]. Without the hydroxylase enzyme, expressed collagens have little to no triple helical stability, and are prone to proteolytic degradation in ways that stable, triple helical collagens are not [94].

S. cerevisiae has been shown to express collagen I [95] and collagen III [36]. Recombinant collagens expressed in *S. cerevisiae* are not secreted into the extracellular

media, and thus during purification the cells must be lysed and treated with proteases (commonly pepsin) to digest the other yeast proteins and change the collagen from its pro-form to triple-helical collagen [21]. *S. cerevisiae* is preferable to use for our research over the yeast *Pichia pastoris*, which requires the integration of any recombinant genes into its genome; in contrast, *S. cerevisiae* allows for easy plasmid transformation, which is much less time consuming than integration. Because of the number of collagen gene variants produced by our lab, it is far easier to use a strain of yeast with the α- and β- subunits of prolyl-4-hydroxylase stably integrated into the genome with the collagen variant genes contained on a plasmid [41]. This way, collagen proline hydroxylation levels can be kept consistent while allowing for the facile construction of many different collagen-expressing yeast mutants.

Collagen expression in yeast is advantageous over "off the shelf" collagen from animal sources because of possible immunogenic effects from using ECM-derived collagen from a host that is not human, as well as an inability to modify the collagen to tailor it to specific needs. Genetically modified yeast allow for the production of recombinant collagen with the same amino acid profile as native human collagen, and without any potential immunogenic effects [34]. Collagen may be expressed in human cells in culture, but expression and purification in this system is costly in both time and money, and results in low yields of harvested collagen [96]. In addition, yeast can be grown to high cell densities and double quickly, resulting in larger quantities of the desired recombinant protein [97]. Previous research into recombinant collagen expression in yeast systems has produced truncated fragments of collagen [98], non-hydroxylated gelatins [99], and full length, hydroxylated collagen molecules [84, 85, 100]. Despite the production of both collagen

fragments and full length collagens, there has been no research into the production of modifiable full-length collagen proteins outside of our group. Short, synthetic collagen-mimetic peptides that can self-assemble into longer fibril-like structures have been made [101], but these only consist of peptides that can assemble into fibrils ~120 nm in length, less than half of the typical 300 nm length of a full collagen molecule [29].

Modular Collagen Gene Design

One major reason that full length, modifiable collagen constructs are not seen is due to the repeating nature of the collagen gene. The Gly-X-Y repeat prevalent in collagen means that any attempt at mutating a specific locus on the gene will be met with incorrect hybridizations in unintended areas along the gene. In the seminal work for this project, our laboratory [36] used an optimized computer algorithm [102] to design fragments of the human collagen III gene that are optimized for expression in yeast. These fragments utilized the degeneracy of the genetic code to create oligonucleotide sequences which favor correct overlaps during gene synthesis and disrupt mismatches. This computationally optimized DNA assembly (CODA) algorithm also took codon redundancy into account and changed the DNA sequence to codons preferentially used in yeast. The final result of this work was a modular system of small oligonucleotides that, when assembled together, form the entire full length collagen gene with an amino acid sequence exactly matching native human collagen III. The amino acid sequence remains unchanged, despite a change in the DNA sequence of the gene due to codon redundancy. Because collagen III is a homotrimer, only one gene sequence needed to be created for modification and engineering. The collagen gene is expressed in yeast, and as long as the genes for prolyl-4-hydroxylase have

been genetically integrated into the yeast genome, the collagen monomers will assemble into the final, triple helical collagen III molecule.

Collagen-mimetic proteins were created using the modular collagen gene assembly approach described in Chan *et al.* to produce full length recombinant human collagen III [36]. Because the triple helical region of collagen is typically problematic to assemble, it was divided into 12 modules in order to assist in gene synthesis and sequence design. Codon redundancy was used to create oligonucleotide sequences which were optimized to yield higher melting temperatures for correct hybridizations over incorrect hybridizations [102]. These optimized oligonucleotides were assembled into 12 primary fragments (PFs) by polymerase chain reaction (PCR). Primary fragments were then PCR-assembled into larger secondary fragments (SFs), which were subsequently PCR-assembled into the full length collagen gene. This allowed for the production of a recombinant protein with the same amino acid sequence as human collagen III ("baseline" collagen, or rCol), but with the codon sequence optimized for use yeast. Our primary fragment strategy also allows for the creation of collagen-mimetic mutants, proteins that share the majority of their amino acid sequences with collagen, but with slight variation [42].

Design of this modular collagen gene system was completed previously by Chan *et al*. [36]. The full length rCol gene was divided into three regions: front (comprising the N-terminal collagen propeptide), middle (comprising the triple helical region of collagen), and back (comprising the C-terminal collagen propeptide). The front and back regions make up their own individual secondary fragments (FSF and BSF, respectively), and are not modulated in our system. The front secondary fragment comprises two primary fragments (FPF 1 and 2), and the back secondary fragment comprises 4 primary fragments (BSF 1-4).

The middle region is divided into 12 primary fragments, which combine to form 3 secondary fragments, which combine to form the entire triple helical domain of collagen [36]. PFs 1-12 refer to PFs in this middle, triple helical region of collagen. The primary fragments were divided into either 10 or 12 DNA oligonucleotides with overlapping 20-30 nucleotide sequences (Integrated DNA Technologies, Corvall, IA). This oligonucleotide-primary fragment-secondary fragment-full length collagen strategy was used to create baseline rCol.

Motivation and Objectives

In order to better tune the physical and biological properties of tissue scaffolds made with collagen, modifications can be made to the sequence of collagen III by using a system of computationally-optimized oligonucleotide sequences which split up the collagen gene into many fragments. These fragments can be easily modified (unlike the full collagen gene) with non-native amino acid sequences to tune the collagen molecule to have desired characteristics. Here, the native collagen III was altered to produce collagen mutants with variable MMP-dependent cleavability. While synthetic scaffolds have been previously described with variable degradability by MMPs [44], the degradation ability of the natural tissue engineering scaffold collagen has never been modulated. By altering the sequence of native collagen III to change the location or number of MMP-cleavable sites it contains, the degradation profile of the collagen scaffold can be decoupled from other physical and biochemical properties of the collagen. This allows for more control and tenability over the collagen matrix as a tissue engineering substrate, while still maintaining the advantages of a natural, biomimetic scaffold.

Previous studies by our laboratory have performed 2D cell adhesion assays using our recombinant collagen [41-43], but the thermal instability of the recombinant collagen required that these assays be performed at 30 °C instead of 37 °C. The apparent melting temperature of our recombinant was measured to be 35.8 ± 0.5 °C [41]. Because this melting temperature is below 37 °C, it was hypothesized that the collagen was denaturing and preventing cell adhesion in assays performed at physiological temperature. Atomic force microscopy (AFM) was used to observe that recombinant collagen incubated at 37 °C contained a variety of different structures, including both intact and denatured collagen, while human collagen was expectedly intact [41]. The use of covalent chemistry was explored to stabilize our recombinant collagen so that it may be used at physiological temperatures for use in cell-based assays.

Therefore, the objectives for this thesis are:

1. Design, express and characterize collagen III variants with modified MMP sites, and examine whether they offer increased degradation over native human collagen III (Chapter 2).
2. Determine conditions to covalently stabilize recombinant collagen expressed in our yeast system to perform cell-based assays at 37 °C (Chapter 3).

Chapter 2: Collagen Mutant Creation and Degradation

Introduction

Previous work in our research group has demonstrated that collagens can be expressed in yeast, with amino acid sequences altered from the native collagen sequence [36, 42, 43]. The Gly-X-Y repeating nature of the collagen triple helix means that changing the collagen-encoding gene is difficult, and cannot be completed through conventional mutation methods [36]. By splitting the collagen gene into smaller primary- and secondary-fragments, the DNA sequence can be reliably modified by altering the desired collagen fragment sequence. The fragments can then be mix-and-matched together to create a full-length collagen gene with the desired sequence changes. The DNA of collagen modified in this way is then expressed in yeast because of the ability of yeast to hydroxylate collagen prolines when the genes for the expression of prolyl-4-hydroxylase have been integrated into the yeast genome [32].

MMPs are necessary for degradation of the ECM [80], and collagen is not readily degradable by non-specific proteases outside of MMPs [69]. Therefore, control over MMP-cleavable sites in collagen would offer the ability to create natural, biomimetic scaffolds for tissue engineering purposes with tunable degradation profiles. The ability of cells to cleave through and infiltrate the matrix would be decoupled from other physical and biochemical properties such as substrate stiffness and porosity. This gives added flexibility into the design of the matrix that has not been possible in a natural ECM mimicking tissue engineering scaffold.

In this work, non-native amino acid sequences were introduced into the human collagen III sequence to produce added or deleted MMP-cleavable sites using our modular

collagen system. PFs were mutated to contain a desired MMP site or MMP site knockout, mix-and-matched into SFs and again into full-length collagen genes to create MMP-variants of full length collagen. MMP-variant collagens were expressed in yeast and characterized using circular dichroism (CD) and digested using matrix metalloproteinase 1 (MMP-1) to examine whether an increased number of MMP cleavage sites does indeed result in enhanced collagen digestion. The formation of 3D collagen hydrogels was then explored with the goal of showing increased MMP-dependent digestion in a thick tissue mimic.

Materials and Methods

Mutagensis of Primary Fragments (PFs) for MMP-Site Mutants

Initial mutagenesis to fabricate MMP- mutant primary fragments was completed by Richard Que and Zhenghao Li. Collagen III PFs were originally designed by Chan *et al.* [36], and all PCR reactions for primary fragments, secondary fragments, and full length collagen variant gene assembly are based on protocols from this work [36]. Site-directed mutagenesis primers (48-60 base pairs long) were constructed to add a non-native MMP site into PFs 2 and 6, as well as remove a native MMP site in PF 10 (MMP knockout). Primers had a 3' overlap of 20 base pairs, and were designed to have higher annealing temperatures in the regions that did not overlap than the overlapping regions. In a 50 µL final volume, 1 µM mutation primer pairs, 10 ng DNA template (TOPO vector containing the PF), 200 µM dNTPs, 2.5 U PfuUltra II polymerase, and 1x PfuUltra II buffer were mixed. The PCR reaction was then set as follows: 5 minutes at 95 ^{o}C, followed by 25 cycles of 1 minute at 95 ^{o}C, 1 minute at 55-75 ^{o}C, 2 minutes at 72 ^{o}C. The reaction was finished off with 1 minute at 50-60 ^{o}C and 10 minutes at 72 ^{o}C. The PCR product was then digested with 20 U

DpnI for 2 hours at 37 °C, run on a 0.7% agarose gel, and extracted. The DNA was ligated into a pCR-BlunII-TOPO vector and transformed into *Escherichia coli* DH5α and grown on LB-Kanamycin. *E. coli* transformation was performed by adding at least 50 ng of PF DNA (in the TOPO vector) to 100 µL of competent DH5α tubes thawed on ice. The cells were incubated on ice for 30 minutes, then heat shocked at 42 °C for 30 seconds. Cells were then immediately placed on ice. 250 µL of Super Optimal broth with Catabolite repression (SOC) media was then added to the cells, which were placed in a 37 °C shaker incubator for 60 minutes. Finally, cells were added to LB-Kanamycin plates to grow for miniprep DNA acquisition. Miniprep was performed using a Qiagen spin miniprep kit (Qiagen, Venlo, Netherlands). Final PF sequences containing the desired mutations were confirmed by sequencing (Genewiz, South Plainfield, NJ). Table 2-1 below shows the mutagenesis primers used to create MMP-mutant primary fragments. DNA and amino acid sequences for all PFs are shown in Supplemental Information.

Table 2-1: Mutagenesis primers used to create MMP mutant primary fragments.

PF number	Primer sequence (5'->3') Forward strand on top, Reverse strand on bottom
PF 2 native DNA template	5'-TAG AGG ATC TGA CGG ACA ACC TGG CCC ACC TGG ACC ACC TGG AAC AGC TGG CTT TCC TGG-3' 5'-GTT GTC CGT CAG ATC CTC TAG CAC CAT CAT TAC CAC GGG CAC CGG CTG CTC CTG GCA AAC-3'
PF 2 MMP site addition	5'-TGC TGG AAT TAC TGG AGC TCG TGG CTT GGC TGG ACC ACC TGG AAC AGC TGG CTT TCC TGG -3' 5'-GAG CTC CAG TAA TTC CAG CAA TAC CCA ATG GAC CAC GGG CAC CGG CTG CTC CTG GCA AAC-3'
PF 6 native DNA template	5'-AGG TGA TAC CGG ACC ACC TGG ACC CCA AGG ATT GCA AGG ACT ACC TGG TAC AGG AGG ACC-3' 5'-CAG GTG GTC CGG TAT CAC CTT TGT CTC CAC CTG GAC CTG TTG GGC CGG GTG GCC CTT GTG-3'
PF 6 MMP site addition	5'-TGG TAT TAC CGG AGC TCG TGG ATT GGC TGG ATT GCA AGG ACT ACC TGG TAC AGG AGG ACC-3' 5'-CAC GAG CTC CGG TAA TAC CAG CAA TTC CCA ATG GAC CTG TTG GGC CGG GTG GCC CTT GTG-3'
PF 7 native DNA template	5'-GGT CCT CCT GGT GCT GCC GGG ACT CCA GGT TTG CAA GGT ATG CCA GGA GAG AGA GG-3' 5'-CCC GGC AGC ACC AGG AGG ACC AGG TGG TCC TGC AGC TCC TTT ACC TCC TTC GGG ACC AG-3'
PF 7 MMP site addition	5'-GGT ATT GCT GGT ATT ACT GGG GCT AGA GGT TTG GCT GGT ATG CCA GGA GAG AGA GG-3' 5'-CCC AGT AAT ACC AGC AAT ACC CAA TGG TCC TGC AGC TCC TTT ACC TCC TTC GGG ACC AG-3'
PF 10 native DNA template	5'-TTA CTG GTG CTA GGG GCT TGG CTG GAC CGC CAG GCA TGC CAG GTC CTA GGG GCA GTC C-3' 5'-AAA CCA TTA GCA CCT GGT TTT CCT GAC TCG CCT TTA ACG CCT TGT GGA CCT GGA CTG C-3'
PF 10 MMP site KO	5'-TTC CTG GTG CTC CTG GTT TAC CTG GAC CGC CAG GCA TGC CAG GTC CTA GGG GCA GTC C-3' 5'-AAA CCA GGA GCA CCA GGA ATT CCA GGT ATT CCC GGT GGA CCT GGT GCG CCG GGC GGA C-3'

Secondary Fragments

Secondary fragment assembly was completed by Richard Que. Primary fragments with and without non-native MMP sites or the MMP knockout sequence were mixed and matched to create secondary fragments with MMP sites in the desired location with the following reactions: In a final volume of 50 µL, 2 nM of each primary fragment, 200 nM of each secondary fragment primer, 200 µM dNTPs, 1.5 U PfuUltraII, and 1x PfuUltra II reaction buffer were mixed together. This reaction was then set in a thermocycler with the following program: 10 minutes at 95 ^{o}C, then 16 cycles of 20 seconds at 95 ^{o}C, 30 seconds at 68 ^{o}C, 45 seconds at 72 ^{o}C, and a final hold at 72 ^{o}C for 5 minutes. The reaction products were run on a 0.7% agarose gel, and the product band of expected size was excised. This product band was then ligated into the pCR-BluntII TOPO vector, transformed into *E. coli* DH5α as before and placed on LB-Kanamycin plates. Colony DNA was sequenced to screen for mutations and the correct sequence. DNA and amino acid sequences for all SFs are shown in Supplementary Information.

Full Length Collagen Genes

Full length collagen mutant genes were created by mixing and matching the appropriate secondary fragments using the following reaction: In a final volume of 50 µL, 2 nM of each secondary fragment with 200 nM of each cloning primer, 200 µM dNTPs, 1 mM $MgSO_4$, 1 U KOD Hot Start DNA Polymerase and 1x KOD reaction buffer (EMD Millipore, Billerica, MA) were mixed together. This reaction was then set in a thermocycler with the following program: 10 minutes at 95 ^{o}C, then 16 cycles of 20 seconds at 95 ^{o}C, 30 seconds at 68 ^{o}C, 1.25 minutes at 72 ^{o}C, and a final hold at 72 ^{o}C for 5 minutes [36]. The reaction

products were run on a 0.7% agarose gel, and the product band of expected size was excised. This product band was then ligated into the pCR-BluntII TOPO vector, transformed into *E. coli* DH5α and placed on LB-Kanamycin plates. Colony DNA was sequenced to screen for mutations and the correct sequence.

Correct full length collagen mutants were then transferred to the CEN/ARS plasmid, described previously by Chan *et al.* [103]. The collagen gene is located after the GAL1 promoter on the plasmid. Therefore, yeast grown with galactose and without competing sugars will be forced to utilize the GAL1 promoter sequence and in turn produce collagen. This resulted in the plasmids YCp-MCOL-0M, YCp-MCOL-1M-PF2, YCp-MCOL-1M-PF6, YCp-MCOL-1M-PF7, YCp-MCOL-2M, and YCp-MCOL-3M. YCp-MCOL-0M produces collagen containing the native MMP site knockout (rCol-0M). YCp-MCOL-1M-PF2, YCp-MCOL-1M-PF6, and YCp-MCOL-1M-PF7 all contain the native MMP site knock out plus a new MMP site on either PF 2, 6, or 7 to produce rCol-1M-PF2, rCol-1M-PF6, or rCol-1M-PF7, respectively. YCp-MCOL-2M contains the native on PF 10 plus an additional site on PF 2 (rCol-2M), and YCp-MCOL-3M contains the native site plus additional sites on PFs 2 and 6 (rCol-3M). Figure 2-1 shows the created collagen mutants and the placement of additional MMP sites along the collagen gene. DNA and amino acid sequences for full length collagen genes are shown in Supplemental Information.

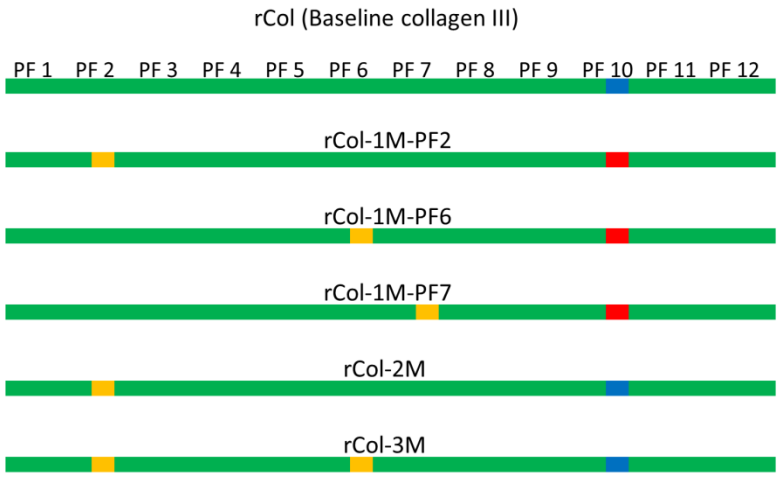
Protein Expression and Purification

Full length collagen genes were transformed into *Saccharomyces cerevisiae* BYα2β2, a strain which contains two integrated copies each of the α- and β-subunits of human prolyl-4-hydroxylase, and used to express the recombinant collagen[41]. To transform

yeast with the collagen gene, yeast were grown to OD_{600} 0.7 in 5 mL of selective SDC(A,U) media. Cells were centrifuged at 3000 RPM for 5 minutes in a Beckman GS-6 centrifuge (Beckman Coulter, Incorporated, Brea, CA). The supernatant was discarded and the cell pellet was resuspended in 1 mL of sterile MilliQ water and transferred to a sterile 1.5 mL centrifuge tube. Cells were centrifuged at 3000 RPM for 5 minutes in a microcentrifuge, and the supernatant discarded. The cell pellet was resuspended in 1 mL sterile 1X TE / LiOAc, made fresh from 10X TE (100 mM Tris, 100 mM EDTA) and 10X LiOAc (1 M LiOAc), then centrifuged for 5 minutes at 3000 RPM. The supernatant was discarded and the cell pellet was resuspended in 250 µL of sterile 1X TE/ LiOAc. 50 µL aliquots of this cell suspension was mixed with 500 ng of DNA, then 300 uL of sterile 40% PEG 4000 (Sigma Aldrich, St. Louis, MO), 1X TE, 1X LiOAc was added. Cells were incubated at 30 °C for 1 hour with occasional shaking. 40 µL of sterile dimethyl sulfoxide (DMSO) was added to the cells, which were then heat shocked at 42 °C for 15 minutes. Cells were then centrifuged at 1.8 kg for 20 seconds, the supernatant discarded, and the cells resuspended in 1 mL 1X TE. The cells were then again centrifuged at 1.8 kg for 20 seconds, the supernatant discarded, and the cells resuspended in 200 µL 1X TE. Finally, cells were plated onto selective media SDC(A) plates and allowed to grow until colonies formed.

To express recombinant collagen, yeast were cultured in selective media as described previously [41]. Yeast were cultured in SDC(A) and SGC(A) media with 5 g/L bacto casamino acids, 1.7 g/L yeast nitrogen base(BD Biosciences, Franklin Lakes, NJ), 5 g/L ammonium sulfate (Thermo Fisher Scientific, Waltham, MA), 100 mg/L adenine sulfate (Sigma), and 20 g/L glucose or galactose, respectively [41]. Yeast were grown in SDC(A) at 30 ^{o}C to OD_{600} between 1-2, as measured by a SpectraMax M2 plate reader (Molecular

A



B

Location	Sequence modified														
PF 10 native MMP site	G	P	L	G	I	A	G	I	T	G	A	R	G	L	A
PF 10 KO	G	P	P	G	I	P	G	I	P	G	A	P	G	L	P
PF 2 native	G	N	D	G	A	R	G	S	D	G	Q	O	G	P	O
PF 6 native	G	P	G	G	D	K	G	T	G	G	P	O	G	P	Q
PF 7 native	G	P	O	G	P	O	G	A	A	G	T	O	G	L	Q

Figure 2-1*: (A) Placements of additional non-native MMP sites and native MMP site knockout in engineered recombinant collagen III. Blue represents the native MMP-cleavable site, yellow represents new, non-native sites, and red indicates a native MMP site knockout. rCol-2M contains the native MMP site on PF 10 as well as an additional site on PF 2. rCol-3M contains the native MMP site and additional sites on PFs 2 and 6. All PFs shown here are middle PFs within collagen's triple helical region. The N- and C-terminal regions and their subsequent PFs are not shown here; they are expressed as part of the full-length collagen molecule and are involved in protein trimerization, but are removed by pepsin digestion after triple helical assembly [43]. (B) DNA sequences modified to create collagen variants. The native MMP site on PF 10 was replaced with the KO sequence where applicable. This table also shows the native sequences on PFs 2, 6 and 7 that were replaced with the MMP-cleavable sequence (from PF 10) in the appropriate collagen mutants.*

Devices, Sunnyvale, CA). Cultures were then re-inoculated (1% v/v) into SGC(A) media, and allowed to grow to OD_{600} of 3. Cells were harvested by centrifugation and frozen at -80 ^{o}C for at least 30 minutes. Expression of collagen and hydroxylase genes in these yeast is controlled by the *GAL1* promotor, so the switch into SGC(A) media caused the expression of the recombinant genes [41].

To purify the recombinant collagen, cell pellets were resuspended in ice cold buffer with 100mM tris-HCl and 400mM NaCl (pH 7.4) to 0.1 g/L. 2 U/mL DNAse (Worthington Biochemicals, Lakewood, NJ), 1 U/mL RNAse If (New England Biotech), 1mM $MgCL_2$ and 1 mM phenylmethylsulfonyl fluoride (PMSF) were added to the cell-containing buffer. Cells were then lysed at 25000 psi with a French press cell disrupter (Thermo Fisher) and EDTA was added to 1mM. To digest unwanted yeast proteins, the pH of the cell solution was dropped to 2 using 4M HCl and incubated with 0.2 mg/mL pepsin (MP Biomedicals, Santa Ana, CA) for 12 hours at 4 ^{o}C. Pepsin digests unwanted proteins, but collagen's triple helical stability prevents it from being digested. After pepsin digestion, centrifugation was used to clear the lysate. 0.5M acetic acid and 3M NaCl (final concentrations) were then added to the collagen-containing supernatant, and allowed to incubate at 4 ^{o}C for 1 hour before being centrifuged. The precipitate was then dissolved in 0.1N HCl, and 200mM tris-HCl pH 8.6 was added to raise the pH. 5M NaCl was added to 3M final concentration to precipitate the collagen, and the sample was incubated at 4 ^{o}C for 1 hour then centrifuged. The pellet was dissolved in 0.1N HCl and dialyzed against 50mM acetic acid at 4 ^{o}C for 4 hours. Collagen concentration was determined by BCA assay (Pierce, Rockford, IL) and size and purity by SDS-PAGE. Circular dichroism was used to verify the triple helical structure of the final purified product.

Circular Dichroism

Circular dichroism (CD) spectroscopy was performed in order to determine the structure and melting temperature of the collagen mutants. Measurements were performed with a J-810 spectropolarimeter (Jasco, Easton, MD) [104]. Ellipticity was determined by placing purified recombinant collagen samples, at 100-300 μg/mL concentration in 50mM acetic acid, in a 1 mm quartz cell and scanning at 10 nm/min. Apparent melting temperature (T_m) was determined by measuring ellipticity at 221 nm while increasing temperature at 1 ^oC/min [105]. The Gibbs-Helmholtz equation was used to fit molar ellipticity vs. temperature [105].

Anti-Collagen Western Blot

7% SDS-PAGE was completed, and the gel and a nitrocellulose membrane (Roche Diagnostics, Indianapolis, IN) incubated in transfer buffer (25 mM tris, 192 mM glycine, 20% methanol) for 5 minutes. Protein transfer was completed, and the membrane was incubated in blocking solution (0.5% milk in tris-buffered saline (TBS), 20 mM tris, 500 mM NaCl, pH 7.5) for 2 hours, then washed twice with TBS for 15 minutes. The membrane was incubated with goat anti-collagen primary antibody (Santa Cruz Biotechnology, Dallas, TX) at 1:1000 dilution in TBS overnight at 4 °C, then washed 3x with TBS + 0.5% tween-20 (Anatrace, Maumee, OH). After washing, the membrane was incubated with rabbit anti-goat secondary antibody (Santa Cruz Biotech) for 2 hours at room temperature, then washed 3x with TBS + 0.5% tween-20. The membrane was incubated in Genius Buffer (100 mM tris, 50 mM $MgCl_2$, 100 mM NaCl, pH 9.5) for 5 minutes, then incubated in developing solution

(200 µL of NBT/BCIP (Roche) in 10 mL of Genuis buffer), washed with MilliQ water and imaged.

MMP Digestion Assay

The protocol for this assay was based on Williams *et al.* [81]. Inactive pro-MMP-1 (R&D Systems, Minneapolis, MN) was activated using *p*-aminophenylmercuric acetate (APMA) (Sigma) at 1mM final concentration in TNC buffer (50mM tris-HCl, 150mM NaCl, 10mM $CaCl_2$, pH 7.5) and incubated at 37 ^oC for 2 hours. This activation with APMA cleaves the MMP's N-terminal pro-peptide domain and allows the active site of the enzyme to bind to and hydrolyze collagen[62]. MMP-1 was added to 1.5 µg of collagen in TNC buffer at a 1:5 molar ratio of MMP-1:collagen in 30 µL reaction volumes. Samples were incubated for 18 hours at 25 ^oC. The reaction was terminated with 25 mM EDTA, and the mixture boiled at 95 ^oC and run on 12% SDS-PAGE. Gels were stained with Imperial Coomassie blue (Thermo Fisher) over night, de-stained with water, and imaged.

Microrheology Glass Slide Silanization

Standard 3 inch x 1 inch x 1mm glass slides (VWR International, Radnor, PA) and 18 mm x 18 mm glass cover slides (VWR) were silanized to increase glass surface hydrophobicity. Slides were submerged in 1% Hellmanex (Hellma Analytics, Müllheim, Germany) and baked at 60 °C for 1 hour. Each slide and cover slip was then washed 3x with MilliQ water and placed in a dust-free environment to dry. Slides were then washed with acetone and placed in a solution of 2% (3-Aminopropyl)triethoxysilane (APTES) in acetone and agitated for 2 minutes. Slides were washed again with acetone and allowed to dry in a

dust-free environment. Complete microrheology slides were made by placing 2 pieces of double-sided tape on top of a glass slide approximately 1 cm apart. A silanated glass cover slip was then placed on top of the tape, creating a 70 μm-deep channel [42]. 10 μL of liquid hydrogel solution was then added by capillary action to each channel, and the channels were sealed using UV-sensitive Norland optical adhesive 81 (Norland Products, Cranbury, NJ) and cured for 1 minute with a handheld UV light.

Polystyrene Microparticle Prepration

Aminated fluorescent polystyrene 1 μm-diameter microparticles (Invitrogen) were PEGylated with ms$(PEG)_{24}$ (Pierce) to prevent particle aggregation. All microparticle work was done in the dark. 100 μL of stock aminated fluorescent particles (2% v/v) were sonicated for 5 minutes, then centrifuged and washed 3x at 6000 RPM with potassium phosphate buffer (pH 7.4). After each centrifugation step, the supernatant was decanted and replaced with potassium phosphate buffer, including the last step after which the the particles were resuspended in 500 μL potassium phosphate buffer. 1 μL of ms$(PEG)_{24}$ solution was diluted into 9 μL potassium phosphate buffer, then 7.29 μL of this diluted PEG was added to the microparticles. Particles were agitated for 2 hours at room temperature, then overnight at 4 °C. Microparticles were then washed 3x with MilliQ water, and finally resuspended in 100 μL MilliQ water.

Similarly, aminated fluorescent particles were coated with either 10 kDa or 40 kDa polyvinylpyrrolidone (PVP) to prevent particle aggregation, with a protocol based on Graf *et al.* [106]. Calculations were completed to adhere 60 PVP molecules to each microsphere. 100 μL of microparticles were sonicated for 5 minutes and washed 3x with MilliQ water, as

above, and finally resuspended in 500 µL water. 103.5 mg of 10 kDa PVP and 414 mg of 40 kDa PVP were each dissolved into separate 1 mL glass vials of MilliQ water by sonication. 500 µL of each of these PVP solutions was added to the 500 µL of particles, and the two different PVP-particles solutions were agitated for 2 hours at room temperature, then overnight at 4 °C. Microparticles were then washed 3x with MilliQ water, and finally resuspended in 100 µL MilliQ water.

Hydrogel Formation in Microrheology slides

2 mg / mL collagen hydrogels were crosslinked using a 20 kDa 4-arm NHS-PEG (Nanocs Incorporated, New York, NY) in 20 µL reaction volumes with 1:50 and 1:500 molar ratios of NHS:lysine, with 1X PBS used as a reaction buffer at pH 7. Fluorescent tracer microparticles prepared as above were used as a diagnostic tool to determine if hydrogels were formed. Particles were sonicated for 5 minutes, then 1 µL of fluorescent particles was used per 20 µL reaction (final concentration of particles was 0.1 % v/v). Hydrogel reagents were mixed together and added to the silanized glass microrheology slides, sealed, and rotated at 4 °C for at least 4 hours in the dark. Fluorescent particle movement was observed with an Axio Observer inverted microscope (Carl Zeiss Microimaging, Inc., Jena, Germany). Hydrogels were determined to be formed if the tracer particles were not moving throughout the channel. If particles were moving (due to random Brownian motion), then no hydrogel had formed.

Microfluidic Device Design and Photolithography

In order to create a system in which small hydrogels could be formed, but also digested with the addition of MMP, a microfluidic platform was developed. SU-8 deposition and photomask fabrication was completed in collaboration with Tiffany Lin. SU-8 2050 (MicroChem, Westborough, MA) was deposited onto a silicon wafer using standard photolithography methods [107] for the fabrication of a master mold, in order to make polydimethylsiloxane (PDMS) microfluidic devices to test collagen hydrogel formation. Figure 2-2 shows a diagram of the device template used to fabricate PDMS microfluidic devices. Microfluidic channels are 500 µm wide and 10 mm long, with a 1.5 mm diameter

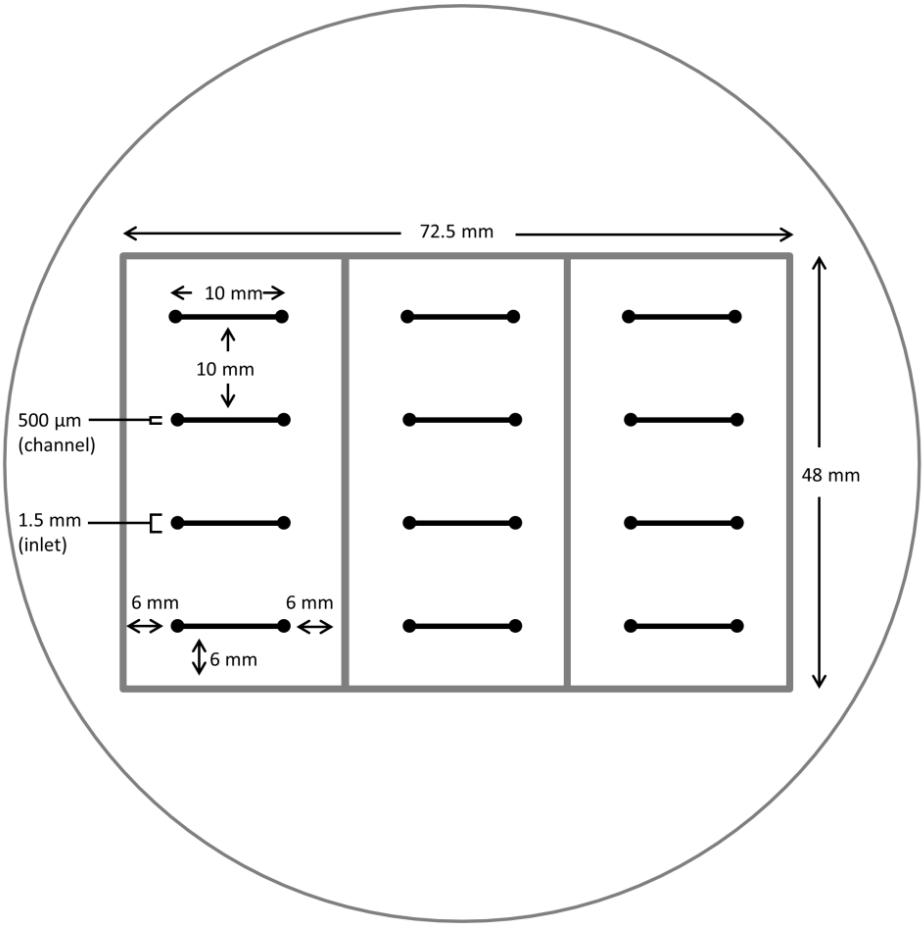


***Figure 2-2**: Microfluidic device photoresist template.*

circle at either end for inlets and outlets to allow fluid to flow through the channels. Fabricated molds are measured to be between 86-90 μm deep. There are 12 channels per photoresist mold, and the design allows for division into 3 PDMS devices. Master molds were silanized using trichloro(1*H*,1*H*,2*H*,2*H*-perfluorooctyl)silane (Sigma). PDMS (Dow Corning Corporation, Midland, MI) was mixed at a 10:1 ratio of base to curing agent, de-gassed, cast into the master mold and allowed to cure at 65 °C overnight. The cured PDMS devices were cut out according to Figure 2-2 including holes for the inlets and outlets, and plasma bonded onto standard 3 inch x 1 inch x 1 mm glass slides (VWR) to create an irreversible seal using a Harrick plasma cleaner (Harrick Plasma, Ithaca, NY). After bonding, glass tubes (7 mm outer diameter, 5 mm inner diameter, 2 cm long) were attached using PDMS above each inlet and outlet. Figure 2-3 below shows the complete microfluidic device. Positive pressure is created by adding 200 μL of liquid to a tube at the arbitrarily defined "inlet" end of the channel and 50 μL of liquid in the "outlet" tube. This

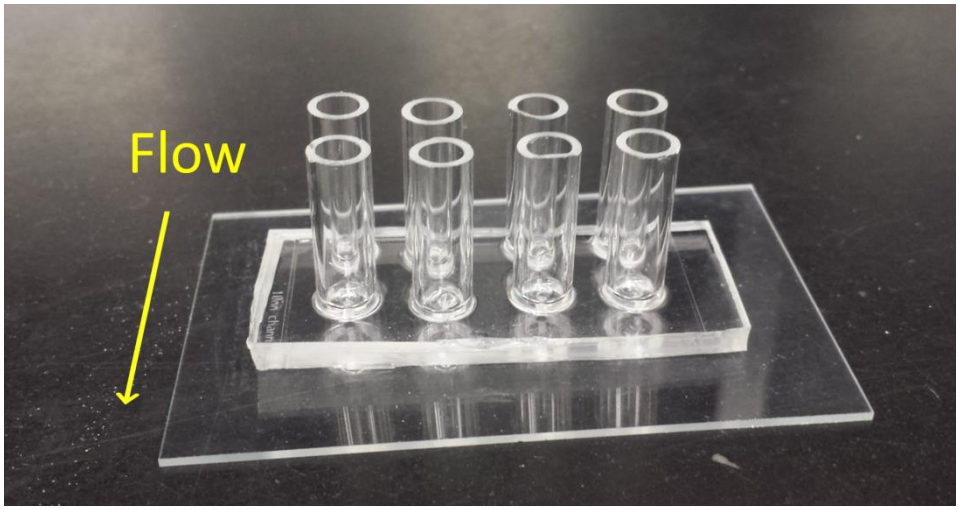


Figure 2-3*: Final, complete PDMS microfluidic device. Each device has 4 channels, which run along the shorter length of the device.*

creates both positive pressure and a continuous liquid-liquid interface which allows for liquid, such as MMP-carrying buffer, to flow through the channel without needing to break through liquid- air interface surface tension.

Results and Discussion

Construction Strategy of MMP Variants of Recombinant Collagen

The ability of cells to invade and remodel their matrix environment is important in tissue engineering applications [65]. However, tunable matrix degradation has never been shown in collagen systems, only in synthetic matrices [63, 82, 108]. To add MMP variations to our recombinant collagen III, either additional non-native MMP sites or to knockout the native MMP site, modifications were made to the appropriate primary fragment (see Figure 2-1). The primary fragments were mix and matched into secondary fragments, then again into the full length, modified collagen gene.

Portions of the native collagen sequence on PFs 2, 6, and 7 were replaced with the MMP-cleavable sequence *GPLGIAGITGARGLA* located natively on PF 10. The locations of the added MMP digestion sites on PFs 2 and 6 were chosen because of their position along the collagen gene; the native MMP site exists around ¾ of the length of the collagen molecule, so PF 2 was chosen because it is close to ¼ of the length of the protein, and PF 6 was chosen because it is close to the middle of the length of the collagen molecule. For the rCol-3M collagen mutant, MMP digestion sites therefore exist at ¼, ½, and ¾ the length of the protein, effectively cutting the collagen into quarters. New MMP sites were also placed into positions which avoided integrin-binding sites, which are needed for cell adhesion.

The native MMP site in human collagen III exists in a region of low triple helical stability [69], which helps collagenases further unwind the collagen molecule to cleave it [79]. Areas of high stability flank the native MMP site in both directions. In order to preserve this motif of High Stability – Low Stability – High Stability, non-native MMP sites on PFs 2 and 6 were added into areas of similar low triple helical stability in between regions of high stability, as determined by apparent melting temperature from Raynal *et al.* [109]. "Low stability" regions were chosen as peptides from the work in Raynal *et al.* with apparent melting temperatures of 39.0 °C or lower, and "high stability" regions were peptides with melting temperatures of 42.0 °C or higher. In order to potentially perform experiments to explore the effects of collagen stability in the regions surrounding the MMP site, a non-native MMP-cleavable site was added into an area of high stability on PF 7, with further areas of high stability on either side of the replaced sequence. This created a High Stability – Low Stability (MMP site) – High Stability region that did not exist previously. The low-stability MMP site therefore replaced the middle high-stability region on PF 7.

An MMP-insensitive sequence was used to create a knockout of the native MMP sequence located on PF 10. The MMP-cleavable sequence *GPLGIAGITGARGLA* was replaced with *GLPGIPGIPGAPGLP*, with each amino acid in the "Y" position of the Gly-X-Y repeat being replaced with a proline (which presumably becomes post-translationally modified into a hydroxyproline). This sequence change has previously been shown to prevent cleavage of the collagen molecule by MMP by stabilizing the triple helix [81] and becoming unrecognizable by MMP-1 [62]. Using the newly added MMP sites onto PFs 2, 6, and 7, as well as the native MMP-site knockout on PF 10, PFs were mixed and matched to design MMP-variant collagens with different MMP-cleavable properties. rCol-1M-PF2 contains a

knockout of the native MMP sequence on PF 10, as well as a non-native MMP site on PF 2. rCol-1M-PF6 and rCol-1M-PF7 follow suit, with additional MMP sites on PFs 6 and 7, respectively, as well as knockouts of the native MMP sequence. rCol-2M contains the native MMP site as well as an additional MMP site on PF 2, and rCol-3M contains additional MMP sites on PFs 2 and 6 to accompany the native site on PF 10. The collagen mutants and locations of MMP sites herein are described in Figure 2-1.

Collagen Characterization

There is precedent from our laboratory that multiple amino acid mutations in the human collagen III sequence can still result in a standard triple helical structure [43], and we confirmed this for the MMP collagen variants as well. Circular dichroism spectroscopy was used to determine the structure and thermostability of the recombinant collagen variants. These CD spectra and temperature stability profiles are shown in Figures 2-4 and 2-5, and are summarized in Table 2-2.

All collagen variants show the typical triple helical structure of collagen, as indicated by the positive ellipticity peaks at 221 nm and troughs at 198 nm [104]. Melting temperatures for the MMP-variant collagens are similar to previously reported melting temperatures for recombinant human collagen in our yeast system [41, 43], as well as other recombinant fibrillar-type collagens expressed in yeast [31]. rCol-0M has a statistically significantly ($p<0.05$) higher melting temperature than both rCol and rCol-3M (statistics performed using ANOVA and subsequent Tukey test). No other pairs of variants had significantly different apparent melting temperatures. These melting temperatures are lower than native human collagen III (39.1 ± 0.1 °C) [42], possibly due to the heterologous

expression of collagen in yeast [41, 110], or to lower hydroxylation levels of our recombinant collagen versus native human collagen [43]. Despite the slightly lower melting temperatures, all collagen variants shown here are triple-helical and mimic native human collagen III. It is not unexpected that there is a difference in thermostability between

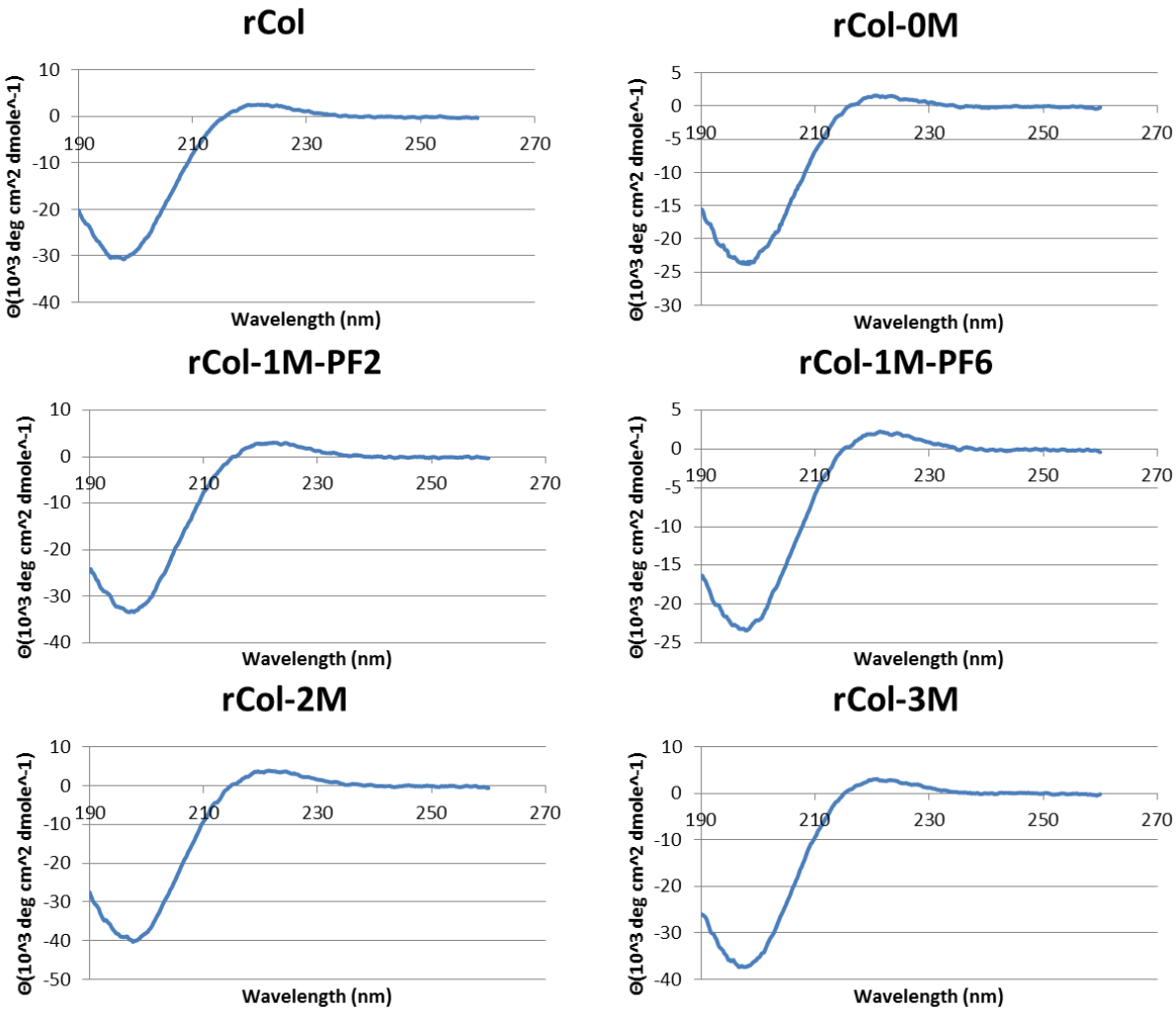


Figure 2-4*: Representative CD spectra wavelength scans of expressed recombinant collagen variants. All collagens in this figure show the characteristic triple helical structure of collagen [104].*

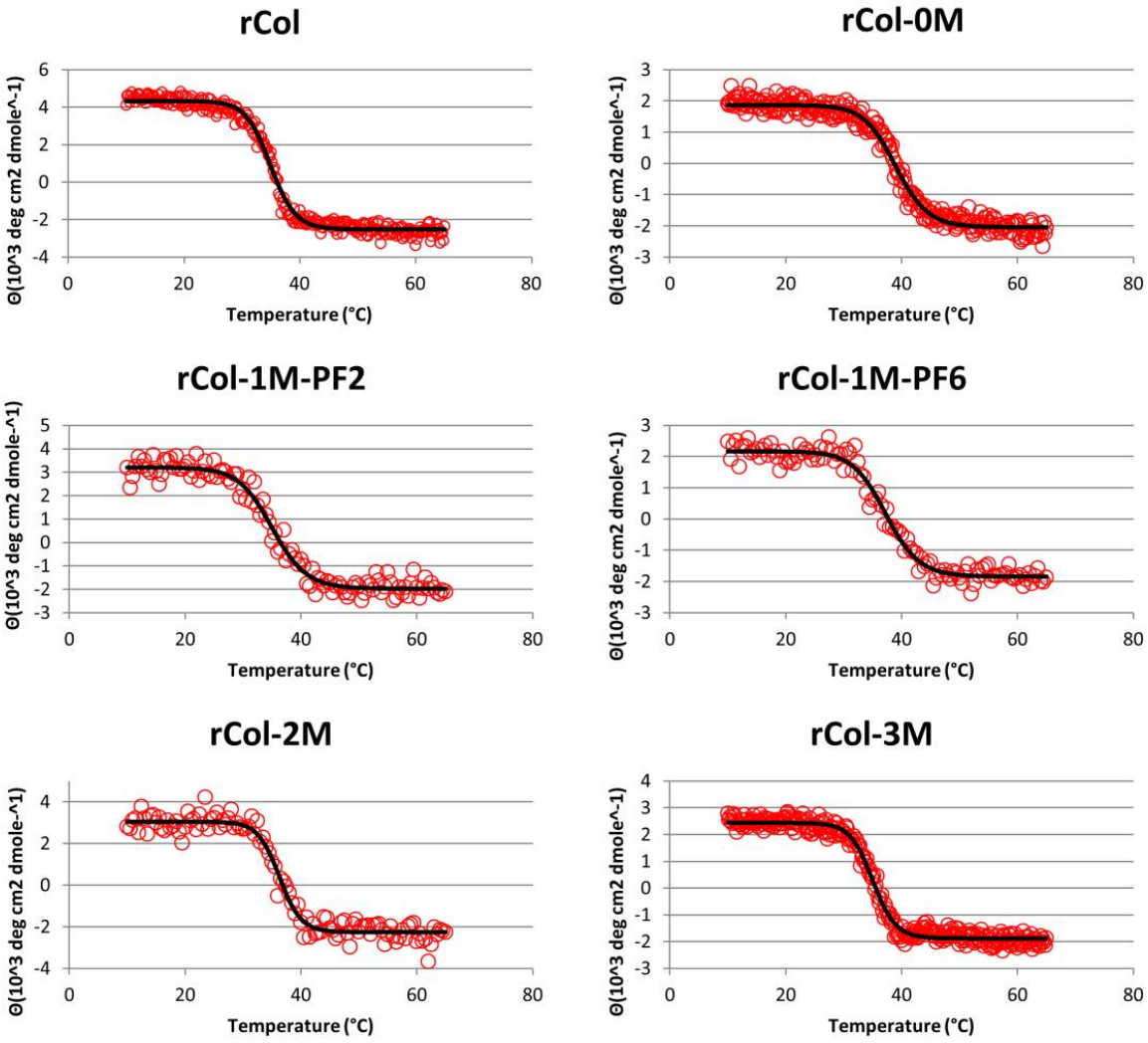


Figure 2-5*: Representative melting temperature scans for expressed recombinant collagen variants.*

Table 2-2: Collagen variant apparent melting temperatures (n=2 expressions).

Collagen variant	Average melting temperature (°C)
rCol	35.3 ± 0.6
rCol-0M	38.8 ± 0.2
rCol-1M-PF2	36.1 ± 1.5
rCol-1M-PF6	36.5 ± 0.4
rCol-2M	35.7 ± 0.9
rCol-3M	35.1 ± 0.1

rCol-0M and rCol-3M. The MMP-cleavable sequence inserted into these collagens is inherently a low-stability site [69, 80, 109], while the MMP-KO sequence has inherently higher stability [81]. It stands to reason that adding in additional sites of low stability would decrease overall stability of the collagen triple helix in the rCol-3M mutant, while introducing a high-stability site in the rCol-0M mutant would increase triple helix stability. However, this needs to be confirmed via further studies that examine the stability of the triple helix at these sites on a local level, and determine their overall effect on the thermostability of the collagen triple helix as a whole.

As noted by Que *et al.* [43], this flexible system of collagen variant fabrication allows for the expression of triple-helical, thermally stable recombinant human collagen, including non-native mutations introduced into the collagen amino acid sequence. rCol-3M collagen contains 19 amino acid mutations from the native sequence, confirming that large numbers of amino acid mutations can be introduced into this collagen system, provided the Gly-X-Y motif typical of collagen remains intact [43]. This previous work had shown that up to 40 non-native amino acid mutations could be introduced and still result in triple helical collagen, further consistent with the ability of the collagen scaffold to tolerate the additional MMP-cleavable sequences that were introduced here, provided the integrity of the triple helix is maintained.

Failure of rCol-1M-PF7 to be Successfully Purified

Despite the robustness of the collagen sequence and its abilities to tolerate mutations, the scaffold cannot handle all introduced mutations successfully. While collagen variants created in this work had MMP-cleavable sites introduced into regions of low

thermostability (mimicking the native MMP site in collagen III), rCol-1M-PF7 had an MMP site introduced into a region of high stability. The objective was to examine the effects of triple helix stability on the ability of MMP to cleave collagen. The rCol-1M-PF7 gene was successfully fabricated and transformed into BYα2β2 yeast. However, rCol-1M-PF7 was never successfully purified, yielding only CD spectra indicative of contaminating proteins, but not collagen. An anti-collagen Western blot was performed after cell breaking, and before further purification, to determine if the collagen was being expressed by the yeast before the purification process began (Figure 2-6). From Figure 2-6 it appears that collagen

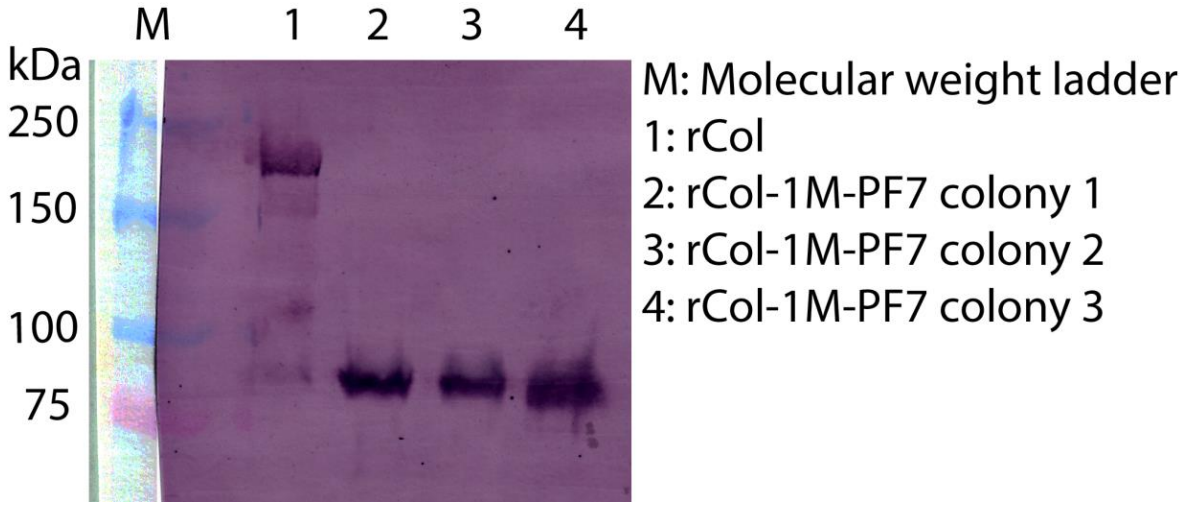


Figure 2-6*: Anti-Collagen Western blot on expressed colonies of rCol-1M-PF7 collagen after breaking via French Press, with no further purification or separation. Lane 1: rCol, used as a positive control. Lanes 2-4: Different transformation colonies of rCol-1M-PF7. The rCol-1M-PF7 colonies do express collagen, but the length of the molecule appears much smaller than expected.*

is being expressed by the rCol-1M-Pf7 yeast colonies, although the molecular weight of the protein is approximately half of what would be expected judging by the positive control

rCol. It is hypothesized that the full-length collagen gene is being expressed, then the protein is being digested near the added MMP site by a protease. Because the non-native MMP site, a sequence of inherently low triple helical stability, was added into an area that natively showed high stability in the collagen sequence, it is possible that the triple helix became unstable and allowed a protease to cleave through the individual collagen monomers. Collagen's resistance to degradation by non-specific proteases comes from its triple helical stability, so it is possible that a loss in triple helix fidelity could lead to proteolytic cleavage. Although protease inhibitors are used during the process of breaking yeast cells with the French Press (PMSF and EDTA), an additional protease-inhibitor cocktail (*SIGMA*FAST, Sigma) was added to the cell suspension, to test if a yeast protease was digesting the rCol-1M-PF7 collagen after the cells were broken. The addition of a new protease-inhibitor cocktail did not prevent the collagen from being digested by an intracellular protease that exists within the yeast cell, but could prevent proteolysis resulting from proteases being released into the cell suspension after the cells are broken.

The Western blot was repeated using the additional protease inhibitor cocktail in the cell suspension during cell breaking, and is shown in Figure 2-7. Despite the additional protease inhibitor, the rCol-1M-PF7 collagen still runs much smaller than expected. For this experiment, rCol-1M-PF7 samples were separated into soluble and insoluble fractions to determine if the collagen was stuck in the pellet, or if it came out into the soluble fraction. This was to help find where in the purification process the collagen was being lost. It is still unclear what causes the shortening of this collagen variant, and if the rCol-1M-PF7 collagen is initially triple-helical. Without a purified sample it is impossible to determine if the mutant ever produced triple-helical collagen, regardless of the action of possible proteases.

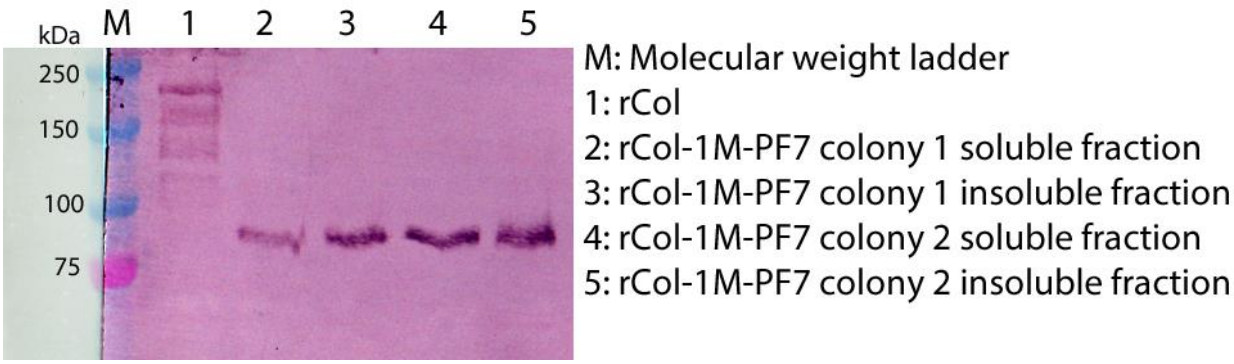


Figure 2-7*: Anti-Collagen Western blot of rCol-1M-PF7 collagen after breaking via French Press with the additional SIGMA*FAST *protease inhibitor cocktail added immediately prior to breaking. After cells were broken, the samples were separated into soluble (supernatant) and insoluble (pellet) fractions by ultracentrifugation at 30,000 RPM for 30 minutes. Lane 1: rCol, used as a positive control. Lanes 2-3: rCol-1M-PF7 transformation colony 1. Lanes 4-5: rCol-1M-PF7 transformation colony 2.*

More studies would be required to test the hypothesis that proteases were to blame for the shortening of the rCol-1M-PF7 collagen. Additional discussion of the failure to purify rCol-1M-PF7 is in the Supplementary Information section. Due to time constraints, the expression of rCol-1M-PF7 collagen was not pursued further.

Digestion of Collagen Variants with MMP-1

Collagen variants were incubated with MMP-1 to test the hypothesis that non-native MMP sites introduced into the collagen backbone could allow for cleavage at novel sites in the collagen molecule. In turn, cleavage of additional MMP sites along the collagen length should result in increased degradation of the biopolymer. Table 2-3 shows the expected molecular weights of each collagen variant if MMP-1 were to cleave each molecule at its

MMP site. In Table 2-3 each digestion product is broken down into its particular portion of the collagen molecule, and its expected size. These values were calculated assuming that the pepsin digestion used in purification of the recombinant collagen cleaved the collagen at both termini, leaving only the triple helical portion of the collagen molecule, located

Table 2-3: Expected protein band size of each collagen mutant after digestion with MMP-1. Proline hydroxylation was assumed to be at the reported level of 31.1 ± 2.4% [41].

Collagen Mutant		Portion of collagen	Length of expected digestion band (kDa)
rCol		N-terminus to PF 10 digestion site	70.2
		PF 10 digestion site to C-terminus	22.6
rCol-0M		N-terminus to C-terminus	92.8
rCol-1M-PF2		N-terminus to PF 2 digestion site	14.6
		PF 2 digestion site to C-terminus	78.1
rCol-1M-PF6		N-terminus to PF 6 digestion site	42.0
		PF 6 digestion site to C-terminus	50.8
rCol-2M		N-terminus to PF 2 digestion site	14.6
		PF 2 digestion site to PF 10 digestion site	55.6
		PF 10 digestion site to C-terminus	22.6
rCol-3M		N-terminus to PF 2 digestion site	14.6
		PF 2 digestion site to PF 6 digestion site	27.4
		PF 6 digestion site to PF 10 digestion site	28.2
		PF 10 digestion site to C-terminus	22.6

between amino acids 168-1196. Collagen's triple helix confers its stability and general resistance to proteases, so it assumed that the non-helical termini of the collagen molecule were digested by pepsin during purification. The reported proline hydroxylation levels of 31.1% [43] in our yeast collagen expression system was also taken into account to calculate the molecular weight of each digested collagen fragment.

Collagen variants were incubated with MMP-1 and run on SDS-PAGE, as shown in Figure 2-8 to determine if MMP variant collagen offer different digestion profiles from native collagen. Digestion of both hCol III and rCol show the characteristic ¾ (TC_A) and ¼

(TC_B) bands at ~140 kDa and ~30 kDa (collagen is known to run higher on SDS-PAGE than its actual molecular weight [111]). rCol-0M predictably shows no degradation from MMP-1. rCol-1M-PF2 shows a similar pattern to rCol, with a ~ ¾ length band showing digestion. This is assumed to be the 78.1 kDa digestion band of rCol-1M-PF2 running larger than anticipated on SDS-PAGE. There should be a ~ ¼ band (14.6 kDa expected size) in the gel as well that may not have been detected using this staining method. rCol-1M-PF6 shows two digested bands roughly half the length of the full collagen molecule, consistent with the MMP-site placement near the middle of the molecule. It is not unreasonable to say that the two digested rCol-1M-PF6 bands are the expected 42.0 and 50.8 kDa fragments running larger on SDS-PAGE. rCol-2M shows three bands: a ~ ¾ length band similar to rCol and rCol-1M-PF2 (indicative of incomplete collagen digestion), a ~75 kDa band representing the molecule between cleavage sites on PF2 and PF10 (expected 55.6 kDa), and a final band at 30 kDa, akin to the TC_B band seen in hCol III and rCol. Finally, rCol-3M shows faint bands at ~80 kDa, ~60 kDa, and ~45 kDa. There are smaller predicted digested bands that do not appear in this SDS-PAGE possibly due to staining sensitivity. Discussion of SDS-PAGE staining and the inability to see small digested bands is in Supplementary Information.

Because of uncertainty of the true sizes of the digested rCol-3M bands due to SDS-PAGE, it is impossible to accurately determine which size bands correspond to which MMP digestion site along the molecule. For instance, if rCol-3M was digested at only the MMP site on PF 6, an expected band of 50.8 kDa would appear. If rCol-3M were digested at the MMP sites on PF 2 and 10, with no digestion at PF 6, an expected band of 55.6 kDa would appear. Similarly, it is impossible to tell if the digested band shown in Figure 2-8 at ~60 kDa is the result of a cut only PF 6, a cut at PFs 2 and 10, or some other combination.

Because of the ambiguity and relatively low sensitivity of SDS-PAGE and of this staining method to identify the digested collagen fragment sizes, other methods of determining fragment sizes were explored. Matrix Assisted Laser Desporption / Ionization-Time of Flight (MALDI-TOF) mass spectrometry was used with rCol baseline collagen to try

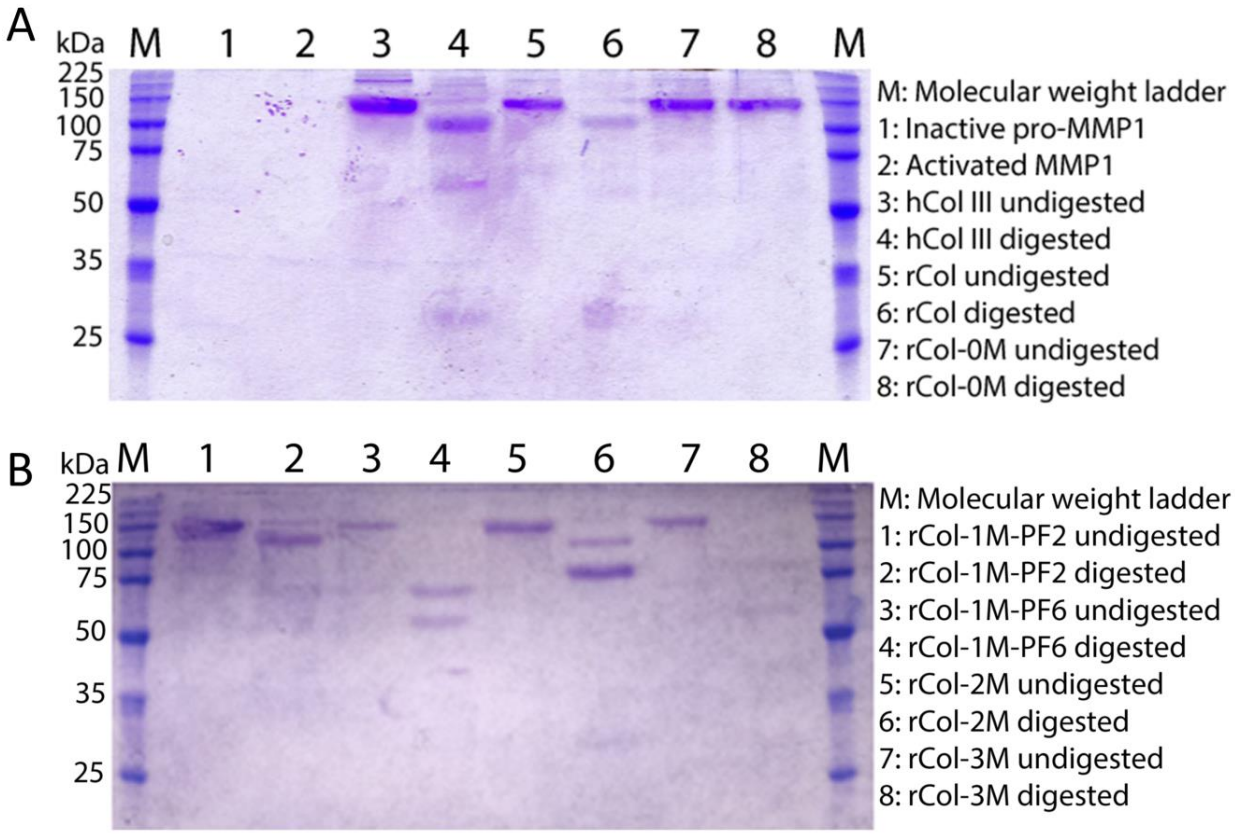


Figure 2-8*: MMP degradation assay of hCol III and each of the expressed MMP-variant collagens. A) Lanes 1-2 are control lanes with either inactive or activated MMP-1. These bands do not appear here using this Coomassie blue staining method. Lanes 3, 5, and 7 are undigested collagen, with no MMP-1 added. Lanes 4, 6, and 8 are collagens that have been digested with MMP-1 for 18 hours. B) Lanes 1, 3, 5, and 7 show undigested collagen, again with no MMP-1 added. Lanes 2, 4, 6, and 8 show digested collagen with MMP-1 added and incubated for 18 hours.*

to identify the sizes of the ¾ and ¼ digested bands, but no reliable signals could be ascertained and contrasted from the sample noise. It is possible that a small amount of contaminating yeast proteins were still in the final collagen product, which further contributed to the noise of the MALDI-TOF scan and prevented a good collagen band signal. A similar attempt was made using Size-Exclusion Chromatography (SEC), but again only noise resulted. It is believed that higher concentrations of recombinant protein per digest would be needed to identify the digested fragment sizes using SEC. Further studies are needed to truly identify and characterize the digested collagen bands more accurately and precisely than the SDS-PAGE analysis shown here.

Despite the lack of appearance of the smallest digested collagen bands, each collagen variant shows digestion consistent with its MMP-site makeup. rCol-0M is not digested at all. rCol, rCol-1M-PF2, and rCol-1M-PF6 are all only digested in one place. rCol-2M is clearly digested in 2 places, and rCol-3M shows bands consistent with incomplete digestion at 3 different sites. Thus, this novel method of creating recombinant collagen variants has resulted in digestion of the collagen molecule in ways never before reported. These collagen variants can offer insights into the mechanism needed to digest the collagen molecule based on the surrounding structure of the triple helix. Given their biomimetic yet tunable nature, the collagen mutants offer novel options for tissue engineering applications that other scaffolds cannot provide.

Time-Point MMP Digestions of Collagen Variants

The partial digestion of rCol-2M and rCol-3M lead to the hypothesis that some of the MMP sites may be digested faster than other sites. To test this hypothesis, MMP-1 digestions were performed on rCol-2M with varying digestion duration. This variant was chosen because it had the highest quality digested bands in Figure 2-8. The time-point digestions of rCol-2M are shown in Figure 2-9. Interestingly, the MMP digestion in Figure 2-9 resulted in full collagen digestion after only 3 hours. In contrast, incomplete digestion

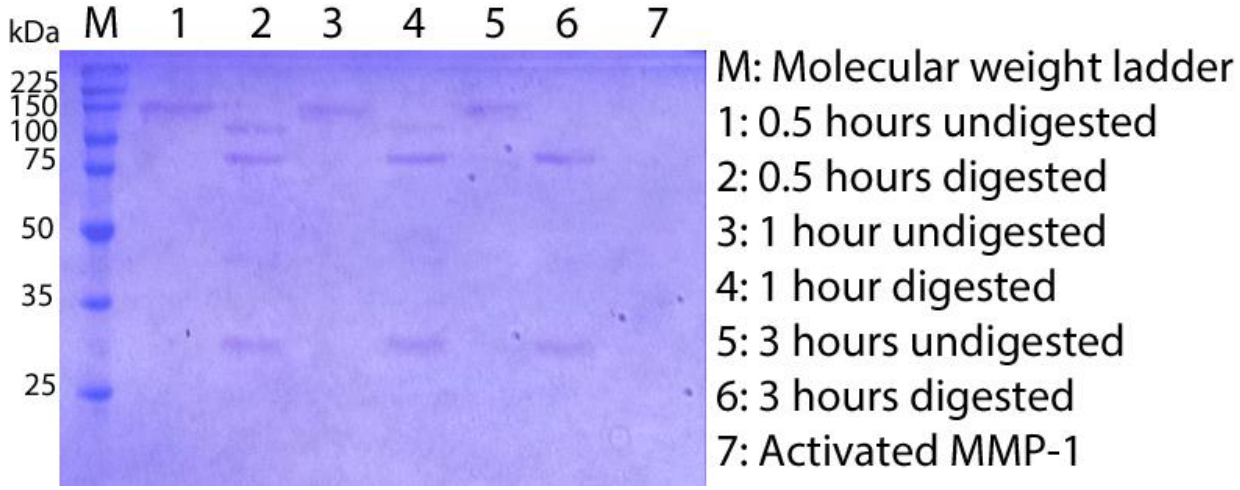


Figure 2-9*: Time-point digestion of rCol-2M. Although a large portion of the collagen seems to be digested after just 30 minutes, progressive digestion, especially of the band at ~110 kDa, can be seen as digestion time increases.*

was seen in Figure 2-8 for this variant after 18 hours of digestion. For unknown reasons it seems that the digestion kinetics are inconsistent. Progressive digestions like Figure 2-9 were completed for the other collagen variants, but either resulted in improper staining of the SDS-PAGE gel or incomplete, unsuccessful digestions. Further testing will need to be

completed to quantify the kinetics of degradation at each of the MMP sites located in the various collagen mutants.

Collagen Hydrogel Formation

MMPs have been implicated in the 3D remodeling of the ECM and other synthetic tissue engineered matrices [8, 44, 112]. Therefore, we initiated studies to create 3D hydrogels using our recombinant collagen, in order to study the ability of the MMP-modified collagen variants to be degraded in 3D. Because of the use of the non-specific protease pepsin during purification, our collagen likely cannot form higher-order collagen structures like fibrils [43], as the telopeptide regions at each terminus of the collagen molecule used to form fibrils [113] are probably removed by the pepsin. Previous studies have explored the incorporation of non-native cysteines into recombinant collagen to allow for the creation of 3D hydrogels [42]. These non-native reactive cysteine sites were not added into the MMP-variant collagens; therefore, the use of a 4-arm 20 kDa NHS-PEG crosslinker was investigated. The collagen sequence contains 38 lysines within its triple helical region (the portion of our collagen remaining after pepsin digestion), providing an ample number of primary amines to react with the NHS on the PEG crosslinker.

Hydrogel formation was tested using glass slides commonly used in particle-tracking microrheology, and 1 μm-diameter fluorescent tracer particles were used as a diagnostic to determine if hydrogels were formed or not. When viewed under a confocal microscope, if the tracer particles were stationary, then the particles were immobilized within a hydrogel. If the tracer particles were moving (via Brownian motion), then a hydrogel had not formed and remained liquid, like its initial reagents. Figure 2-10 shows

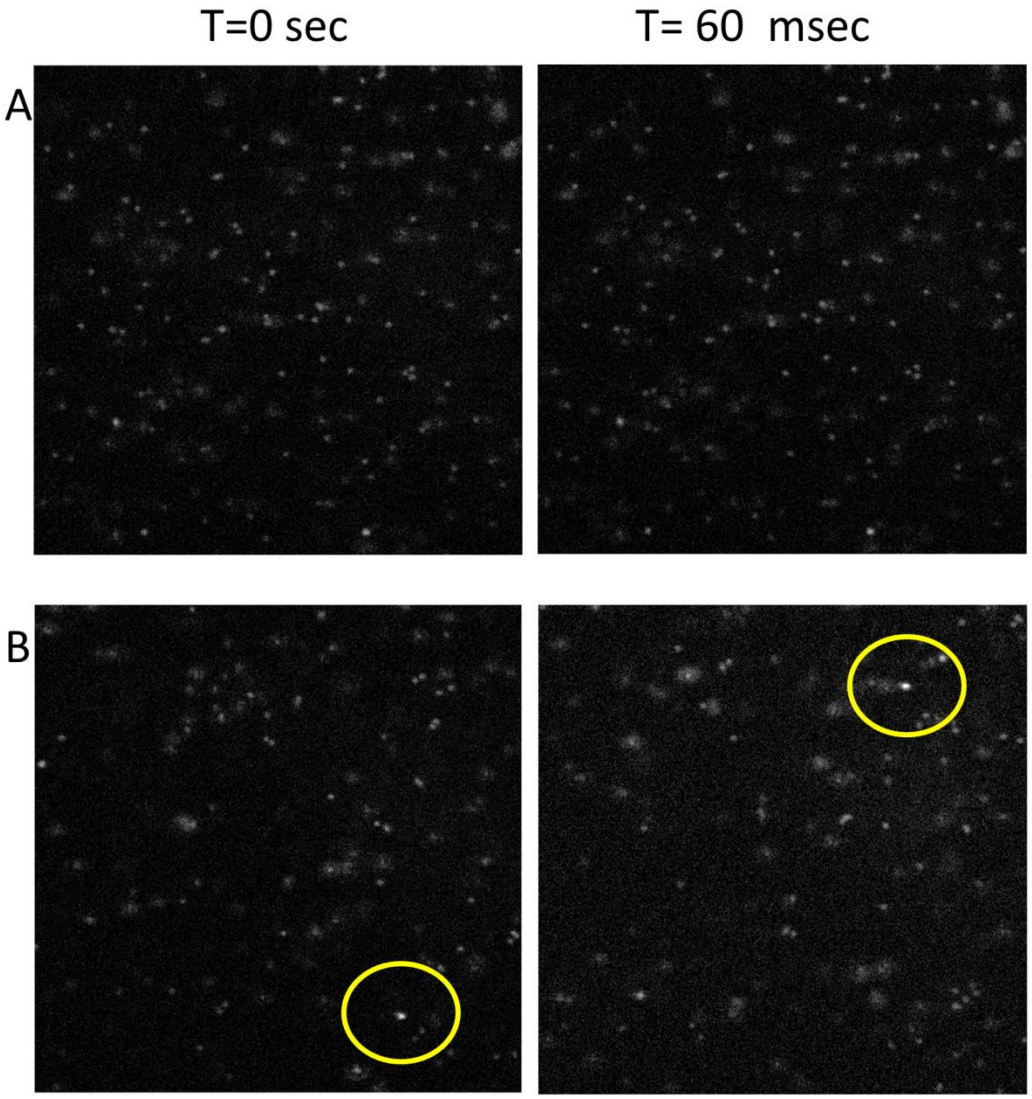


***Figure 2-10**: Representative images of tracer particles used to determine the formation of collagen hydrogels. Images on the left show the field of view at t = 0 seconds, images on the right show the same field of view 60 msec after. A) Example of tracer particles in a formed hydrogel. No difference can be seen between the left and right images, as the particles have not moved due to immobilization in a collagen hydrogel. B) Tracer particles in an unformed, incomplete hydrogel. Notice how the images are different, as particles have moved between frames. The yellow circle tracks one particular particle and clearly shows its movement.*

examples of how hydrogel formation was determined. Different ratios of 4-arm NHS-PEG crosslinker were tested in glass microrheology slides. Molar ratios of 1:1, 1:50, 1:100, 1:500, and 1:1000 NHS:lysine were tested in 2 mg/mL rCol collagen for hydrogel formation. 1:50 and 1:500 molar ratios of NHS:lysine were shown to consistently form hydrogels in the glass microrheology slide system. No other molar ratio tested resulted in hydrogel formation. Judging only the movement or immobilization of the tracer particles does not directly give the rheological properties of the gel in the same way passive particle rheology can (see below), but these experiments still offer insight into the stiffness of the matrix. This method was adopted because passive rheology tests could not be performed on the hydrogels because of microparticle aggregation (discussed later in this chapter). Que *et al.* [42] showed hydrogel formation using recombinant collagen with a maximum stiffness of 0.3 Pa using passive particle microrheology. Because the particles in that system were immobilized within the matrix (as also seen in the hydrogel formation here), it can be inferred that any hydrogel reactions which allowed particle movement must have a stiffness lower than 0.3 Pa.

Because of the enclosed nature of the glass slides, the hydrogels formed in this system could not be used for further experiments. This led to the development of a PDMS microfluidic device that could be used to form hydrogels. This device also allows the addition of MMP-1 to the hydrogel to determine the ability of the MMP-1 to degrade and digest a 3D collagen gel. The hypothesis for this experiment was that the passive rheological properties of a collagen hydrogel could be used to determine the amount of degradation of the gel due to the addition of MMP-1. It was expected that an increased number of MMP sites, like those in rCol-2M and rCol-3M, would result in more digestion of

the hydrogel than using collagen with only 1 or fewer sites. The microfluidic device shown in Figures 2-2 and 2-3 was designed for this project to create hydrogels within its microfluidic channels (dimensions are shown in Materials and Methods). Once formed, the inlets and outlets could be used to flow an enzyme solution through the channel which could digest the matrix.

Particle-tracking rheology was initially explored to test this hypothesis. The loss and storage moduli of a viscous fluid or hydrogel can be determined using the Brownian motion of small fluorescent beads [114], and thus passive microrheology was an attractive method to determine the amount of digestion of the collagen hydrogels given that fluorescent tracer particles were already being used to determine if hydrogel formation had occurred. Monodispersed fluorescent tracer particles are needed to perform particle tracking rheology [115, 116]. However, despite numerous attempts to monodisperse the particles throughout the collagen hydrogels, the particles always aggregated, preventing the measurement of the hydrogel's rheological properties. 1X PBS, 50mM sodium bicarbonate [117], and 5mM HEPES [42] buffers were all tested, but they failed to mitigate microparticle aggregation. Fluorescent microparticles were also coated in the amphiphilic polymer polyvinylpyrrolidone (PVP), which is commonly used to stabilize and solubilize colloids in solution, but particle aggregation still occurred. It is hypothesized that small yeast proteins or other contaminates may be forming salt bridges with the particles and causing them to aggregate, but further testing is needed to determine if that is indeed the cause. When particles were added to vendor bought human collagen III, particles remained monodispersed, indicating that there could be contaminating proteins in our recombinant collagen that cause particle aggregation. Particles were monodispersed in control reactions

with all hydrogel reagents (particles, buffer, crosslinker) except recombinant collagen and did not show aggregation.

Hydrogel formation as determined by the movement or immobilization of fluorescent tracer particles was also quite inconsistent in the microfluidic device designed for this project. For all buffers and reaction conditions tested, a roughly 50% success rate was found for hydrogel formation using 2 mg/mL rCol collagen and the 4-arm 20 kDa NHS-PEG crosslinker within the microfluidic device channels. This suggests there is stochasticity to hydrogel formation using this crosslinker in this system, although the reason is unclear. Future studies could explore the use of other crosslinkers to create collagen hydrogels, or perhaps incorporate non-native reactive sites like the cysteines introduced by Que *et al.* [42] to offer flexibility in reaction conditions.

Conclusions

Collagen variants with varying digestibilities were created by incorporating non-native MMP sites into the collagen sequence, or by deleting the native MMP site. With the exception of rCol-1M-PF7, all new collagens created were triple helical and had similar melting temperatures to human collagen III. Through degradation with MMP-1, it was shown that an increase in MMP sites resulted in more complete collagen digestion. Collagen variants were thus created which offer more complete protein digestion than native human collagen. Other groups have incorporated MMP-cleavable sequences into synthetic scaffolds [44, 63, 64, 82], but never before has the collagen sequence itself been modified with MMP sequence modulations. Although more work is ultimately needed (e.g.,

showing degradation in 3D hydrogels), this work opens to the door for enhanced proteolytic degradation of collagen matrices using this recombinant, MMP-variant collagen.

Future Directions

Instead of using passive particle tracking rheology to evaluate the digestion of 3D collagen hydrogels, a fluorescence-based assay could be developed using the microfluidic device designed for this project. The microfluidic device allows for flow through the channel containing a collagen hydrogel through positive pressure due to differences in inlet and outlet volumes. This is the reason for the design of the device, and it can be exploited in a number of different ways. A collagen hydrogel could be formed, and then fluorescently labeled MMP enzymes added to a channel's inlet and allowed to flow through the hydrogel within the channel. These fluorescently labeled enzymes could be tracked as they move through and digest the matrix, potentially shining light on the ability of the 3D matrix to be degraded. The kinetics of matrix degradation could be extrapolated by how quickly the enzyme moves through the matrix from one end of the channel to the other, or at least compared between recombinant collagen variants and native collagen [118]. Similarly, a collagen hydrogel could be formed, then MMP allowed to digest the matrix for a given time. A fluorescent dye in solution, such as DAPI (for example) could be flowed through the channel, saturating the hydrogel with dye. Fluorescence Recovery After Photobleaching (FRAP) could then be employed to determine the fluidity of the digested collagen matrix [119]. The hypothesis for this experiment is that if matrices formed with collagen containing more MMP sites are degraded more than matrices formed with collagen with fewer sites, then gels formed from collagen with more MMP sites will show faster recovery

after photobleaching. A faster recovery time means the dye in the matrix has a higher diffusion coefficient [120], indicating increased matrix digestion. This experiment would also rely heavily on the use of the control rCol-0M, which as shown previously cannot be degraded by MMP-1. rCol-0M would be used a negative control against which to compare diffusion coefficients of the other MMP-modified collagen variants matrices.

Chapter 3: Stabilization of Recombinant Collagen on Solid Surfaces for Cell Adhesion

Introduction

The creation of recombinant human collagen is for tissue engineering and regenerative medicine purposes. Therefore, it follows that materials which are meant to mimic the extracellular matrix of the human body should be stable at physiological temperatures. In human collagen, stability of the triple helix is conferred multiple different ways. The location of the small amino acid glycine in the middle of the triple helix, hydroxylation of lysine residues [23], and glycosylation [121] all contribute to the stability of the collagen triple helix. However, the primary cause of collagen stability involves the post-translational hydroxylation of "Y" position prolines in the Gly-X-Y repeating sequence by the enzyme prolyl-4-hydroxylase [23, 32, 40].

Although human cells have the ability to hydroxylate prolines for collagen stabilization, yeast do not, and thus must be genetically modified with hydroxylase genes [95]. Our recombinant yeast system for expressing human collagen relies on two copies each of the α2 and β1 subunits of prolyl-4-hydroxylase being stably integrated into the yeast genome for the synthesis of the hydroxylase enzyme [36]. With increasing hydroxylation levels generally comes increasing thermal stability [122], and so two copies of each hydroxylase subunit were integrated into yeast: to increase hydroxylation, and therefore stability of the collagen. [41]. Proline hydroxylation levels of our rCol recombinant collagen are 31.1 ± 2.4%, with a melting temperature to 35.8 ± 0.5 °C [41]. In contrast, native human collagen III has hydroxylation levels of 45.0 ± 1.4% (unpublished experimental data), with an apparent melting temperature of 39.1 ± 0.1 °C. Although the

hydroxylation level and melting temperature of our recombinant collagen are lower than those of human collagen III, they are still comparable enough to warrant 2D cell-based studies.

Two-dimensional cell adhesion studies have been carried out on recombinant collagen expressed in our yeast system [41-43], but the studies needed to be performed at 30 °C instead of 37 °C because of the thermal instability of the collagen. Cell adhesion to collagen III occurs through integrin-binding sites [123], and these integrin binding sites require a stable collagen triple helix for cells to attach via integrin-mediated binding [124]. At 37 °C, it was noticed that cells incubated on the recombinant collagen would detach easily during wash steps, while cells incubated on native human collagen III remained intact and adhered to the surface. This suggests that the collagen was either denaturing or desorbing off of the surface of the polystyrene plate, prompting a move from 37 °C down to 30 °C. However, the cell morphology appeared more rounded and less spread out than seen on positive controls at 37 °C [41]. It is well established that physiological temperature is 37 °C, and in order to have a material useful for tissue engineering purposes, the material must be stable at 37 °C. One method of stabilizing collagen is with the use of 1-Ethyl-3-[3 dimethylaminopropyl] carbodiimide hydrochloride (EDC), a zero-length crosslinker, which is aided by *N*-hydroxysuccinimide (NHS) [125]. EDC / NHS chemistry links carboxylates to primary amines, such as those found on lysines within the collagen molecule [126]. This crosslinking reaction has been shown previously to stabilize collagen [127, 128], and because of its biocompatibility and effectiveness, it is an attractive option for chemically crosslinking collagen to a surface to stabilize it [129, 130].

Therefore, EDC / NHS chemistry was applied to baseline recombinant human collagen III, as well as the MMP-variant collagens rCol-0M and rCol-3M. MMP-cleavable sites exist in a locally-unwound portion of the collagen triple helix [131], and so despite the similar apparent melting temperatures (see Chapter 2) it was hypothesized that an increased number of MMP sites would destabilize the rCol-3M mutant. This makes rCol-3M an attractive target for stabilization tests. Our work focuses on the covalent conjugation of recombinant human collagen onto carboxylated polystyrene plates using EDC / NHS chemistry, and the subsequent adhesion of cells onto this stabilized collagen.

Materials and Methods

Recombinant Collagen Design, Expression and Purification

See Chapter 2: Materials and Methods for the design, expression, and purification of the recombinant human collagen variants. rCol, rCol-0M and rCol-3M were all used for these experiments. rCol was tested because it could be easily compared to previous cell adhesion work. rCol-0M was used as a thermostable control, given its high apparent melting temperature of 38.8 °C (see Chapter 2). rCol-3M was used as a less thermostable control, with its apparent melting temperature of 35.1 °C.

HT-1080 Cell Culture

HT-1080 human fibrosarcoma cells (ATCC, Manassas, VA) were cultured in Dulbecco's Modified Eagle Medium (DMEM) high-glucose (Sigma-Aldrich, St. Louis, MO) + 10% Fetal Bovine Serum (FBS) (Life Technologies, Carlsbad, CA) on tissue culture plates (BD Biosciences, Franklin Lakes,NJ), and supplemented with 100 U/mL penicillin and 100

μg/mL streptomycin (Life Technologies) at 37 ^{o}C and 5% CO_2. Cells were passaged using trypsin/ EDTA (Life Technologies).

Calcein AM Cell Adhesion Assay

Human collagen III (hCol III) (BD), bovine serum albumin (BSA) (Thermo Fisher Scientific, Waltham, MA), recombinant "baseline" collagen III (rCol), native MMP site-knockout collagen (rCol-0M), and 3x MMP-site collagen (rCol-3M) collagen were all used in this assay. Two different types of experimental plates were used: untreated polystyrene 24-well plates and carboxylated 24-well plates.

For the untreated 24-well plates (BD): 237.5 μL of 20 μg/mL protein in phosphate-buffered saline (PBS) was incubated for 24 hours at 4 ^{o}C. For the carboxylated 24-well plates (BD): 1-Ethyl-3-[3 dimethylaminopropyl] carbodiimide hydrochloride (EDC) and *N*-hydroxysuccinimide (NHS) (Thermo) were dissolved in 2-(*N*-morpholino)ethanesulfonic acid (MES) buffer (EMD Millipore, Billerica, MA) pH 5.4 and added at a 100:100:1 ratio of EDC:NHS:Lysine on protein to be incubated. EDC is used to crosslink carboxyl groups to primary amines, like those found in lysines. This EDC-COOH reaction was incubated at room temperature for 20 minutes, then washed 3x with PBS. 237.5 μL of protein at 20 μg/mL was added to each EDC/NHS reacted well and incubated at room temperature for 2 hours.

For both the untreated and carboxylated plates: the protein solution was decanted and wells washed 3x with PBS. Each well was then blocked with 237.5 μL of 1% BSA at room temperature for 1 hour and washed 3X with PBS. The crosslinking reaction was terminated by adding 200 μL of 50mM Tris to each well for 1 hour at room temperature.

Plates were then washed 3x with PBS and $5.93x10^4$ HT-1080 cells were seeded into each well in DMEM+0.1% BSA, and allowed to adhere for 4 hours at 37 ^oC, 5% CO_2. Cells were imaged with an Olympus IX-51 microscope and washed gently with phenol-free DMEM (Life Technologies), then imaged again in the phenol-free DMEM. The media was gently decanted and 237.5 µL of 4µM calcein AM dye (BD) in PBS+0.1% BSA was added to the cells, which were then incubated for 1 hour at 37 ^oC, 5% CO_2. The dye solution was gently decanted and cells were lysed with 200 µL CelLytic M (Sigma) for 30 minutes at room temperature in the dark. Fluorescence was read at 494/517 nm (Abs/Em) with 515 nm automatic cutoff using a SpectraMax M2 plate reader (Molecular Devices, Sunnyvale, CA). A total of 3 independent experiments were performed, with wells in triplicate. Statistics were performed using 2-way ANOVA and Tukey test to determine statistical significance.

Results and Discussion

Quantitative Cell Adhesion Analysis

The recombinant collagen expressed in our yeast system was previously shown to have lower thermal stability relative to native human collagen. As a result, cell-based adhesion assays had to be performed at 30 °C instead of 37 °C [41-43]. 37 °C is above the apparent melting temperatures for our recombinant collagens, which leads to denaturing of the collagen triple helix and decreases the ability of cells to adhere to the collagen [41]. To increase the stability of our recombinant collagen for use with cell-based assays at physiological temperatures, EDC / NHS chemistry was used to crosslink recombinant collagen to cell culture plates featuring a carboxylated surface. Calcein AM was used to quantify the amount of cell adhesion of the HT-1080 cells incubated on recombinant

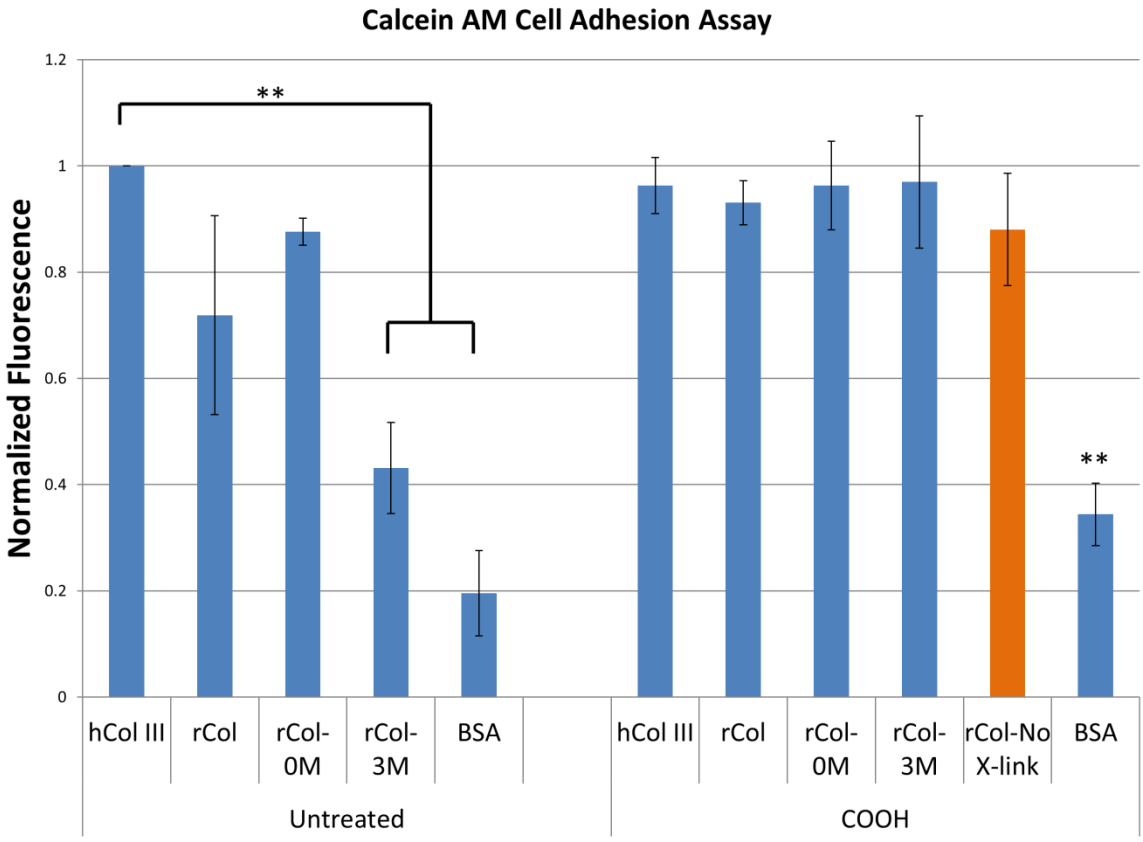


Figure 3-1*: Normalized fluorescence readings from calcein AM cell adhesion assay. Fluorescence readings correspond to the number of HT-1080 cells adhered to the plate after incubation with each different protein at 37 °C and washing to remove poorly adhered cells. Human collagen III (hCol III) on untreated plates is used as the positive control for this assay, and each experiment's fluorescence values are normalized to this value for that particular experiment. "rCol-No xlink" refers to baseline collagen that was incubated on the carboxylated plate, but not crosslinked like the other proteins on that plate (shown in orange). BSA is a negative control. N=3 independent experiments with wells in triplicate, error bars denote standard deviation, ** denotes p<0.01. Experiments were compared using ANOVA and Tukey test.*

collagen that had been chemically crosslinked to the surface of the cell culture plate versus collagen that was simply adsorbed onto a polystyrene cell culture plate. Experiments were performed at 37 °C to determine if the collagen had been sufficiently stabilized to support cell adhesion.

The results of the calcein AM cell adhesion assay are shown in Figure 3-1. rCol was used in order to compare the results of these experiments to previously reported cell adhesion data from our lab [41-43]. rCol-0M and rCol-3M were chosen for use in this experiment because of their high and low apparent melting temperatures (38.8 and 35.1 °C, respectively). The untreated plate experimental group replicates the previous cell adhesion studies performed using this recombinant collagen [41-43], with the exception that the cell adhesion was performed at 37 °C here. Figure 3-1 shows clearly that crosslinking the recombinant collagen (rCol, rCol-0M or rCol-3M) to a carboxylated plate stabilizes the protein and better allows cells to adhere to the collagen.

Although there visually seems to be an obvious difference between the cell adhesion with cells incubated on hCol III vs. rCol and rCol-0M collagen, the p-values for these comparisons do not show statistical significance ($p=0.051$ and 0.061, respectively). There is a significant difference between the untreated hCol III and untreated rCol-3M wells. These differences are removed when the recombinant collagen is covalently crosslinked to a carboxylated plate using EDC / NHS chemistry, as there is no statistically significant difference between adhesion on human collagen III incubated on untreated plates vs. any of the crosslinked recombinant collagen wells. On the untreated plates, there is a significant difference in the cell adhesion on rCol and rCol-0M versus the negative control BSA ($p<0.05$), but there is no such difference between rCol-3M and BSA. However, when the

protein samples are covalently stabilized, a significant difference arises between any collagen variation and BSA ($p<0.01$). "rCol-No X-link" does not show any significant differences from the positive control of human collagen III, or from the proteins that were covalently crosslinked to the surface of the carboxylated plate.

Qualitative Cell Adhesion Analysis

As a more qualitative analysis of the effectivness of the EDC / NHS stabilization of our recombinant collagen, images were taken of the HT-1080 fibrosarcoma cells before and after washing to compare cell morphology, and are shown in Figure 3-2. Figure 3-2A shows the cell adhesion onto proteins that were incubated for 24 hours on non-treated polystyrene plates. Here, the cells are shown after incubation at 37 ^{o}C instead of 30 °C used previously. Cells incubated on hCol III, rCol, and rCol-0M are adhered, spread out, and appear healthy, in approximate accordance with their respective fluoresnce levels in Figure 3-1. The less thermostable rCol-3M mutant showed significantly less adhesion when the protein was simply adhered to the surface in Figure 3-1, and those results are replicated in Figure 3-2: there are noticably less cells in the image, and the cells appear more balled-up and less spread out on the surface. The BSA negative controls also show this balled-up morphology, suggesting that 37 °C is too high of a temperature for the rCol-3M mutant collagen to maintain stability.

Figure 3-2B shows the cell adhesion onto proteins that were stably crosslinked on the surface of the carboxylated plates. The fluorescence readings in Figure 3-1 show that the cells adhered better to the crosslinked collagens than to the protein that was merely incubated to the surface of the non-treated plates. The cell images in Figure 3-2 show

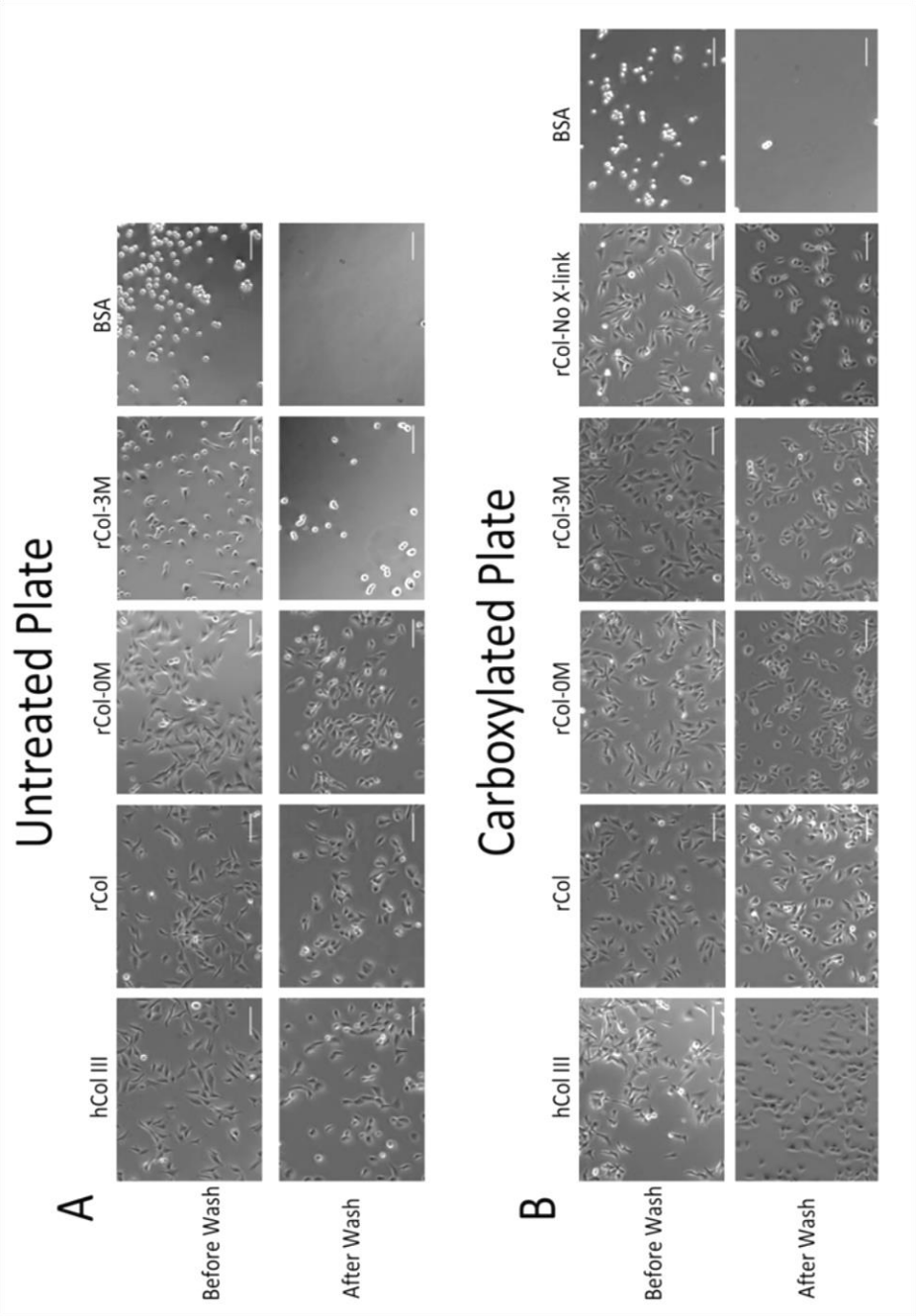


Figure 3-2*: Cell images before and after washing in the calcein AM cell adhesion assay, performed at 37 °C. A) Untreated plate cell images. B) Carboxylated plate cell images. hCol III is the positive control. BSA is the negative control. As before, "rCol-No X-link" refers to rCol collagen that was incubated on the carboxylated plate, but not chemically crosslinked to it. Scale bar = 50 µm.*

improved cell morphology, particularly in the rCol-3M image, which corresponds to the increased cell adhesion seen in Figure 3-1. Although there are no significant differences in cell adhesion for the hCol III, rCol, and rCol-0M groups on the carboxylated plate versus the untreated plate, crosslinking clearly results in an improvement in cell adhesion. This, combined with the significantly increased adhesion and obviously improved morphology in the rCol-3M wells, indicates that the EDC / NHS crosslinking has stabilized the collagen to prevent it from denaturing, allowing cells to more easily adhere and interact with it.

Despite the similar levels of cell adhesion on the covalently-stabilized collagen, minor differences can still be seen in cell morphology when comparing cells incubated on hCol III versus the recombinant collagens. Cells incubated on the recombinant collagen variants seem lightly more rounded and less spread out compared to the cells incubated on hCol III, which is consistent with prior cell adhesion experiments on collagen expressed in our system [41]. It was noted previously that these minor morphology differences suggest slight molecular changes could exist between hCol III and the recombinant collagens [41]. Improvements in the collagen purification protocol may help to decrease these small molecular changes.

Baseline collagen that was incubated, but not crosslinked, onto the carboxlated plate showed increased cell adhesion and improved morphology. The crosslinking reaction, and subsequently the incubation for the non-crosslinked baseline in this experiment, was done in PBS at pH 7.4. At this pH collagen III is positively charged (isoelectric point of 8.66, as calculated using the ExPASy tool at http://web.expasy.org/compute_pi/) so the positively charged collagen simply interacted with the negatively charged carboxyl groups on the

surface of the plate and stabilized the protein. However, although this ionic interaction did serve to prevent the collagen from becoming denatured at 37 °C for this short 4 hour time frame, it is unlikely that this interaction would be as stable as the covalently crosslinked samples in longer term experiments. More experiments, particularly of a longer duration, could be used to verify this. Future experiments should focus on exposing the collagen to 37 °C for longer than 4 hours to determine how thermal stability is affected by the ionic interactions with the carboxylated plate versus chemical crosslinking. More discussion of cell adhesion assay techniques and protocols are in Supplementary Information.

The EDC/NHS chemistry used clearly stabilized the collagen mutants enough for cells to adhere well above their respective apparent melting temperatures. However little is still known about how much the collagen was crosslinked and the effects of the crosslinking on cells' abilities to interact with the collagen. For instance, a slightly loosened collagen triple helix is needed for MMP-dependent collagen cleavage [131] and EDC/NHS crosslinking has been previously shown to stabilize collagen to the point of making it resistant to enzymatic degradation [132]. Further studies, particularly those involving 3D collagen hydrogel formation, should explore the stability of the collagen matrix and the ability of collagenases to cleave it. If EDC / NHS chemistry is used as the crosslinking method, the ratio of EDC:NHS:lysines within the matrix can be varied to tune the matrix stiffness and crosslinking levels. These studies must be careful that crosslinking levels are not so high that the collagen triple helix is stabilized to the point that MMPs are unable to digest the hydrogel because the matrix has become too dense or stabilized. MMP-1 has been shown to locally unwind the collagen triple helix before hydrolysis of the collagen

chain [79], so too much of an increase in collagen stability could have the unintended effect of preventing collagenase efficacy.

There were no significant differences seen in the quantitative or qualitative amounts of cell adhesion between hCol III and rCol, which is in direct contrast with what has been seen previously [41]. This could be due to slight differences in the collagen purification protocol (final collagen samples are now dialyzed against 50 mM acetic acid instead of 40 mM sodium acetate), improved lab purification techniques, or perhaps different colonies of transformed yeast. Recent experiments appear to show our recombinant collagen is more stable that previously reported, stemming from a newly transformed yeast colony [Que, Pham, unpublished data]. Given the results of these cell adhesion experiments, and of recent unpublished experiments in our lab, future studies may be proposed to use AFM to examine the triple helical structure of our recombinant collagen now and compare it to previous AFM studies [41] to see if our recent collagen is now more triple helical and stable at 37 °C.

Conclusions

Our recombinant collagen III has previously shown to have decreased thermostability compared to native human collagen III, which makes cell adhesion assays performed at 37 °C impossible because of collagen denaturation [41]. Cell assays using this collagen had to be performed at 30 °C to eliminate the possibility of the collagen denaturing and inhibiting cell adhesion. Therefore, EDC / NHS chemistry was used to covalently stabilize recombinant collagen onto carboxylated tissue culture plates, to be used at 37 °C. Cells showed improved adhesion and morphology at 37 °C on rCol, rCol-0M, and rCol-3M

collagens after covalent stabilization. Using this stabilization protocol our recombinant collagen can be used in cell-based assays at 37 °C, and there is no longer a limitation to only perform these assays at 30 °C.

References

[1] H. Järveläinen, A. Sainio, M. Koulu, T. N. Wight, and R. Penttinen, "Extracellular Matrix Molecules: Potential Targets in Pharmacotherapy," *Pharmacological Reviews,* vol. 61, pp. 198-223, 2009.

[2] R. A. F. Clark, "Fibrin and Wound Healing," *Annals of the New York Academy of Sciences,* vol. 936, pp. 355-367, 2001.

[3] H. K. Kleinman and G. R. Martin, "Matrigel: Basement membrane matrix with biological activity," *Seminars in Cancer Biology,* vol. 15, pp. 378-386, 2005.

[4] G. S. Schultz and A. Wysocki, "Interactions between extracellular matrix and growth factors in wound healing," *Wound Repair and Regeneration,* vol. 17, pp. 153-162, 2009.

[5] A. R. Gillies and R. L. Lieber, "Structure and function of the skeletal muscle extracellular matrix," *Muscle and Nerve,* vol. 44, pp. 318-331, 2011.

[6] C. Frantz, K. M. Stewart, and V. M. Weaver, "The extracellular matrix at a glance," *Journal of Cell Science,* vol. 123, pp. 4195-4200, 2010.

[7] F. Rosso, A. Giordano, M. Barbarisi, and A. Barbarisi, "From Cell-ECM Interactions to Tissue Engineering," *Journal of Cellular Physiology,* vol. 199, pp. 174-180, 2004.

[8] M. P. Lutolf, J. L. Lauer-Fields, H. G. Schmoekel, A. T. Metters, F. E. Weber, G. B. Fields*, et al.*, "Synthetic matrix metalloproteinase-sensitive hydrogels for the conduction of tissue regeneration: Engineering cell-invasion characteristics," *Proceedings of the National Academy of Sciences,* vol. 100, pp. 5413-5418, 2003.

[9] L. Moroni, J. R. de Wijn, and C. A. van Blitterswijk, "Integrating novel technologies to fabricate smart scaffolds," *Journal of Biomaterials Science, Polymer Edition,* vol. 19, pp. 543-572, 2008.

[10] K. N. Chua, W. S. Lim, P. Zhang, H. Lu, J. Wen, S. Ramakrishna*, et al.*, "Stable immobilization of rat hepatocyte spheroids on galactosylated nanofiber scaffold," *Biomaterials,* vol. 26, pp. 2537-2547, 2005.

[11] J. F. Mano, G. A. Silva, H. S. Azevedo, P. B. Malafaya, R. A. Sousa, S. S. Silva*, et al.*, "Natural origin biodegradable systems in tissue engineering and regenerative medicine: present status and some moving trends," *Journal of The Royal Society Interface,* vol. 4, pp. 999-1030, 2007.

[12] F. Yang, C. G. Williams, D.-a. Wang, H. Lee, P. N. Manson, and J. Elisseeff, "The effect of incorporating RGD adhesive peptide in polyethylene glycol diacrylate hydrogel on osteogenesis of bone marrow stromal cells," *Biomaterials,* vol. 26, pp. 5991-5998, 2005.

[13] B. P. Chan and K. W. Leong, "Scaffolding in tissue engineering: general approaches and tissue-specific considerations," *European Spine Journal,* vol. 17, pp. 467-479, 2008.

[14] S. F. Badylak, D. O. Freytes, and T. W. Gilbert, "Extracellular matrix as a biological scaffold material: Structure and function," *Acta Biomaterialia,* vol. 5, pp. 1-13, 2009.

[15] C. H. Lee, A. Singla, and Y. Lee, "Biomedical applications of collagen," *International Journal of Pharmaceutics,* vol. 221, pp. 1-22, 2001.

[16] V. R. King, A. Alovskaya, D. Y. T. Wei, R. A. Brown, and J. V. Priestley, "The use of injectable forms of fibrin and fibronectin to support axonal ingrowth after spinal cord injury," *Biomaterials,* vol. 31, pp. 4447-4456, 2010.
[17] J. A. Burdick and G. D. Prestwich, "Hyaluronic Acid Hydrogels for Biomedical Applications," *Advanced Materials,* vol. 23, pp. H41-H56, 2011.
[18] N. Annabi, S. M. Mithieux, E. A. Boughton, A. J. Ruys, A. S. Weiss, and F. Dehghani, "Synthesis of highly porous crosslinked elastin hydrogels and their interaction with fibroblasts in vitro," *Biomaterials,* vol. 30, pp. 4550-4557, 2009.
[19] S. V. Madihally and H. W. T. Matthew, "Porous chitosan scaffolds for tissue engineering," *Biomaterials,* vol. 20, pp. 1133-1142, 1999.
[20] J. A. Rowley, G. Madlambayan, and D. J. Mooney, "Alginate hydrogels as synthetic extracellular matrix materials," *Biomaterials,* vol. 20, pp. 45-53, 1999.
[21] Y. Chunlin, P. J. Hillas, J. A. Báez, M. Nokelainen, J. Balan, J. Tang*, et al.*, "The Application of Recombinant Human Collagen in Tissue Engineering," *BioDrugs,* vol. 18, pp. 103-119, 2004.
[22] M. G. Patino, M. E. Neiders, S. Andreana, B. Noble, and R. E. Cohen, "Collagen: An Overview," *Implant Dentistry,* vol. 11, pp. 280-285, 2002.
[23] M. D. Shoulders and R. T. Raines, "Collagen Structure and Stability," *Annual Review of Biochemistry,* vol. 78, pp. 929-958, 2009.
[24] J. Myllyharju and K. I. Kivirikko, "Collagens, modifying enzymes and their mutations in humans, flies and worms," *Trends in Genetics,* vol. 20, pp. 33-43, 2004.
[25] R. E. Shadwick, "Elastic energy storage in tendons: Mechanical differences related to function and age," *Journal of Applied Physiology,* vol. 68, pp. 1033-1040, 1990.
[26] K. E. Kadler, C. Baldock, J. Bella, and R. P. Boot-Handford, "Collagens at a glance," *Journal of Cell Science,* vol. 120, pp. 1955-1958, 2007.
[27] J. Myllyharju and K. I. Kivirikko, "Collagens and collagen-related diseases," *Annals of Medicine,* vol. 33, pp. 7-21, 2001.
[28] R. D. B. Fraser, T. P. MacRae, and E. Suzuki, "Chain conformation in the collagen molecule," *Journal of Molecular Biology,* vol. 129, pp. 463-481, 1979.
[29] K. Gelse, E. Pöschl, and T. Aigner, "Collagens—structure, function, and biosynthesis," *Advanced Drug Delivery Reviews,* vol. 55, pp. 1531-1546, 2003.
[30] A. Lamberg, T. Helaakoski, J. Myllyharju, S. Peltonen, H. Notbohm, T. Pihlajaniemi*, et al.*, "Characterization of Human Type III Collagen Expressed in a Baculovirus System: PRODUCTION OF A PROTEIN WITH A STABLE TRIPLE HELIX REQUIRES COEXPRESSION WITH THE TWO TYPES OF RECOMBINANT PROLYL 4-HYDROXYLASE SUBUNIT," *Journal of Biological Chemistry,* vol. 271, pp. 11988-11995, 1996.
[31] D. R. Olsen, S. D. Leigh, R. Chang, H. McMullin, W. Ong, E. Tai*, et al.*, "Production of Human Type I Collagen in Yeast Reveals Unexpected New Insights into the Molecular Assembly of Collagen Trimers," *Journal of Biological Chemistry,* vol. 276, pp. 24038-24043, 2001.
[32] A. Vuorela, J. Myllyharju, R. Nissi, T. Pihlajaniemi, and K. I. Kivirikko, *Assembly of human prolyl 4-hydroxylase and type III collagen in the yeast Pichia pastoris: formation of a stable enzyme tetramer requires coexpression with collagen and assembly of a stable collagen requires coexpression with prolyl 4-hydroxylase* vol. 16, 1997.

[33] K. Vuori, T. Pihlajaniemi, M. Marttila, and K. I. Kivirikko, "Characterization of the human prolyl 4-hydroxylase tetramer and its multifunctional protein disulfide-isomerase subunit synthesized in a baculovirus expression system," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 89, pp. 7467-7470, 1992.

[34] D. Olsen, C. Yang, M. Bodo, R. Chang, S. Leigh, J. Baez*, et al.*, "Recombinant collagen and gelatin for drug delivery," *Advanced Drug Delivery Reviews,* vol. 55, pp. 1547-1567, 2003.

[35] M. Tandon, M. Wu, T. P. Begley, J. Myllyharju, A. Pirskanen, and K. Kivirikko, "Substrate specificity of human prolyl-4-hydroxylase," *Bioorganic & Medicinal Chemistry Letters,* vol. 8, pp. 1139-1144, 1998.

[36] S. W. P. Chan, S.-P. Hung, S. K. Raman, G. W. Hatfield, R. H. Lathrop, N. A. Da Silva*, et al.*, "Recombinant Human Collagen and Biomimetic Variants Using a De Novo Gene Optimized for Modular Assembly," *Biomacromolecules,* vol. 11, pp. 1460-1469, 2010.

[37] H. Notbohm, M. Nokelainen, J. Myllyharju, P. P. Fietzek, P. K. Müller, and K. I. Kivirikko, "Recombinant Human Type II Collagens with Low and High Levels of Hydroxylysine and Its Glycosylated Forms Show Marked Differences in Fibrillogenesis in Vitro," *Journal of Biological Chemistry,* vol. 274, pp. 8988-8992, 1999.

[38] L. Sipilä, H. Ruotsalainen, R. Sormunen, N. L. Baker, S. R. Lamandé, M. Vapola*, et al.*, "Secretion and Assembly of Type IV and VI Collagens Depend on Glycosylation of Hydroxylysines," *Journal of Biological Chemistry,* vol. 282, pp. 33381-33388, 2007.

[39] L. C. Abraham, E. Zuena, B. Perez-Ramirez, and D. L. Kaplan, "Guide to collagen characterization for biomaterial studies," *Journal of Biomedical Materials Research Part B: Applied Biomaterials,* vol. 87B, pp. 264-285, 2008.

[40] B. Brodsky and J. A. M. Ramshaw, "The collagen triple-helix structure," *Matrix Biology,* vol. 15, pp. 545-554, 1997.

[41] A. Jabaiah, X. Wang, S. K. Raman, R. Ragan, N. A. Da Silva, and S.-W. Wang, "Nanoscale architecture and cellular adhesion of biomimetic collagen substrates," *Journal of Biomaterials Applications,* 2013.

[42] R. Que, A. Mohraz, N. A. Da Silva, and S.-W. Wang, "Expanding Functionality of Recombinant Human Collagen Through Engineered Non-Native Cysteines," *Biomacromolecules,* 2014.

[43] R. A. Que, S. W. P. Chan, A. M. Jabaiah, R. H. Lathrop, N. A. Da Silva, and S.-W. Wang, "Tuning cellular response by modular design of bioactive domains in collagen," *Biomaterials,* vol. 53, pp. 309-317, 2015.

[44] M. Bracher, D. Bezuidenhout, M. P. Lutolf, T. Franz, M. Sun, P. Zilla*, et al.*, "Cell specific ingrowth hydrogels," *Biomaterials,* vol. 34, pp. 6797-6803, 2013.

[45] H. Nagase and J. F. Woessner, "Matrix Metalloproteinases," *Journal of Biological Chemistry,* vol. 274, pp. 21491-21494, 1999.

[46] J. D. Mott and Z. Werb, "Regulation of matrix biology by matrix metalloproteinases," *Current Opinion in Cell Biology,* vol. 16, pp. 558-564, 2004.

[47] C. M. Alexander, E. J. Hansell, O. Behrendtsen, M. L. Flannery, N. S. Kishnani, S. P. Hawkes*, et al.*, "Expression and function of matrix metalloproteinases and their inhibitors at the maternal-embryonic boundary during mouse embryo implantation," *Development,* vol. 122, pp. 1723-1736, 1996.

[48] T. H. Vu and Z. Werb, "Matrix metalloproteinases: effectors of development and normal physiology," *Genes & Development,* vol. 14, pp. 2123-2133, 2000.
[49] C. Colnot, Z. Thompson, T. Miclau, Z. Werb, and J. A. Helms, "Altered fracture repair in the abscence of MMP9," *Development,* vol. 130, pp. 4123-4133, 2003.
[50] G. Bergers, R. Brekken, G. McMahon, T. H. Vu, T. Itoh, K. Tamaki*, et al.*, "Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis," *Nature Cell Biology,* vol. 2, pp. 737-744, 2000.
[51] S. Lee, S. M. Jilan, G. V. Nikolova, D. Carpizo, and M. Luisa Iruela-Arispe, "Processing of VEGF-A by matrix metalloproteinases regulates bioavailability and vascular patterning in tumors," *Journal of Cell Biology,* vol. 169, pp. 681-691, 2005.
[52] P. E. Van Den Steen, P. Proost, A. Wuyts, J. Van Damme, and G. Opdenakker, "Neutrophil gelatinase B potentiates interleukin-8 tenfold by aminoterminal processing, whereas it degrades CTAP-III, PF-4, and GRO-α and leaves RANTES and MCP-2 intact," *Blood,* vol. 96, pp. 2673-2681, 2000.
[53] G. Hashimoto, I. Inoki, Y. Fujii, T. Aoki, E. Ikeda, and Y. Okada, "Matrix metalloproteinases cleave connective tissue growth factor and reactivate angiogenic activity of vascular endothelial growth factor 165," *Journal of Biological Chemistry,* vol. 277, pp. 36288-36295, 2002.
[54] T. Twardowski, A. Fertala, J. P. R. O. Orgel, and J. D. San Antonio, "Type I Collagen and Collagen Mimetics as Angiogenesis Promoting Superpolymers," *Current Pharmaceutical Design,* vol. 13, pp. 3608-3621, 2007.
[55] W. C. Parks, C. L. Wilson, and Y. S. López-Boado, "Matrix metalloproteinases as modulators of inflammation and innate immunity," *Nature Reviews Immunology,* vol. 4, pp. 617-629, 2004.
[56] S. Werner and R. Grose, "Regulation of Wound Healing by Growth Factors and Cytokines," *Physiological Reviews,* vol. 83, pp. 835-870, 2003.
[57] A. Boire, L. Covic, A. Agarwal, S. Jacques, S. Sherifi, and A. Kuliopulos, "PAR1 is a matrix metalloprotease-1 receptor that promotes invasion and tumorigenesis of breast cancer cells," *Cell,* vol. 120, pp. 303-313, 2005.
[58] L. Blavier and Y. A. DeClerck, "Considering the critical interface between tumor cells and stromal cells in the search for targets for anticancer therapy," *Cancer Cell,* vol. 7, pp. 408-409, 2005.
[59] B. C. Sheu, S. M. Hsu, H. N. Ho, H. C. Lien, S. C. Huang, and R. H. Lin, "A novel role of metalloproteinase in cancer-mediated immunosuppression," *Cancer Research,* vol. 61, pp. 237-242, 2001.
[60] H. Kataoka, H. Uchino, T. Iwamura, M. Seiki, K. Nabeshima, and M. Koono, "Enhanced tumor growth and invasiveness in vivo by a carboxyl-terminal fragment of α1-proteinase inhibitor generated by matrix metalloproteinases: A possible modulatory role in natural killer cytotoxicity," *American Journal of Pathology,* vol. 154, pp. 457-468, 1999.
[61] R. A. Dean, J. H. Cox, C. L. Bellac, A. Doucet, A. E. Starr, and C. M. Overall, "Macrophage-specific metalloelastase (MMP-12) truncates and inactivates ELR + CXC chemokines and generates CCL2, -7, -8, and -13 antagonists: Potential role of the macrophage in terminating polymorphonuclear leukocyte influx," *Blood,* vol. 112, pp. 3455-3464, 2008.

[62] Z. Yu, R. Visse, M. Inouye, H. Nagase, and B. Brodsky, "Defining Requirements for Collagenase Cleavage in Collagen Type III Using a Bacterial Collagen System," *Journal of Biological Chemistry,* vol. 287, pp. 22988-22997, 2012.
[63] D. S. Ferreira, Y.-A. Lin, H. Cui, J. A. Hubbell, R. L. Reis, and H. S. Azevedo, "Molecularly Engineered Self-Assembling Membranes for Cell-Mediated Degradation," *Advanced Healthcare Materials,* vol. 4, pp. 602-612, 2015.
[64] S. Sokic, M. C. Christenson, J. C. Larson, A. A. Appel, E. M. Brey, and G. Papavasiliou, "Evaluation of MMP substrate concentration and specificity for neovascularization of hydrogel scaffolds," *Biomaterials Science,* vol. 2, pp. 1343-1354, 2014.
[65] M. P. Lutolf, F. E. Weber, H. G. Schmoekel, J. C. Schense, T. Kohler, R. Muller*, et al.*, "Repair of bone defects using synthetic mimetics of collagenous extracellular matrices," *Nat Biotech,* vol. 21, pp. 513-518, 2003.
[66] M. C. Giano, D. J. Pochan, and J. P. Schneider, "Controlled biodegradation of Self-assembling β-hairpin Peptide hydrogels by proteolysis with matrix metalloproteinase-13," *Biomaterials,* vol. 32, pp. 6471-6477, 2011.
[67] K. M. Galler, L. Aulisa, K. R. Regan, R. N. D'Souza, and J. D. Hartgerink, "Self-Assembling Multidomain Peptide Hydrogels: Designed Susceptibility to Enzymatic Cleavage Allows Enhanced Cell Migration and Spreading," *Journal of the American Chemical Society,* vol. 132, pp. 3217-3223, 2010.
[68] M. T. Salmela, S. L. F. Pender, M. L. Karjalainen-Lindsberg, P. Puolakkainen, T. T. MacDonald, and U. Saarialho-Kere, "Collagenase-1 (MMP-1), matrilysin-1 (MMP-7), and stromelysin-2 (MMP-10) are expressed by migrating enterocytes during intestinal wound healing," *Scandinavian Journal of Gastroenterology,* vol. 39, pp. 1095-1104, 2004.
[69] G. B. Fields, "A MODEL FOR INTERSTITIAL COLLAGEN CATABOLISM BY MAMMALIAN COLLAGENASES," *Journal of Theoretical Biology,* vol. 153, pp. 585-602, 1991.
[70] M. Moreau, I. Brocheriou, L. Petit, E. Ninio, M. J. Chapman, and M. Rouis, "Interleukin-8 Mediates Downregulation of Tissue Inhibitor of Metalloproteinase-1 Expression in Cholesterol-Loaded Human Macrophages: Relevance to Stability of Atherosclerotic Plaque," *Circulation,* vol. 99, pp. 420-426, 1999.
[71] M. Egeblad and Z. Werb, "New functions for the matrix metalloproteinases in cancer progression," *Nature Reviews Cancer,* vol. 2, pp. 161-174, 2002.
[72] T. E. Curry and K. G. Osteen, "The Matrix Metalloproteinase System: Changes, Regulation, and Impact throughout the Ovarian and Uterine Reproductive Cycle," *Endocrine Reviews,* vol. 24, pp. 428-465, 2003.
[73] S. Iyer, R. Visse, H. Nagase, and K. R. Acharya, "Crystal Structure of an Active Form of Human MMP-1," *Journal of Molecular Biology,* vol. 362, pp. 78-88, 2006.
[74] H. Nagase and R. Visse, "Triple Helicase Activity and the Structural Basis of Collagenolysis," in *Extracellular Matrix Degradation*. vol. 2, W. C. Parks and R. P. Mecham, Eds., ed: Springer Berlin Heidelberg, pp. 95-122, 2011.
[75] A. Page-McCaw, A. J. Ewald, and Z. Werb, "Matrix metalloproteinases and the regulation of tissue remodelling," *Nature reviews. Molecular cell biology,* vol. 8, pp. 221-233, 2007.

[76] W. Bode, F.-X. Gomis-Rüth, and W. Stöckler, "Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'," *FEBS Letters,* vol. 331, pp. 134-140, 1993.
[77] H. G. Welgus, J. J. Jeffrey, and A. Z. Eisen, "The collagen substrate specificity of human skin fibroblast collagenase," *Journal of Biological Chemistry,* vol. 256, pp. 9511-9515, 1981.
[78] C. Overall, "Molecular determinants of metalloproteinase substrate specificity," *Molecular Biotechnology,* vol. 22, pp. 51-86, 2002.
[79] L. Chung, D. Dinakarpandian, N. Yoshida, J. L. Lauer-Fields, G. B. Fields, R. Visse*, et al.*, "Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis," *The EMBO Journal,* vol. 23, pp. 3020-3030, 2004.
[80] S. W. Manka, F. Carafoli, R. Visse, D. Bihan, N. Raynal, R. W. Farndale*, et al.*, "Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 109, pp. 12461-12466, 2012.
[81] K. E. Williams and D. R. Olsen, "Matrix metalloproteinase-1 cleavage site recognition and binding in full-length human type III collagen," *Matrix Biology,* vol. 28, pp. 373-379, 2009.
[82] J. Patterson and J. A. Hubbell, "Enhanced proteolytic degradation of molecularly engineered PEG hydrogels in response to MMP-1 and MMP-2," *Biomaterials,* vol. 31, pp. 7836-7845, 2010.
[83] S. I. Miekka, R. Y. Forng, R. G. Rohwer, C. MacAuley, R. E. Stafford, S. L. Flack*, et al.*, "Inactivation of viral and prion pathogens by γ-irradiation under conditions that maintain the integrity of human albumin," *Vox Sanguinis,* vol. 84, pp. 36-44, 2003.
[84] J. Myllyharju, M. Nokelainen, A. Vuorela, and K. I. Kivirikko, "Expression of recombinant human type I-III collagens in the yeast Pichia pastoris," *Biochemical Society Transactions,* vol. 28, pp. 353-357, 2000.
[85] M. Nokelainen, H. Tu, A. Vuorela, H. Notbohm, K. I. Kivirikko, and J. Myllyharju, "High-level production of human type I collagen in the yeast Pichia pastoris," *Yeast,* vol. 18, pp. 797-806, 2001.
[86] F. Ruggiero, J. Y. Exposito, P. Bournat, V. Gruber, S. Perret, J. Comte*, et al.*, "Triple helix assembly and processing of human collagen produced in transgenic tobacco plants," *FEBS Letters,* vol. 469, pp. 132-136, 2000.
[87] J. Veijola, T. Pihlajaniemi, and K. I. Kivirikko, "Co-expression of the alpha subunit of human prolyl 4-hydroxylase with BiP polypeptide in insect cells leads to the formation of soluble and insoluble complexes. Soluble alpha-subunit-BiP complexes have no prolyl 4-hydroxylase activity," *Biochemical Journal,* vol. 315, pp. 613-618, 1996.
[88] C. Merle, S. Perret, T. Lacour, V. Jonval, S. Hudaverdian, R. Garrone*, et al.*, "Hydroxylated human homotrimeric collagen I in Agrobacterium tumefaciens-mediated transient expression and in transgenic tobacco plant," *FEBS Letters,* vol. 515, pp. 114-118, 2002.
[89] A. M. Whatmore, "Streptococcus pyogenes sclB encodes a putative hypervariable surface protein with a collagen-like repetitive structure," *Microbiology,* vol. 147, pp. 419-429, 2001.

[90] Y. Xu, D. R. Keene, J. M. Bujnicki, M. Höök, and S. Lukomski, "Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices," *Journal of Biological Chemistry,* vol. 277, pp. 27312-27318, 2002.
[91] D. D. Buechter, D. N. Paolella, B. S. Leslie, M. S. Brown, K. A. Mehos, and E. A. Gruskin, "Co-translational incorporation of trans-4-hydroxyproline into recombinant proteins in bacteria," *Journal of Biological Chemistry,* vol. 278, pp. 645-650, 2003.
[92] F. Ruggiero and M. Koch, "Making recombinant extracellular matrix proteins," *Methods,* vol. 45, pp. 75-85, 2008.
[93] O. Pakkanen, E.-R. Hämäläinen, K. I. Kivirikko, and J. Myllyharju, "Assembly of Stable Human Type I and III Collagen Molecules from Hydroxylated Recombinant Chains in the Yeast Pichia pastoris: EFFECT OF AN ENGINEERED C-TERMINAL OLIGOMERIZATION DOMAIN FOLDON," *Journal of Biological Chemistry,* vol. 278, pp. 32478-32483, 2003.
[94] T. Pihlajaniemi, R. Myllylä, and K. I. Kivirikko, "Prolyl 4-hydroxylase and its role in collagen synthesis," *Journal of Hepatology,* vol. 13, Supplement 3, pp. S2-S7, 1991.
[95] P. D. Toman, G. Chisholm, H. McMullin, L. M. Giere, D. R. Olsen, R. J. Kovach*, et al.*, "Production of Recombinant Human Type I Procollagen Trimers Using a Four-gene Expression System in the Yeast Saccharomyces cerevisiae," *Journal of Biological Chemistry,* vol. 275, pp. 23303-23309, 2000.
[96] I. Majsterek, E. McAdams, E. Adachi, S. T. Dhume, and A. Fertala, "Prospects and limitations of the rational engineering of fibrillar collagens," *Protein Science,* vol. 12, pp. 2063-2072, 2003.
[97] S. Kyle, A. Aggeli, E. Ingham, and M. J. McPherson, "Production of self-assembling biomaterials for tissue engineering," *Trends in Biotechnology,* vol. 27, pp. 423-433, 2009.
[98] P. R. Vaughan, M. Galanis, K. M. Richards, T. A. Tebb, J. A. M. Ramshaw, and J. A. Werkmeister, "Production of Recombinant Hydroxylated Human Type III Collagen Fragment in Saccharomyces cerevisiae," *DNA and Cell Biology,* vol. 17, pp. 511-518, 1998.
[99] M. W. T. Werten, T. J. van den Bosch, R. D. Wind, H. Mooibroek, and F. A. de Wolf, "High-yield secretion of recombinant gelatins by Pichia pastoris," *Yeast,* vol. 15, pp. 1087-1096, 1999.
[100] S. Gomes, I. B. Leonor, J. F. Mano, R. L. Reis, and D. L. Kaplan, "Natural and genetically engineered proteins for tissue engineering," *Progress in Polymer Science,* vol. 37, pp. 1-17, 2012.
[101] F. W. Kotch and R. T. Raines, "Self-assembly of synthetic collagen triple helices," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 103, pp. 3028-3033, 2006.
[102] L. S. Z. Larsen, C. D. Wassman, G. W. Hatfield, and R. H. Lathrop, "Computationally Optimised DNA Assembly of synthetic genes," *International journal of bioinformatics research and applications,* vol. 4, pp. 324-336, 2008.
[103] S. W. P. Chan, J. Greaves, N. A. Da Silva, and S.-W. Wang, "Assaying proline hydroxylation in recombinant collagen variants by liquid chromatography-mass spectrometry," *BMC Biotechnology,* vol. 12, pp. 51-51, 2012.
[104] N. J. Greenfield, "Using circular dichroism spectra to estimate protein secondary structure," *Nat. Protocols,* vol. 1, pp. 2876-2890, 2007.

[105] N. J. Greenfield, "Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions," *Nat. Protocols,* vol. 1, pp. 2527-2535, 2007.
[106] C. Graf, D. L. J. Vossen, A. Imhof, and A. van Blaaderen, "A General Method To Coat Colloidal Particles with Silica," *Langmuir,* vol. 19, pp. 6693-6700, 2003.
[107] G. Liu, Y. Tian, and Y. Kan, "Fabrication of high-aspect-ratio microstructures using SU8 photoresist," *Microsystem Technologies,* vol. 11, pp. 343-346, 2005.
[108] K. Kyburz and K. Anseth, "Synthetic Mimics of the Extracellular Matrix: How Simple is Complex Enough?," *Annals of Biomedical Engineering,* vol. 43, pp. 489-500, 2015.
[109] N. Raynal, S. W. Hamaia, P. R.-M. Siljander, B. Maddox, A. R. Peachey, R. Fernandez*, et al.*, "Use of Synthetic Peptides to Locate Novel Integrin α2β1-binding Motifs in Human Collagen III," *Journal of Biological Chemistry,* vol. 281, pp. 3821-3831, 2006.
[110] S. W. P. Chan, "Fabrication of Modular Human Collagen and Collagen Variants in Yeast," 3502568 Ph.D., University of California, Irvine, 2012.
[111] Z. Deyl and I. Mikšík, "Advanced separation methods for collagen parent α-chains, their polymers and fragments," *Journal of Chromatography B: Biomedical Sciences and Applications,* vol. 739, pp. 3-31, 2000.
[112] D. Rodríguez, C. J. Morrison, and C. M. Overall, "Matrix metalloproteinases: What do they not do? New substrates and biological roles identified by murine models and proteomics," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research,* vol. 1803, pp. 39-54, 2010.
[113] K. E. Kadler, D. F. Holmes, J. A. Trotter, and J. A. Chapman, "Collagen fibril formation," *Biochemical Journal,* vol. 316, pp. 1-11, 1996.
[114] F. Gittes, B. Schnurr, P. D. Olmsted, F. C. MacKintosh, and C. F. Schmidt, "Microscopic Viscoelasticity: Shear Moduli of Soft Materials Determined from Thermal Fluctuations," *Physical Review Letters,* vol. 79, pp. 3286-3289, 1997.
[115] T. G. Mason and D. A. Weitz, "Optical Measurements of Frequency-Dependent Linear Viscoelastic Moduli of Complex Fluids," *Physical Review Letters,* vol. 74, pp. 1250-1253, 1995.
[116] T. H. Larsen and E. M. Furst, "Microrheology of the Liquid-Solid Transition during Gelation," *Physical Review Letters,* vol. 100, p. 146001, 2008.
[117] P. J. Stahl, N. H. Romano, D. Wirtz, and S. M. Yu, "PEG-Based Hydrogels with Collagen Mimetic Peptide-Mediated and Tunable Physical Cross-Links," *Biomacromolecules,* vol. 11, pp. 2336-2344, 2010.
[118] Susanta K. Sarkar, B. Marmer, G. Goldberg, and Keir C. Neuman, "Single-Molecule Tracking of Collagenase on Native Type I Collagen Fibrils Reveals Degradation Mechanism," *Current Biology,* vol. 22, pp. 1047-1056, 2012.
[119] M. C. Branco, D. J. Pochan, N. J. Wagner, and J. P. Schneider, "Macromolecular diffusion and release from self-assembled β-hairpin peptide hydrogels," *Biomaterials,* vol. 30, pp. 1339-1347, 2009.
[120] T. L. Meyvis, S. De Smedt, P. Van Oostveldt, and J. Demeester, "Fluorescence Recovery After Photobleaching: A Versatile Tool for Mobility and Interaction Measurements in Pharmaceutical Research," *Pharmaceutical Research,* vol. 16, pp. 1153-1162, 1999.
[121] J. G. Bann, D. H. Peyton, and H. P. Bächinger, "Sweet is stable: glycosylation stabilizes collagen," *FEBS Letters,* vol. 473, pp. 237-240, 2000.

[122] J. Engel and H. Bächinger, "Structure, Stability and Folding of the Collagen Triple Helix," in *Collagen*. vol. 247, J. Brinckmann, H. Notbohm, and P. K. Müller, Eds., ed: Springer Berlin Heidelberg, pp. 7-33, 2005.
[123] J. K. Kim, Y. Xu, X. Xu, D. R. Keene, S. Gurusiddappa, X. Liang*, et al.*, "A Novel Binding Site in Collagen Type III for Integrins α1β1 and α2β1," *Journal of Biological Chemistry,* vol. 280, pp. 32512-32520, 2005.
[124] J. Emsley, C. G. Knight, R. W. Farndale, and M. J. Barnes, "Structure of the Integrin α2β1-binding Collagen Peptide," *Journal of Molecular Biology,* vol. 335, pp. 1019-1028, 2004.
[125] M. J. B. Wissink, M. J. A. van Luyn, R. Beernink, F. Dijk, A. A. Poot, G. H. M. Engbers*, et al.*, "Endothelial Cell Seeding on Crosslinked Collagen: Effects of Crosslinking on Endothelial Cell Proliferation and Functional Parameters," *Thrombosis and Haemostasis,* vol. 84, pp. 325-331, 2000.
[126] R. Usha, K. J. Sreeram, and A. Rajaram, "Stabilization of collagen with EDC/NHS in the presence of l-lysine: A comprehensive study," *Colloids and Surfaces B: Biointerfaces,* vol. 90, pp. 83-90, 2012.
[127] R. Walton, D. Brand, and J. Czernuszka, "Influence of telopeptides, fibrils and crosslinking on physicochemical properties of Type I collagen films," *Journal of Materials Science: Materials in Medicine,* vol. 21, pp. 451-461, 2010.
[128] L. H. H. Olde Damink, P. J. Dijkstra, M. J. A. van Luyn, P. B. van Wachem, P. Nieuwenhuis, and J. Feijen, "Cross-linking of dermal sheep collagen using a water-soluble carbodiimide," *Biomaterials,* vol. 17, pp. 765-773, 1996.
[129] P. Angele, J. Abke, R. Kujat, H. Faltermeier, D. Schumann, M. Nerlich*, et al.*, "Influence of different collagen species on physico-chemical properties of crosslinked collagen matrices," *Biomaterials,* vol. 25, pp. 2831-2841, 2004.
[130] I. Rault, V. Frei, D. Herbage, N. Abdul-Malak, and A. Huc, "Evaluation of different chemical methods for cros-linking collagen gel, films and sponges," *Journal of Materials Science: Materials in Medicine,* vol. 7, pp. 215-221, 1996.
[131] C. M. Stultz, "Localized Unfolding of Collagen Explains Collagenase Cleavage Near Imino-poor Sites," *Journal of Molecular Biology,* vol. 319, pp. 997-1003, 2002.
[132] J. M. Lee, H. H. L. Edwards, C. A. Pereira, and S. I. Samii, "Crosslinking of tissue-derived biomaterials in 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC)," *Journal of Materials Science: Materials in Medicine,* vol. 7, pp. 531-541, 1996.
[133] C. L. Parr, R. A. B. Keates, B. C. Bryksa, M. Ogawa, and R. Y. Yada, "The structure and function of Saccharomyces cerevisiae proteinase A," *Yeast,* vol. 24, pp. 467-480, 2007.
[134] E. Kominami, H. Hoffschulte, L. Leuschel, K. Maier, and H. Holzer, "The substrate specificity of proteinase B from Baker's yeast," *Biochimica et Biophysica Acta (BBA) - Enzymology,* vol. 661, pp. 136-141, 1981.

Supplementary Information

Possible Reasons for Failure of rCol-1M-PF7 to Purify

Because the additional protease inhibitor cocktail could only affect proteases that had been released into the cell suspension after the cells were broken, it is possible that endogenous yeast proteases were able to unravel the collagen triple helix and cleave it, resulting in the shortened proteins shown in Figures 2-6 and 2-7. If, indeed, the addition of the non-native MMP-cleavable site into the area of high stability in the collagen triple helix resulted in a loss of fidelity, it is possible that the proteolysis of the collagen could result in shortened bands as seen in Figures 2-6 and 2-7. The anti-collagen antibody used in the Western blots binds near the C-terminus of the collagen strand. The molecular weight of the collagen molecule from the MMP digestion site on PF 7 to the C-terminus of the protein would be 72.4 kDa, assuming, of course, that the collagen was digested at the MMP cleavage site. Given that collagen, because of its linear nature, has a tendency to run larger than expected on SDS-PAGE (and, thus, western blots), it is possible that a 72.4 kDa collagen band is shown in the rCol-1M-PF7 lanes in Figures 2-6 and 2-7. If the non-native MMP site resulted in a loosely-wound collagen triple helix, there are a number of potential protease sites in or near the MMP site sequence, as determined by the protease calculator at http://web.expasy.org/peptide_cutter/. Yeast proteinases A and B can both cleave amino acid sequences found within the MMP site [133, 134], so it is possible that some endogenous yeast protease (whether it be yeast proteinases A, B, or something else) was able to degrade and cleave the collagen, resulting in the shortened protein bands seen here.

Yeast proteinase A is able to cleave between the I-A and L-A amino acids found in the MMP site, and yeast proteinase B is able to cleave the L-G, R-G, and G-L bonds in the MMP site.

SDS-PAGE Staining to Detect Digested Collagen Fragments

It is important to note the difficulty of staining these SDS-PAGE gels to identify digested collagen bands. Staining with imidazole and zinc sulfate resulted only in the staining of large collagen bands (>75 kDa). Bands smaller than 75 kDa were not shown using the imidazole / zinc sulfate staining method. In stark contrast, silver staining resulted only in bands below 35 kDa being stained, with none of the larger bands appearing. The most consistent staining method was the vendor-bought Imperial Blue stain from Thermo, which still had its own drawbacks. After staining the SDS-PAGE overnight, digested collagen bands could be seen clearly among high background staining. Any attempts to de-stain the gel to decrease background staining, however, resulted in a loss of the digested bands, especially those smaller than 100 kDa. The small digested collagen bands appear to migrate out of the gel to the surface during the staining process, making them fragile and prone to being smeared or washed away during de-staining. To combat this protein migration, digested collagen was run on SDS-PAGE, then stained with a homemade Coomassie Blue R-250 dye with 45% methanol, 10% glacial acetic acid in the staining solution to fix the proteins. This homemade Coomassie Blue stain resulted in small digested collagen bands that were fixed to the gel, but overall did not stain the collagen very well. Bands were hard to discern amid high background staining, and attempts to de-stain with 45% methanol, 10% glacial acetic acid only resulted in decreased staining of the gel in every facet- both background and protein-bound stain were decreased, showing very light,

faint collagen bands once background stain was eliminated. One final staining method was performed with the gel being fixed for 3 hours in 45% methanol, 10% glacial acetic acid, then stained overnight with the vendor-bought Imperial Blue stain. However, despite the methanol fixative, small bands that were stained with the Imperial Blue stain still rubbed off and disappeared if the gel was de-stained or imaged on a scanner. This indicates that the Imperial Blue stain is not compatible with the fixing of collagen bands in the SDS-PAGE gel, even though a homemade Coomassie Blue stain was successfully used to stain a gel (albeit with very light protein staining). Ultimately, it was decided that the best images could be obtained by staining the SDS-PAGE gels of collagen digestions by staining with the Imperial Blue stain overnight, and carefully taking a picture of the gel with a simple camera. That is the reason for the blue background in Figures 2-8 and 2-9: those gels could not be further de-stained for fear that the digested bands would disappear. Future studies should be wary of this limitation of SDS-PAGE staining. This also reiterates the need for another method of detection and identification of digested collagen fragments, such as mass spectrometry or size-exclusion chromatography.

Cell Adhesion Assay Execution Discussion

Light washing by pipet was used to remove any un- or weakly- adhered cells from the surface of the protein. This adds some inherent possibility for human error, as washings may not be consistent from well to well or day to day. This potential source of error and inconsistency can be mitigated by gently centrifuging the cell-containing plates upside down for a short time. Centrifugation of the entire plate at once keeps the force applied to each well constant, so there is reduced variability between wells. However,

during early experiments, the usual method of using PCR sealing tape to seal the wells and prevent them from leaking media when turned upside down proved unsuccessful because of the inability for the tape to seal to the edges of the 24 well plate, unlike the 96 well plate used in previous experiments [41]. 24 well microplates were used instead of the usual 96 well microplates because BD does not produce carboxylated plates in the 96 well size, forcing us to use the next smallest size (24 wells). The reagent volumes were adjusted and scaled up using the ratio of 24 well plate surface area: 96 well plate surface area, which is 5.93. Future experiments may adjust these reaction and incubation volumes to more appropriately fit the 24 well plate system and to make the numbers seem "prettier" for posterity's sake, but the reaction volumes presented here are scientifically sound, although they may seem odd at first glance.

Modular Collagen Gene PF, SF, and Full Gene Sequences

rCol "Baseline" full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTAATGATG
GTGCTAGAGGATCTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC
GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACAAAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG

CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA
GGGTCCGCCCGGCGCACCAGGCCCGTTGGGAATAGCTGGAATTACTGGTGCTAGGGGCTTGGCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol “Baseline” full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQELDCPNPEIPFGECCAVCPQPPTAPTR
PPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGICESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSP
GSPGYQGPPGEPGQAGPSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGAPGLKGENGLPG
ENGAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGI
NGSPGGKGEMGPAGIPGAPGLMGARGPPGPAGANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGER
GAPGFRGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGESGRPGPPGPSGPRGQPGVMGFPG
PKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKGDTGPPGPQGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGA
PGERGPPGLAGAPGLRGGAGPPGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPPGPAGQP
GDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPGVAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAG
FPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGNTGAPGSPGVSGPKGDAGQGEKGSPGAQGPPGAPGPLGIAGITGARGLAGPPGMPGPRGSP
GPQGVKGESGKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPGPPGPVGPAGKSGDRGESG
PAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGSPGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGP
RGNRGERGSEGSPGHPGQPGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPARNC
RDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMDGGFQFSYGNPELPEDVLDVQLAFLRL
LSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEF
GVDVGPVCFL-

rCol “Baseline” primary fragments

MPF 1 DNA sequence

CCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATCTGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCG
GACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCCGGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAG
TGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAGCCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTG
CCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTTTCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTG
AAACTGGTGCT

MPF 1 amino acid sequence

PQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAGPSGPPGP
PGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA

MPF 2 DNA sequence

ATAAAAGGCCCTGCGGGAATACCAGGTTTTCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTG
GTTTAAAAGGAGAAAATGGCTTGCCAGGCGAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCC
AGGAGCAGCCGGTGCCCGTGGTAATGATGGTGCTAGAGGATCTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGT
CCCGGCGCCAAAGGTGAAGTTGGTCCAGCAGGTTCACCCGGTTCAAATGGTGCT

MPF 2 amino acid sequence

IKGPAGIPGFPGMKGHRGFDGRNGEKGETGAPGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGA
RGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEVGPAGSPGSNGA

MPF 3 DNA sequence

CCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGCAGGTTCACCCGGTTCAAATGGTGCTCCTG
GACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATTAACGGTAGCCCAGGAGGTAAAGGCGAAAT
GGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAGCTAACGGTGCTCCAGGTCTAAGAGGTGGA
GCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGA

MPF 3 amino acid sequence

PPGPPGTAGFPGSPGAKGEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGI
PGAPGLMGARGPPGPAGANGAPGLRGGAGEPGKNGAKGEPGPRG

MPF 4 DNA sequence

GGTCCAGCTGGAGCTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAAC
GTGGAGAAGCCGGCATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGC
TGCAGGCGAAAGAGGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGA
CCGGCCGGACCACGTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCA

MPF 4 amino acid sequence

GPAGANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERG
APGFRGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVP

MPF 5 DNA sequence

CCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCACGTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAG
GTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAAACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGG
TCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGCCCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGA
GGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAA

MPF 5 amino acid sequence

PGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGESGRPGPPGPS
GPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGE

MPF 6 DNA sequence

CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACAAAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGA

MPF 6 amino acid sequence

PKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKGDTGPPGPQGLQGLPGTGGPPGE
NGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRG

MPF 7 DNA sequence

GGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCGCCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAG
CTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGACCACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGA
GAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGGTGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGA
CCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCTGGTGATAAA

MPF 7 amino acid sequence

GAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGPPGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLG
SPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPPGPAGQPGDK

MPF 8 DNA sequence

CCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCTGGTGATAAAGGTG
AAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAGCTGGATTCCCAGG
TGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGTTGCAGGTCCACCA
GGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAA

MPF 8 amino acid sequence

PGADGVPGKDGPRGPTGPIGPPGPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQN
GEPGGKGERGAPGEKGEGGPPGVAGPPGGSGPAGPPGPQGVKGE

MPF 9 DNA sequence

GAAAAAGGAGAGGGTGGACCACCAGGAGTTGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAG
GTTCTCCTGGAGGTCCAGGAGCCGCTGGTTTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTC
TGGATCTCCTGGTAAAGACGGTCCACCTGGACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAA
CCCGGTGAAAAAGGTTCACCA

MPF 9 amino acid sequence

EKGEGGPPGVAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGK
DGPPGPAGNTGAPGSPGVSGPKGDAGQPGEKGSP

MPF 10 DNA sequence

GGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCAGGGTCCGCCCGGCGCACCAGGCCCGT
TGGGAATAGCTGGAATTACTGGTGCTAGGGGCTTGGCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGTCCAGGTCCACAAGGCGTTAAAGGCGA
GTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGATTGGCTGGTACTGCTGGTGAGCCAGGT
AGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGT

MPF 10 amino acid sequence

GSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPLGIAGITGARGLAGPPGMPGPRGSPGPQGVKGESGKPG
ANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPG

MPF 11 DNA sequence

GGTTTGCCTGGATTGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCA
AGGGAGATAGAGGTGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGA
TAGAGGAGAGTCAGGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGT
GAAACAGGTGAGAGAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGT

MPF 11 amino acid sequence

GLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPGPPGPVGPAGKSGDRGESG
PAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPG

MPF 12 DNA sequence

AGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGAGAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTA
ATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAGTCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGG

ACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGACCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCT
GGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTTGTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTG
AAAAAGCTGGTGGATTC

MPF 12 amino acid sequence

RGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGSPGPAGQQGAIGSPGPAGPRGPVGPSGPPGKD
GTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGF

rCol "Baseline" secondary fragments

FSF DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATAC

FSF amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQY

MSF 1 DNA sequence

CCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATCTGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCG
GACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCCGGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAG
TGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAGCCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTG
CCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTTTCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTG
AAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGCGAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGG
TAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTAATGATGGTGCTAGAGGATCTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCT
GGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGCAGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTC
AGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATTAACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGC
TCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAGCTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGT
GCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGGCATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTG
GTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGAGGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAA
AGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCACGTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATG
AGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAAACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCC
CCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGCCCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGG
TGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAA

MSF 1 amino acid sequence

PQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAGPSGPPGP
PGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGAPGLKGEN
GLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAKGEVGPAG
SPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAGANGAPGL
RGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGFRGPAGPN
GIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGESGRPGPPG
PSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGE

MSF 2 DNA sequence

CCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCACGTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAG
GTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAAACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGG

TCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGCCCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGA
GGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAGGACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACA
AAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGGAGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCC
AAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCGCCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGA
TTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGACCACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAG
GTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGGTGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGA
TGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCTGGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATT
GCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAGCTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTG
GTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGTTGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGG
TCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGTTTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAAT
GGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTGGACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTAT
CTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCAGGGTCCGCCCGGCGCACCAGGC

MSF 2 amino acid sequence

PGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGESGRPGPPGPS
GPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKGDTGPPGPQG
LQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGPPGPEGGKGA
AGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPPGPAGQPGDK
GEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPGVAGPPGGSG
PAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGNTGAPGSPGV
SGPKGDAGQPGEKGSPGAQGPPGAPG

MSF 3 DNA sequence

GAAAAAGGAGAGGGTGGACCACCAGGAGTTGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAG
GTTCTCCTGGAGGTCCAGGAGCCGCTGGTTTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTC
TGGATCTCCTGGTAAAGACGGTCCACCTGGACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAA
CCCGGTGAAAAAGGTTCACCAGGAGCTCAGGGTCCGCCCGGCGCACCAGGCCCGTTGGGAATAGCTGGAATTACTGGTGCTAGGGGCTTGGCTGGAC
CGCCAGGCATGCCAGGTCCTAGGGGCAGTCCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGG
TCCGCCTGGGCCTCAGGGTTTGCCTGGATTGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGAC
GGATCCCCTGGTGGCAAGGGAGATAGAGGTGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAG
CCGGCAAATCTGGTGATAGAGGAGAGTCAGGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCC
TAGAGGTGACAAGGGTGAAACAGGTGAGAGAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGA
CCAGCTGGTCAACAAGGTGCTATTGGTAGTCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTG
GTCATCCAGGTCCAATAGGTCCTCCTGGACCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCC
TGGACCACCAGGTGCTCCGGGGCCCTGTTGTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTC

MSF 3 amino acid sequence

EKGEGGPPGVAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGK
DGPPGPAGNTGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPLGIAGITGARGLAGPPGMPGPRGSP
GPQGVKGESGKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPG
APGAPGHPGPPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGF
PGNPGAPGSPGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHP
GQPGPPGPPGAPGPCCGGVGAAAIAGIGGEKAGGF

BSF DNA sequence

GGGCCCTGTTGTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGG
ACTTTAAGATTAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCC
AGCTAGGAACTGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATT
AAGGTGTTCTGTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGA
AGAAGCATGTTTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGC
ATTCTTGCGTTTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAG
AAGGCATTAAAGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAAC

ACACGGGTGAATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCC
AGATCAAGAATTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTG

BSF amino acid sequence

GPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPARNCR
DLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGESMD
GGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEFKAE
GNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL

rCol-0M full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTAATGATG
GTGCTAGAGGATCTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC
GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACAAAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG
CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA
GGGTCCGCCCGGCGCACCAGGtCCaccGGGAATAcCTGGAATTcCTGGTGCTcctGGtTTacCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol-0M full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAG
PSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA
PGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAK
GEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAG
ANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGES
GRPGPPGPSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKG
DTGPPGPQGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGP
PGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPP
GPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPG
VAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGN
TGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPPGIPGIPGAPGLPGPPGMPGPRGSPGPQGVKGES
GKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPG
PPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGS
PGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
GAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPAR
NCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGE
SMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEF
KAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL-

rCol-0M MMP KO MPF 10 DNA sequence

GGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCAGGGTCCGCCCGGCGCACCAGGtCCac
cGGGAATAcCTGGAATTcCTGGTGCTcctGGtTTacCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGTCCAGGTCCACAAGGCGTTAAAGGCGA
GTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGATTGGCTGGTACTGCTGGTGAGCCAGGT
AGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGT

rCol-0M MMP KO MPF 10 amino acid sequence

GSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPPGIPGIPGAPGLPGPPGMPGPRGSPGPQGVKGESGKPG
ANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPG

rCol-1M-PF2 full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTccattgG
GTatTgctGGAatTactGGAgctCgTGGCttggCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC

GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACAAAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG
CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA
GGGTCCGCCCGGCGCACCAGGtCCaccGGGAATAcCTGGAATTcCTGGTGCTcctGGtTTacCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol-1M-PF2 full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAG
PSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA
PGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGPLGIAGITGARGLAGPPGTAGFPGSPGAK
GEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAG
ANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGES
GRPGPPGPSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKG
DTGPPGPQGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGP
PGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPP
GPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPG
VAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGN
TGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPPGIPGIPGAPGLPGPPGMPGPRGSPGPQGVKGES
GKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPG
PPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGS
PGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
GAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPAR
NCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGE
SMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEF
KAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL-

rCol-1M-PF2 MMP added PF 2 DNA sequence

ATAAAAGGCCCTGCGGGAATACCAGGTTTTCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTG
GTTTAAAAGGAGAAAATGGCTTGCCAGGCGAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCC
AGGAGCAGCCGGTGCCCGTGGTCCATTGGGTATTGCTGGAATTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGT
CCCGGCGCCAAAGGTGAAGTTGGTCCAGCAGGTTCACCCGGTTCAAATGGTGCT

rCol-1M-PF2 MMP added PF 2 amino acid sequence

IKGPAGIPGFPGMKGHRGFDGRNGEKGETGAPGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGA
RGPLGIAGIDGQPGPPGPPGTAGFPGSPGAKGEVGPAGSPGSNGA

rCol-1M-PF6 full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTAATGATG
GTGCTAGAGGATCTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC
GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCATTGGGAATTGCTGGTATTACCGGAGCTCGTGGATTGGCTGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG
CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA
GGGTCCGCCCGGCGCACCAGGtCCaccGGGAATAcCTGGAATTcCTGGTGCTcctGGtTTacCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol-1M-PF6 full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAG
PSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA
PGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAK
GEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAG
ANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGES
GRPGPPGPSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPLGIAG
ITGARGLAGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGP
PGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPP
GPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPG
VAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGN
TGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPPGIPGIPGAPGLPGPPGMPGPRGSPGPQGVKGES
GKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPG
PPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGS
PGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
GAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPAR
NCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGE
SMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEF
KAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL-

rCol-1M-PF6 MMP added PF 6 DNA sequence

CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCATTGGGAATTGCTGGTATTACCGGAGCTCGTGGATTGGCTGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGA

rCol-1M-PF6 MMP added PF 6 amino acid sequence

PKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPLGIAGITGARGLAGLQGLPGTGGPPGE
NGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRG

rCol-1M-PF7 full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTAATGATG
GTGCTAGAGGATCTGACGGACAACCTGGCCCACCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC

GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACAAAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CAttgGGTattgctGGTattactGGGgctagaGGTTTGgctGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG
CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA
GGGTCCGCCCGGCGCACCAGGtCCaccGGGAATAcCTGGAATTcCTGGTGCTcctGGtTTacCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol-1M-PF7 full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAG
PSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA
PGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGNDGARGSDGQPGPPGPPGTAGFPGSPGAK
GEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAG
ANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGES
GRPGPPGPSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKG
DTGPPGPQGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGP
PGPEGGKGAAGPLGIAGITGARGLAGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPP
GPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPG
VAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGN
TGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPPGIPGIPGAPGLPGPPGMPGPRGSPGPQGVKGES
GKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPG
PPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGS
PGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
GAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPAR
NCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGE
SMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEF
KAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL-

rCol-1M-PF7 MMP added PF 7 DNA sequence

GGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCGCCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAG
CTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGACCAttgGGTattgctGGTattactGGGgctagaGGTTTGgctGGTATGCCAGGAGA
GAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGGTGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGA
CCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCTGGTGATAAA

rCol-1M-PF7 MMP added PF7 amino acid sequence

GAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGPPGPEGGKGAAGPLGIAGITGARGLAGMPGERGGLG
SPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPPGPAGQPGDK

rCol-2M full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTccattgG
GTatTgctGGAatTactGGAgctCgTGGCttggCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC
GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAGGTGGAGACAAAGGTGATACCGGACCACCTGGACCCCAAGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG
CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA
GGGTCCGCCCGGCGCACCAGGCCCGTTGGGAATAGCTGGAATTACTGGTGCTAGGGGCTTGGCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol-2M full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAG
PSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA
PGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGPLGIAGITGARGLAGPPGTAGFPGSPGAK
GEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAG
ANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGES
GRPGPPGPSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPGGDKG
DTGPPGPQGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGP
PGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPP
GPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPG
VAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGN
TGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPLGIAGITGARGLAGPPGMPGPRGSPGPQGVKGES
GKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPG
PPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGS
PGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
GAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPAR
NCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGE
SMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEF
KAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL-

rCol-3M full length collagen DNA sequence

ATGATGTCTTTCGTTCAAAAGGGGAGCTGGTTGTTGTTGGCTCTACTCCACCCGACCATTATTTTGGCCCAGCAGGAAGCTGTTGAAGGTGGATGTT
CTCATTTGGGTCAAAGCTACGCGGATAGAGATGTTTGGAAGCCAGAACCATGTCAAATCTGTGTCTGTGATAGCGGTTCGGTCCTGTGTGATGACAT
TATTTGTGACGATCAGGAGCTCGACTGTCCTAACCCTGAGATCCCATTCGGTGAGTGTTGTGCTGTTTGTCCGCAACCGCCCACTGCTCCAACTCGT
CCTCCCAATGGTCAAGGTCCCCAAGGTCCAAAAGGTGATCCGGGGCCACCAGGCATACCAGGTAGAAATGGTGACCCAGGAATACCCGGACAGCCAG
GTTCCCCAGGGAGTCCAGGACCTCCCGGAATTTGTGAGTCTTGTCCAACTGGACCCCAGAACTACAGCCCGCAATACGATTCTTACGACGTCAAATC
TGGTGTTGCTGTTGGTGGGCTAGCTGGTTATCCAGGTCCCGCTGGTCCTCCCGGACCGCCTGGTCCACCAGGTACATCTGGACATCCTGGCTCACCC
GGTTCACCAGGTTATCAAGGTCCACCTGGTGAGCCCGGACAAGCTGGTCCAAGTGGTCCTCCTGGTCCGCCAGGAGCTATCGGTCCATCTGGTCCAG
CCGGTAAGGATGGTGAATCTGGTAGACCAGGCAGACCTGGCGAAAGGGGCTTGCCAGGTCCACCTGGAATAAAAGGCCCTGCGGGAATACCAGGTTT
TCCTGGCATGAAAGGTCACAGAGGTTTTGATGGTAGAAATGGCGAAAAGGGTGAAACTGGTGCTCCTGGTTTAAAAGGAGAAAATGGCTTGCCAGGC
GAAAATGGAGCTCCAGGCCCAATGGGTCCAAGGGGTGCTCCAGGAGAGAGGGGTAGACCTGGTTTGCCAGGAGCAGCCGGTGCCCGTGGTccattgG
GTatTgctGGAatTactGGAgctCgTGGCttggCTGGACCACCTGGAACAGCTGGCTTTCCTGGCAGTCCCGGCGCCAAAGGTGAAGTTGGTCCAGC
AGGTTCACCCGGTTCAAATGGTGCTCCTGGACAAAGGGGAGAACCCGGACCTCAGGGTCATGCTGGTGCTCAAGGTCCTCCCGGTCCACCCGGAATT
AACGGTAGCCCAGGAGGTAAAGGCGAAATGGGTCCTGCCGGAATACCTGGTGCTCCAGGCTTAATGGGTGCGAGAGGCCCTCCTGGTCCAGCTGGAG
CTAACGGTGCTCCAGGTCTAAGAGGTGGAGCCGGTGAACCAGGAAAGAATGGTGCTAAAGGCGAACCCGGACCAAGGGGAGAACGTGGAGAAGCCGG
CATACCCGGAGTACCCGGTGCTAAAGGTGAGGATGGCAAAGATGGTTCCCCTGGTGAGCCTGGGGCTAATGGATTGCCTGGTGCTGCAGGCGAAAGA
GGAGCTCCTGGATTCAGAGGACCCGCAGGACCAAATGGAATACCAGGTGAAAAAGGCCCAGCAGGTGAAAGAGGTGCGCCCGGACCGGCCGGACCAC
GTGGAGCAGCAGGAGAACCCGGCAGAGATGGAGTTCCAGGAGGTCCAGGAATGAGAGGTATGCCAGGTTCGCCAGGTGGTCCTGGATCTGATGGTAA
ACCCGGTCCACCGGGAAGCCAAGGTGAGTCAGGCAGACCCGGTCCTCCTGGCCCCTCAGGACCAAGGGGTCAACCAGGAGTTATGGGCTTTCCTGGC
CCAAAAGGCAACGATGGCGCACCTGGTAAAAATGGTGAAAGAGGCGGACCCGGTGGTCCCGGCCCACAAGGACCTCCTGGAAAGAATGGTGAAACAG
GACCACAAGGGCCACCCGGCCCAACAGGTCCAttgGGAattgctGGTatTACCGGAgCtCgTGGAttggctGGATTGCAAGGACTACCTGGTACAGG
AGGACCTCCCGGAGAGAATGGTAAACCAGGTGAACCTGGTCCAAAAGGTGATGCTGGAGCTCCAGGAGCTCCGGGTGGAAAAGGTGATGCAGGCGCG
CCAGGTGAACGTGGTCCACCAGGTTTGGCAGGAGCACCAGGATTGAGAGGAGGAGCTGGTCCACCTGGTCCCGAAGGAGGTAAAGGAGCTGCAGGAC
CACCTGGTCCTCCTGGTGCTGCCGGGACTCCAGGTTTGCAAGGTATGCCAGGAGAGAGAGGCGGTTTGGGTTCACCCGGACCAAAAGGTGATAAGGG
TGAACCTGGTGGTCCTGGCGCAGACGGTGTTCCCGGTAAAGATGGTCCCAGAGGACCAACTGGACCAATTGGACCTCCCGGACCCGCAGGTCAGCCT
GGTGATAAAGGTGAAGGTGGTGCACCCGGATTACCAGGAATTGCTGGCCCAAGGGGTAGTCCAGGTGAAAGGGGAGAAACAGGTCCACCAGGCCCAG
CTGGATTCCCAGGTGCCCCTGGACAAAATGGTGAGCCTGGTGGTAAAGGTGAAAGAGGGGCTCCAGGCGAAAAAGGAGAGGGTGGACCACCAGGAGT
TGCAGGTCCACCAGGAGGTTCCGGACCAGCAGGACCTCCGGGTCCCCAGGGTGTTAAAGGAGAAAGAGGTTCTCCTGGAGGTCCAGGAGCCGCTGGT
TTTCCAGGAGCTAGGGGATTGCCAGGACCGCCCGGATCTAATGGTAATCCTGGACCACCCGGTCCTTCTGGATCTCCTGGTAAAGACGGTCCACCTG
GACCGGCGGGAAATACTGGTGCACCGGGTTCACCCGGTGTATCTGGACCAAAGGGAGATGCTGGTCAACCCGGTGAAAAAGGTTCACCAGGAGCTCA

GGGTCCGCCCGGCGCACCAGGCCCGTTGGGAATAGCTGGAATTACTGGTGCTAGGGGCTTGGCTGGACCGCCAGGCATGCCAGGTCCTAGGGGCAGT
CCAGGTCCACAAGGCGTTAAAGGCGAGTCAGGAAAACCAGGTGCTAATGGTTTATCGGGAGAAAGAGGTCCGCCTGGGCCTCAGGGTTTGCCTGGAT
TGGCTGGTACTGCTGGTGAGCCAGGTAGAGATGGTAATCCCGGTTCTGATGGATTACCCGGCAGGGACGGATCCCCTGGTGGCAAGGGAGATAGAGG
TGAAAATGGTTCCCCAGGTGCTCCCGGTGCTCCAGGGCATCCGGGACCTCCCGGTCCTGTAGGACCAGCCGGCAAATCTGGTGATAGAGGAGAGTCA
GGTCCTGCAGGTCCCGCAGGAGCTCCTGGCCCCGCCGGATCTAGGGGTGCACCAGGTCCCCAAGGCCCTAGAGGTGACAAGGGTGAAACAGGTGAGA
GAGGTGCTGCGGGCATTAAAGGTCATAGAGGCTTTCCAGGTAATCCTGGTGCACCTGGTAGTCCTGGACCAGCTGGTCAACAAGGTGCTATTGGTAG
TCCTGGCCCTGCAGGTCCTCGTGGACCAGTTGGACCAAGCGGACCACCTGGTAAAGATGGAACTTCTGGTCATCCAGGTCCAATAGGTCCTCCTGGA
CCAAGGGGCAATAGAGGAGAAAGAGGCTCTGAAGGTAGCCCTGGTCATCCAGGACAGCCAGGACCTCCTGGACCACCAGGTGCTCCGGGGCCCTGTT
GTGGTGGTGTTGGTGCTGCTGCTATTGCTGGCATCGGCGGTGAAAAAGCTGGTGGATTCGCTCCATATTATGGTGATGAGCCCATGGACTTTAAGAT
TAATACGGATGAGATCATGACGTCCCTGAAATCGGTCAATGGGCAAATTGAATCGTTGATTTCACCCGATGGTTCCAGAAAGAATCCAGCTAGGAAC
TGTCGTGACCTAAAATTCTGTCACCCCGAATTAAAATCCGGTGAATATTGGGTCGATCCAAATCAAGGCTGTAAGTTGGACGCTATTAAGGTGTTCT
GTAACATGGAAACTGGCGAAACGTGTATTTCTGCGAATCCCCTCAATGTTCCTAGAAAACACTGGTGGACTGATAGTTCTGCTGAGAAGAAGCATGT
TTGGTTCGGTGAATCCATGGACGGTGGTTTTCAATTTAGCTACGGTAATCCAGAATTGCCAGAGGATGTTTTGGACGTTCAATTGGCATTCTTGCGT
TTGTTGTCATCAAGGGCGTCGCAGAATATAACTTATCATTGTAAGAACTCCATCGCCTATATGGATCAAGCATCCGGCAACGTGAAGAAGGCATTAA
AGTTGATGGGTTCTAATGAGGGCGAATTTAAGGCTGAAGGTAACAGCAAGTTCACCTATACTGTTCTGGAGGATGGTTGTACTAAACACACGGGTGA
ATGGTCTAAGACGGTGTTCGAATACAGAACGAGGAAGGCTGTGAGATTGCCCATTGTTGACATAGCTCCATACGATATAGGTGGTCCAGATCAAGAA
TTTGGGGTTGATGTTGGTCCGGTTTGTTTCTTGTAA

rCol-3M full length collagen amino acid sequence

MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCDDIICDDQ
ELDCPNPEIPFGECCAVCPQPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIPGQPGSPGSPGPPGI
CESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPAGPPGPPGPPGTSGHPGSPGSPGYQGPPGEPGQAG
PSGPPGPPGAIGPSGPAGKDGESGRPGRPGERGLPGPPGIKGPAGIPGFPGMKGHRGFDGRNGEKGETGA
PGLKGENGLPGENGAPGPMGPRGAPGERGRPGLPGAAGARGPLGIAGITGARGLAGPPGTAGFPGSPGAK
GEVGPAGSPGSNGAPGQRGEPGPQGHAGAQGPPGPPGINGSPGGKGEMGPAGIPGAPGLMGARGPPGPAG
ANGAPGLRGGAGEPGKNGAKGEPGPRGERGEAGIPGVPGAKGEDGKDGSPGEPGANGLPGAAGERGAPGF
RGPAGPNGIPGEKGPAGERGAPGPAGPRGAAGEPGRDGVPGGPGMRGMPGSPGGPGSDGKPGPPGSQGES
GRPGPPGPSGPRGQPGVMGFPGPKGNDGAPGKNGERGGPGGPGPQGPPGKNGETGPQGPPGPTGPLGIAG
ITGARGLAGLQGLPGTGGPPGENGKPGEPGPKGDAGAPGAPGGKGDAGAPGERGPPGLAGAPGLRGGAGP
PGPEGGKGAAGPPGPPGAAGTPGLQGMPGERGGLGSPGPKGDKGEPGGPGADGVPGKDGPRGPTGPIGPP
GPAGQPGDKGEGGAPGLPGIAGPRGSPGERGETGPPGPAGFPGAPGQNGEPGGKGERGAPGEKGEGGPPG
VAGPPGGSGPAGPPGPQGVKGERGSPGGPGAAGFPGARGLPGPPGSNGNPGPPGPSGSPGKDGPPGPAGN
TGAPGSPGVSGPKGDAGQPGEKGSPGAQGPPGAPGPLGIAGITGARGLAGPPGMPGPRGSPGPQGVKGES
GKPGANGLSGERGPPGPQGLPGLAGTAGEPGRDGNPGSDGLPGRDGSPGGKGDRGENGSPGAPGAPGHPG
PPGPVGPAGKSGDRGESGPAGPAGAPGPAGSRGAPGPQGPRGDKGETGERGAAGIKGHRGFPGNPGAPGS
PGPAGQQGAIGSPGPAGPRGPVGPSGPPGKDGTSGHPGPIGPPGPRGNRGERGSEGSPGHPGQPGPPGPP
GAPGPCCGGVGAAAIAGIGGEKAGGFAPYYGDEPMDFKINTDEIMTSLKSVNGQIESLISPDGSRKNPAR
NCRDLKFCHPELKSGEYWVDPNQGCKLDAIKVFCNMETGETCISANPLNVPRKHWWTDSSAEKKHVWFGE
SMDGGFQFSYGNPELPEDVLDVQLAFLRLLSSRASQNITYHCKNSIAYMDQASGNVKKALKLMGSNEGEF
KAEGNSKFTYTVLEDGCTKHTGEWSKTVFEYRTRKAVRLPIVDIAPYDIGGPDQEFGVDVGPVCFL-