Role for Cell-Surface Collagen of Streptococcus pyogenes in Infections

Aubrey J. Ellison, Felix Dempwolff, Daniel B. Kearns, and Ronald T. Raines*

ABSTRACT: Group A Streptococcus (GAS) displays cell-surface proteins that resemble human collagen. We find that a fluorophore-labeled collagen mimetic peptide (CMP) labels GAS cells but not Escherichia coli or Bacillus subtilis cells, which lack such proteins. The CMP likely engages in a heterotrimeric helix with endogenous collagen, as the nonnatural D enantiomer of the CMP does not label GAS cells. To identify a molecular target, we used reverse genetics to "knock-in" the GAS genes that encode two proteins with collagen-like domains, Scl1 and Scl2, into B. subtilis. The fluorescent CMP labels the cells of these B. subtilis strains. Moreover, these strains bind tightly to a surface of mammalian collagen. These data are consistent with streptococcal collagen forming triple helices with damaged collagen in a wound bed and thus have implications for microbial virulence.

KEYWORDS: collagen, cell-surface collagen, collagen mimetic peptides, Streptococcus pyogenes, group A Streptococcus, microbial virulence

For decades, collagen was thought to be a protein found only in animals. Recently, however, collagen has been discovered on the surface of several microbes, including Streptococcus spp., Bacillus anthracis, Clostridium difficile, Legionella pneumococci, and Burkholderia spp.1-3 The most thoroughly characterized of these microbial collagens are those from group A Streptococcus pyogenes (GAS).

GAS is a Gram-positive bacterium that is responsible for a wide range of ailments, ranging from mild infections such as pharyngitis (i.e., "strep throat") and impetigo (which is a localized skin infection) to serious infections such as rheumatic fever, scarlet fever, and necrotizing fasciitis ("flesh-eating disease").4,5 GAS infections can manifest in a variety of host environments. Accordingly, much research has been done to reveal molecular interactions between the bacterium and host.6-9 Of particular interest in this context are bacterial cell-surface proteins known as adhesins, which are involved in adherence and colonization during infection.9

Pathogenic bacteria like GAS often display variety and redundancy in their adhesins.5,7,10 The "streptococcal collagen-like" adhesins Scl1 and Scl2 of GAS, which were discovered by Björck,11 Musser,12,13 and their co-workers, are no exception. The Scl1 and Scl2 proteins are anchored on the bacterial surface via a membrane-binding domain. Immediately adjacent to that domain is a collagen-like sequence that contains canonical Xaa-Yaa-Gly repeats. Finally, on the N terminus is a variable domain. The collagen-like sequence forms a triple helix that holds the variable domain in a trimeric complex, leading the proteins to adopt a lollypop-like shape (Figure 1).14-16 These Scl proteins have been implicated in binding to a wide variety of factors on the surface of host cells, including fibronectin, laminin, integrins αβ1 and α1β1, thrombin-
activatable fibrinolysis inhibitor, LDL, ApoB, and Factor H.\textsuperscript{3,4,7\textendash}25 In addition to binding to host factors, Scl proteins can play a role in biofilm formation, neutrophil evasion, and protection against host defense peptides.\textsuperscript{3,26,27} Notably, GAS in which Scl1 is knocked out shows diminished virulence in a mouse infection model.\textsuperscript{12}

Known interactions between Scl proteins and host cells occur through the variable domain.\textsuperscript{3} The collagen-like domain of Scl proteins has been thought to act merely as a strut that orients the variable domain outward from the cell surface.\textsuperscript{3,4,14,18,28} We reasoned that the collagen-like sequences of Scl1 and Scl2 could also mediate interactions with the extracellular matrix (ECM) of a host.

Collagen is the major component of the ECM and the most abundant protein in the human body, comprising a third of all extracellular matrix (ECM) of a host.

The thermostability of mammalian collagen is enhanced greatly by the preorganization endowed by its prevalent proline and 4-hydroxyproline residues.\textsuperscript{29,30} In contrast, the collagen domains of Scl1 and Scl2 contain few proline residues and no hydroxylated proline. Their thermostability derives instead from relatively weak noncovalent interactions.\textsuperscript{41,42} Strikingly, the collagen domains of Scl proteins have been thought to act merely as a strut that orients the variable domain outward from the cell surface.\textsuperscript{3,14,18,28} We reasoned that the collagen-like sequences of Scl1 and Scl2 could also mediate interactions with the extracellular matrix (ECM) of a host.

The collagen-like domain of Scl proteins has a T\textsubscript{m} value near 37 °C (where T\textsubscript{m} refers to the temperature at the midpoint of the thermal transition between native and denatured states).\textsuperscript{14,42} Thus, at human body temperature, half of these domains might be single strands that could adhere to damaged collagen.

To test our hypothesis, we used a fluorophore-labeled CMP as a probe for collagen strands on the surface of live GAS cells. To identify the target for cell-surface binding, we expressed the genes that encode \textit{S. pyogenes} Scl1 and Scl2 in \textit{Bacillus subtilis}, which is also a Gram-positive bacterium but lacks a collagen-like protein and has numerous divergencies in its cell envelope.\textsuperscript{43,44} We then probed the surface of these transgenic microbes, as well as their ability to adhere to mammalian collagen. Our findings suggest new strategies for detecting and antagonizing GAS infections.

**RESULTS AND DISCUSSION**

**Scl1 and Scl2 Bind to Collagen Mimetic Peptides.** We synthesized a CMP conjugate to a rhodamine fluorophore. This conjugate, Red-1-CMP, had the sequence Ac-Lys-(Rhodamine Red)-(Ser-Gly)\textsubscript{3}-(L-Pro-L-Pro-Gly)\textsubscript{7}. We chose this sequence because a (L-Pro-L-Pro-Gly)\textsubscript{7} peptide does not form a stable homotrimetric helix (T\textsubscript{m} = 6–7 °C)\textsuperscript{35} but does anneal to damaged collagen.\textsuperscript{31,33,35} We used this conjugate to probe the surface of GAS cells. Specifically, we grew bacterial cells for 24 h and incubated them with Red-1-CMP for 1 h at 37 °C. Then, we washed the cells and visualized them using confocal microscopy. Red fluorescence would be indicative of a binding interaction between CMP and the cell surface. To show that this interaction is unique to GAS, we also evaluated binding to \textit{B. subtilis} and \textit{Escherichia coli}, which is a Gram-negative bacterium that does not produce collagen-like proteins. We found that GAS was the only bacterium that exhibited red fluorescence, indicative of an interaction with Red-1-CMP (Figure 2).

**Figure 2.** Confocal microscopy images of live GAS, \textit{E. coli}, and \textit{B. subtilis} cells that were incubated for 1 h at 37 °C with Red-1-CMP (3 \textmu M) or Red-\textit{d}-CMP (3 \textmu M) or left untreated and were washed (3X). Scale bar: 10 \textmu m.

In addition to the bacterial controls, we also employed a CMP control. If adherence to the cell surface is due to triple-helix formation between Red-1-CMP and a bacterial protein then the enantiomeric \textit{d}-CMP should not label the cell. If, however, the labeling is due to electrostatic or hydrophobic interactions then \textit{d}-CMP should behave the same as l-CMP. Accordingly, we synthesized Red-\textit{d}-CMP, which is Ac-Lys-(Rhodamine Red)-(Ser-Gly)\textsubscript{3}-(\textit{d}-Pro-\textit{d}-Pro-Gly)\textsubscript{7}. We found that GAS cells were not labeled with Red-\textit{d}-CMP (Figure 2). This selectivity for l-CMP over d-CMP is consistent with an interaction that requires collagen triple-helix formation.

**Effect of Temperature.** Our initial conception of adherence by a GAS cell to damaged collagen took into consideration the T\textsubscript{m} value of the collagen-like triple helix of the Scl proteins.\textsuperscript{3,4,28} We reasoned that, at low temperatures, the Scl proteins would be engaged in a triple helix (Figure 1) and therefore would not be able to bind to Red-1-CMP. Nonetheless, the collagen-like domains of Scl1 (50 Xaa-Yaa-Gly repeats) and Scl2 (38 repeats) are much longer than the CMP (7 repeats), and the integrity of their triple helices might not be complete throughout. To assess that integrity on the surface of live cells, we repeated the Red-1-CMP labeling experiment on cells incubated at different temperatures. We observed significant labeling at all temperatures with microscopy (Figure 3). We quantified the binding with cytometry. Again, we found that GAS cells were labeled by Red-1-CMP at all temperatures (Figure 4). These data suggest that segments of the collagen-like domains of the Scl adhesins are not triple-helical on live cells.

**Scl1 and Scl2 B. subtilis Knock-Ins.** GAS has variety of adhesins that enable binding to a host.\textsuperscript{17\textendash}25 Some of these adhesins can bind to collagen.\textsuperscript{46\textendash}50 Knowing that Red-1-CMP binds to GAS cells (Figures 2\textendash}4), we next sought to demonstrate that the binding is due to collagen-like proteins Scl1 and Scl2 rather than to another factor. Our strategy was to isolate these two adhesins and see if the CMP-labeling of GAS cells was recapitulated in another background. Having tested
B. subtilis in which no labeling was observed (Figures 2 and 4), we created inducible B. subtilis knock-ins of Scl1 or Scl2.51 We evaluated the production of Scl1 and Scl2 in B. subtilis by assessing mRNA production with qPCR. Upon induction, we observed a 3.7-fold increase in the expression of Scl1 and a 4.0-fold increase in the expression of Scl2 (Figure 5).

Upon repeating the CMP-labeling experiment with the knock-in B. subtilis cells, we expected to see cells fluoresce red if the interaction is indeed due to the heterologous collagen-like proteins. By detecting labeling with confocal microscopy (Figure 6) and flow cytometry (Figure 7), we did indeed observe labeling by Red-LCMP and more so upon induction. Thus, Scl1 and Scl2 confer the phenotype in B. subtilis cells that was observed in GAS cells. This concurrence is strong evidence that these collagen-like proteins are responsible for the interaction between GAS and a CMP. Again, the near-absence of binding by the enantiomeric Red-DCMP is indicative of the binding by Red-LCMP being due to triple-helix formation.

Scl1 and Scl2 Bind to a Collagen Surface. Although we have shown that the collagen-like proteins on the surface of GAS can interact with a single-stranded CMP, our data do not report on how these proteins interact with a collagen surface, like that of human skin. To evaluate adhesion to a collagen surface, we employed an assay using a crystal violet stain.52,53 This assay was performed by growing cells on a surface of mammalian collagen. The surface was then washed, leaving only the adhered bacteria. Then, the surface was submerged in an aqueous solution of crystal violet (0.1% w/v). The excess
there is an increase in bacterial adherence to the collagen surface. Being responsible for the binding of bacterial cells to a collagen surface, Scl2 knock-in cells to uncoated and collagen-coated wells, as determined by the absorbance of extracted crystal violet at 620 nm. Scl2 knock-ins showed significantly greater adherence to the collagen-coated surface than did wild-type *B. subtilis* (Figure 8). In addition, significant adherence occurred only in wells coated with collagen. Apparently, the enhanced bacterial adherence is due to an interaction between the displayed collagen-like proteins and the collagen surface. As seen with CMP-labeling (Figures 6 and 7), there is an increase in bacterial adherence to the collagen surface upon induction. These data support Scl proteins as being responsible for the binding of bacterial cells to a collagen surface.

**CONCLUSION**

The collagen-like domains of Scl1 and Scl2 on GAS cells can both bind to synthetic collagen mimetic peptides and adhere to mammalian collagen. These attributes could enhance microbial virulence as well as serve as the basis of new means to detect streptococcal cells and antagonize their interaction with host organisms. This deployment of collagen mimetic peptides could also be efficacious with the many other microbes that display collagen on their cell surface.

**MATERIALS AND METHODS**

**General.** Amino acid derivatives and HOBt were from Chem-Impex International (Wood Dale, IL). Fmoc-Gly-2-chlorotrityl resin was from EMD Millipore (La Jolla, CA). The Rhodamine Red-X, succinimidyl ester, 5-isomer was from Thermo Fischer Scientific (Waltham, MA). Nuclei lysis solution and protein precipitation solution were from Promega (Fitchburg, WI). The miRCURY RNA isolation kit was from Qiagen (Hilden, Germany). All other reagents were from Sigma-Aldrich (St. Louis, MO) and were used without further purification.

DMF was dried with a Glass Contour system from Pure Process Technology (Nashua, NH) and passed through an associated isocyanate “scrubbing” column to remove any amines. Water was purified with an Arium Pro system from Sartorius (Göttingen, Germany).

PCR was performed with a 1:30 min extension time (annealing temperature of 58 °C, 30 cycles).

All procedures were performed in air at ambient temperature (∼22 °C) and pressure (1.0 atm) unless indicated otherwise.

**Instrumentation.** Solid-phase peptide synthesis was performed at the University of Wisconsin—Madison Biotechnology Center with a Prelude peptide synthesizer from Protein Technologies (Tucson, AZ). Synthetic peptides were purified by HPLC with a Prominance instrument from Shimadzu (Kyoto, Japan) equipped with a VarioPrep 250/21 C18 column from Macherey—Nagel (Düren, Germany). The molecular mass was determined by matrix-assisted laser desorption/ionization—time-of-flight (MALDI—TOF) mass spectrometry on an α-cyano-4-hydroxycinnamic acid matrix with a Voyager DE-Pro instrument from Thermo Fischer Scientific at the Biophysics Instrumentation Facility of the University of Wisconsin—Madison. Purity analyses were performed with an Acquity UPLC H-Class system from Waters (Waltham, MA) that was equipped with an Acquity photodiode array detector, an Acquity quaternary solvent manager, an Acquity sample manager with a flow-through needle, an Acquity UPLC BEH C18 column (2.1 × 50 mm, 1.7-μm particle size), and Empower 3 software. Cells were imaged with an Eclipse Ti inverted confocal microscope from Nikon (Melville, NY) at the Biochemistry Optical Core of the University of Wisconsin—Madison. Flow cytometry was performed with an LSR Fortessa flow cytometer from BD Biosciences (San Jose, CA) at the Carbone Cancer Center at the University Wisconsin—Madison. qPCR was performed with a QuantStudio 7 instrument from Thermo Fisher Scientific. Absorbance was measured with an Infinite M1000 plate reader from Tecan (Männedorf, Switzerland).

**Peptide Synthesis.** Red-LCMP, which is Ac-Lys-(Rhodamine Red)-(Ser-Gly)3-(Pro-L-Pro-Gly)7, was synthesized as described previously. Red-oCMP, which is Ac-Lys(Rhodamine Red)-(Ser-Gly)3-(o-Pro-d-Pro-Gly)7, was synthesized as follows.

**Ac-Lys-(Ser-Gly)3-(o-Pro-d-Pro-Gly)7.** Using the Fmoc-o-Pro-d-Pro-Gly-OM tripeptide synthesized in solution without chromatography and an Fmoc-d-Pro-OH monomer, Ac-Lys-
Ser-Gly)3-(D-Pro-D-Pro-Gly)7 was synthesized by two additions of monomer followed by six segment-condensations of tripeptide on preloaded Fmoc-Gly-2-chlorotrityl resin (0.19 mmol/g). Fmoc-deprotection was achieved by treatment with piperidine (20% v/v) in DMF. The peptide was cleaved from the resin with 96.5:2.5:1.0 TFA/H2O/TIPSH (5 mL), precipitated from aqueous EtOH (70% v/v) and the tube was inverted gently. The solution was mixed gently. The resulting solution was dissolved in 2.0 mL of DMSO. Rhodamine Red-X, acetomitrile containing 0.1% v/v TFA) to yield 0.5 mg (11%) of Red-LCMP and Red-DCMP was assessed with UPLC as being >95% (Figure S1).

For cytometry, cells were washed, resuspended, and exposed to a laser at 488 and 561 nm. The purity of Red-LCMP and Red-DCMP was assessed with UPLC as being >95% (Figure S1).

**Cell Growth.** Cultures of group A S. pyogenes (SF 370) were grown for 24 h in a brain heart infusion broth at 37 °C. Cultures of E. coli (RP437) were grown for 16 h in LB medium at 37 °C on a platform shaker. Cultures of B. subtilis (OI 1085) were grown overnight in LB medium at 37 °C on a platform shaker. B. subtilis knock-in cultures were grown in LB medium. Knock-in cultures were induced after 4 h by the addition of ITPG to a final concentration of 1.0 mM. All experiments with knock-in cultures were done after 8 h of growth.

**Cell Labeling for Microscopy and Cytometry.** Aliquots (1.0 mL) of bacterial culture were treated with either Red-1-CMP or Red-1CMP to a final concentration of 3.25 µM. Mixtures were allowed to incubate at 37 °C for 1 h. Cultures were pelleted, and the supernatant was removed and resuspended in water. This procedure was repeated three times to wash the cultures. Finally, cells were resuspended in 1.0 mL of water.

For microscopy, cells were spotted on a thin layer of LB agarose on a microscope slide. Images were merged from brightfield and fluorescence (560 nm laser with a 595/50 filter) microscopy. Three biological replicates of each experiment were performed, and recorded images were representative.

For cytometry, cells were washed, resuspended, and exposed to 1.0 µL of a solution of SYTO 9 (5 mM) in DMSO. Cells (50000) were counted by using 488 and 561 nm lasers with S30/30 nm and S86/15 nm filters with a flow cytometer.

**Creation of B. subtilis Knock-Ins of Sc11 and Sc12.** To achieve heterologous expression, the coding sequence for Sc11 (SPy1983) and Sc12 (SPy1054) from the M1 serotype of GAS were integrated at the amyE locus in the B. subtilis chromosome (with translation of the mRNA being ensured by the addition of the tufA ribosome-binding site) and controlled expression with the IPTG-inducible hyspank promoter. The presence of the genes was verified by PCR, and chromosomal integration was confirmed by plating on starch agar. Integration of the construct leads to the disruption of the gene coding for amylose and hence to the loss of the ability to degrade starch.

To isolate genomic DNA, cells from 1.0 mL of a turbid culture were harvested by centrifugation (2 min at maximum rpm). Cells were suspended in 480 µL of 50 mM EDTA (pH 8.0). Lysozyme (60 µL of a 10 mg/mL solution) was added, and the resulting solution was mixed gently. The solution was then incubated for 30 min at 37 °C, and 600 µL of nuclei lysis solution (Promega A7943) was added. The resulting mixture was incubated for 2–3 min at room temperature, 200 µL of protein precipitation solution (Promega A7951) was added, and the mixture was vortexed for 20 s. The mixture was incubated on ice for 5 min and then clarified by centrifugation for 3 min at maximum rpm. A 900 µL aliquot of the supernatant was transferred to a new tube containing 600 µL of isopropl alcohol, and the tube was inverted gently. The resulting precipitate was collected by centrifugation for 2 min at maximum rpm. The pellet was washed with 500 µL of aqueous EtOH (70% v/v) and collected by centrifugation for 1 min at maximum rpm. The dry pellet was suspended in 100 µL of H2O, and the resulting solution was incubated for 30 min at 37 °C. The isolated genomic DNA was stored at −20 °C.

To generate an IPTG-inducible construct of Sc11 (i.e., plasmid pFK1), the spy1983 gene was PCR-amplified from genomic DNA by using primers 5′-TCAGCATGTTAGTTGTTTCTTGAGCTTTCGCT-3′ and 5′-TCAAGCTTTTTAAGGAGTATTTGAGCTACCAAAGCCCAT-AATC-3′ and was digested with HindIII and SpeI. The ensuing fragment was ligated into the HindIII/SpeI sites of plasmid pDP111, which contains the hyspank promoter, the lacI gene encoding the LacI repressor, and a kanamycin-resistance cassette between the arms of the amyE gene. The tufA ribosome-binding site was added to the spy1983 gene within the forward primer to increase its expression, and the resulting construct was transformed into DK1042 B. subtilis cells45 to generate strain DK4458 (amyE::hyspank-RBS<sub>amyE</sub>sc11 kan).

To generate an IPTG-inducible construct of Sc12 (i.e., plasmid pFK3), the spy1054 gene was PCR-amplified from genomic DNA by using primers 5′-TCAGCATGTTAGTTGTTTCTTGAGCTTTCGCT-3′ and 5′-TCAAGCTTTTTAAGGAGTATTTGAGCTACCAAAGCCCAT-AATC-3′ and was digested with HindIII and SpeI. The ensuing fragment was ligated into the HindIII/SpeI sites of pDP111, which contains the hyspank promoter, the lacI gene encoding the LacI repressor, and a kanamycin-resistance cassette between the arms of the amyE gene. The tufA ribosome-binding site was added to the spy1054 gene within the forward primer to increase its expression, and the construct was transformed into DK1042 B. subtilis cells45 to generate strain DK4460 (amyE::hyspank-RBS<sub>amyE</sub>sc12 kan).

**qPCR Analysis of B. subtilis Knock-Ins.** Prior to qPCR, RNA was isolated from cultures using a miRCURY RNA isolation kit, and complementary DNA was prepared from the RNA by reverse transcription using an Applied Biosystems high-capacity cDNA reverse transcription kit. Samples were prepared in duplicate and added to Quanta PerfeCTa SYBR Green Fastmix Low Rox master mix. These samples were amplified by using an Applied Biosystems QuantStudio 7 kit.
Relative gene expression was calculated with the 2\(^{-\Delta\Delta C_T}\) method, normalized against RspJ, RspE, and yoxA.

qPCR primers were designed by using the program Primer3 and are listed in Table S1. cDNA and primers were tested against each other and analyzed by gel electrophoresis on a gel to ensure purity and integrity.

**Bacterial Adherence Assays.** Cultures (2.0 mL) of the B. subtilis knock-ins and wild-type O1 1085 cells were grown in the wells of a 24-well microplate, which were coated with rat-tail collagen (product A1142802 from Thermo Fisher Scientific). Gene expression was induced after 4 h of growth. All cultures were removed from the wells, which were washed three times with water. An aqueous solution (1.0 mL) of crystal violet (0.1% w/v) was added to the wells. After 3 min, the crystal violet solution was removed, and the wells were washed three times with water. Ethanol (500 μL) was added to the wells, and an aliquot (200 μL) from each well was transferred to a 96-well plate; the absorbance at 620 nm was recorded with a plate reader. Data were analyzed with Prism soft ware from GraphPad (San Diego, CA). The significance was assessed by applying an unpaired Student’s t test to paired comparators.

### ASSOCIATED CONTENT

#### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsinfecdis.0c00073.

Table S1 and Figure S1 (PDF)

### AUTHOR INFORMATION

**Corresponding Author**

Ronald T. Raines  —  Department of Chemistry and Department of Biochemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States; Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States;  [orcid.org/0000-0001-7164-1719](https://orcid.org/0000-0001-7164-1719); Email: rtraines@mit.edu

**Authors**

Aubrey J. Ellison  —  Department of Chemistry, University of Wisconsin—Madison, Madison, Wisconsin 53706, United States

Felix Dempwolff  —  Department of Biology, Indiana University, Bloomington, Indiana 47405, United States

Daniel B. Kearns  —  Department of Biology, Indiana University, Bloomington, Indiana 47405, United States

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acsinfecdis.0c00073

### Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

We are grateful to Professor Douglas B. Weibel for advice and contributive discussions. We thank Emily R. Garnett for help with qPCR analysis and Drs. Rachael T. Sheridan, Elle K. Grevestd, and Darrell R. McCaslin for facility support. This work was supported by Grants R56 AR044276 and R35 GM131783 (NIH). A.J.E. was supported by Chemistry–Biology Interface (CBI) Training Grant T32 GM008505 (NIH). The UWCCC Flow Cytometry Laboratory was supported by Grant P30 CA014520 (NIH).

### REFERENCES


(28) In some serotypes, the collagen sequence contains a binding motif for collagen-binding integrins on the host cell surface (see: refs 17, 19, 22, and 24).


