Published in final edited form as:

J Struct Biol. 2018 September ; 203(3): 247–254. doi:10.1016/j.jsb.2018.05.002.

Effects of flexibility of the $\alpha 2$ chain of type I collagen on collagenase cleavage

Arya Mekkat^{1,#}, Erik Poppleton^{2,#}, Bo An², Robert Visse³, Hideaki Nagase³, David L. Kaplan², Barbara Brodsky^{2,*}, and Yu-Shan Lin^{1,*}

¹Department of Chemistry, Tufts University, Medford MA USA

²Department of Biomedical Engineering, Tufts University, Medford MA USA

³Kennedy Institute of Rheumatology, University of Oxford, Oxford UK

Abstract

Cleavage of collagen by collagenases such as matrix metalloproteinase 1 (MMP-1) is a key step in development, tissue remodeling, and tumor proliferation. The abundant heterotrimeric type I collagen composed of two $\alpha 1(I)$ chains and one $\alpha 2(I)$ chain is efficiently cleaved by MMP-1 at a unique site in the triple helix, a process which may be initiated by local unfolding within the peptide chains. Atypical homotrimers of the $\alpha 1(I)$ chain, found in embryonic and cancer tissues, are very resistant to MMP cleavage. To investigate MMP-1 cleavage, recombinant homotrimers were constructed with sequences from the MMP cleavage regions of human collagen chains inserted into a host bacterial collagen protein system. All triple-helical constructs were cleaved by MMP-1, with $\alpha 2(I)$ homotrimers cleaved efficiently at a rate similar to that seen for $\alpha 1(II)$ and $\alpha 1(III)$ homotrimers, while $\alpha 1(I)$ homotrimers were cleaved at a much slower rate. The introduction of destabilizing Gly to Ser mutations within the human collagenase susceptible region of the $\alpha 2(I)$ chain did not interfere with MMP-1 cleavage. Molecular dynamics simulations indicated a greater degree of transient hydrogen bond breaking in $\alpha 2(I)$ homotrimers compared with $\alpha 1(I)$ homotrimers at the MMP-1 cleavage site, and showed an extensive disruption of hydrogen bonding in the presence of a Gly to Ser mutation, consistent with chymotrypsin digestion results. This study indicates that $\alpha 2(I)$ homotrimers are susceptible to MMP-1, proves that the presence of an $\alpha 1(I)$ chain is not a requirement for $\alpha 2(I)$ cleavage, and supports the importance of local unfolding of $\alpha 2(I)$ in collagenase cleavage.

Keywords

Type I collagen; collagenase; MMP-1; $\alpha 1(I)$ chain; $\alpha 2(I)$ chain; cleavage sites; molecular dynamics; triple helix; homotrimer

*Corresponding authors.

#Equal contributions to this work

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Collagen, the major structural protein in the extracellular matrix (ECM), is a stable protein with a long turnover time. Controlled degradation of collagen at precise times and locations is critical during development, tissue remodeling, and wound healing, while unregulated degradation results in pathological consequences in osteoarthritis and tumor progression (Nagase et al., 2006). The key structural feature of all collagens is the triple helix, a conformation promoted by its (Gly-Xaa-Yaa)_n repeating amino acid sequence and high imino acid content. The triple-helix structure consists of three polyproline II-like chains which are supercoiled around a common axis, and this tightly wound conformation is resistant to most proteases. The most abundant collagens (types I, II, and III) self-associate to form characteristic fibrils with a 67-nm axial periodicity, and the highly specific degradation of native fibrillar collagens is mediated by a subset of the matrix metalloproteinase (MMP) family of zinc-dependent proteases, MMP-1, MMP-8, MMP-13, and MMP-14, which are collectively known as collagenases. These collagenases target a single site within the triple helix of fibrillar collagens, splitting the molecule into two fragments which can then be further degraded by these MMPs or by nonspecific proteases (Chung et al., 2004). Elucidation of the detailed mechanism leading to the initial collagenase cleavage of native collagen is important for understanding physiological development and remodeling of tissues as well as serving as a basis for therapeutic inhibition in many common diseases.

Collagenases, such as MMP-1, consist of a hemopexin-like domain involved in binding and a catalytic domain which mediates hydrolysis of the collagen chain (Amar et al., 2017; Manka et al., 2012; Nagase et al., 2006). Intriguingly, the catalytic cleft of MMP-1 is too small to accommodate the entire triple helix (Bode, 1995), and molecular dynamics (MD) studies (Stultz, 2002), together with the low imino acid content adjacent to the cleavage site (Fields, 1991), support the presence of transient unfolding such that a single chain may be accessible to catalysis. Experiments with catalytically inactive variants and isolated catalytic domains suggest that the hemopexin domain plays a role in unwinding the triple helix prior to hydrolysis (Chung et al., 2004). Some fibrillar collagens are homotrimers; the compositions of type II and type III collagens are [α 1(II)]₃ and [α 1(III)]₃, respectively. In contrast, the highly abundant type I collagen is a heterotrimer composed of two α 1(I) chains and one α 2(I) chain, and it appears the α 1(I) and α 2(I) chains play distinct roles during the MMP cleavage process. Experimental and computational studies suggest that the α 2(I) chain is more unfolded than the α 1(I) chain near the cleavage site, and may be bound initially by the hemopexin domain (Chung et al., 2004; Stultz, 2002). Although heterotrimeric type I collagen is the predominant natural form, α 1(I) homotrimers have been observed in fetal tissues and are secreted by cancer cells. Studies have indicated that α 1(I) homotrimers are highly resistant to MMP cleavage, a feature which may have important physiological and pathological consequences in cell motility and tumor growth (Han et al., 2010; Makareeva et al., 2010). In contrast to naturally occurring α 1(I) homotrimers, α 2(I) homotrimers have never been observed in a natural setting.

This study seeks to better define the collagenase cleavage of fibrillar collagens, and, in particular, the role of the $\alpha 2(I)$ chain of type I collagen, through a combined experimental and computational study. Experimentally, the collagenase cleavage region from a human collagen chain is inserted within a bacterial collagen protein. The Scl2 protein of *Streptococcus pyogenes* includes an N-terminal globular trimerization domain (denoted V) followed by a 79-triplet triple-helix domain (denoted CL) (Yoshizumi et al., 2009). Previously, introduction of the sequence surrounding the MMP cleavage site from human type III collagen between two bacterial CL domains was shown to produce a recombinant VCLCL protein that could be cleaved by MMP-1 at the original site (Yu et al., 2012). This is extended here to produce chimeric homotrimeric collagen proteins containing the human collagen sequences from collagenase cleavage regions of all major fibrillar collagen chains: $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$ and $\alpha 1(III)$. In addition, Gly to Ser replacements were introduced within the inserted $\alpha 2(I)$ sequences of the MMP cleavage region, to examine the effect of local triple-helix disruption on MMP-1 cleavage. MD simulations suggest that local triple-helix unwinding correlates with MMP cleavage of native collagens.

2. Materials and Methods

2.1 Construction, expression and purification of recombinant bacterial constructs with inserts from human collagen MMP cleavage sites

All the proteins were expressed using the cold-shock vector system, as previously described. The constructs were designed as shown in Figure 1A. Four or six triplet sequences at the MMP cleavage site from the human $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$ or $\alpha 1(III)$ chain were inserted between two CL domains (the (Gly-X-Y)₇₉ collagen-like domain from the *S. pyogenes* Scl2 strain). For the $\alpha 1(I)$, $\alpha 2(I)$ and $\alpha 1(II)$ chains, the six-triplet sequences were residues 769–786 of the triple-helix domain, while the four-triplet sequences were residues 772–783 (Supplementary Figure 1A). For the $\alpha 1(III)$ chain, the six-triplet insertion covered residues 775–792 of the triple helix, while the four-triplet insertion included residues 778–789 of the triple helix. The natural trimerization domain V was included at the N-terminus, as well as an N-terminal 8× His tag for purification. The MMP cleavage site is denoted by a dashed line in Figure 1A. The constructs were obtained by inserting the annealed oligos in between the DNA sequence of the two CL domains of the V-CL-CL protein, using the *SmaI* and *ApaI* restriction sites. The inserted human MMP cleavage sites were flanked by sequences of the bacterial collagen, GKD-GKD-GQP-GKP at the N-terminus and GPR-GEQ-GPT-GPT at the C-terminus (Yu et al., 2012). All the constructs were confirmed by DNA sequencing and then transformed into *E. coli* BL21 for protein expression.

Single colonies were selected from LB + 50 µg/ml Ampicillin agar plates and used to inoculate 30 ml tubes of LB + 50 µg/ml Amp broth, which was grown overnight at 37°C. The 30 ml of culture was then used to inoculate 500 ml of LB + Amp. Once the 500 ml cultures reached an optical density of 0.8 under 600 nm UV light, transcription of the VCLCL-MMP gene was induced with the addition of 1 mM of Isopropyl β -D-1-thiogalactopyranoside (IPTG) and dropping the incubation temperature to 18°C. The protein was expressed overnight and the cells were pelleted by centrifuging at 8,000 rpm and 4°C for 30 minutes and the LB + Amp discarded. The cells were then re-suspended in 25 ml lysis buffer (20

mM Na₃PO₄ and 500 mM NaCl, pH 7.4) and sonicated on ice at 30% amplitude with a 10 sec on 20 sec off pattern for 5 min. The cell lysate was centrifuged in a Sorvall RC6+ refrigerated superspeed centrifuge (Thermal Scientific, Amarillo, TX) at 10,000 rpm (using an F21-8×50Y rotor) for 20 min at 4°C and the pellet discarded. The supernatant was applied to a Ni-NTA agarose column (25 ml) at room temperature. After washing the column with wash buffer (20 mM Na₃PO₄, 500 mM NaCl, 20 mM imidazole, pH 7.4), buffers with increasing concentrations of imidazole (50 mM, 100 mM, 125 mM, and 400 mM) were used to stepwise elute the proteins. The purity was determined by SDS-PAGE and the concentration was measured by absorbance at 280 nm with $\epsilon = 9129 \text{ MM}^{-1}\text{cm}^{-1}$. The purified VCLCL-MMP was then dialyzed into either PBS or TN buffer (20 mM Tris, 150 mM NaCl, pH 7.5) for further experimentation.

2.2. Circular dichroism (CD) spectroscopy

CD data were obtained using an AVIV Model 420 spectropolarimeter (Aviv Associates Inc., Lakewood, NJ). Before CD scanning, proteins were kept in 1 mm cuvettes at 4°C for at least 24 hours. CD spectra were collected from 195 nm to 260 nm at 0.5 nm intervals with an averaging time of 5 seconds, and each scan was repeated three times to obtain an average. The thermal transition was obtained by monitoring the CD signal at 220 nm with increasing temperature from 0°C to 60°C in 0.33°C steps. Proteins were equilibrated at each temperature point for 2 min and the temperature was increased with an average rate of 0.1 °C/min.

2.3 MMP-1 digestion

ProMMP-1 was obtained as previously described (Yu et al., 2012). The substrates were dialyzed against TN buffer. CaCl₂ was added to 10 mM, and the substrates (20 μM) were incubated with 0.05% Brij 35 and 10 nM MMP-1 at 25 °C for 15, 30, 60, 180 min and overnight (~18 hr). The reaction was stopped by adding EDTA to a final concentration of 25 mM. The samples were run on SDS-PAGE (NuPAGE 4-12% Bis-Tris gel, ThermoFisher), and stained with Coomassie Brilliant Blue R250. The digested samples were also analyzed using a Bruker MicroFlex MALDI-TOF mass spectrometer (Billerica, MA) which verified the cleavage site occurred within the inserted human collagen sequence.

The SDS-PAGE gels were scanned and analyzed in ImageJ (<https://imagej.nih.gov/ij/index.html>) to quantitate collagen degradation. The densities of the top protein bands on each lane of the gels were measured using the Analyze→Gel function of the software. Peaks were integrated to calculate the peak area, and the fraction cleaved was calculated as: $(A_0 - A_t)/A_0$, where A_0 was the density of uncleaved protein at time 0 and A_t was the density of the uncleaved protein at certain time point t . During trypsin and chymotrypsin digestion, the V domain of the VCLCL was cleaved almost instantly, yielding the smaller sized CLCL; the molecular weight difference (10 kDa) was used to normalize the density value between A_0 and A_t .

To obtain the best comparisons for the MMP-1 cleavage of the recombinant proteins containing 6 triplets of α1(I), α2(I), α1(II) or α1(III) human chains, data were collected for the 4 recombinant protein samples at 5 time points, and all 20 samples were run on the same

SDS-PAGE, which was then scanned and analyzed. Because of this experimental design, it was not possible to carry out the digestion in triplicate for each sample, but the experiments were repeated multiple times. The trends among the four protein samples were always the same. However, different experiments could not be averaged to obtain error bars, because of the variations inherent in SDS-PAGE, as well as the use of different protein preparations, and different MMP-1 preparations (in which enzymatic activity decreased with time).

2.4 Trypsin and chymotrypsin digestion

Purified bacterial collagens (20 μ M) in TN buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.5) were incubated with 0.01 mg/ml (430 nM) trypsin at 25°C for 1, 2, 4, 6 hr and overnight (~18 hr). The reaction was stopped by adding phenylmethylsulphonyl fluoride (PMSF) to 1 mM and cleavage was analyzed using SDS-PAGE, followed by gel scanning.

For chymotrypsin digestion assays, the proteins were dialyzed into TN buffer. Digestion was performed using a 1:20 enzyme to protein weight ratio. For each recombinant protein, 20 μ l 1 mg/ml collagen was mixed with 2 μ l chymotrypsin (0.5 mg/ml) in the Tris-NaCl buffer and 0.5 μ l 1M CaCl₂, for a total reaction volume of 50 μ l. The assay was performed at 25°C in a temperature-controlled water bath. 10 μ l was removed from the reaction mixture at 0.5, 1, 2, 3 hr as well as an overnight (~18 hr) time point and neutralized with 1 μ l of 100 mM phenylmethylsulfonyl fluoride (PMSF) in 4 μ l of LDS sample loading buffer and 1.6 μ l of DTT reducing agent. Cleavage was visualized using SDS-PAGE, followed by gel scanning.

2.5 Molecular dynamics simulations of MMP cleavage site

Molecular dynamics (MD) simulations were performed for triple-helical structures containing the collagen sequences used for the experimental studies. The sequences of six triplets (residues 769–786) were retrieved from the UniProt database from the human α 1(I) (P02452), α 2(I) (P08123), and α 1(II) (P02458) chains and residues 775–792 from the α 1(III) (P02461) chain (Apweiler et al., 2004); each was flanked by four triplets from the bacterial collagen insertion site (underlined below) and three GPO stabilizing triplets at both ends. The sequences used in the wild-type (WT) simulations are:

- α 1(I): (GPO)₃-GKDGKDGQPGKP-GTPGPQGIAGQRGVVGLP-GPRGEQGPTGPT-(GPO)₃
- α 2(I): (GPO)₃-GKDGKDGQPGKP-GTPGPQGLLGAPGILGLP-GPRGEQGPTGPT-(GPO)₃
- α 1(II): (GPO)₃-GKDGKDGQPGKP-GPPGPQGLAGQRGIVGLP-GPRGEQGPTGPT-(GPO)₃
- α 1(III): (GPO)₃-GKDGKDGQPGKP-GAPGPLGIAGITGARGLA-GPRGEQGPTGPT-(GPO)₃

The initial triple-helical structures for the WT collagen peptides were built using the Triple-Helical collagen Building Script (THeBuScr) (Rainey and Goh, 2004). All peptide chains were capped with an acetyl group in the N-terminus and with an NH₂ group in C-terminus. In addition to the native collagen sequences, Gly to Ser substitutions in the MMP-1 cleavage

region of the $\alpha 2(I)$ chain homotrimers were also constructed using the UCSF Chimera package (Pettersen et al., 2004) to make residue substitutions.

All simulations were performed with GROMACS 4.6.7 (Hess et al., 2008) using the GROMOS 54a7 force field (Schmid et al., 2011) and SPC water model (van Gunsteren et al., 1983). For each triple helix, the starting structure was energy-minimized in vacuum using the steepest descent algorithm with an initial step size of 0.01 nm and a force tolerance of 1.0 kJ/mol/nm for a maximum of 2,000 steps. All heavy atoms of the triple helix were position restrained with a force constant of 1,000 kJ/mol/nm² and no bonds were constrained. Periodic boundary conditions were not applied in this step. Coulombic and van der Waals cutoffs were implemented, both of 999.0 nm.

Next, the structure was solvated in a rectangular box of water such that no peptide atom was closer than 1.5 nm to the edges of the box and the long axis of the triple helix was parallel to the z -axis of the box. The system was then further energy-minimized using the steepest descent algorithm with an initial step size of 0.01 nm and a force tolerance of 1.0 kJ/mol/nm for a maximum of 5,000 steps with flexible water and no position restraints or bond constraints. Periodic boundary conditions in x , y , and z directions were applied. Particle-Mesh Ewald (PME) electrostatics (Essmann et al., 1995) was used with a short-range cutoff of 0.8 nm. The Fourier spacing was set to 0.12 nm and the cubic interpolation for PME was implemented. The relative strength of the Ewald-shifted direct potential at the short-range electrostatic cutoff was 1×10^{-5} . A cutoff of 0.8 nm was used for the van der Waals interactions and long-range dispersion correction for energy was applied.

Next, in the first stage of equilibration, with all heavy atoms of the triple helix position-restrained with a force constant of 1000 kJ/mol/nm² and using the rigid water model, the system was annealed from 5 K to 300 K over 20 ps. The system was then equilibrated at 300 K for 30 ps. Both steps were performed with a 2 fs time step using the leap-frog algorithm in an NVT ensemble with a v -rescale thermostat using a time constant of 0.1 ps and one temperature-coupling group (i.e. the whole system). Initial velocities were generated at 5 K using a random seed. Then, the system underwent the second stage of equilibration using the leap-frog algorithm for 500 ps with a 2 fs time step. The triple-helix backbone atoms were position-restrained with a force constant of 1000 kJ/mol/nm² and the rigid water model was used. Equilibration was done in an NPT ensemble. The Berendsen barostat was used with a time constant of 2.0 ps and a reference pressure of 1.0 bar; pressure was isotropically coupled with isothermal compressibility of water at $4.5 \times 10^{-5} \text{ bar}^{-1}$. The Nose-Hoover thermostat was implemented with two temperature-coupling groups of protein and non-protein (Cheng and Merz, 1996; Hoover, 1985; Nose, 1984). For both protein and non-protein, a time constant of 1.0 ps and a reference temperature of 300 K was used. For both these equilibrations and the production run, all protein bond lengths were constrained using the fourth order expansion of the LINCS algorithm (Hess et al., 1997). For the rigid water, its geometry was constrained using SETTLE. Periodic boundary conditions in x , y , and z directions were applied. PME was used for electrostatics; a cutoff of 0.8 nm was used for the van der Waals interactions and long-range dispersion correction for energy was applied.

Lastly, a 200 ns production run using the leap-frog algorithm with a 2 fs time step was performed. Only the first and last C α atoms on each chain of the triple helix (six atoms total) were position-restrained with a force constant of 10 kJ/mol/nm² and the rigid water model was used. Production was done using an NPT ensemble at 1.0 bar and 300 K.

3. Results

3.1 MMP-1 cleavage of chimeric bacterial collagen constructs containing the collagenase cleavage sites from human collagen chains

Six triplets surrounding the collagenase cleavage sites from chains of the major fibrillar collagens, $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$, and $\alpha 1(III)$ were inserted within the triple-helix domain of the bacterial collagen ScI2 protein, as illustrated Figure 1A. The recombinant proteins, designated as VCLCL- $\alpha 1(I)$ MMP, VCLCL- $\alpha 2(I)$ MMP, VCLCL- $\alpha 1(II)$ MMP and VCLCL- $\alpha 1(III)$ MMP were expressed in *E. coli* at 22°C, and the N-terminal His-tagged proteins purified through a Ni-NTA column. Circular dichroism (CD) spectroscopy confirmed the triple-helix nature of the expressed proteins, giving a maximum at 220 nm and a minimum near 198 nm (Figure 1B). Monitoring of the CD maximum at 220 nm with increasing temperature gave a sharp thermal transition for all purified recombinant proteins, with a melting temperature of around 36°C (Figure 1C), similar to that seen for the VCLCL control (Yoshizumi et al., 2009). An additional set of four recombinant collagen proteins was generated containing only four triplets from the collagenase cleavage site (Supplementary Figure 1A). Similar to their six-triplet counterparts, VCLCL constructs with four-triplet insertion also formed stable triple helices with similar melting temperatures (data not shown).

As previously reported, the wild-type bacterial collagen VCLCL showed no cleavage by MMP-1, even at high concentrations and long incubations (Yu et al., 2012). The chimeric recombinant proteins with an insertion of four triplets from human collagen also showed no MMP-1 cleavage (Supplementary Figure 1B). In contrast, native homotrimeric proteins with an insertion of six human collagen triplets at the collagenase cleavage site all showed cleavage within the inserted human sequence, converting the original protein (57 kDa) to 33-kDa and 23-kDa fragments corresponding to the VCL and CL fragments (Figure 2A). MMP-1 cleaved homotrimers with 6-triplet insertions from human $\alpha 1(II)$, $\alpha 1(III)$, or $\alpha 2(I)$ chains at a similar rate, with less than 20% of the intact VCLCL protein remaining after an overnight incubation (Figure 2B). Plotting of initial rate vs. substrate concentration allowed estimates of the kinetic parameters, giving $K_m = 63 \mu M$ for $\alpha 2(I)$. The homotrimer construct with an $\alpha 1(I)$ inclusion was cleaved by MMP-1 at a slower rate compared with the other three constructs; after an overnight incubation, a significant amount (>50%) of the original construct remained in multiple repeated experiments (Figure 2B, purple curve). Because of the slow rate of digestion, it was not possible to obtain meaningful kinetic parameters for this protein as high concentrations of collagen substrates, which are needed for data points for a Michaelis–Menten plot, cause the protein to aggregate. It is possible that the early fast cleavage occurs on a small amount of denatured $\alpha 1(I)$ -containing homotrimer collagen, which would be highly susceptible to MMP-1 cleavage. The expressed proteins studied here contain Pro, rather than Hyp in the Yaa position of the Gly-Xaa-Yaa sequence, because of

the absence of prolyl hydroxylase in the bacterial expression system, but that this does not appear to influence MMP-1 cleavage (Yu et al., 2012).

The tightness of the triple helix at the MMP cleavage site was explored experimentally through trypsin digestion. It is known that the native tightly folded triple helix is resistant to non-collagenase enzymes such as trypsin, while denatured collagen is highly susceptible (Bruckner and Prockop, 1981; Kang et al., 2001). The control VCLCL construct is largely trypsin resistant after an overnight digestion, but there is a small fraction digested immediately (within 2 min.) that may represent collagen denatured during the preparation and purification (Yu et al., 2012). The $\alpha 1(\text{II})$ and $\alpha 1(\text{III})$ -containing homotrimer recombinant proteins show a moderate susceptibility to trypsin, with the intact chain completely degraded after an overnight incubation (Figure 3, red and blue), suggesting a greater degree of looseness in the inserted MMP cleavage site. The $\alpha 1(\text{I})$ homotrimer shows an initial small loss of intensity, which may correspond to a fraction of denatured collagen, but then little degradation, even after an overnight incubation in trypsin (Figure 3, purple), suggesting a tighter triple helix than seen for the type II and III homotrimers. The $\alpha 2(\text{I})$ chain insert in the recombinant collagens lacks the Arg or Lys residues required for trypsin cleavage (Figure 1A), which is consistent with the low susceptibility of VCLCL- $\alpha 2(\text{I})$ MMP homotrimers (Figure 3, green). The rapid digestion by trypsin of a small amount of $\alpha 2(\text{I})$ homotrimers again may reflect a small amount of denatured collagen in the preparation. The sequence of the $\alpha 2(\text{I})$ insert contains hydrophobic residues which are susceptible to chymotrypsin digestion, and $\alpha 2(\text{I})$ homotrimers were completely digested by chymotrypsin within several hours (see Figure 5B), supporting a looseness in its MMP region.

3.2 Molecular dynamics simulations of homotrimers containing the collagenase cleavage region

To correlate the structural properties of the triple helices with their enzyme susceptibility, molecular dynamics (MD) simulations were carried out on model triple helices with the sequences of the six triplets in the MMP cleavage region. The $\alpha 1(\text{I})$ homotrimer showed a high level of interchain peptide hydrogen bonding (Gly NH $\cdots$ OC Xaa) at all times during the simulation (Figure 4A, top left). A very stable triple helix with regular hydrogen bonding along the whole sequence is consistent with the high trypsin resistance observed for the $\alpha 1(\text{I})$ homotrimer. The $\alpha 2(\text{I})$ homotrimers showed a somewhat lower average hydrogen bond occupancy as well as more variability in the percent of hydrogen bonding during the simulation (Figure 4A, top right). During the simulations of the $\alpha 2(\text{I})$ homotrimers, there were several deep drops of the average interchain hydrogen bond occupancy with local minima approaching approximately 0.60 in the MMP cleavage region, suggesting that as many as 4–5 hydrogen bonds in this region may be transiently disrupted. This transient unfolding occurs largely C-terminal to the cleavage site, where the hemopexin domain of MMP-1 is known to bind (Manka et al., 2012).

MD simulations of the homotrimers indicated average hydrogen bond occupancies of 0.92 in the $\alpha 1(\text{I})$ homotrimer, 0.88 in the $\alpha 1(\text{III})$ homotrimer, 0.76 in the $\alpha 2(\text{I})$ homotrimer, and 0.67 in the $\alpha 1(\text{II})$ homotrimer (Figure 4B). The highest hydrogen bond occupancy seen for the $\alpha 1(\text{I})$ homotrimer is consistent with its increased trypsin resistance and increased

MMP-1 resistance compared with other constructs. The order of the average hydrogen bond occupancy, $\alpha 1(I) > \alpha 1(III) > \alpha 2(I) > \alpha 1(II)$, models the order of MMP-1 cleavage rates seen in Figure 2B.

Even though the $\alpha 1(I)$ and $\alpha 1(II)$ sequences differ in only three residues within the six triplets of the MMP-1 cleavage region (Figure 1A), the $\alpha 1(I)$ homotrimer is more resistant to trypsin and MMP-1 cleavage than the $\alpha 1(II)$ homotrimer. MD simulations of the $\alpha 1(II)_3$ homotrimer indicated a lower average hydrogen bond occupancy than the $\alpha 1(I)_3$ homotrimer (Figure 4B, red vs. purple), as well as dips in the occupancy throughout simulation time (Figure 4A, bottom left), indicative of transient hydrogen bond breaking and reforming. MD simulations on hybrid $\alpha 1(I)/\alpha 1(II)$ homotrimers where either an I776L or V782I substitution from the $\alpha 1(II)$ sequence was introduced within the $\alpha 1(I)$ sequence showed a substantial disruptive influence of the I776L replacement, while the V782I substitution had minimal impact on hydrogen bond occupancy (Supplementary Figure 2). This was further supported when the relative stability of the triplets in the sequence was analyzed (Persikov et al., 2005) as the GIA triplet of the $\alpha 1(I)$ sequence was more stable than the GLA triplet of the $\alpha 1(II)$ sequence. However, the loss of hydrogen bonding due to the I776L substitution is considerably less than that seen for the $\alpha 1(II)$ homotrimer compared with the $\alpha 1(I)$ homotrimer, suggesting other cooperative effects.

Experimental studies using the recombinant bacterial system were carried out to investigate the effect of a Gly to Ser replacement within the inserted $\alpha 2(I)$ collagenase cleavage region on triple-helix structure, and on MMP-1 digestion and trypsin susceptibility. The six triplets included in the construct for the human $\alpha 2(I)$ chain are ⁷⁶⁹GTP-GPQ-GLL-GAP-GIL-GLP⁷⁸⁶ (see Figure 1A). Previous studies indicated that the N-terminal triplet is not required for cleavage (Yu et al., 2012), and five recombinant constructs were designed to each include one Gly to Ser replacement at each of the other five sites (residues G772, G775, G778, G781, and G784) within the inserted sequence. Only three recombinant proteins with mutations were successfully expressed and purified, representing those with Ser at positions 772, 778, and 784. For Gly to Ser mutations at the other two sites (775 and 781), the recombinant proteins were heavily truncated and appeared to be unstable, so that not enough protein could be purified for further study. The recombinant proteins with Gly to Ser mutations were designed, expressed and purified as described above, and again showed the presence of a predominantly triple-helix structure in CD spectra (data not shown). The replacement of a Gly by a Ser within the (Gly-Xaa-Yaa)_n repeating sequence is known to perturb and destabilize the triple helix in the context of relatively short peptides (Beck et al., 2000; Bryan et al., 2011; Xiao et al., 2011), and is expected to result in a local perturbation within larger collagen molecules. Such a local disruption was probed enzymatically, using chymotrypsin, since the inserted six triplets of the $\alpha 2(I)$ chain do not have trypsin cleavage sites, as noted above. The native $\alpha 2(I)$ chain was quite susceptible to chymotrypsin, and the cleavage was even faster when Gly to Ser mutations were introduced (Figure 5B). To investigate the effect of the disruption or loosening of the triple helix on collagenase activity, MMP-1 cleavage was compared for the control and the three mutant collagens. Similar or slightly increased MMP-1 digestion was observed in the presence of the mutations (Figure 5A).

MD studies were also carried out on the same mutant sequences. The introduction of a Gly to Ser mutation at G772, G778, and G784 within the six-triplet human $\alpha 2(I)$ sequence decreased the average hydrogen bond occupancy significantly from 0.76 in the wild-type to 0.50, 0.43, and 0.64 for each mutation, respectively (Figure 5C) and hydrogen bonds were particularly disrupted C-terminal to the mutation (Supplementary Figure 3). These observations in the MD simulations are consistent with the higher MMP-1 digestion and chymotrypsin susceptibility for these mutants compared to the wild-type sequence (Figures 5B–C).

4. Discussion

The high-resolution structure of a collagen triple-helical peptide in complex with MMP-1 was reported recently (Manka et al., 2012), but the mechanism of collagenase cleavage of native triple-helical collagen is still not fully understood. The sequence around the unique collagenase cleavage site in fibrillar collagens is unusually low in imino acids and high in hydrophobic residues (Fields, 1991), which may mediate the localized unwinding of the triple helix required prior to MMP-1 cleavage (Chung et al., 2004). The ability to incorporate the collagenase sensitive region of human fibrillar collagen $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$, and $\alpha 1(III)$ chains within a bacterial collagen construct allows monitoring and manipulation of the cleavage process. Experimental results on these recombinant proteins are correlated here with MD simulations that report the triple-helix hydrogen bonding.

An earlier report on recombinant bacterial collagen with an inserted type III collagen sequence showed that six triplets around the cleavage sites in human collagens are sufficient to confer enzyme susceptibility, while four triplets are not (Yu et al., 2012). Here, this is extended to show that four triplets from the $\alpha 1(I)$, $\alpha 2(I)$, and $\alpha 1(II)$ collagenase cleavage site do not result in cleavage, while digestion does occur when six triplets are inserted

The $\alpha 1(I)$ homotrimer bacterial chimeric construct shows decreased MMP-1 cleavage compared with the $\alpha 1(II)$ and $\alpha 1(III)$ homotrimer constructs. This slow cleavage is consistent with the observation that naturally occurring $\alpha 1(I)$ homotrimers are very resistant to MMP-1 cleavage (Han et al., 2010). MD studies of $\alpha 1(I)$ homotrimer show a very high percentage of hydrogen bonding throughout the simulations, suggesting a high degree of rigidity within the cleavage region which could relate to its MMP-1 resistance. Compared to the $\alpha 1(I)$ homotrimer, the $\alpha 1(II)$ and $\alpha 1(III)$ homotrimers exhibit smaller hydrogen bond occupancy, consistent with the higher MMP-1 and trypsin susceptibilities observed experimentally.

For the first time, an $\alpha 2(I)$ sequence in the collagenase cleavage region has been studied as a homotrimer. In contrast to naturally occurring $\alpha 1(I)$ homotrimers, $\alpha 2(I)$ homotrimers have never been observed in a natural or pathological setting (Han et al., 2010). While purified $\alpha 1(I)$ chains will renature to form a stable homotrimeric triple helix, isolated $\alpha 2(I)$ chains are not capable of such renaturation, perhaps due to their overall lower imino acid content (Tkocz and Kuhn, 1969). The cleavage of $\alpha 2(I)$ within the native type I collagen heterotrimer by collagenase has been shown to depend on the presence of a cleavable $\alpha 1(I)$ chain; wild-type $\alpha 2(I)$ chains present in heterotrimers with mutant non-susceptible $\alpha 1(I)$

chains were not digested by MMP-1 in transgenic mice studies (Wu et al., 1990). The efficient MMP-1 cleavage of $\alpha 2(I)$ homotrimers reported here suggests that the presence of $\alpha 1(I)$ chain is not essential for $\alpha 2(I)$ cleavage within a triple helix. The transient breaking of hydrogen bond in the MMP cleavage region may help explain the effective MMP-1 cleavage of $\alpha 2(I)$ homotrimers. These results are consistent with a previous computational study on the type I heterotrimer which found that the $\alpha 2(I)$ chain near the MMP cleavage site has energetically preferred local unfolding relative to the $\alpha 1(I)$ chains (Nerenberg and Stultz, 2008).

MD simulations to investigate the hydrogen bonding in the native type I heterotrimer were undertaken, but were limited by the uncertainty about the chain register (Bella, 2016). Studies have led to conflicting predictions of the register of the $\alpha 1(I)$, $\alpha 1(I)$ and $\alpha 2(I)$ chains (Bender et al., 1982; Brondijk et al., 2012; Hofmann et al., 1978; Orgel et al., 2006), and it has not been shown that all molecules have one unique register. Our MD simulations indicated that the register of the three chains had a marked effect on the preservation of hydrogen bonding within the MMP cleavage region, varying from high hydrogen bonding occupancy in $\alpha 1(I)\alpha 1(I)\alpha 2(I)$ to much lower occupancy at time points in the simulation of $\alpha 2(I)\alpha 1(I)\alpha 1(I)$ (Supplementary Figure 4). As a result, it was not possible to compare our $\alpha 1(I)$ homotrimer or $\alpha 2(I)$ homotrimer simulation results with that for heterotrimeric type I collagen.

The concept that flexibility/mobility is needed in the triple helix to mediate MMP-1 cleavage was investigated in previous studies by increasing the rigidity through the replacement of residues near the scissile bond by Pro residues (Williams and Olsen, 2009; Wu et al., 1990). An altered $\alpha 1(I)$ chain with Ile776Pro or Gln774Pro/Ala777Pro mutations introduced into a murine fibroblast cell line prevented collagenase cleavage (Wu et al., 1990). Similar Pro replacements were introduced near the cleavage site of human $\alpha 1(III)$ collagen expression in a yeast expression system, and again the increased rigidity eliminated collagenase cleavage (Williams and Olsen, 2009). The mutations described here perturb the triple helix in the opposite direction. Rather than increasing rigidity, the replacement of Gly by Ser has been shown to disrupt the triple helix, dramatically decreasing the stability of peptides (Beck et al., 2000; Bryan et al., 2011), and leading to a small decrease in global stability within recombinant bacterial collagen (Cheng et al., 2010; Chhum et al., 2016) and OI collagens (Makareeva et al., 2008). MD simulations confirm the local unfolding effect of the Gly to Ser replacements in the system studied here. This unfolding is also experimentally seen in the increased chymotrypsin susceptibility of recombinant proteins with $\alpha 2(I)$ chain mutations. In contrast to Pro replacements which eliminated collagenase susceptibility, the Gly to Ser replacement allowed MMP-1 cleavage similar or slightly increased to the control collagen substrate. This supports a mechanism where flexibility of the triple helix is important and indicates that increasing the flexibility beyond its normal amount does not affect the cleavage.

The bacterial recombinant collagens studied here provide new insights into the cleavage of $\alpha 2(I)$ chains by collagenases. In addition, the ability to create artificial homotrimeric collagenase cleavage sites and to introduce variants such as Gly mutations present an opportunity to modulate MMP activity for custom biomaterials purposes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grants #EB011620 (to BB and DLK) and #GM60048 (to BB) and by a Tufts start-up fund (to Y.-S. L.). We thank Dr. David Wilbur from Tufts Chemistry for allowing us to access the MALDI-TOF MS equipment and Dr. Yimin Qiu for helpful discussions. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the NIH.

References

- Amar S, Smith L, Fields GB. Matrix metalloproteinase collagenolysis in health and disease. *Biochim Biophys Acta*. 2017; 1864:1940–1951. [PubMed: 28456643]
- Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS. UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res*. 2004; 32:D115–119. [PubMed: 14681372]
- Beck K, Chan VC, Shenoy N, Kirkpatrick A, Ramshaw JAM, Brodsky B. Destabilization of osteogenesis imperfecta collagen-like model peptides correlates with the identity of the residue replacing glycine. *Proc Natl Acad Sci USA*. 2000; 97:4273–4278. [PubMed: 10725403]
- Bella J. Collagen structure: new tricks from a very old dog. *Biochem J*. 2016; 473:1001–1025. [PubMed: 27060106]
- Bender E, Silver FH, Hayashi K, Trelstad RL. Type-I Collagen Segment Long Spacing Banding-Patterns - Evidence That the Alpha-2 Chain Is in the Reference or a Position. *J Biol Chem*. 1982; 257:9653–9657. [PubMed: 6179935]
- Bode W. A helping hand for collagenases: the haemopexin-like domain. *Structure*. 1995; 3:527–530. [PubMed: 8590012]
- Brondijk THC, Bihan D, Farndale RW, Huizinga EG. Implications for collagen I chain registry from the structure of the collagen von Willebrand factor A3 domain complex. *Proc Natl Acad Sci USA*. 2012; 109:5253–5258. [PubMed: 22440751]
- Bruckner P, Prockop DJ. Proteolytic-Enzymes as Probes for the Triple-Helical Conformation of Procollagen. *Anal Biochem*. 1981; 110:360–368. [PubMed: 7015914]
- Bryan MA, Cheng HM, Brodsky B. Sequence Environment of Mutation Affects Stability and Folding in Collagen Model Peptides of Osteogenesis Imperfecta. *Biopolymers*. 2011; 96:4–13. [PubMed: 20235194]
- Cheng AL, Merz KM. Application of the Nose-Hoover chain algorithm to the study of protein dynamics. *J Phys Chem*. 1996; 100:1927–1937.
- Cheng HM, Rashid S, Yu ZX, Yoshizumi A, Brodsky B. Investigation of Collagen Glycine Substitution Mutations Leading to Disease in a Bacteria Collagen System and Collagen Like Peptides. *Biophys J*. 2010; 98:446a–446a.
- Chhum P, Yu HT, An B, Doyon BR, Lin YS, Brodsky B. Consequences of Glycine Mutations in the Fibronectin-binding Sequence of Collagen. *J Biol Chem*. 2016; 291:27073–27086. [PubMed: 27799304]
- Chung LD, Dinakarandian D, Yoshida N, Lauer-Fields JL, Fields GB, Visse R, Nagase H. Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. *EMBO J*. 2004; 23:3020–3030. [PubMed: 15257288]
- Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A Smooth Particle Mesh Ewald Method. *J Chem Phys*. 1995; 103:8577–8593.
- Fields GB. A model for interstitial collagen catabolism by mammalian collagenases. *J Theor Biol*. 1991; 153:585–602. [PubMed: 1666905]

- Han S, Makareeva E, Kuznetsova NV, DeRidder AM, Sutter MB, Losert W, Phillips CL, Visse R, Nagase H, Leikin S. Molecular mechanism of type I collagen homotrimer resistance to mammalian collagenases. *J Biol Chem*. 2010a; 285:22276–22281. [PubMed: 20463013]
- Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINC: A linear constraint solver for molecular simulations. *J Comput Chem*. 1997; 18:1463–1472.
- Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput*. 2008; 4:435–447. [PubMed: 26620784]
- Hofmann H, Fietzek PP, Kuhn K. Role of Polar and Hydrophobic Interactions for Molecular Packing of Type-I Collagen - 3-Dimensional Evaluation of Amino-Acid Sequence. *J Mol Biol*. 1978; 125:137–165. [PubMed: 731689]
- Hoover WG. Canonical Dynamics - Equilibrium Phase-Space Distributions. *Physical Review A*. 1985; 31:1695–1697.
- Kang MS, Lim BK, Seong IS, Seol JH, Tanahashi N, Tanaka K, Chung CH. The ATP-dependent CodWX (HslVU) protease in *Bacillus subtilis* is an N-terminal serine protease. *EMBO J*. 2001; 20:734–742. [PubMed: 11179218]
- Makareeva E, Han S, Vera JC, Sackett DL, Holmbeck K, Phillips CL, Visse R, Nagase H, Leikin S. Carcinomas contain a matrix metalloproteinase-resistant isoform of type I collagen exerting selective support to invasion. *Cancer Res*. 2010; 70:4366–4374. [PubMed: 20460529]
- Makareeva E, Mertz EL, Kuznetsova NV, Sutter MB, DeRidder AM, Cabral WA, Barnes AM, McBride DJ, Marini JC, Leikin S. Structural heterogeneity of type I collagen triple helix and its role in osteogenesis imperfecta. *J Biol Chem*. 2008; 283:4787–4798. [PubMed: 18073209]
- Manka SW, Carafoli F, Visse R, Bihan D, Raynal N, Farndale RW, Murphy G, Englund JJ, Hohenester E, Nagase H. Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. *Proc Natl Acad Sci U S A*. 2012; 109:12461–12466. [PubMed: 22761315]
- Nagase H, Visse R, Murphy G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc Res*. 2006; 69:562–573. [PubMed: 16405877]
- Nerenberg PS, Stultz CM. Differential unfolding of alpha 1 and alpha 2 chains in type I collagen and collagenolysis. *J Mol Biol*. 2008; 382:246–256. [PubMed: 18644377]
- Nose S. A Molecular-Dynamics Method for Simulations in the Canonical Ensemble. *Mol Phys*. 1984; 52:255–268.
- Orgel JP, Perumal S, Antipova O, Irving TC. The molecular structure and arrangement of collagen type I. *Matrix Biol*. 2006; 25:S76–S76.
- Persikov AV, Ramshaw JA, Brodsky B. Prediction of collagen stability from amino acid sequence. *J Biol Chem*. 2005; 280:19343–19349. [PubMed: 15753081]
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004; 25:1605–1612. [PubMed: 15264254]
- Rainey JK, Goh MC. An interactive triple-helical collagen builder. *Bioinformatics*. 2004; 20:2458–2459. [PubMed: 15073022]
- Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, van Gunsteren WF. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Eur Biophys J*. 2011; 40:843–856. [PubMed: 21533652]
- Stultz CM. Localized unfolding of collagen explains collagenase cleavage near imino-poor sites. *J Mol Biol*. 2002; 319:997–1003. [PubMed: 12079342]
- Tkocz C, Kuhn K. Formation of Triple-Helical Collagen Molecules from Alpha1 or Alpha2 Polypeptide Chains. *Eur J Biochem*. 1969; 7:454. [PubMed: 5776239]
- van Gunsteren WF, Berendsen HJ, Hermans J, Hol WG, Postma JP. Computer simulation of the dynamics of hydrated protein crystals and its comparison with x-ray data. *Proc Natl Acad Sci U S A*. 1983; 80:4315–4319. [PubMed: 6576339]
- Williams KE, Olsen DR. Matrix metalloproteinase-1 cleavage site recognition and binding in full-length human type III collagen. *Matrix Biol*. 2009; 28:373–379. [PubMed: 19398000]

- Wu H, Byrne MH, Stacey A, Goldring MB, Birkhead JR, Jaenisch R, Krane SM. Generation of Collagenase-Resistant Collagen by Site-Directed Mutagenesis of Murine Pro-Alpha-1(I) Collagen Gene. *Proc Natl Acad Sci USA*. 1990; 87:5888–5892. [PubMed: 2165607]
- Xiao JX, Madhan B, Li YJ, Brodsky B, Baum J. Osteogenesis Imperfecta Model Peptides: Incorporation of Residues Replacing Gly within a Triple Helix Achieved by Renucleation and Local Flexibility. *Biophys J*. 2011; 101:449–458. [PubMed: 21767498]
- Yoshizumi A, Yu Z, Silva T, Thiagarajan G, Ramshaw JA, Inouye M, Brodsky B. Self-association of streptococcus pyogenes collagen-like constructs into higher order structures. *Protein Sci*. 2009; 18:1241–1251. [PubMed: 19472339]
- Yu Z, Visse R, Inouye M, Nagase H, Brodsky B. Defining requirements for collagenase cleavage in collagen type III using a bacterial collagen system. *J Biol Chem*. 2012; 287:22988–22997. [PubMed: 22573319]

Highlights

The type I collagen heterotrimeric triple helix is readily cleaved by MMP-1, but the homotrimer of the $\alpha 1(I)$ chain of type I collagen is only slowly cleaved. A recombinant bacterial collagen system was used to create homotrimers with the $\alpha 2(I)$ cleavage region which proved highly susceptible to collagenase. Experimental studies and molecular dynamics simulations showed extensive disruption of hydrogen bonding in $\alpha 2(I)$ homotrimers that correlated with their collagenase susceptibility.

Figure 1.

A. Schematic of the recombinant bacterial construct His-VCLCL, highlighting the insertion of six Gly-Xaa-Yaa triplets surrounding the collagenase cleavage site in human $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$, and $\alpha 1(III)$ collagen chains. **B.** CD spectrum of the four recombinant constructs shown in part A. **C.** CD thermal melting profiles of the four recombinant constructs in part A.

Figure 2.

MMP-1 digestion of recombinant bacterial constructs containing the six triplets from the collagenase cleavage region of human $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$, and $\alpha 1(III)$ chains. **A.** SDS-PAGE of the products after MMP-1 digestion for varying lengths of time (in hours; ON = overnight), showing the VCLCL recombinant collagens are cleaved within the inserted sequence to give VCL and CL fragments. **B.** Plot of the fraction of the original VCLCL band cleaved as a function of time.

Figure 3.

Trypsin digestion of recombinant bacterial constructs containing the six triplets from the collagenase cleavage region of human $\alpha 1(I)$, $\alpha 2(I)$, $\alpha 1(II)$, and $\alpha 1(III)$ chains. **A.** SDS-PAGE of the products after different digestion times (in hours; ON = overnight). **B.** Plot of the fraction of the original protein that was digested as a function of time. Note that the sequence designed between the V domain and the adjacent CL domain is very susceptible to trypsin (Yoshizumi et al., 2009), so that any exposure to trypsin cleaves VCLCL to CLCL. Local unfolding of the triple helix within the inserted collagenase cleavage sequence will lead to trypsin cleavage at that site and loss of the CLCL band.

Figure 4.

A. Interchain NH...CO hydrogen bonds in the $\alpha 1(I)_3$, $\alpha 2(I)_3$, $\alpha 1(II)_3$, and $\alpha 1(III)_3$ wild-type homotrimers. The human sequence is colored in black while the sequence from the bacterial VCLCL region is colored in grey. Hydrogen bonds in the localized region of interest are indicated by arrows and colored based on the hydrogen bond occupancy, ranging from red (0.0) to blue (1.0); hydrogen bonds not used in the calculations are colored in grey. A hydrogen bond is defined as donor acceptor distance <3.5 Å and hydrogen donor acceptor angle $<30^\circ$. The hydrogen bond occupancy shown here is calculated using the average of three 200 ns MD runs. The graph beneath each hydrogen bond schematic depicts the average occupancy of the interchain NH...CO hydrogen bonds in the systems over the simulation time for the runs, denoted as dashed, dotted, and dash-dotted lines; the average of the three

runs is shown in a solid line. A running average with a sliding window of 1 ns is used. **B.** Occupancy of interchain NH $\cdots$ CO hydrogen bonds of the $\alpha 1(I)_3$, $\alpha 2(I)_3$, $\alpha 1(II)_3$, and $\alpha 1(III)_3$ homotrimers. The average hydrogen bond occupancy is calculated as an average of three runs over 200 ns with sample standard deviation as an error bar. 3.3 Mutations within the $\alpha 2(I)$ MMP sequence

Figure 5.

Digestion of recombinant bacterial constructs with Gly to Ser mutations at sites within the inserted human collagen $\alpha 2(I)$ sequence (G772S, G778S, and G784S) by **A.** MMP-1 and **B.** chymotrypsin. **C.** Average occupancy of interchain NH...CO hydrogen bonds for $\alpha 2(I)$ homotrimers as well as for $\alpha 2(I)$ homotrimers containing the Gly to Ser mutations, calculated as an average of three runs over 200 ns with sample standard deviation as an error bar. Student's t-test was used to analyze statistically significant difference of average hydrogen bond occupancies of the mutant systems from the wild-type; ** for $p < 0.01$, *** for $p < 0.001$, **** for $p < 0.0001$