Formation of multimers of bacterial collagens through introduction of specific sites for oxidative cross-linking

Violet Stoichevska[1], Bo An[2], Yong Y. Peng[1], Sezin Yigit[2], Aditya V. Vashi[1], David L. Kaplan[2], Jerome A. Werkmeister[1], Geoff J. Dumsday[1] and John A.M. Ramshaw[1]

[1] CSIRO Manufacturing, Bayview Avenue, Clayton, VIC 3169, Australia.

[2] Department of Biomedical Engineering, Tufts University, Medford, MA 02155, USA

Abstract

A range of non-animal collagens has been described, derived from bacterial species, which form stable triple-helical structures without the need for secondary modification to include hydroxyproline in the sequence. The non-animal collagens studied to date are typically smaller than animal interstitial collagens, around one quarter the length and do not pack into large fibrillar aggregates like those that are formed by the major animal interstitial collagens. A consequence of this for biomedical products is that fabricated items, such as collagen sponges, are not as mechanically and dimensionally stable as those of animal collagens. In the present study, we examined the production of larger, polymeric forms of non-animal collagens through introduction of tyrosine and cysteine residues that can form selective cross-links through oxidation. These modifications allow the formation of larger aggregates of the non-animal collagens. When Tyr residues were incorporated, gels were obtained. And with Cys soluble aggregates were formed. These materials can be formed into sponges that are more stable than those formed without these modifications.

Keywords: recombinant collagen; bacterial collagen; cross-linking; disulfide bond; dityrosine bond

Introduction

Collagens are the major structural proteins in the extracellular matrices of animals. Collagens have been successful in a wide variety of biomedical applications, where they are safe and widely accepted.[1] Collagens are defined by a characteristic triple-helix structural motif that requires a $(Gly\text{-}Xaa\text{-}Yaa)_n$ repeating sequence.[2] The amino acids found in the Xaa and Yaa positions are frequently proline, where Pro in the Yaa position is post-translationally modified to hydroxyproline (Hyp), which enhances triple-helical stability.[2,3] In humans, a family of at least 28 collagen types is present, each with type-specific biological and structural functions.[4] The most abundant of these collagens are the interstitial, fibril-forming collagens, particularly type I collagen, which is the main collagen used in biomedical materials.[1]

One concern about the use of collagen as a biomedical material has been the possible transmission of animal-derived diseases. However, the $(Gly\text{-}Xaa\text{-}Yaa)_n$ repeating sequence has recently been identified in a wide range of non-animal species, especially in bacterial species.[5,6] The triple-helical conformation has been identified in a limited number of examples,[6] but can be inferred for several others, especially those that contain around 30 or more contiguous (Gly-Xaa-Yaa) triplets. These non-animal collagens provide recombinant products as an alternative biomedical material that eliminates the potential risks of transmission of animal-derived diseases.

The most studied of these non-animal, collagen-like, triple-helical domains is the Scl2 gene from *Streptococcus pyogenes*.[7-13] This collagen is comprised of an N-terminal globular domain of 74 amino acids, called the V-domain, followed by an uninterrupted collagen, CL, domain of 78 Gly-X-Y triplets (234 amino acids), followed by an anchoring domain that is typically removed from the protein expression constructs.[9-11,13] Chemically, the CL-domain has no hydroxyproline, yet is still stable with a T_m of around 37 °C.[9]

Bacterial collagens could provide new options for biomedical materials.[8,11-12,14-17] These collagens could be readily produced in good yield through fermentation in *E. coli* as there is no requirement for proline hydroxylation,[13] and can then be purified in large scale without the need for complex chromatographic steps.[18] Further, the CL-domain of the *S. pyogenes* Scl2 gene has been shown to be non-cytotoxic and non-immunogenic.[11] Interestingly, this collagen seems to have few, if any, interaction sites for mammalian cells and binding proteins and so behaves as a 'blank slate'. This allows for design of specific molecules that include new functions, including for example, matrix metalloproteinase sites,[19] integrin sites,[12] a fibronectin binding site[20] and a heparin binding site.[21] These substitutions and insertions are not limited to one per molecule, but two or more can be introduced.[20-21]

Despite all these positive features, one difficulty has been reported. This was that freeze dried sponges of the collagen were not very stable when solution-based chemical crosslinking was directly attempted, with the sponge disintegrating and dispersing in the solution.[11] Vapour diffusion of glutaraldehyde allowed cross-linking,[11] but this approach is not easily controlled for different amounts of cross-linking. Stable hydrogels can be prepared by modification of the collagen before crosslinking. For example, this can be achieved by reaction with poly(ethylene glycol)-acrylate-*N*-hydroxysuccinimide groups along the triple-helical domain and photo-crosslinking along with poly(ethylene glycol) diacrylate.[12]

In the present study we have examined a genetic modification approach to generate stable collagen sponges. This was achieved through the modification of the *S. pyogenes* collagen-like gene by introduction of Tyr or Cys residues at either or both of the N- and C-terminals of the CL-domain. This approach leaves the CL-domain itself unmodified, allowing for new functions to be introduced, for example for integrin binding.[12] These modifications to the gene sequence introduced residues that are not normally found within the CL domain

of bacterial collagens. Both Tyr and Cys can readily be included in crosslink formation through oxidation, leading to dityrosine and disulfide bridge crosslinks, respectively. Locating these modifications at the ends of the collagen domains leads to specific end to end crosslinks, to make small aggregates or polymers of the collagen molecules.

Material and Methods

Gene constructs

The DNA sequence for the fragment of the *scl2.28* allele (Q8RLX7) encoding the combined globular (V-domain) and collagen-like (CL-domain) portions of the Scl2.28 protein, but lacking the C-terminal attachment domain was as previously described.[10-11] In addition, a sequence for a His_6-tag was introduced at the N-terminal of the sequence and an enzyme cleavage sequence, LVPRGSP, was inserted between the V and CL domains. This sequence is for thrombin digestion, but also allows digestion by trypsin, pepsin or papain.[18] In the present study pepsin cleavage was used. A triplet sequence Gly-Lys-Tyr was included at the C-terminal of the CL domain, followed by a stop codon, with *Nde*I and *Bam*HI restriction sites for cloning, giving the construct V-Cl-tyr. A further construct included an additional Tyr residue, in a Gly-Tyr-Pro sequence, between the N-terminal globular domain (V) and the following enzyme cleavage sequence, as well as one following $(Gly\text{-}Xaa\text{-}Yaa)_n$ collagen-like domain (CL) sequence, giving the construct V-tyr-CL-tyr. Other constructs included one with a Cys residue included at the C terminal of the CL domain, as a Gly-Tyr-Cys sequence, followed by a stop codon and restriction sites, construct V-Cl-cys, and also one with a second Cys residue between the N-terminal globular V- domain and the following enzyme cleavage sequence, as well as one following CL domain (CL), as an extra Gly-Cys-Pro triplet sequence, as construct V-cys-CL-cys (Figs. 1 and 2). These sequences were synthesised commercially with codon optimisation for *E. coli* expression (GeneArt®, Regensburg, Germany) or could also be obtained through site directed mutagenesis.

A further construct was based on the V-CLCL dimer structure previously described.[10] In this construct, a Gly-Cys-Pro sequence was added between the V and first CL domain, and a further Gly-Cys-Pro sequence was added after the second CL domain, giving construct V-

cys-CLCL-cys (Figs. 1 and 2). Both these changes were introduced using a site directed mutagenesis kit (QuikChange II, Agilent).

Expression

Each of the various gene constructs was cloned into *E. coli* and expressed as triple-helical protein using shake-flasks. DNA sequences were sub-cloned into the *E. coli* expression vector system pColdIII (Takara Bio, Shiga, Japan), ~~using the unique sites 5' *Nde*I and 3' *Bam*HI. PCR~~ followed by colony screening ~~was then used to detect positive clones. These clones were grown up in 100ml culture volumes~~ and Qiagen midi preps ~~carried out to expand the vector quantity~~ as previously described.[13,18,20] ~~For~~ Expression ~~of selected positive clone was transformed~~ in~~to the~~ *E. coli* host BL21-DE3~~. Cells were grown~~ in either 2 x YT ~~Media (16 g tryptone, 10 g yeast extract and 5 g NaCl per litre) with ampicillin (50 µg/ml)~~ or LB Media ~~(10 g tryptone, 5 g yeast extract and 10 g NaCl per litre) with ampicillin (1000 µg/ml) at 37 °C for 24 h and cell culture optical density at A600 reached around 3-4 (YT) or around 1 (LB).~~ was also as previously described.[13,18,20] The culture was then incubated at 25 °C and 1 mM isopropyl beta-D-thiogalactopyranoside (Sigma-Aldrich Co.) added to induce protein expression. After 10 h incubation at 25°C, the temperature was decreased to 15°C for a further 14 h incubation. After 24 h incubation, cells were harvested by centrifugation. All chemicals used in these and subsequent procedures were of the highest grade readily available.

Purification

For extraction~~, each 1 gram of~~ wet cell paste~~, derived from the above expression,~~ was extracted by sonication, ~~in 40 mM sodium phosphate buffer, pH8.0, 1 mM phenylmethanesulfonyl fluoride, using 20 ml buffer. The~~ the cell lysate mixture clarified by

centrifugation ~~and (20k x g for 40 min)~~ and the clear supernatant containing the triple-helical protein retained, as previously described.[13,21] Protein purification on ~~Clarified supernatant was taken to 20 mM sodium phosphate 300 mM NaCl and 30 mM imidazole buffer, pH8.5, and absorbed onto a~~ Ni charged HyperCel-Sepharose metal ion affinity resin (Pall Life Sciences) ~~Elution was by the same buffer, but containing 500 mM imidazole. Eluted fractions containing recombinant protein were pooled, concentrated and exchanged into 20 mM sodium phosphate buffer, pH8.0,~~ and concetration using ~~a 10 kDa~~ cross-flow filtration ~~membrane (Pall Life Sciences)~~ was also as previously described.[11,13] The V-domain was removed as required by 0.02 mg/ml pepsin, followed by further purification ~~in 20 mM sodium phosphate buffer, pH8.0, on~~ a Sephacryl S200 26/60 column (GE Healthcare).[11] Protein purity was assessed by SDS-PAGE,[22] with samples heated at 90 °C for 1 min and separated using NuPAGE (Invitrogen) 4-12% Bis-Tris gels with MES running gel buffer, at 180 V for 60 min, followed by staining with Coomassie Blue R-250.[13]

Formation of tyrosine-based, dityrosine crosslinks

A solution of purified protein, 1 mg/ml, containing one or more introduced Tyr residues, was adjusted to pH 7.2 in 20 mM sodium phosphate buffer, 0.15 M NaCl and 2mM $[Ru^{II}(bpy)_3]Cl_2$ (Sigma-Aldrich Co.) and between 10 - 50mM sodium persulfate (SPS) (Sigma-Aldrich Co.) added. Crosslinking was performed by irradiation for 3 times at 20 sec at a distance of 10 - 15 mm with a LED dental curing lamp (430-480 nm, peak wavelength 455nm ± 10nm, 1200 mW/cm^2 at source, 3M ESPE^{TM} S10 LED curing Light).[23] Changes in sample composition were examined by SDS-PAGE.[22]

Formation of cysteine—based, disulfide bond crosslinks

A solution of purified protein, containing introduced Cys residues, was adjusted to pH 7.2 in 20 mM sodium phosphate buffer, 0.15 M NaCl. Either air was bubbled though the solution for 30 mins, with the volume topped up with water as needed to maintain the initial volume, or alternatively, the sample was allowed to stand in air for up to 7 days. Changes in sample composition were examined by SDS-PAGE under non-reducing conditions.[22]

Mass spectroscopy

MALDI organic sample matrix was prepared by making a saturated sinapinic acid solution in 50% (v/v) acetonitrile and 0.3% (v/v) trifluoroacetic acid. A 6 µl aliquot of 1 µg/ml sample was mixed with 24 µl of organic sample matrix, 1 µl solution was plated onto a 96 spot target plate and allowed to dry. MALDI-TOF mass spectra were acquired on a Microflex LT system (Bruker Corporation, Billerica, MA) with 50% laser intensity using standard LP (linear positive) 60 kDa method provided by the software.[24]

Dynamic light scattering

Dynamic light scattering (DLS) measurements were performed using a DynaPro Titan instrument (Wyatt Technology Corp., Santa Barbara, CA) equipped with a temperature controller using Eppendorf UVette cuvettes with 10 mm path length. Protein concentrations of the samples (in PBS) were adjusted to 1 mg/ml. Samples were centrifuged at 14,000 x g for 10 min and filtered, when needed, through 0.2 µm Whatman Anotop 10 syringe filters before measurement. Samples were measured at 75% laser intensity. Twenty acquisitions were taken for every sample with each acquisition lasting 60 s. To obtain the hydrodynamic radius (Rh), the intensity autocorrelation functions were analyzed by Dynamic software.[24] Samples were reduced, when needed, by incubation with 20 mM tris(2-carboxyethyl)phosphine (TCEP) (Sigma-Aldrich Co.) for > 5 min.

~~***Scanning electron microscopy***~~

~~Scanning electron microscopy (SEM) was performed on freeze dried samples that had been iridium coated using a Polaron range SC5750 sputter coater and then imaged at 2 kV using a Philips XL30 field emission scanning electron microscope.~~[25]

Qualitative assessment of sponge stability

Various samples of crosslinked collagens were freeze dried, typically from 2-4 mg/ml solutions in 10 mM sodium phosphate, pH 7.2, to obtain collagen sponges. The stability of these sponges was assessed by immersion in either phosphate buffered saline and/or 50% EtOH, with limited agitation. The physical state of the sponges was observed after 5 min.

Results

Expression and purification

A selection of constructs was made containing various Tyr and Cys insertions (Figure 1). These insertions were typically made through addition of an extra Gly-Xaa-Yaa triplet into the parent *S. pyogenes* Scl2 monomer or dimer sequences, adjacent to the triple helical domain. Segments of the nucleic acid and protein sequences indicating the inserted sequences are shown in Figure 2. All constructs were readily expressed in *E. coli*, using the pCold vector system. The high yields were evident from SDS-PAGE of the initial expression, for example for the V-cys-CL-cys construct (Figure 3A). These various expressed proteins could subsequently be purified by IMAC chromatography (Figure 3), and the V-domain could be efficiently removed when required by proteolysis (Figure 3A, Figure 3B). The resistance of the CL domain to proteolysis, or CD spectroscopy[10] confirmed that

the CL-domains of the constructs were all triple-helical. However, on SDS-PAGE gels, the observed bands are single chains and not triple helical structures.

Tyrosine-based dityrosine crosslink formation

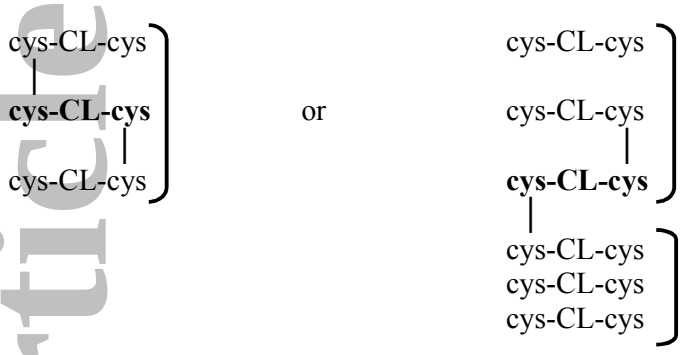
For constructs with a single Tyr insertion, for example V-CLCL-tyr, cross-linking was observed, with all material forming soluble higher molecular weight aggregates that do not enter the SDS-PAGE gel (Figure 4A). When the V-domain had been removed, for example in CL-tyr, again cross-linking occurs giving a soluble product, with the majority of material again being too large to enter the gel (Figure 4B). The extent of crosslinking will depend on the amount of sodium persulfate that is used.[25]

Similar crosslinking of the construct that had two introduced Tyr residues and which lacked the V-domain, tyr-CL-tyr, gave a noticeably different result (Figure 5). Crosslinking of this construct gave a stable gel, suggesting that substantial crosslinking to form a network had occurred (Figure 5B). Examination of samples by SDS-PAGE was not readily performed due to sample insolubility. ~~SEM examination of this material after freeze drying showed a denser surface layer and then open pockets with irregular surfaces that may support cell attachment if required (Figure 5C).~~

Cysteine-based disulfide bond crosslink formation

Oxidation of Cys residues to form disulfide bonds is readily achieved by exposure to air and without the need for addition of further reagents, as is required for the Tyr cross-linking. SDS-PAGE indicated that various cross-linked forms were being produced (Figure 6). For the construct, V-CL-cys, with or without the V domain, dimer and trimer formation could be intramolecular within a single triple helix molecule or intermolecular between two triple helices. For cys-CL-cys, and V-cys-CLCL-cys constructs, again dimers would most probably

arise from intramolecular crosslinking, but trimers could contain either all intramolecular or some intermolecular crosslinks such that, for example, either



could be occurring or many other potential permutations. Higher multimers will always involve one or more intermolecular crosslinks (Figure 6). Thus for the cys-CL-cys construct, standing in air for 1 week led to all the monomer chains forming disulfide bonds giving dimer, trimer and higher molecular weight components on SDS-PAGE (Figure 6A). For the V-cys-CLCL-cys construct, similar results were found. A fully reduced sample was not noticeably different in mobility to a V-CLCL construct that did not have the Cys insertions (Figure 6B). On standing for more than 24 h, oxidation to form disulfide bonds had commenced and higher molecular weight bands were present, although monomer chains were also still present. However, as with cys-CL-cys, after 1 week little, if any, monomer chains remained and all the chains were present as higher molecular weight components, including material that remained in the gel pocket (Figure 6B). Comparison with a sample after 1 week incubation in the same PBS, but with 1mM of H_2O_2 as an additive, showed no visible differences in the gel (data not shown).

MALDI-TOF mass spectra showed individual chains rather than trimeric molecules (Figure 7). For the cys-CL-cys construct (Figure 7A) monomer up to tetramer peaks, at molecular weights of 23608, 47250, 70878 and 94802 were observed, although the intensities dropped with increasing molecular weight. The observed molecular weights suggested that

the pepsin cleavage step had cleaved off the V-domain in the vicinity of the Ile-Gln bond prior to the Cys containing triplet. While dimer and trimer peaks could arise from either intra- or intermolecular crosslinking, the presence of a tetramer peak can only arise from intermolecular crosslinking. For the oxidised V-cys-CLCL-cys construct, 3 principal peaks were seen at molecular weights of 57482, 115062 and 173488 kDa, the calculated molecular weight for the single V-cys-CLCL-cys is 57641kDa, which is close to mass of the first peak representing the single chain, the second and third peaks with 2 and 3 times of the mass of a single chain represent the dimer and trimer chains, Again, the signal decreases significantly with the increase in molecular weight, such that higher multimers, if present, would not be detected. As noted above, the dimers and trimers could arise from either intramolecular or intermolecular cross-linking, but MALDI cannot distinguish this for the CL dimer constructs.

DLS does show that intramolecular and probably intermolecular crosslinking is occurring for the V-cys-CLCL-cys construct (Table 1). For oxidised material, the predominant component was a broad signal with a hydrodynamic radius of around 628 nm, indicating that a substantial proportion of polymeric material was present. After reduction by TCEP, this polymeric material substantially disappeared and the major component, by mass, was a narrow signal with a hydrodynamic radius of around 4 nm, with a small amount of residual higher molecular weight material, <10% of the mass (Table 1). The signal for the major reduced material is probably monomeric, triple helical molecules. In both these samples, the autocorrelation curves were good (data not shown). If samples were filtered through 0.2μM membrane (Whatman Anotop 10 Syringe Filter), then only low molecular size material was recovered in low yield, result in a random fluctuation and noisy autocorrelation curve which likely indicate insufficient amounts of analytes are present to

give a good DLS signal (data not shown). This suggests that intermolecular aggregates had been formed on oxidation.

Qualitative assessment of sponge stability

The various samples of freeze dried, crosslinked collagens that were examined, including both Tyr and Cys crosslinked materials, were all stable in either phosphate buffered saline and/or 50% EtOH test solutions, so long as there was not vigorous agitation.

Discussion

There has been concern that animal collagen-based biomedical products may allow transmission of various diseases, especially bovine spongiform encephalopathy ('mad cow disease'). One approach to resolve this problem has been the development of recombinant forms of animal collagens. Initially, yeast host systems were explored to produce human recombinant collagen in commercial quantities.[26] These systems are complicated as prolyl hydroxylation is needed to generate materials with adequate stability.[3] More recently, plant based expression system have also been developed that show promise for producing collagen in good yields.[27]

Recently, the discovery of a range of collagen-like molecules in bacterial species[5] that are stable at around 37°C, but without the need for Hyp has led to the suggestion that these proteins may be suitable for alternative collagen-based biomedical materials.[11-12] This bacterial system has the further advantage as a non-animal collagen that could be produced in a fully non-animal system if required. However, it was previously reported that sponges made from bacterial collagen lacked the stability of those made from animal collagens.[11] A possible reason for this difference is that bacterial collagens do not form fibrillar aggregates. Soluble animal collagens readily form fibrillar structures where the native staggered overlap structure is re-formed in vitro.[2,4] On the other hand, bacterial collagens do not form large fibrillar aggregate structures, and only limited association is observed.[10] A key reason for this striking differences maybe the lack of Hyp in the bacterial collagens. Previously it has been shown that unhydroxylated, triple-helical collagen I, produced as a recombinant product in transgenic tobacco, does not form into banded fibrils like native, hydroxylated collagens.[28] The previous observation of poor sponge cohesion was based on the finding that the bacterial collagen sponges rapidly dispersed when immersed in 50% EtOH, as a step to move to 100% EtOH for 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) initiated crosslinking.

While this does not constitute a quantitative test, it does allow empirical observations for sponge stability. In the present study, all freeze dried, crosslinked materials remained intact when immersed in 50% EtOH and/or phosphate buffered saline. The extent of the present crosslinking was limited, and could potentially allow rapid resorption of material in biomedical applications, for example in wound dressings. In some applications, however, additional crosslinking may be required. The stability of these present constructs is sufficient to allow standard solution crosslinking methods to be employed, such as with glutaraldehyde or EDC, for a range of extent of crosslinking, so that designed variations in resorption rates are possible. Previously, vapour phase crosslinking by glutaraldehyde had been used,[11] but this approach appears to go rapidly to completion and does not readily allow a graded series with different stabilities to be achieved (data not shown).

In the present study this extra physical stability for the sponges was achieved by introducing Gly-Xaa-Yaa collagen triplets containing either Tyr or Cys residues to the N- and C-terminal of the CL triple-helical domain. These locations allow site specific crosslink formation, and any cross-links do not interfere with any additional functionalities, such as integrin or fibronectin binding sites, that have been introduced into the structure.[12,20-21] Although the locations are the same, the effects of crosslinking are different for the Tyr and Cys insertions. For the Cys insertions, no gel formation was observed, but rather soluble intermolecular crosslinked aggregates were formed. The entanglements that these were able to form were sufficient to stabilise the freeze dried sponge material. Similar results were obtained with a single Tyr insertion. The present study used a photo-oxidation method that included the coloured $[Ru^{II}(bpy)_3]Cl_2$, [23] but it is possible that other approaches including a modified Fenton reaction or an enzymatic method[29] may also be suitable. When the V-domain was present it is possible that there was some involvement of the two Tyr residues from this domain. The tertiary structure of this domain from strain scl2.3 has been reported[30]

and the present strain scl2.28 has a very similar sequence and the same sequence length to that for scl2.3.[30] The structure shows that one Tyr, at position 52 in scl2.3, is buried in an internal α-helical, 3 coil bundle and would not be accessible for crosslink formation. The other Tyr, at position 45, is located at the junction of a partially unstructured turn and link sequence and the second outer α-helical coil, contributing to hydrophobic interactions, so is also unlikely to be in a conformation to enable involvement in crosslink formation. On the other hand, when Tyr residues were introduced at both the N- and C-terminals of the triple helical domain and the V-domain removed, a gel was formed indicating a more extensive network than had been obtained using constructs that had two Cys insertions.

The difference observed between the cys-CL-cys and tyr-CL-tyr constructs, with the Tyr construct forming a gel, could arise from the extent of crosslinking. However, the disulfide crosslink formation appeared essentially complete after 7 days in air, yet no gel was formed. It is more likely that steric considerations are important. Intramolecular crosslinks in homotrimer triple helices are readily formed, for example in native type III collagen,[30] or in triple helix forming peptidyl constructs,[32] so it is probable that intramolecular crosslinks are preferentially formed, leading to small aggregates an no extended network formation. For Tyr residues, the extent of cross-linking would be expected to be high. Thus, tyrosyl radicals can be generated by a variety of methods,[25,31-32] including the photo-crosslinking approach used in the present study. These radicals react poorly with oxygen, but combine readily to form dimers with 3,3-dityrosine formation,[33] leading to new, visible blue fluorescence that is characteristic of the generation of dityrosine.[35] In the present study, Tyr crosslinking led to gel formation; this has also been seen previously with soluble type I collagen.[36] The gel formation contrasts to the Cys crosslinked materials where no gels were formed and where it is known that intramolecular crosslinking readily occurs. This suggests that in the present constructs with Tyr insertions, steric considerations may limit, or possibly preclude, the

formation of intramolecular crosslinks, as has been raised for globular proteins.[37] Any steric interference in intramolecular crosslink formation would lead to extensive intermolecular crosslinking and network formation, rather than formation of smaller soluble aggregates, giving gels.

In the present study, the Cys and Tyr residues have been introduced specifically at the ends of the triple-helical CL domain. This ensures that the subsequent crosslinking does not interfere with any function domain that has been introduced into this CL domain.[12, 20-21] This does not preclude inclusion of substituted or inserted triplets containing these amino acids within the helical domain, nor the use of chemical methods to introduce the sulfhydryl or phenolic functionalities into the helical domain. For example, sulfhydryl entities can be introduced through reaction of Lys residues with 2-iminothiolane (methyl 4-mercaptobutyrimidate; Traut's Reagent),[38] while phenolic entities can be added by reaction of Asp or Glu residues with tyramine.[25] Both examples would then allow oxidative crosslinking to give more extensively modified materials.

Bacterial collagen provide a range of opportunities for new biomedical applications, especially after modifications to introduce specific functional sites. For example, in addition to single substitutions,[16] a variety of binding domains and functions have been introduced, including various integrin binding domains,[12,21] a fibronectin binding motif,[20] a heparin binding domain,[21] and MMP cleavage sites.[17,19] These various designed proteins, especially those with designed cell functions, include materials for wound repair,[12,16] as coatings, for polyurethanes[14] and for titanium materials[15] and as a bioactive material for cartilage repair.[17] (Parmar *et al* 2015), with many other opportunities also possible.

Conclusions

The introduction of new Tyr or Cys residues proximal to the triple-helical domain of bacterial collagens, where they are typically not found, allowed for easy oxidative cross-linking. For Cys residues, the extent of intermolecular crosslinking may not be as great as intramolecular crosslinking. However, the extent of intermolecular Tyr crosslinking was more extensive, perhaps due to steric constrains, such that extended networks formed leading to hydrogel formation. In both cases, the presence of these readily formed, direct molecule to molecule crosslinks, led to freeze dried sponges for biomedical applications that were more readily handled and further modified that those produced by the original bacterial collagen lacking these Tyr and Cys additions.

Acknowledgments

We thank Professor Barbara Brodsky for helpful discussions~~, Dr Yen B Truong for assistance with SEM~~ and Linda Howell for assistance with collagen production. This work was supported in part through NIH grant #EB011620.

References

1. Ramshaw JAM, Werkmeister JA, Glattauer V. Collagen-based biomaterials. Biotechnol Genet Eng Rev 1996;13:335-382.
2. Brodsky B, Ramshaw JAM. The collagen triple-helix structure. Matrix Biol 1997; 15:545-554.
3. Rosenbloom J, Harsch M, Jimenez S. Hydroxyproline content determines the denaturation temperature of chick tendon collagen. Arch Biochem Biophys 1973;158:478-484.
4. Ricard-Blum S. The collagen family. Cold Spring Harb Perspect Biol 2011; 3:a004978.
5. Rasmussen M, Jacobsson M, Björck L Genome-based identification and analysis **of** collagen-related structural motifs in bacterial and viral proteins. J Biol Chem 2003;278:32313-32316.
6. Yu Z, An B, Ramshaw JAM, Brodsky B. Bacterial collagen-like proteins that form triple-helical structures. J Struct Biol 2014;186:451-461.
7. Xu Y, Keene DR, Bujnicki JM, Höök M, Lukomski S. Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices. J Biol Chem 2002;277:27312-27318.
8. Han, R., Zwiefka, A., Caswell, C.C., Xu, Y., Keene, D.R., Lukomska, E., Zhao, Z., Höök, M., Lukomski, S. Assessment of prokaryotic collagen-like sequences derived from streptococcal Scl1 and Scl2 proteins as a source of recombinant GXY polymers. Appl Microbiol Biotechnol 2006;72:109-115.
9. Mohs A, Silva T, Yoshida T, Amin R, Lukomski S, Inouye M, Brodsky B Mechanism of stabilization of a bacterial collagen triple helix in the absence of hydroxyproline. J Biol Chem 2007;282:29757-29765.

10. Yoshizumi A, Yu Z, Silva T, Thiagarajan G, Ramshaw JAM, Inouye M, Brodsky B. Self-association of *Streptococcus pyogenes* collagen-like constructs into higher order structures. Protein Sci 2009;18:1241-1251.

11. Peng YY, Yoshizumi A, Danon SJ, Glattauer V, Prokopenko O, Mirochnitchenko O, Yu Z, Inouye M, Werkmeister JA, Brodsky B, Ramshaw JAM. A *Streptococcus pyogenes* derived collagen-like protein as a non-cytotoxic and non-immunogenic cross-linkable biomaterial. Biomaterials 2010;31:2755-2761.

12. Cosgriff-Hernandez E, Hahn MS, Russell B, Wilems T, Munoz-Pinto D, Browning MB, Rivera J, Höök M. Bioactive hydrogels based on designer collagens. Acta Biomater 2010;6:3969-3977.

13. Peng YY, Howell L, Stoichevska V, Werkmeister JA, Dumsday GJ, Ramshaw JAM. Pilot production of a collagen-like protein from *Streptococcus pyogenes* for biomedical applications. Microbial Cell Fact 2012; 11:146.

14. Browning MB, Dempsey D, Guiza V, Becerra S, Rivera J, Russell B, Höök M, Clubb F, Miller M, Fossum T, Dong JF, Bergeron AL, Hahn M, Cosgriff-Hernandez E. Multilayer vascular grafts based on collagen-mimetic proteins. Acta Biomater 2012;8:1010-1021.

15. Bronk JK, Russell BH, Rivera JJ, Pasqualini R, Arap W, Höök M, Barbu EM. A multifunctional streptococcal collagen-mimetic protein coating prevents bacterial adhesion and promotes osteoid formation on titanium. Acta Biomater 2014; 10:3354-3362.

16. Cereceres S, Touchet T, Browning MB, Smith C, Rivera J, Höök M, Whitfield-Cargile C, Russell B, Cosgriff-Hernandez E. Chronic Wound Dressings Based on Collagen-Mimetic Proteins. Adv Wound Care 2015;4:444-456.

17. Parmar PA, Chow LW, St-Pierre JP, Horejs CM, Peng YY, Werkmeister JA, Ramshaw JAM, Stevens MM. Collagen-mimetic peptide-modifiable hydrogels for articular cartilage regeneration. Biomaterials 2015;54:213-225.

18. Peng YY, Stoichevska V, Madsen S, Howell L, Dumsday GJ, Werkmeister JA, Ramshaw JAM. A simple cost-effective methodology for large-scale purification of recombinant non-animal collagens. Appl Microbiol Biotechnol 2014;98:1807-1815.

19. 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.

20. An B, Abbonante V, Yigit S, Balduini A, Kaplan DL, Brodsky B. Definition of the native and denatured type II collagen binding site for fibronectin using a recombinant collagen system. J Biol Chem 2014;289:4941-4951.

21. Peng YY, Stoichevska V, Schacht K, Werkmeister JA, Ramshaw JAM. Engineering multiple biological functional motifs into a blank collagen-like protein template from Streptococcus pyogenes. J Biomed Mater Res A 2014;102:2189-2196.

22. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227:680-85.

23. Vashi AV, Ramshaw JAM, Glattauer V, Elvin CM, Lyons RE, Werkmeister JA. Controlled surface modification of tissue culture polystyrene for selective cell binding using resilin-inspired polypeptides. Biofabrication 2013;5:035005.

24. An B, DesRochers TM, Qin G, Xia X, Thiagarajan G, Brodsky B, Kaplan DL. The influence of specific binding of collagen-silk chimeras to silk biomaterials on hMSC behavior. Biomaterials 2013;34:402-12.

~~25. Truong YB, Glattauer V, Lang G, Hands K, Kyratzis IL, Werkmeister JA, Ramshaw JAM. A comparison of the effects of fibre alignment of smooth and textured fibres in electrospun membranes on fibroblast cell adhesion. Biomed Mater 2010;5:25005.~~

25. Elvin CM, Vuocolo T, Brownlee AG, Sando L, Huson MG, Liyou NE, Stockwell PR, Lyons RE, Kim M, Edwards GA, Johnson G, McFarland GA, Ramshaw JA, Werkmeister JA. A highly elastic tissue sealant based on photopolymerised gelatin. Biomaterials 2010;31:8323-8331.

26. Báez J, Olsen D, Polarek JW. Recombinant microbial systems for the production of human collagen and gelatin. Appl Microbiol Biotechnol 2005;69:245-252.

27. Brodsky B, Kaplan DL. Shining light on collagen: expressing collagen in plants. Tissue Eng Part A 2013;19:1499-1501.

28. Perret S, Merle C, Bernocco S, Berland P, Garrone R, Hulmes DJ, Theisen M, Ruggiero F. Unhydroxylated triple helical collagen I produced in transgenic plants provides new clues on the role of hydroxyproline in collagen folding and fibril formation. J Biol Chem. 2001;276:43693-43698.

29. Qin G, Rivkin A, Lapidot S, Hu X, Preis I, Arinus SB, Dgany O, Shoseyov O, Kaplan DL. Recombinant exon-encoded resilins for elastomeric biomaterials. Biomaterials 2011;32:9231-9243.

30. Squeglia F, Bachert B, De Simone A, Lukomski S, Berisio R. The crystal structure of the streptococcal collagen-like protein 2 globular domain from invasive M3-type group A Streptococcus shows significant similarity to immunomodulatory HIV protein gp41. J Biol Chem. 2014;289:5122-5133.

31. Glanville RW, Allmann H, Fietzek PP. Cysteine residues mark the C-terminal triple-helical non-triplehelical junction in type III collagen. Hoppe Seylers Z Physiol Chem 1976; 357:1663-1665.

32. Barth D, Kyrieleis O, Frank S, Renner C, Moroder L. The role of cystine knots in collagen folding and stability; Part II. Conformational properties of (Pro-Hyp-Gly)n model trimers with N- and C-terminal collagen type III cystine knots. Chemistry 2003; 9:3703-3714.

33. Hunter EPL, Desrosiers MF, Simic MG. The effect of oxygen, antioxidants, and superoxide radical on tyrosine phenoxyl radical dimerization. Free Radical Biol Med 1989;6;581–585.

34. Winterbourn CC, Parsons-Mair HN, Gebicki S, Gebicki JM, Davies MJ. Requirements for superoxide-dependent tyrosine hydroperoxide formation in peptides Biochem J 2004;381:241–248.

35. Lehrer SS, Fasman GD. Ultraviolet irradiation effects in poly-L-tyrosine and model compounds. Identification of bityrosine as a photoproduct. Biochemistry 1967;6:757-767.

36. LaBella F, Waykole P, Queen G. Formation of insoluble gels and dityrosine by the action of peroxidase on soluble collagens. Biochem Biophys Res Commun 1968;30:333-338.

37. Nwamba CO, Chilaka FC. A proposed significance of the δ region and its implications in the mechanism of cooperativity in haemoglobin. Med Hypotheses Res 2010;6:25-35.

38. Traut RR, Bollen A, Sun TT, Hershey JW, Sundberg J, Pierce LR. Methyl 4-mercaptobutyrimidate as a cleavable cross-linking reagent and its application to the Escherichia coli 30S ribosome. Biochemistry 1973;12:3266-3273.

Table 1: Dynamic Light Scattering for V-cys-CLCL-cys samples.

	Hydrodynamic radius (nm)	Percentage of intensity	Percentage of mass
Oxidised sample with no filtration:			
Peak 1	60.6	2.4	7.7
Peak 2	628.2	97.6	92.3
Reduced (TCEP) sample with no filtration:			
Peak 1	3.9	1.9	90.6
Peak 2	243.4	98.1	9.4

Legends to Figures

Figure 1 Schematic representations of the original (A) V-CL monomer sequence, and modifications to this sequence. The structures each comprise 3 chains with the collagen domain having a triple-helical structure. (B) Addition of a C-terminal Tyr, (C) addition of an N- and a C-terminal Tyr, (D) addition of a C-terminal Cys, (E) addition of an N- and a C-terminal Cys, (F) addition of a C-terminal Tyr to a V-CL-CL dimer and (G) addition of an N- and a C-terminal Cys to a V-CL-CL dimer. Removal of the V-domain, when required, was by pepsin digestion.

Figure 2 Sequence data showing the DNA and protein sequences surrounding Tyr and Cys insertions (underlined). (A) Addition of C-terminal Tyr to CL monomer or CLCL dimer, (B) addition of C-terminal Cys (and Tyr) to CL monomer, (C) addition of N-terminal Tyr to CL monomer, (D) addition of N-terminal Cys to CL monomer, (E) addition of C-terminal Cys to CL dimer, and (F) addition of N-terminal Cys to CL dimer.

Figure 3 Examples of expression and purification of various constructs shown using a combination of SDS-PAGE gels. (A) Expression and purification of V-cys-CL-cys. Lane 1, a non-induced control, Lane 2, expression showing the new protein, Lane 3, partial purification using IMAC, Lane 4, cys-CL-cys after proteolytic removal of the V-domain, (B) Lane 1, expressed and purified of tyr-CL-tyr, (C) Lane 1, expressed and purified of V-cys-CLCL-cys. MW indicates molecular weight standards in kDa. On SDS-PAGE, the observed bands are single chains and not triple helical structures.

Figure 4 Examples various constructs that include Tyr residues shown using a combination of SDS-PAGE gels. (A) V-CLCLtyr, showing Lane 1, after purification ~~on IMAC~~, and Lane 2, after photo-crosslinking, and (B) CL-tyr, showing Lane 3, after purification on IMAC, Lane 4, after photo-crosslinking. MW indicates molecular weight standards in kDa. On SDS-PAGE, the observed bands are single chains and not triple helical structures

Figure 5 Photo-crosslinking of purified tyr-CL-tyr. (A) Collagen solution with added $[Ru^{II}(bpy)_3]Cl_2$ prior to crosslinking, (B) a collagen gel formed after photo-crosslinking. ~~, and (C) SEM of freeze dried gel material, Bar = 10 μm.~~

Figure 6 Examples various constructs that include Cys residues shown using a combination of SDS-PAGE gels. (A) cys-CL-cys, showing Lane 1, after purification ~~in the presence of β- mercaptoethanol~~, and Lane 2, after standing in air for 1 week, and (B) V-cys-CLCL-cys, showing Lane 3, a control V-CLCL lacking Cys insertions, Lane 4, in the presence of 20 mM TCEP, Lane 5, after preparation and standing in air without reducing agent for 1 day, and Lane 6, and subsequently standing in air for 1 week. MW indicates molecular weight standards in kDa. On SDS-PAGE, the observed bands are single chains and not triple helical structures

Figure 7 MALDI spectra for (A) cys-CL-cys and (B) V-cys-CLCL-cys constructs, showing the presence of cross-linked components.

Table 1: Dynamic Light Scattering for V-cys-CLCL-cys samples.

	Hydrodynamic radius (nm)	Percentage of intensity	Percentage of mass
Oxidised sample with no filtration:			
Peak 1	60.6	2.4	7.7
Peak 2	628.2	97.6	92.3
Reduced (TCEP) sample with no filtration:			
Peak 1	3.9	1.9	90.6
Peak 2	243.4	98.1	9.4

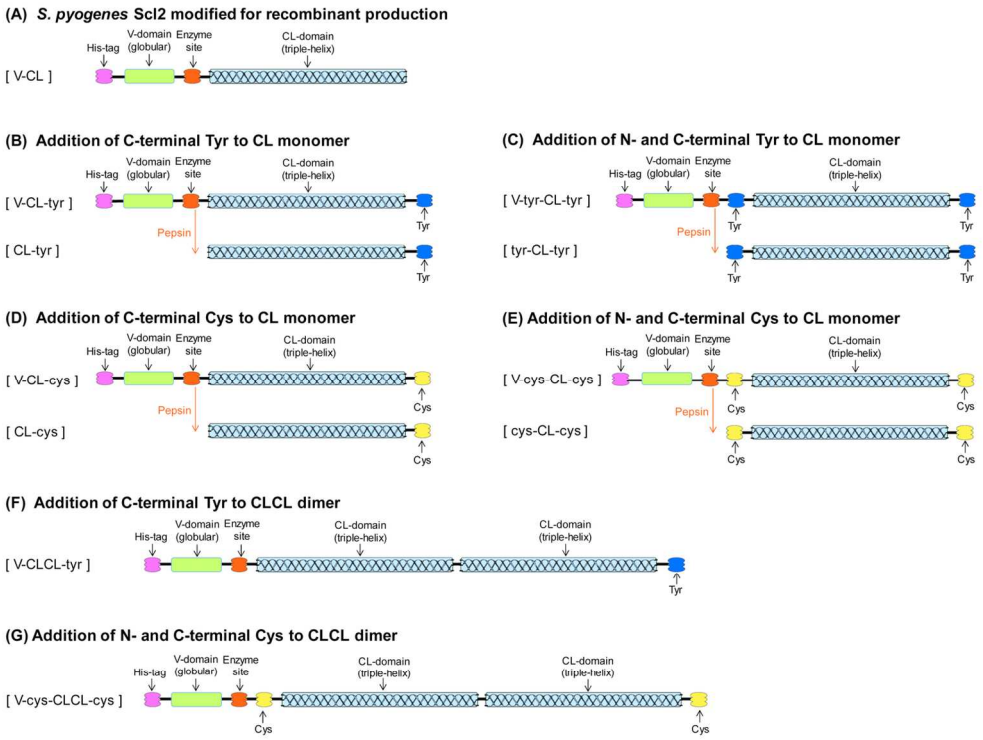


Figure 1 Schematic representations of the original (A) V-CL monomer sequence, and modifications to this sequence. The structures each comprise 3 chains with the collagen domain having a triple-helical structure. (B) Addition of a C-terminal Tyr, (C) addition of an N- and a C-terminal Tyr, (D) addition of a C-terminal Cys, (E) addition of an N- and a C-terminal Cys, (F) addition of a C-terminal Tyr to a V-CL-CL dimer and (G) addition of an N- and a C-terminal Cys to a V-CL-CL dimer. Removal of the V-domain, when required, was by pepsin digestion.
133x103mm (300 x 300 DPI)

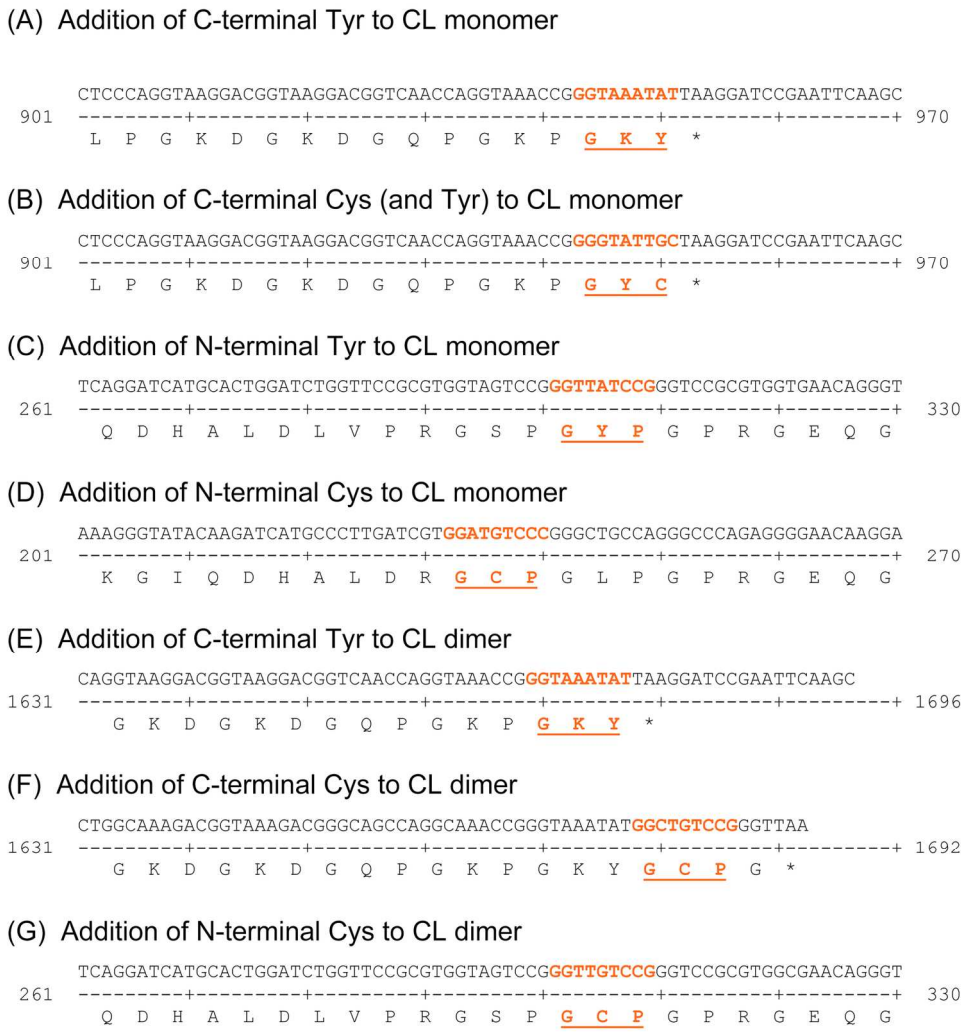


Figure 2 Sequence data showing the DNA and protein sequences surrounding Tyr and Cys insertions (underlined). (A) Addition of C-terminal Tyr to CL monomer or CLCL dimer, (B) addition of C-terminal Cys (and Tyr) to CL monomer, (C) addition of N-terminal Tyr to CL monomer, (D) addition of N-terminal Cys to CL monomer, (E) addition of C-terminal Cys to CL dimer, and (F) addition of N-terminal Cys to CL dimer.
155x168mm (300 x 300 DPI)

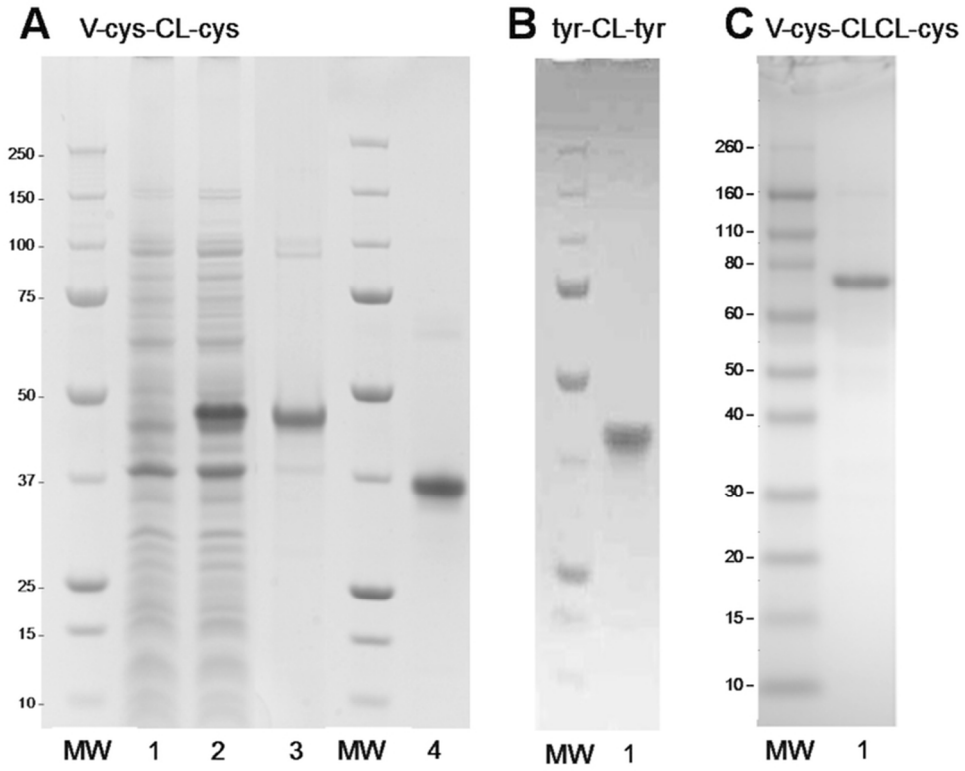


Figure 3 Examples of expression and purification of various constructs shown using a combination of SDS-PAGE gels. (A) Expression and purification of V-cys-CL-cys. Lane 1, a non-induced control, Lane 2, expression showing the new protein, Lane 3, partial purification using IMAC, Lane 4, cys-CL-cys after proteolytic removal of the V-domain, (B) Lane 1, expressed and purified of tyr-CL-tyr, (C) Lane 1, expressed and purified of V-cys-CLCL-cys. MW indicates molecular weight standards in kDa. On SDS-PAGE, the observed bands are single chains and not triple helical structures.
73x60mm (300 x 300 DPI)

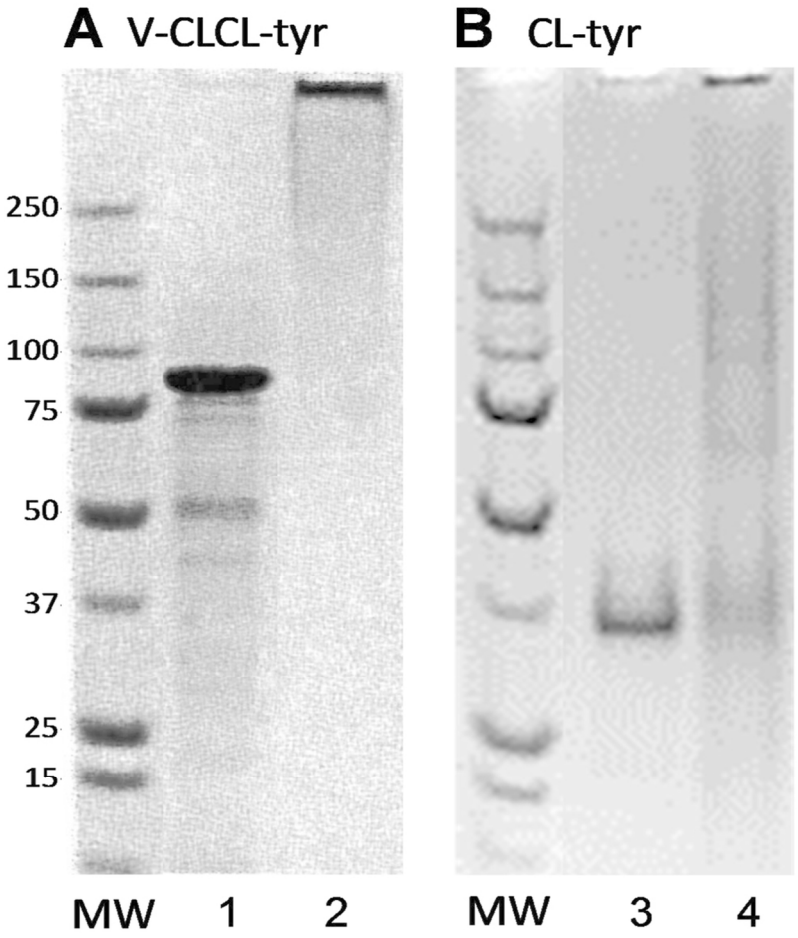


Figure 4 Examples various constructs that include Tyr residues shown using a combination of SDS-PAGE gels. (A) V-CLCLtyr, showing Lane 1, after purification, and Lane 2, after photo-crosslinking, and (B) CL-tyr, showing Lane 3, after purification on IMAC, Lane 4, after photo-crosslinking. MW indicates molecular weight standards in kDa. On SDS-PAGE, the observed bands are single chains and not triple helical structures.
106x106mm (300 x 300 DPI)

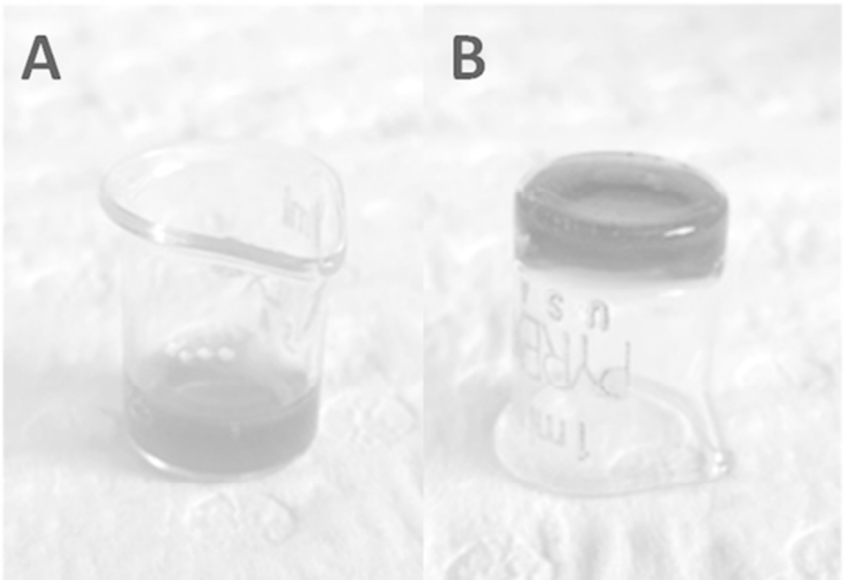


Figure 5 Photo-crosslinking of purified tyr-CL-tyr. (A) Collagen solution with added [RuII(bpy)3]Cl2 prior to crosslinking, (B) a collagen gel formed after photo-crosslinking.
58x33mm (300 x 300 DPI)

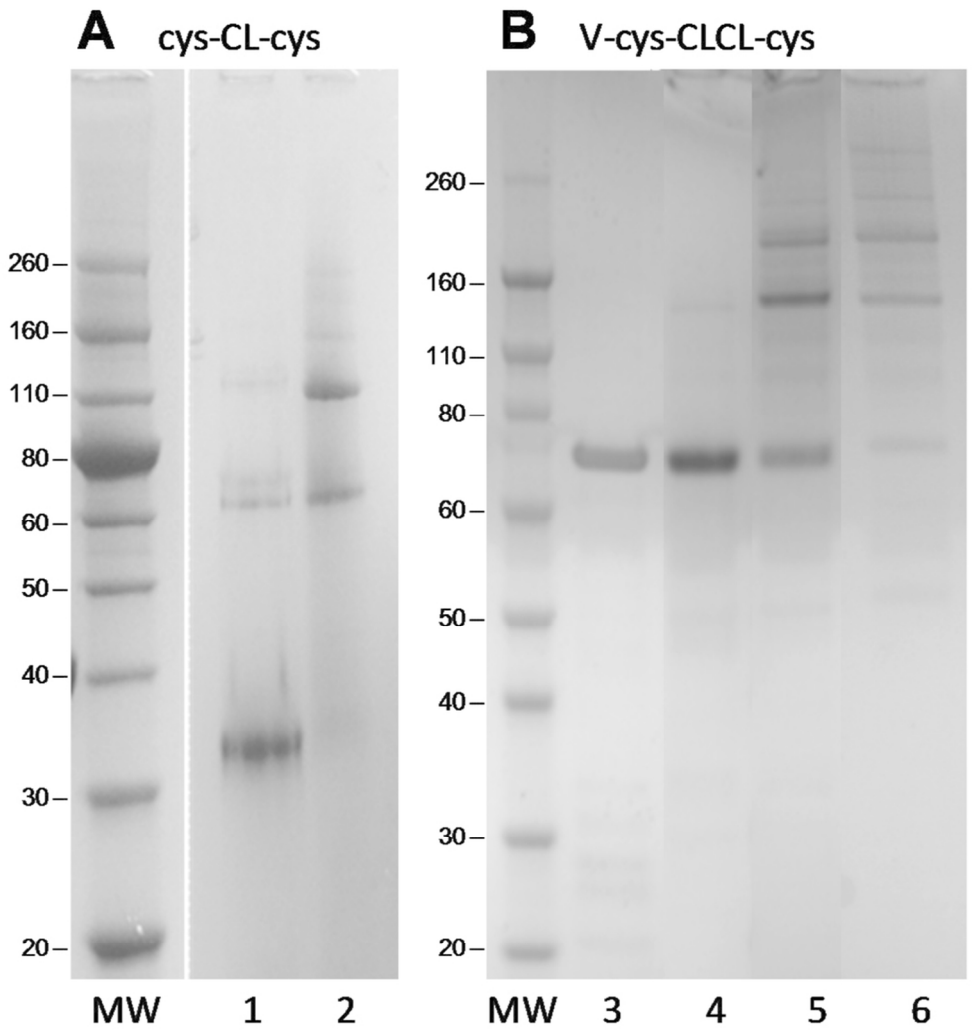


Figure 6 Examples various constructs that include Cys residues shown using a combination of SDS-PAGE gels. (A) cys-CL-cys, showing Lane 1, after purification in the presence of β–mercaptoethanol, and Lane 2, after standing in air for 1 week, and (B) V-cys-CLCL-cys, showing Lane 3, a control V-CLCL lacking Cys insertions, Lane 4, in the presence of 20 mM TCEP, Lane 5, after preparation and standing in air without reducing agent for 1 day, and Lane 6, and subsequently standing in air for 1 week. MW indicates molecular weight standards in kDa. On SDS-PAGE, the observed bands are single chains and not triple helical structures.
93x98mm (300 x 300 DPI)

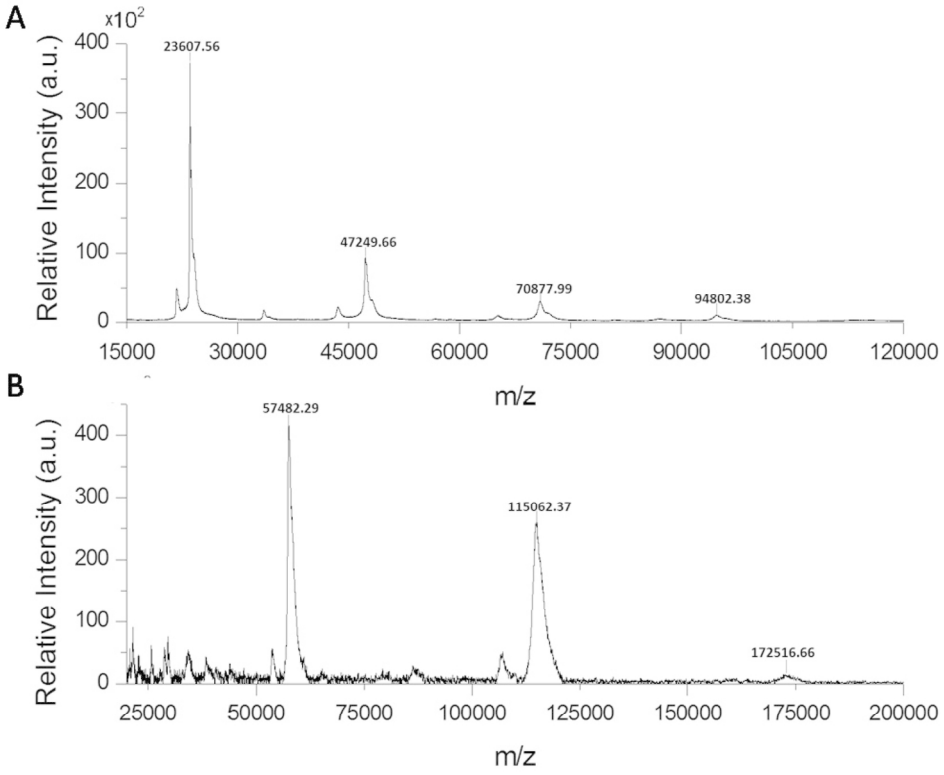


Figure 7 MALDI spectra for (A) cys-CL-cys and (B) V-cys-CLCL-cys constructs, showing the presence of cross-linked components.
74x62mm (600 x 600 DPI)

Author/s:
Stoichevska, V; An, B; Peng, YY; Yigit, S; Vashi, AV; Kaplan, DL; Werkmeister, JA; Dumsday, GJ; Ramshaw, JAM

Title:
Formation of multimers of bacterial collagens through introduction of specific sites for oxidative crosslinking

Date:
2016-09-01

Citation:
Stoichevska, V., An, B., Peng, Y. Y., Yigit, S., Vashi, A. V., Kaplan, D. L., Werkmeister, J. A., Dumsday, G. J. & Ramshaw, J. A. M. (2016). Formation of multimers of bacterial collagens through introduction of specific sites for oxidative crosslinking. JOURNAL OF BIOMEDICAL MATERIALS RESEARCH PART A, 104 (9), pp.2369-2376. https://doi.org/10.1002/jbm.a.35772.