**Engineering specific chemical modification sites into a collagen-like protein
from *Streptococcus pyogenes***

Violet Stoichevska, Yong Y. Peng, Aditya V. Vashi, Jerome A. Werkmeister,

Geoff J. Dumsday and John A.M. Ramshaw

CSIRO Manufacturing, Bayview Avenue, Clayton 3168, Australia.

Correspondence to:

John Ramshaw

john.ramshaw@csiro.au and/or johnamramshaw@gmail.com

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version record](#). Please cite this article as [doi:10.1002/jbm.a.35957](https://doi.org/10.1002/jbm.a.35957).

Abstract

Recombinant bacterial collagens provide a new opportunity for safe biomedical materials. They are readily expressed in *E. coli* in good yield and can be readily purified by simple approaches. However, recombinant proteins are limited in that direct secondary modification during expression is generally not easily achieved. Thus, inclusion of unusual amino acids, cyclic peptides, sugars, lipids and other complex functions generally needs to be achieved chemically after synthesis and extraction. In the present study we have illustrated that bacterial collagens that have had their sequences modified to include cysteine residue(s), which are not normally present in bacterial collagen-like sequences, enable a range of specific chemical modification reactions to be produced. Various model reactions were shown to be effective for modifying the collagens. The ability to include alkyne (or azide) functions allows the extensive range of substitutions that are available via 'click' chemistry to be accessed. When bifunctional reagents were used, some crosslinking occurred to give higher molecular weight polymeric proteins, but gels were not formed.

Keywords: Bacterial collagen, chemical modification, crosslinking, recombinant

INTRODUCTION

Collagens are typically found throughout the animal kingdom, where they are major structural proteins and also have key functional roles in molecular and cellular interactions. All collagens are characterised by a specific tertiary structure, the triple-helix, in which three left-handed, polyproline II-like chains are wound into a right-handed super-helix.¹ A consequence of this specific, tightly wound tertiary structure is that every third residue in the amino acid sequence is a Gly, as only Gly is small enough to fit properly within the triple-helical structure. This, therefore, leads to a characteristic amino acid sequence of (Gly-Xaa-Yaa)_n for collagens. There are 28 human (mammalian) collagens that all have distinct functional roles in the extracellular matrix.² This characteristic sequence is also found in a range of other proteins that also contain the triple-helical motif.¹ More recently, collagens have been identified in several non-animal species, including bacteria,³⁻⁵ fungi,⁶ viruses⁷ and phage.⁸ These proteins all show extended (Gly-Xaa-Yaa)_n sequences and in selected instances have been clearly shown to fold into the triple-helical structure.^{5,8}

The most abundant collagens, for example type I, type II and type III collagens, have been used successfully in a wide variety of biomedical applications.⁹ Despite this track record, there still remain concerns about the safety of these products. The potential for transmission of animal diseases such as spongiform encephalopathy (mad cow disease) is a specific concern. There have been promising outcomes with a variety of recombinant collagen products, using for example, yeast, plant and transgenic expression systems,¹⁰ but the major issue with all these systems is the need to also include and express functional prolyl-4-hydroxylase activity in the constructs.¹⁰ The opportunity to use bacterially derived collagens provides an alternative strategy as these proteins are readily expressed as recombinant products and do not need introduction of prolyl-hydroxylase.¹¹ These collagens have been shown to be non-cytotoxic and non-immunogenic,¹² have been produced in

excellent yields¹¹ and have been readily purified in large scale.¹³ A key feature is that they do not include or require hydroxyproline as a secondary modification, an essential modification for the stability of mammalian collagens.¹⁴ They have been shown, nevertheless, to be stable for production and fabrication, with melting temperatures typically around 37 °C.⁵

In some cases, such as the Scl2 protein from *Streptococcus pyogenes*,³ these recombinant bacterial collagens behave like a ‘blank slate’ and do not show any of the typical molecular and cellular interactions that have been identified for animal collagens.¹⁵⁻¹⁷ This allows for specific gene designs to include one or more introduced function. Examples, which have been shown to function as expected, include the addition of conformation dependent, triple helical domains for integrins,^{16,18} heparin¹⁷ and fibronectin¹⁹ binding, and MMP degradation sites.²⁰ It is also possible to add non-triple helical entities, either at the ends of the triple-helical domain or as an inset between triple-helical segments. For example, a (GlyAlaGlyAlaGlySer)_n silk-like motif has been added to a bacterial collagen and was shown to be functional as the purified protein annealed to a silk film.²¹ A variety of other binding motifs could also be considered for inclusion in designed constructs. However, not all potential modifications can be as readily incorporated. The bacterial collagens are not usually the subject of secondary modifications, although some, such as that from *Bacillus anthracis*, are glycosylated.²²

Some chemical modifications, such as glycosylation can be achieved by chemical methods on purified proteins. A chemical approach for various modifications adds significantly to the versatility of the bacterial collagens, as constructs can be made not only with triple-helical functional sites but also with other functional sites not found in the collagen triple-helix. This can provide a wide range of functions all associated with a single delivery molecule. So beyond triple-helical domains, function that could be added by chemical modification include functional sugars, for example galactose residues to enhance

certain cell interactions.²³ This also applies to inclusion of cyclic peptides, peptides with D-amino acids, lipids and other chemical functions.

The additional modifications could be introduced at random locations on the triple-helix through a general strategy, or example targeting lysine residues. However, this would provide a dispersed mixture of products, some of which could interfere with functions in the triple helix.^{16,18} Higher levels of modification can lead to loss of stability, although strategies have been developed that can fully reverse this effect.^{16,18} A preferable approach would be to direct the modifications to selected locations, either within the triple-helix, or preferably at the ends of the triple helix. In the present study, we have examined the potential for specific chemical modification at defined locations through incorporation of one or more Cys residues within the construct design.^{24,25} Cys does not normally occur within bacterial collagen structures and so allows for very specific reaction sites. Each introduced Cys residue provides three modification sites in the triple-helical molecule. Cys has long-standing, well established reactions with a variety of reagents in a specific manner²⁶ that have been used extensively with peptides and globular proteins. Typically, there are few non-specific reactions.^{26,27} Their use with fibrous proteins is less well documented, especially for collagens where Cys is only rarely found²⁸ and there is a particularly extensive and ordered hydration shell^{29,30} that may limit reactivity. The potential to modify Cys residues included in a collagen has been examined using different reaction pathways..

MATERIALS AND METHODS

Clone production

The construct used in this study to illustrate potential chemical reactions was based around a single, modified bacterial collagen structure,²⁴ although other bacterial collagen options^{5,13}

201 AAAGGGTATACAAGATCATGCCCTTGATCGTGATGTCCCGGGCTGCCAGGGCCCAGAGGGGAACAAGGA 270
K G I O D H A L D R G C P G L P G P R G E O G

CTCCCAGGTAAGGACGGTAAGGACGGTCAACCAGGTAACCGGGGTATTGCTAAGGATCCGAATTC AAGC

901 -----+-----+-----+-----+-----+-----+-----+-----+-----+----- 970

L P G K D G K D G Q P G K P **G Y C** *

The DNA construct was cloned into *E. coli* expression vector system pColdIII using the unique sites 5' *Nde*I and 3' *Bam*HI. For expression, a selected positive clone was transformed into the *E. coli* host BL21-DE3. Cells were grown in 2 x YT Media with ampicillin (50 µg/ml) at 37 °C for 24 h and cell culture optical density at A600nm reached around 3-5. For expression in shake flasks, the culture was then incubated at 25 °C and 1 mM isopropyl beta-D-thiogalactopyranoside was 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.¹² After 24 h total incubation, cells were harvested by centrifugation and extracted by sonication into pH 8 phosphate buffer as previously described.¹² The cell lysate mixture was clarified by centrifugation (12,000 x g for 40 min) and the clear supernatant containing the triple helical protein was retained.

For purification, 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.³² The product was recovered and exchanged into 20 mM sodium phosphate buffer, pH8.0, using a 10 kDa cross-flow filtration membrane apparatus (Pall Life Sciences).¹³ The pH was adjusted to pH 2.5 with acetic acid and the V-cCLc treated with 0.01 mg/mL pepsin for 16 h at 4 °C to remove the V-domain. Final purification was achieved on a Sephacryl S200 26/60 column (GE Healthcare).³² All steps were performed at 4 °C. Purity was examined by SDS-PAGE³³ using 4-12% NU-PAGE gels. Reduction of purified proteins to eliminate disulfide bonds used 5mM tris(2-carboxyethyl)phosphine (TCEP) prior to modification reactions. Protein stability was examined by susceptibility to proteolysis and CD spectroscopy.³¹

Modification through reaction with a maleimidyl peptide sequence

FLAG-peptide (Sigma) which was reacted at a 5 fold excess with NHS-maleimidyl (PEG)₂ (ThermoScientific) following the manufacturer's recommendations and at low pH in pH 6.8 sodium phosphate buffer, to preference reaction with the N-terminal α -amino group. A solution of purified, reduced cCLc protein at 5 mg/ml was adjusted to pH 7.4 in 20 mM sodium phosphate buffer, 0.15 M NaCl and was reacted for 2 h with the maleimidyl PEG₂-FLAG peptide conjugate. The modified protein was recovered by dialysis and freeze drying. FLAG peptide addition was shown by SDS-PAGE followed by Western blotting with an anti-

FLAG antibody and using a horseradish conjugated–anti-mouse secondary antibody, as previously described,³⁴ except that SuperSignal™ West Pico Chemiluminescent Substrate (Thermo Scientific) was used as the detection system.

Modification through reaction with a bromoacetyl peptide sequence

A solution of purified, reduced cCLc protein at 5 mg/ml was adjusted to pH 8.5 in 20 mM sodium phosphate buffer and was then reacted with excess bromoacetyl-Gly-Arg-Arg-Arg (Mimotopes).³⁵ The modified protein was recovered by dialysis and freeze drying. The reaction with bromoacetyl-Gly-Arg-Arg-Arg was assessed by increased binding of fluorescein isothiocyanate labelled heparin (FITC-heparin; Molecular Probes) as previously described.¹⁷ Modified protein and control (unmodified) samples were coated onto a tissue culture plastic plates and left overnight at 4 °C. Samples were washed 3 times with 3% w/v BSA and then held with 3% w/v BSA at room temperature for 5 h. FITC-heparin was added and incubated for > 4 h at 4 °C in the dark. After extensive washing with PBS, the fluorescence was examined with a plate reader (PHERAstar).

Modification through reaction with a vinylsulfone labelled dye

A solution of purified, reduced cCLc protein at 5 mg/ml was adjusted to pH 8.0 in 20 mM sodium phosphate buffer, 0.15 M NaCl and was then reacted with a 2 fold molar excess of QXL 610 vinylsulfone (Anaspec), using a concentrated stock (15.6mM) in dimethylsulfoxide following the manufacturer's recommendations. After 1 h reaction, excess dye was quenched with 10 mM DTT.³⁶ Excess reagents were removed by gel filtration, as used above in protein purification. The presence of labelling was examined by visible light spectroscopy, at an expected principal extinction maximum at 594 nm.

Modification through reactions enabling 'click' chemistry coupling

A solution of purified, reduced cCLc protein at 5 mg/ml was adjusted to pH 7.0 in 100 mM potassium phosphate buffer and was then reacted with a 2-fold molar excess of alkyne-PEG₄-maleimide (Sigma) for 16 h at room temperature with gentle agitation, following the manufacturer's recommendations. The product of this reaction was then reacted with a 6 fold excess of 6-FAM-Azide (6-fluorescein azide, Jena Bioscience), in 100 mM potassium phosphate buffer, also following the manufacturer's recommendations. The 'click' coupling was achieved by adding fresh 5mM sodium ascorbate and a pre mixed solution of 0.1 mM CuSO₄ and 0.5 mM tris(3-hydroxypropyltriazolylmethyl)amine (Sigma) under N₂ for 1 h.³⁷ Labelled protein was separated by SDS-PAGE and visualised under UV light.

Crosslinking through reaction with bis-vinylsulfone PEG

A solution of purified, reduced cCLc protein at 5 mg/ml was adjusted to pH 8.0 in 20 mM sodium phosphate buffer, 0.15 M NaCl and was reacted with a 2 fold molar excess of either di-vinylsulfone PEG (MW 3500) or di-vinylsulfone PEG (MW 5000) (JenKem Technology) for 16 h. Cross-linked protein was recovered by dialysis and freeze drying and examined by SDS-PAGE.

Crosslinking through reaction with bis-maleimidyl PEG

A solution of purified, reduced cCLc protein at 5 mg/ml was adjusted to pH 6.8 in 20 mM sodium phosphate buffer, 0.15 M NaCl and 10 mM EDTA. Protein was reacted with either bis maleimidyl (PEG)₂ (Pierce), bis maleimidyl (PEG)₃ (Pierce), bis maleimidyl (PEG-10,000) (Sunbright DE100MA; NOF Corporation) or bis maleimidyl (PEG-20,000) (Sunbright DE200MA; NOF Corporation) The reaction was allowed to proceed for 2 h.

Cross-linked protein was recovered by dialysis and freeze drying and was analysed by SDS-PAGE.

RESULTS

In the present study a single representative construct containing additional Cys residues was made. It expressed well in shake flasks and was readily purified, principally by IMAC. The yield was similar to that found in other shake flask trials, based on SDS-PAGE gel staining, but full determination of yield would need studies in bioreactors as previously described.¹¹

The stability of the product, demonstrated by a resistance to proteolysis and determined by CD, was slightly reduced at 36.5 °C compared to the 37.6 °C for unmodified protein or 37.5 °C for a construct where there are two Cys residue inclusions, adjacent to each other, were at the C-terminal of the protein.³¹ If further Cys residues were included that led to a more problematic loss in stability, then changes to other residues in the triple helix to enhance stability are possible.^{18,38} Examination by SDS-PAGE, in addition to showing purity, also showed that the introduction of Cys residues did not lead to a major alteration of the electrophoretic mobility of reduced protein (Fig. 2).

Other related structures, based on *S. pyogenes* and other bacterial collagen sequences, have also been described that would have been suitable for this study, including constructs that include substitutions in the triple-helix, and multiple insertions, including of adjacent Cys residues.^{25,31} In the present study, the N-terminal Cys residue was placed in the Xaa position in the first triplet prior to the following (Gly-Xaa-Yaa)_n triple helical repeat, although the number of residues preceding it after the proteolytic removal of the V-domain was not determined. The Xaa position in the triple helix has greater accessibility,^{38,39} which could enhance reactivity, although Cys provides slightly greater stability when in the Yaa

position.^{38,40} The Cys at the C-terminal was in the Yaa position and was the terminal residue of the construct, ensuring accessibility.

The cCLc protein was used for various modification reactions. The ratio of the various reagents to collagen followed, as a minimum, those recommended in the literature or in the manufacturer's instructions to generally provide full reaction yields for peptides and globular proteins. No detailed study on the minimum requirement for effective modification or effective yield was done for this fibrillar protein.

The reaction with a maleimide group was illustrated by the reaction of cCLc collagen with a maleimidyl-PEG₂-FLAG peptide conjugate. This reaction was suitable for modification of Cys-containing collagens, as shown by electroblotting using an anti-FLAG antibody where the modified collagen was readily observed (Fig. 3).

An alternate reaction for modification of Cys residues was through reaction with bromoacetyl groups, or iodoacetyl groups, as frequently used for carboxymethylation of proteins.²⁶ In the present study a bromoacetyl peptide, bromoacetyl-Gly-Arg-Arg-Arg was used. The successful reaction by this route was shown by the increase in the heparin binding capability of the modified protein, using fluorescently labelled (FITC) heparin. This showed that fluorescent intensity increased in comparison to a control cCLc sample where no peptide had been added. Thus, fluorescent data showed that when bromoacetyl-peptide (3 and 6 fold excess) had been reacted with the cCLc protein, fluorescent intensity increased in comparison to the control cCLc sample where no peptide was added, with the fluorescent intensity of both the 3-fold and 6-fold excess samples increasing from 100 to 200 arbitrary instrument units. As previously observed¹⁷ unmodified bacterial collagen can bind heparin, probably non-specifically, and a high background of 100 arbitrary instrument units was also observed in the present study.

A third alternative reaction for Cys residues was the reaction with vinyl sulphone derivatives. An example of this reaction was demonstrated using a vinyl sulphone-labelled dye. The spectrum of the reaction product (data not shown) illustrated the effectiveness of the dye reactions, with new peaks being observed, with maxima around 595 nm and 630 nm that gave the product a blue colour, as expected from the manufacturer's description.

The previous examples allow direct attachment of peptides or other entities. A further alternative was to add an intermediate, specific chemical entity that allows further diverse reactions to be considered. An example of this was 'click' chemistry, in which either an azido group or an alkyne group could be attached to a Cys residue, as neither of these functionalities can presently be readily and specifically incorporated directly in recombinant proteins. In the present study, an alkyne function was added. The effectiveness of the maleimide-based alkyne addition was demonstrated by further reaction, using copper catalysed 'click' chemistry, with a fluorescent azide derivative. The product of this reaction was readily observed under UV light after separation of the protein by SDS-PAGE (Fig. 4).

The previous examples have all demonstrated mono-functional modification reactions. The same chemical reactions were also suitable for inclusion in bi-functional (or multifunctional) reagents. This was illustrated by reactions with bis vinyl sulfone PEGs and with bis-maleimidyl-PEGs. These reactions allowed a more specific approach to crosslinking than the various reactions based on reaction with Lys residues in the CL triple-helical domain.¹⁶ The reaction of cCLc with two bis vinyl sulfone PEG reagents, with PEG molecular weights of 3500 and 5000 Da, showed crosslinking, but an extensively crosslinked, high molecular weight derivative was not seen in the gel-pocket (Fig. 5). Also, no gel-like materials were formed. Thus, although the recommendation for full reaction with a globular protein were used, the extent of reaction was less with the fibrillar protein, possibly due to poor accessibility of reactive sites in the quite large structures that initially form and which

become less readily available for higher polymeric products to form, with molecular entanglements possibly playing a role. The reaction with the higher molecular weight reagent shows as slightly slower, higher molecular weight dimer and trimer bands (Fig. 5). No gelation was observed, as has been observed previously through oxidative photo-crosslinking of Tyr residue substitutions.²⁴ Crosslinking by oxidation of introduced Cys residues, however, did not lead to gelation either.²⁴ It was shown, nevertheless, that the formation of small polymers was sufficient to assist in fabrication as it led to more stable sponge formation.²⁴ Addition of further Cys residues may enable gel formation if this were necessary for a particular application.

A similar limited extent of reaction was also seen for the reaction of reduced cCLc with various bis-maleimidyl-PEG derivatives. At the recommended excess of reagent, the reactions also gave varying, limited degrees of crosslinking (Fig. 6), with no gel formation. For reagents with 2 and 3 PEG units, crosslinking was seen, with the PEG₃ crosslinks being more extensive (Fig. 6). In both cases the paired dimer bands suggest that additional mono-functional reaction rather than crosslinking had also occurred. The presence of mono-functional reaction was more readily seen with the PEG-10,000 and PEG-20,000 reagents, where additional bands were seen between the monomer, dimer and polymer bands (Fig. 6). Again, a lack of accessibility due to the size of the collagen, entanglements and the possibility of the accumulation of mono-functional blocking reactions that did not proceed further may explain the limited cross-linking that was observed.

DISCUSSION

The introduction of one or more Cys residues into constructs of bacterial collagens, either proximal to the triple-helical domain or within the triple-helical domain²⁵ is ideal for specific

modifications of recombinant proteins as bacterial collagen Cys residues are rare or absent and can readily be coded into DNA constructs. The Scl2 collagen from *S. pyogenes* behaves as a 'blank slate' as it shows few, if any, interactions with mammalian cell systems.¹⁶ Certain functions can be introduced by recombinant methods within the triple helical domain, based on known functions in mammalian collagens.⁴¹ However, many functions cannot be readily introduced by recombinant methods and so a chemical approach is necessary. The inclusion of multiple functions in a single molecule¹⁷ will enable a complex single molecule to provide multiple simultaneous reactivities. The present modification and chemical approach, therefore, augments the opportunities that are available solely through recombinant means, providing new materials that can be made fully animal free.

The present approach illustrates that selective and location specific modifications to be achieved that cannot otherwise be readily accomplished using recombinant approaches. The present studies illustrate the suitability of various chemical reactions.²⁶ Although the reactions were chosen to examine and illustrate the chemistry involved, in some cases, the products could be useful of themselves. For example, dye-modified protein could be useful for tracking the bacterial collagen as no monoclonal antibody is readily available yet. Also, the crosslinking reactions could be useful in sponge fabrication and stability.²⁴ Practical examples could include, for example, the addition of peptides that contain D-amino acids or other non-coded amino acids, functional peptides that are cyclic or branched,⁴² as well as a range of non-peptidyl ligands, such as sugars, lipids or more complex ligands.

The present study used a representative example, cCLc, to illustrate several of the chemical modification reactions that could be used to achieve successful reactions. Thus, reactions using maleimide, bromoacetyl and vinyl sulphone derivatives have all been shown to be appropriate as modification strategies. The modified collagens could allow single to

multiple sites to be included, with 3 copies per molecule per Cys insertion as these collagens are all homotrimers. Various other constructs beyond cCLc have been described.^{25,31}

Other methods for modification or for cross-linking of bacterial collagens are possible. Most notable would be the reaction with the many Lys residues that are found along the triple-helical CL domain. Low degrees of modification will not show specificity of location as it is unlikely that any one Lys residue would be significantly more reactive than any other. For high degrees of reaction, the loss of the Lys amino functionalities may lead to loss of triple helical stability, as Lys residues contribute significantly to triple-helical stability through involvement in ion pair bonding in the absence of stabilisation by hydroxyproline.⁴³ Previously, crosslinking has been achieved through reactions at Lys residues.¹⁶ The use of bi-functional maleimide or vinyl sulphone reagents allowed a different approach to crosslinking, but this did not lead to gel formation.

The availability of maleimidyl reagents that allow active site precursors for 'click' chemistry to be readily attached to the Cys-modified bacterial collagens presents the greatest opportunity for production of a wide range of derivatives for functional evaluation. Whether either an alkyne or an azide is attached to the protein, large numbers (>100) of partner reagents are readily available for further reactions.^{44,45}

CONCLUSION

Modification of recombinant bacterial collagens to include introduced Cys residues in specific locations provided an effective approach for further, secondary modification of the expressed proteins at specified locations. This particularly allows the introduction of functionalities that cannot be readily introduced through recombinant methods. These include inclusion of non-normal amino acids, cyclic and branched entities and non-protein

components such as sugars. The three reagents evaluated, maleimidyl, bromoacetyl and vinyl sulphone derivatives, all proved suitable for successful modification of the modified recombinant bacterial proteins. In combination with functions that can be introduced through recombinant means, this additional approach allows for the creation of more complex molecules with multiple functions.

ACKNOWLEDGMENTS

The authors would like to thank Drs L. Lu and G. Lovrecz for providing anti-FLAG mAb.

REFERENCES

1. Brodsky B, Ramshaw JAM. The collagen triple-helix structure. *Matrix Biol.* 1997;15:545-554.
2. Ricard-Blum S. The collagen family. *Cold Spring Harb Perspect Biol.* 2011;3:a004978.
3. 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.
4. 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.
5. Yu Z, An B, Ramshaw JAM, Brodsky B. Bacterial collagen-like proteins that form triple-helical structures. *J Struct Biol.* 2014;186:451-461.
6. Tuntevski K, Durney BC, Snyder AK, Lasala PR, Nayak AP, Green BJ, Beezhold DH, Rio RV, Holland LA, Lukomski S. *Aspergillus* collagen-like genes (acl): identification, sequence polymorphism, and assessment for PCR-based pathogen detection. *Appl Environ Microbiol.* 2013;79:7882-7895.
7. Luther KB, Hülsmeier AJ, Schegg B, Deuber SA, Raoult D, Hennet T. Mimivirus collagen is modified by bifunctional lysyl hydroxylase and glycosyltransferase enzyme. *J Biol Chem.* 2011;286:43701-43709.
8. Ghosh N, McKillop TJ, Jowitt TA, Howard M, Davies H, Holmes DF, Roberts IS, Bella J. Collagen-like proteins in pathogenic *E. coli* strains. *PLoS One.* 2012;7:e37872.
9. Ramshaw JAM, Werkmeister JA, Glattauer V. Collagen-based biomaterials. *Biotechnol Genet Eng Rev.* 1996;13:335-382.
10. Werkmeister JA, Ramshaw JAM. Recombinant protein scaffolds for tissue engineering. *Biomed Mater.* 2012;7:012002.
11. 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.
12. 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.
13. 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.
  14. Rosenbloom J, Harsch M, Jimenez S. Hydroxyproline content determines the denaturation temperature of chick tendon collagen. *Arch Biochem Biophys*. 1973;158:478-484.
  15. Han R, Zwiefka A, Caswell CC, Xu Y, Keene DR, 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.
  16. 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.
  17. 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*. 2015;102:2189-2196.
  18. 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.
  19. 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.
  20. 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.
  21. 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;34:402-412.
  22. Sylvestre P, Couture-Tosi E, Mock M. A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. *Mol Microbiol*. 2002;45:169-178.

23. Pipia GG, Long MW. Human hematopoietic progenitor cell isolation based on galactose-specific cell surface binding. *Nat Biotechnol.* 1997;15:1007-1011.
24. Stoichevska V, An B, Peng YY, Yigit S, Vashi AV, Kaplan DL, Werkmeister JA, Dumsday GJ, Ramshaw JAM. Formation of multimers of bacterial collagens through introduction of specific sites for oxidative cross-linking. *J Biomed Mater Res A.* 2016; doi: 10.1002/jbm.a.35772. [Epub ahead of print]
25. Ramshaw JAM, Werkmeister JA, Stoichevska V, Peng YY. Modified bacterial collagen-like proteins. 2015. WO2015031950-A1
26. Needleman SB. Protein sequence determination. Berlin & New York: Springer-Verlag; 1970.
27. Lundblad RL. Chemical Reagents for Protein Modification. Boca Raton: CRC Press; 2004.
28. Ramshaw JAM, Shah NK, Brodsky B. Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides. *J Struct Biol.* 1998;122:86-91.
29. Privalov PL, Tiktopulo EI. Thermal conformational transformation of tropocollagen. I. Calorimetric study. *Biopolymers.* 1970;9:127-139.
30. Bella J, Brodsky B, Berman HM. Hydration structure of a collagen peptide. *Structure.* 1995;3:893-906.
31. Peng YY, Stoichevska V, Howell L, Madsen S, Werkmeister JA, Dumsday GJ, Ramshaw JAM. Preparation and characterization of monomers to tetramers of a collagen-like domain from *Streptococcus pyogenes*. *Bioengineered.* 2014;5:378-385.
32. 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.
33. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 1970;227:680-685.
34. Ramshaw JAM, Werkmeister JA. Electrophoresis and electroblotting of native collagens. *Anal Biochem.* 1988;168:82-87.
35. Schelté P, Boeckler C, Frisch B, Schuber F. Differential reactivity of maleimide and bromoacetyl functions with thiols: application to the preparation of liposomal diepitope constructs. *Bioconjug Chem.* 2000;11:118-123.
36. Bayburt TH, Vishnivetskiy SA, McLean MA, Morizumi T, Huang CC, Tesmer JJ, Ernst OP, Sligar SG, Gurevich VV. Monomeric rhodopsin is sufficient for normal

- rhodopsin kinase (GRK1) phosphorylation and arrestin-1 binding. *J Biol Chem.* 2011;286:1420-1428.
37. Presolski SI, Hong VP, Finn MG. Copper-catalyzed azide-alkyne click chemistry for bioconjugation. *Curr Protoc Chem Biol.* 2011;3:153-162.
38. Persikov AV, Ramshaw JAM, Brodsky B. Prediction of collagen stability from amino acid sequence. *J Biol Chem.* 2005;280:19343-19349.
39. Simon-Lukasik KV, Persikov AV, Brodsky B, Ramshaw JAM, Laws WR, Alexander Ross JB, Ludescher RD. Fluorescence determination of tryptophan side-chain accessibility and dynamics in triple-helical collagen-like peptides. *Biophys J.* 2003;84:501-508.
40. Persikov AV, Ramshaw JAM, Kirkpatrick A, Brodsky B. Amino acid propensities for the collagen triple-helix. *Biochemistry.* 2000;39:14960-14967.
41. Sweeney SM, Orgel JP, Fertala A, McAuliffe JD, Turner KR, Di Lullo GA, Chen S, Antipova O, Perumal S, Ala-Kokko L, Forlino A, Cabral WA, Barnes AM, Marini JC, San Antonio JD. Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates. *J Biol Chem.* 2008;283:21187-21197.
42. van Hagen PM, Breeman WAP, Bernard HF, Schaar M, Mooij CM, Srinivasan A, Schmidt MA, Krenning EP, de Jong M. Evaluation of a radiolabelled cyclic DTPA-RGD analogue for tumour imaging and radionuclide therapy. *Int J Cancer.* 2000;90:186-198.
43. 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.
44. Evans RA. The rise of azide-alkyne 1,3-dipolar 'click' cycloaddition and its application to polymer science and surface modification. *Aust J Chem.* 2007;60:384-395.
45. Feast GC, Lepitre T, Mulet X, Conn CE, Hutt OE, Savage GP, Drummond CJ. The search for new amphiphiles: synthesis of a modular, high-throughput library. *Beilstein J Org Chem.* 2014;10:1578-1588.

- FIGURE 1.** A schematic to show the constructs used in the present study with introduced Cys residues, present in each of the three chain of the collagen molecule at the N-and C-terminal ends of the collagen domain. The amino acid sequences of the inserts are given. Removal of the V-domain was by pepsin digestion.
- FIGURE 2.** SDS-PAGE collage of (A) *S. pyogenes* V-CL collagens, where (1) V-CL, (2) V-cCLc. Molecular weight standards (S).
- FIGURE 3.** SDS-PAGE, with (A) protein staining with Coomassie Blue and (B) immunostaining with an anti-Flag antibody, where Lanes are (1) molecular weight standards, (2) FLAG-(PEG)₂-cCLc (3) as (2) but double the amount loaded (4) unmodified cCLc and (5) a recombinant FLAG-labelled control protein in an *E. coli* expression lysate.
- FIGURE 4.** SDS-PAGE of cCLc reacted with alkyne-PEG₄-maleimide followed by 'click' coupling with the fluorescent dye 6-fluorescein azide, and observed under UV light.
- FIGURE 5.** SDS-PAGE of cCLc, (1) unmodified (2) after reaction bis vinyl sulfone PEG3500 and (3) after reaction bis vinyl sulfone PEG5000. Molecular weight standards (S). Arrows suggest possible reaction products, where CL is the collagen domain, Peg3.5k is from bis vinyl sulfone PEG3500 reaction and Peg5.0k is from bis vinyl sulfone PEG5000 reaction.

FIGURE 6. SDS-PAGE collage of (1) cCLc (2) after reaction bis-maleimidyl-PEG₂, (3) after reaction bis-maleimidyl-PEG₃, (4) after reaction bis-maleimidyl-PEG10,000 and (5) after reaction bis-maleimidyl-PEG20,000. Molecular weight standards (S). Arrows suggest possible reaction products, where CL is the collagen domain, Peg10k is from bis-maleimidyl-PEG10,000 reaction, Peg20k is from bis-maleimidyl-PEG20,000 reaction and (Peg)₂ is from bis-maleimidyl-PEG₂ reaction.

Figure 1

35x13mm (300 x 300 DPI)

Accepted A

Figure 2

122x174mm (300 x 300 DPI)

Figure 3
76x68mm (300 x 300 DPI)

Acc

Figure 4

69x82mm (300 x 300 DPI)

Figure 5
88x88mm (300 x 300 DPI)

AC

Figure 6

72x59mm (300 x 300 DPI)

AcceJ

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Stoichevska, V;Peng, YY;Vashi, AV;Werkmeister, JA;Dumsday, GJ;Ramshaw, JAM

Title:

Engineering specific chemical modification sites into a collagen-like protein from *Streptococcus pyogenes*

Date:

2017-03-01

Citation:

Stoichevska, V., Peng, Y. Y., Vashi, A. V., Werkmeister, J. A., Dumsday, G. J. & Ramshaw, J. A. M. (2017). Engineering specific chemical modification sites into a collagen-like protein from *Streptococcus pyogenes*. JOURNAL OF BIOMEDICAL MATERIALS RESEARCH PART A, 105 (3), pp.806-813. <https://doi.org/10.1002/jbm.a.35957>.

Persistent Link:

<http://hdl.handle.net/11343/292149>