Published in final edited form as:

Methods Mol Biol. 2020 ; 2136: 163–179. doi:10.1007/978-1-0716-0467-0_12.

Expression and Purification of Collagen-Like Proteins of Group A Streptococcus

Slawomir Lukomski¹, Dudley H. McNitt^{2,3}

¹Department of Microbiology, Immunology, and Cell Biology, West Virginia University School of Medicine, Morgantown, WV, USA

²Department of Microbiology, Immunology, and Cell Biology, West Virginia University School of Medicine, Morgantown, WV, USA

³Division of Rheumatology and Immunology, Department of Medicine, Vanderbilt University, Nashville, TN, USA

Abstract

Prokaryotic proteins with extended collagen domain are found in many bacterial species that are pathogenic to humans and animals. The collagen domain is often fused to additional ligand-binding domains and plays both structural and functional roles in modular “bacterial collagens.” Here, we describe the step-by-step expression and purification of the recombinant streptococcal collagen-like proteins, rScI, using the *Strep*-tag II system. The integrity and structural characterization of recombinant collagen-like proteins is very important for defining their function.

Keywords

Streptococcal collagen-like protein; Prokaryotic collagens; Recombinant protein production; Affinity purification; Protein characterization

1 Introduction

The collagen structure has evolved in metazoans as an essential component of extracellular matrix [1], but collagen-like proteins have also emerged in prokaryotes, likely by convergent evolution, and are associated with pathogenesis in numerous bacterial species [2] including pathogenic streptococci [3]. The collagen-like proteins 1 and 2, ScI1 and ScI2 [4–8], have been studied extensively, and ScI1 is involved in Group A Streptococcus pathogenesis, while

slukomski@hsc.wvu.edu.

NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS

The copyright law of the United States [Title 17, United States Code] governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use,” that use may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgement, fulfillment of the order would involve violation of copyright law. No further reproduction and distribution of this copy is permitted by transmission or any other means.

the role of Scl2 is less clear. Both Scl proteins are homotrimeric and contain two functional extracellular domains: the cell-surface adjacent collagen (CL) domain and attached outermost globular sequence-variable V domain. The Scl-CL domain forms a triple helix [9, 10], which can bind human collagen receptors, integrins $\alpha 2\beta 1$ and $\alpha 11\beta 1$ [11, 12]. The Scl1-V domain binds selected extracellular matrix proteins [13], as well as human plasma components [14–17], thus performing important adhesive and immune defense functions [3]. The crystal structure of the V domain has been solved as a six-helix bundle, comprised of two conserved antiparallel α -helices, in each monomer, connected by surface-exposed loops [18]. The Scl1-V loops and characteristically distributed surface potential are critical determinants in ligand binding, and the V domain has to be attached to the triple-helical CL domain for ligand-binding function [19, 20].

Collagens are also desirable biomaterials, used in medicinal, cosmetic, and food industries that are extracted from animal sources by harsh procedures (recently reviewed in [21]). In contrast, designed recombinant Scl proteins, rScls, have been expressed in large quantities in *Escherichia coli* by several laboratories [22–26]. The collagen domain in rScl constructs was modified to support new biological functions by inserting binding motifs for heparin, integrins, fibronectin, or matrix metalloproteinases cleavage sites [27–30]. Those collagen mimetics were prepared in forms of surface coatings, hydrogels, or silk-like chimeras [31–33] to support human cell growth and wound healing [34–36].

Here, we describe the expression and purification of the recombinant rScl proteins, using the *Strep*-tag II system. rScl constructs are expressed in the *E. coli* periplasm and are extracted and purified by affinity chromatography under mild near physiological conditions. We also include two “quality control” techniques to detect the presence of collagen triple helix and characteristic domain organization, as a preliminary assessment of rScl functionality.

2 Materials

Prepare all solutions using ultrapure water. Unless otherwise noted, all reagents can be stored at room temperature. For experiments that come into direct contact with bacteria, proper sterile technique should be followed. We utilize the original *Strep*-tag II system (IBA, Lifesciences) [37] to produce rScl proteins. Currently, additional strategies of cloning, expression, and purification using this system are offered, including 6xHis-tag, Flag-tag, GST-tag, etc.; those will require alternative inducing agents and purification resins, but several steps and notes in the materials and methods section will still apply in the production of rScl proteins described here.

2.1 Protein Expression

2.1.1 General Equipment

1. Erlenmeyer flasks (50 mL for *small-scale inductions*).
2. Sterile inoculation loops.
3. Bunsen burner or open flame source.
4. Vortex mixer.

5. Spectrophotometer.
6. Shaking bacterial culture incubator: 30 °C or 37 °C.
7. Microcentrifuge and standard tabletop centrifuge.
8. Centrifuge tubes/bottles (1.5 mL, 35 mL, and 500 mL).

2.1.2 Bacterial Growth Reagents

1. LB-Miller broth (LB-Miller): 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl. A final concentration of 100 µg/mL ampicillin is added to the broth during protein expression for the selection of the pASK-IBA2 plasmid (IBA Lifesciences). Ampicillin stock (100 mg/mL) is made in sterile water and filtered.
2. LB-Miller agar: 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, 15 g/L agar. A final concentration of 100 µg/mL ampicillin is used for the selection of the pASK-IBA2 vector (IBA Lifesciences).
3. 2 mg/mL anhydrotetracycline (AHTC) stock solution in dimethylformamide (DMF). Store at −20 °C.
4. Frozen bacterial culture: frozen glycerol stocks of *E. coli* colony (ies) harboring clones of *scf* sequences in pASK-IBA2 vector.

2.2 Protein Purification

2.2.1 General Equipment

1. Microcentrifuge or centrifuge.
2. Chromatography column.
3. Support stand with clamps.
4. 15-mL or 50-mL conical tubes.
5. Vortex mixer.
6. 1.5-mL microcentrifuge tube or 35-mL centrifuge tubes.
7. Mini-gel SDS–PAGE running tank.
8. SDS-acrylamide mini-gel casting apparatus.

2.2.2 Affinity Chromatography

1. Buffer P: 100-mM Tris/HCl pH 8.0, 500-mM sucrose, 1-mM ethylenediaminetetraacetic acid (EDTA). High sucrose buffer P is used to release the periplasmic content, including recombinant proteins that are expressed in the periplasm. This approach is referred to as the “buffer P method” throughout the text. The use of buffer P or alternative protein extraction buffer, for example, cell lysis buffer, is determined experimentally.

2. Cell Lytic Buffer: CellLytic B® (Sigma) with 1-mM phenylmethylsulfonyl fluoride (PMSF) and 50 U/mL DNase I. If CellLytic B is not available for purchase, a lab made reagent may be used: 50-mM Tris/HCl pH 8.0, 50-mM NaCl, 2-mM magnesium chloride, 2% Triton X-100, 10-mM β -mercaptoethanol, 0.2 mg/mL lysozyme, 1-mM phenylmethylsulfonyl fluoride (PMSF), 50 U/mL DNase I. This protocol is used if recombinant protein cannot be recovered efficiently with buffer P. This approach is referred to as the “cell lytic method” throughout the text. The use of cell lytic buffer or alternative protein extraction buffer, for example, buffer P, is determined experimentally.
3. 50 U/mL DNase IP, for use in the cell lytic buffer method (see above). Store at 4 °C.
4. 1-mM PMSF, for use in the cell lytic buffer method (see above). Store at 4 °C.
5. Tris–EDTA (TE): 10-mM Tris–HCl, 1-mM EDTA.
6. 2 mg/mL egg white avidin solution, made in sterile water. Store at 4 °C.
7. Strep-Tactin® Sepharose suspension (IBA Lifesciences). Store at 4 °C.
8. Buffer W: 100-mM Tris/HCl pH 8.0, 150-mM NaCl, 1-mM EDTA (*see* Note 1).
9. Buffer E: 100-mM Tris/HCl pH 8.0, 150-mM NaCl, 1-mM EDTA, 2.5-mM desthiobiotin. Store at 4 °C.
10. Buffer R: 100-mM Tris/HCl, 150-mM NaCl, 1-mM EDTA, 1-mM hydroxy-azophenyl-benzoic acid (HABA) pH 8.0. Store at 4 °C.
11. SDS–Polyacrylamide 4% stacking gel: 3.04-mL deionized water, 0.48-mL 1.5-M Tris–HCl pH 8.8, 0.4-mL 40% acrylamide, 0.04-mL 10% SDS, 0.04-mL 10% APS, 2- μ L TEMED.
12. SDS–Polyacrylamide 10% resolving gel: 8-mL deionized water, 2.5-mL 1.5-M Tris–HCl pH 8.8, 2.5-mL 40% acrylamide (37.5:1 acrylamide/bis-acrylamide ratio), 0.1-mL 10% sodium dodecyl sulfate (SDS), 0.1-mL 10% ammonium persulfate (APS), 5- μ L TEMED.
13. SDS–Polyacrylamide 12% resolving gel: 4.3-mL deionized water, 2.5-mL 1.5-M Tris–HCl pH 8.8, 3.0-mL 40% acrylamide, 0.1-mL 10% SDS, 0.1-mL 10% APS, 5- μ L TEMED.
14. SDS Polyacrylamide 15% resolving gel: 3.55-mL deionized water, 2.5-mL 1.5-M Tris–HCl pH 8.8, 3.75-mL 40% acrylamide, 0.1-mL 10% SDS, 0.1-mL 10% APS, 5- μ L TEMED.
15. SDS–PAGE gel running buffer (1 $\times$): 25-mM Tris, 192-mM Glycine, 0.1% SDS.
16. Laemmli sample loading buffer (4 $\times$): 0.25-M Tris–HCl pH 6.8, 8% SDS, 40% glycerol, 20% 2-mercaptoethanol, 0.004% bromophenol blue.

¹-If contaminating bands appear with eluted protein of interest, the higher NaCl concentration Buffer W, up to 700 mM, should be tested to increase the stringency washes. Increasing the NaCl concentration may decrease the target protein yield in the process.

17. 25-mM HEPES pH 7.4.
18. Tris–buffer saline (TBS; 1×): 25-mM Tris–HCl, 0.15-M NaCl.
19. Phosphate-buffered saline (PBS): 9.33-mM potassium phosphate, 136-mM NaCl.

2.3 Protein Characterization

1. UV quartz cuvette, 1-mm path length.
2. Circular dichroism (CD) spectrophotometer (e.g., JASCO 810 CD Spectropolarimeter).
3. Nitrogen tank with connected hosing.
4. 1× PBS (*see* Subheading 2.2, **item 19**).
5. Cuvette cleaning solution: 30% concentrated HCl and 70% ethanol.

3 Methods

3.1 Recombinant Protein Expression and Purification

The following steps outline the expression and purification of rScl proteins using the *Strep*-tag II system, specifically, the protein products of the *scl* genes cloned in pASK-IBA2 vector (IBA Lifesciences) [37]. In this system, the rScl coding sequence is fused to the OmpA signal peptide, which results in the accumulation of recombinant protein in the periplasm of *E. coli* (strain BL21 deficient in Lon protease is recommended); in addition to periplasmic fraction, substantial amounts of rScl proteins could be recovered from the culture supernatant (*see* Note 2) [22]. rScl constructs have also been produced intracellularly, if the *scl* gene was cloned in pASK-IBA3 vector (*see* Note 3).

The protein expression section is broken into three different subsections: 1-mL small-scale induction, 10-mL medium-scale induction, and 1- to 2-L large-scale induction. *Small-scale inductions* serve as a quick screening method to identify the *E. coli* colonies, which produce the recombinant protein of interest. The *10-mL medium-scale inductions* are performed with selected colonies and serve as a method to determine the best conditions for protein expression and purification, such as incubation temperature and protein extraction conditions. *Large-scale inductions* are used to produce large quantities of recombinant protein once the optimal expression and extraction conditions are determined.

²-Recombinant streptococcal collagen-like proteins can be expressed intracellularly (pASK-IBA3 vector); intracellular expression was found to result in a higher yield of the same protein (~20 mg/L) compared to periplasmic expression (~2 mg/L). When expressed in the periplasm (pASK-IBA2 vector), additional amounts of rScl proteins (~5 mg/L) are secreted into the culture medium and can be recovered by precipitation with ammonium sulfate [22]. Extracellular protein fraction precipitated with ammonium sulfate (saturation needs to be established experimentally) and is resuspended in 25 mM Tris–HCl, 1 mM EDTA pH 8.0 buffer before application on the Strep-Tactin sepharose column. rScl proteins produced intracellularly or recovered by precipitation from culture medium maintained triple-helical conformation and lollipoplike organization [22].

³-IBA Lifesciences currently offers numerous additional expression plasmids with alternative affinity tags (e.g., 6xHis-tag and Flag-tag) that will require different inducing agents, extraction methods, and affinity-purification resins. However, in any alternative purification method for recombinant expression, several steps and notes outlined here will still be beneficial.

3.1.1 Protein Expression—Small-Scale Induction

1. Streak frozen glycerol stock of *E. coli* colony(ies) harboring clones of *scI* sequences in pASK-IBA2 vector onto LB-Miller agar with 100 µg/mL ampicillin. Grow overnight at 37 °C.
2. The next day, inoculate 3-mL LB-Miller broth with 100 µg/mL ampicillin, with an overnight culture. Grow overnight at 37 °C with shaking set at 225 rpm (*see* Note 4).

(*Optional*): Frozen stocks can be used to directly inoculate 3 mL of LB-Miller broth with 100 µg/mL ampicillin.
3. The next day, for each bacterial culture being tested, setup two 1.5-mL Eppendorf tubes with 1-mL LB-Miller broth/100 µg/mL ampicillin. Inoculate both tubes with 100 µL of overnight culture. Grow for 2 h at 37 °C, shaking at 225 rpm. One tube will serve as a noninduced control.
4. After 2 h, induce protein expression in one tube of each pair by adding to the culture 1 µL of a 1:10 dilution of AHTC stock in sterile water. Continue to grow both tubes for an additional 2 h at 37 °C, at 225 rpm.
5. After 2 h, pellet the bacteria in the microcentrifuge at $>14,000 \times g$ for 3 min. Discard the supernatant.
6. Suspend cell pellets in 500-µL water.
7. Remove 10–20 µL of bacterial cell suspension, add 4× gel loading buffer (to final 1 ×), and boil samples for 5 min.
8. Resolve proteins by SDS-PAGE to determine if expression of the rScI protein is induced; verify expected rScI molecular weight (the molecular weight varies across rScI constructs originating from different GAS strains) and select a colony which is the best rScI producer (Fig. 1). Depending on experiment, standard or gradient (4–20%) gels were used to assess protein mass. Three commonly used concentrations of 10%, 12%, and 15% acrylamide resolving gels are used, combined with the 4% acrylamide stacking gel (10-mL volume is enough to pour 2 mini-gels).

3.1.2 Protein Expression—10-mL Medium-Scale Induction

1. Streak frozen glycerol stock of *E. coli* colony(ies) harboring clones of *scI* sequences in pASK-IBA2 vector onto LB-Miller agar with 100 µg/mL ampicillin. Grow overnight at 37 °C.
2. The next day, inoculate 3-mL LB-Miller broth with 100 µg/mL ampicillin, with an overnight colony. Grow overnight at 30 °C or 37 °C (*see* Note 5) with shaking set at 225 rpm.

⁴While selecting several of *E. coli* colony(ies) harboring clones of *scI* sequences for sequencing, the same colonies can be screened to test for protein expression, using a 1.5-mL *small-scale induction* protocol. This will enable an initial assessment of protein induction and yields, as well as the size of induced bands.

(*Optional*): Frozen stocks can be used to directly inoculate 3 mL of LB-Miller broth with 100 µg/mL ampicillin.

3. The next day, for each bacterial culture being tested, setup an Erlenmeyer flask with 50 mL of LB-Miller broth with 100 µg/mL ampicillin. Inoculate the broth with 1-mL overnight culture. Grow at 30 °C or 37 °C, shaking at 225 rpm, until the OD₆₀₀ reaches 0.3–0.5.

4. Once culture reaches desired OD₆₀₀, add AHTC to culture (1 µL of AHTC stock per 10-mL culture). Grow an additional 2–3 h at 30 °C or 37 °C.

(*Optional*): 10 mL of culture can be transferred to a 35-mL sterile centrifuge tube, prior to AHTC addition, to serve as a noninduced control.

5. After 2- to 3-h incubation, transfer 10 mL of culture to a 35-mL centrifuge tube. Pellet bacterial cells by centrifugation for 10 min at 6500 × *g*, discard supernatant.

6. Two different rScl extraction methods can be utilized at this point: **Steps 7–10** use the buffer P method and **steps 11–15** use the cell lytic method. Both protocols should be initially tested to establish, which method results in greater yield of soluble protein (Figs. 2 and 3).

7. Suspend cell pellet in 100-µL buffer P.

8. Incubate on ice for 15 min.

9. Microcentrifuge for 5 min at >14,000 × *g*.

10. Remove the supernatant and suspend the cellular debris in 500-µL TE buffer; save samples for analysis.

11. Freeze cell pellet for at least 30 min at –20 °C. Pellets can be frozen overnight as well.

12. Thaw pellet and suspend bacterial cells in 200 µL of CellLytic B®.

13. Vortex seven times (7 cycles of 1 min on and 1 min off).

14. Centrifuge 10 min at 6500 × *g*.

15. Remove the supernatant and suspend the cellular debris in 500-µL TE buffer; save samples for analysis.

16. Analyze 10–15 µL of the supernatant and 2–5 µL of the cellular debris from both methods by SDS–PAGE.

⁵ 10-mL medium-scale induction protocols are used to test for the optimal conditions of protein expression. Variables for subsequent production typically include the optimal expression temperature (30 °C or 37 °C) and extraction method (buffer P vs. cell lytic B). Both temperatures and extraction buffers should be tested to establish the optimal condition (Figs. 2 and 3). Lower induction temperatures are recommended when recombinant proteins tend to aggregate.

3.1.3 Protein Expression—Large-Scale Induction

1. Streak frozen glycerol stock of *E. coli* colony(ies) harboring clones of *scf* sequences in pASK-IBA2 vector that were selected in protocols onto LB-Miller agar with 100 µg/mL ampicillin. Grow overnight at 37 °C.
2. The next day, inoculate 50 mL of LB-Miller broth with 100 µg/mL ampicillin with overnight culture. Grow overnight at 30 °C or 37 °C (conditions established with *10-mL induction protocol*, see Note 5) with shaking set at 225 rpm.
(*Optional*): Frozen stocks can be used to directly inoculate 50 mL of LB-Miller broth with 100 µg/mL ampicillin.
3. The next day, inoculate 1–2 L of LB-Miller/100 µg/mL ampicillin, with 10 mL of the overnight culture per 1 L of medium. Grow at 30 °C or 37 °C, shaking at 225 rpm, until the optical density (OD₆₀₀) reaches 0.5–0.8.
4. Once culture reaches desired OD₆₀₀, add AHTC to culture (100 µL of AHTC stock per 1-L culture). Grow for additional 3 h at 30 °C or 37 °C.
5. After 3 h of induction, transfer culture to 500-mL centrifuge bottles. Centrifuge culture sample at 6500 × *g*, for 15 min, at 4 °C, to pellet bacterial cells, discard culture supernatant.
6. Pending the preferred protein extraction method, two different approaches can be used to extract protein: **steps 7–16** describe the buffer P method and **steps 17–25** describe the cell lytic method.
7. Buffer P method: Prewarm the PMSF stock (small crystals typically form at –20 °C), shaking gently for 5 min at RT.
8. Suspend cell pellets in 5 mL of buffer P per each 500 mL of original culture.
9. Combine two samples of cell suspensions (10 mL per 1 L of original culture) into a 35-mL centrifuge tube; add 100-µL prewarmed PMSF stock and 100-µL DNase I per each 10 mL of cell suspension.
10. Incubate on ice for 10 min.
11. Remove from ice and gently rock at RT for 15 min.
12. Centrifuge for 10 min at 17,500 × *g*.
13. Save the supernatant for purification. Suspend the cellular debris in 3–5 mL TE buffer. Store cell debris at 4 °C or –20 °C for analysis by SDS–PAGE (see Note 7).
14. Add 100 µL of avidin stock, per 10-mL sample, to the supernatant. Incubate for 1 h at 4 °C (can be overnight) (see Note 8).

⁷ Substantial amounts of rScl protein may still be in the cellular debris, based on SDS–PAGE analysis. In such case, centrifuge the sample at 15,000 × *g* for 10 min and discard supernatant. Treat pellet using the cell lytic method to extract proteins from cellular debris.

⁸ Egg white avidin is added to bind and mask biotinylated *E. coli* proteins, as well as biotin, that would interfere with purification on Strep-Tactin® sepharose resin (which also binds biotin and biotinylated proteins). The avidin-precipitate fraction may contain rScl

15. Centrifuge sample at $17,500 \times g$ for 15 min at 4 °C. A small white pellet may be present after centrifugation. Transfer supernatant to a 15- or 50-mL conical tube. If avidin pellet is present, dissolve it in 500 μ L of 10 mM Tris-HCl, pH 8.0.
16. Use supernatant during *Affinity Purification* (Subheading 3.2).
17. Cell lytic method: Freeze bacterial cell pellet for at least 1 h at -20 °C, or overnight.
18. Thaw cells and suspend each pellet in 10-mL Cell Lytic B®.
19. Add 100- μ L PMSF and 100 μ L-DNase I, per 500-mL bottle.
20. Combine all pellets into a 35-mL centrifuge tube. Shake gently at RT for 15 min.
21. Centrifuge at 4 °C, $17,500 \times g$, for 10 min.
22. Save the supernatant for the next step (**step 23**). Suspend the cellular debris in 3- to 5-mL TE buffer. Store cell debris at 4 °C or -20 °C.
23. Add 100 μ L of avidin, per 500-mL bottle, to the supernatant. Incubate with avidin for 1 h at 4 °C (can be overnight) (*see* Note 8).
24. Centrifuge sample at $17,500 \times g$ for 15 min at 4 °C. A small white pellet may be present after centrifugation. Transfer supernatant to a 15- or 50-mL conical tube. If avidin pellet is present, dissolve in 500 μ L 10 mM Tris-HCl, pH 8.0.
25. Use supernatant in **step 7**, *Affinity Purification* (Subheading 3.2).

3.2 Protein Purification–Affinity Purification

After isolation of the supernatant fraction, prepare the Strep-Tactin® sepharose column for affinity purification.

1. Add 1-mL Strep-Tactin® sepharose slurry to the column (*see* Note 9).
2. Equilibrate resins with 2 column volume Buffer W (*see* Note 1).
3. Apply 10–15 mL of the final processed supernatant, after either buffer P or cell lytic method, to the column, collecting the flow through (*see* Note 10).
4. Wash the column 5 $\times$ with 1 column volume Buffer W (*see* Note 1). Collect, at least, the first and last washes in Eppendorf tubes for SDS–PAGE analysis to assess if rScl protein is being washed off of the column.
5. Elute the protein with Buffer E, collecting each elution fraction. Elution volumes need to be optimized to maximize protein concentrations and volumes vary between rScl constructs (Fig. 4) (*see* Note 11). Standard starting elution pattern is 6 elution fractions, 500 μ L each.

protein, which can be recovered on the Strep-Tactin® column. Addition of the avidin fraction can be done after all elutions, washes, and flow through fractions have been processed and analyzed by SDS–PAGE.

⁹Depending on the level of rScl expression, a volume different than standard 1 mL of Strep-Tactin sepharose per 1 L of culture may be used for packing the column.

¹⁰Similarly to Note 9, the volume of the supernatant added to the column may be adjusted to minimize protein in the flow through.

6. Save the collected elutions, washes, and flow through for analysis by SDS–PAGE.
7. Regenerate the Strep-Tactin® resin by washing the column 3× with 5 column volumes of Buffer R. Column will turn from white to red after the washes.
8. Remove Buffer R via washing 3 × with 5 column volumes of Buffer W. Column should turn from red to white.
9. Store column in 2 column volumes of Buffer W, at 4 °C, until future use.
10. Analyze equal volumes of each collected samples by SDS–PAGE to assess yield of rScl protein in elutions, amount left in flow through, washes, and protein presence in avidin pellet (Fig. 4).
11. Repeat purification **steps 1–9**, using the remaining supernatant sample(s). Purify also samples obtained by combining those flow through washes. Dissolved avidin fractions, which can contain rScl (determined via SDS–PAGE analysis), can be run after remaining sample and flow through are processed.
12. After all fractions have been run on affinity column and analyzed by SDS–PAGE, combine fractions with similar band intensities (protein concentrations). Typically, there should be 2–3 different final samples (a high, medium, and low concentration).
13. Once all elutions are combined, dialyze the proteins for either immediate usage or storage. Use a 3000–7000 low-molecular-weight cutoff membrane because rScl proteins are not globular. Dialyze proteins for 36–48 h at 4 °C.

Immediate use: proteins are typically dialyzed against either 1× TBS pH 7.4 or 1× PBS pH 7.5, when used in either ELISA-binding assays, circular dichroism, or other assays. See Subheadings 3.3 (*see* Note 12).

Storage: proteins are stored in 25-mM HEPES pH 7.4. After dialysis, proteins are aliquoted into working volumes (50–500 µL) and then frozen in liquid nitrogen (*see* Note 13). Frozen samples are stored at –20 °C.

3.3 Protein Characterization (Quality Control)

Circular dichroism (CD) spectroscopy (reviewed in [38]) was used to detect the collagen triple helix in rScl proteins [9, 22]. This is indicated by the presence of molar ellipticity maximum at 220 nm [39] at 25 °C, which is lost at 50 °C when the collagen triple-helix unfolds (Fig. 5). In addition to CD spectroscopy, rScl proteins are characterized by electron microscopy of rotary shadowed protein samples. Rotary shadowing is a technique which

¹¹Different rScl constructs will elute differently. Some rScl variants will still remain on the column following standard elution protocol, thus, requiring more than standard volume of elution buffer E (Fig. 4(i)), whereas other rScls elute completely within elutions 1–2 (Fig. 4(ii)). The elution volumes are typically altered after the initial standard purification and adjusted until a set condition is established.

¹²When not in immediate use, proteins should be kept at 4 °C for no longer than 1 week.

¹³If an apparatus for freezing tubes with liquid nitrogen is not readily available, an alternative is to freeze samples using a Dewar. Load samples into a storage box and then lower box into liquid nitrogen until it stops bubbling. Bring samples back out of nitrogen and then immediately store at –20 °C.

shows the domain organization of proteins (reviewed in [40]). rScl proteins have a characteristic “lollipop-like” two-domain organization, with the globular V-domain head and CL-domain stalk [9, 22] (Fig. 6). Outlined below is a general protocol for the preparation and conditions used in CD spectroscopy and rotary shadowing of rScl proteins.

3.3.1 Circular Dichroism Spectroscopy

1. Recombinant streptococcal collagen-like proteins should be dialyzed against 1× PBS for 24–48 h at 4 °C (see Note 14).
2. Set up the CD Spectrophotometer to perform a wavelength scan, from 240 nm to 200 nm, using the following conditions: pitch: 0.2 nm, mode: continuous, speed: 10 nm/min, response: 1 s, bandwidth: 1 nm, accumulation: 5, nitrogen flow rate: 3–5 λ/min, temperature: 25 °C.
3. Fill UV quartz cuvette with sample and place into the cell of CD spectrophotometer. Start measurement readings.
4. After performing a wavelength scan at 25 °C, use the same settings on the instrument, except change the temperature to 50 °C to assess loss of the collagen triple-helix structure (Fig. 5). This is important because, depending on the contribution of noncollagen domains within the collagen-like protein, the ellipticity maximum at 220 nm may not be well defined.
5. Remove cuvette from the instrument, discard sample, and clean cuvette by repeating rinses with the cleaning solution.
6. Dry the cuvette using a low flow of 100% nitrogen gas until all liquid is removed from the cuvette.

Acknowledgments

We are grateful to all previous and current members of Lukomski’s laboratory for their contributions to the development of the methods and protocols presented here. We also thank our numerous collaborators, especially Doug Keene for performing rotary shadowing and EM imaging over the years. This work was supported by NIH grants AI50666 and AI083683, as well as West Virginia University HSC Bridge Grant Funding to S.L.

References

1. Exposito JY, Valcourt U, Cluzel C et al. (2010) The fibrillar collagen family. *Int J Mol Sci* 11(2):407–426. 10.3390/ijms11020407 [PubMed: 20386646]
2. Doxey AC, McConkey BJ (2013) Prediction of molecular mimicry candidates in human pathogenic bacteria. *Virulence* 4(6):453–466. 10.4161/viru.25180 [PubMed: 23715053]
3. Lukomski S, Bachert BA, Squeglia F et al. (2017) Collagen-like proteins of pathogenic streptococci. *Mol Microbiol* 103(6):919–930. 10.1111/mmi.13604 [PubMed: 27997716]
4. Lukomski S, Nakashima K, Abdi I et al. (2000) Identification and characterization of the scl gene encoding a group A *Streptococcus* extracellular protein virulence factor with similarity to human collagen. *Infect Immun* 68 (12):6542–6553. 10.1128/iai.68.12.6542-6553.2000 [PubMed: 11083763]

¹⁴The correct buffer is key, as certain commonly used buffers have lower wavelength limits than others, and some buffers will result in background signals which mask the CD spectrum of the rScl protein. The 1 × PBS outlined has a wavelength limit of 200 nm, enabling the evaluation of the collagen triple helix at 220 nm.

5. Lukomski S, Nakashima K, Abdi I et al. (2001) Identification and characterization of a second extracellular collagen-like protein made by group A *Streptococcus*: control of production at the level of translation. *Infect Immun* 69(3):1729–1738. 10.1128/iai.69.3.1729-1738.2001 [PubMed: 11179350]
6. Rasmussen M, Björck L (2001) Unique regulation of SclB - a novel collagen-like surface protein of *Streptococcus pyogenes*. *Mol Microbiol* 40(6):1427–1438 [PubMed: 11442840]
7. Rasmussen M, Edén A, Björck L (2000) SclA, a novel collagen-like surface protein of *Streptococcus pyogenes*. *Infect Immun* 68 (11):6370–6377. 10.1128/iai.68.11.6370-6377.2000 [PubMed: 11035747]
8. Whatmore AM (2001) *Streptococcus pyogenes* sclB encodes a putative hypervariable surface protein with a collagen-like repetitive structure. *Microbiology* 147(Pt 2):419–429. 10.1099/00221287-147-2-419 [PubMed: 11158359]
9. Xu Y, Keene DR, Bujnicki JM et al. (2002) Streptococcal Scl1 and Scl2 proteins form collagen-like triple helices. *J Biol Chem* 277 (30):27312–27318. 10.1074/jbc.M201163200 [PubMed: 11976327]
10. Xu C, Yu Z, Inouye M et al. (2010) Expanding the family of collagen proteins: recombinant bacterial collagens of varying composition form triple-helices of similar stability. *Biomacromolecules* 11(2):348–356. 10.1021/bm900894b [PubMed: 20025291]
11. Caswell CC, Barczyk M, Keene DR et al. (2008) Identification of the first prokaryotic collagen sequence motif that mediates binding to human collagen receptors, integrins $\alpha 2\beta 1$ and $\alpha 11\beta 1$. *J Biol Chem* 283 (52):36168–36175 [PubMed: 18990704]
12. Humtsoe JO, Kim JK, Xu Y et al. (2005) A streptococcal collagen-like protein interacts with the $\alpha 2\beta 1$ integrin and induces intracellular signaling. *J Biol Chem* 280(14):13848–13857 [PubMed: 15647274]
13. Caswell CC, Oliver-Kozup H, Han R et al. (2010) Scl1, the multifunctional adhesin of group A *Streptococcus*, selectively binds cellular fibronectin and laminin, and mediates pathogen internalization by human cells. *FEMS Microbiol Lett* 303(1):61–68. 10.1111/j.1574-6968.2009.01864.x [PubMed: 20002194]
14. Caswell CC, Han R, Hovis KM et al. (2008) The Scl1 protein of M6-type group A *Streptococcus* binds the human complement regulatory protein, factor H, and inhibits the alternative pathway of complement. *Mol Microbiol* 67 (3):584–596. 10.1111/j.1365-2958.2007.06067.x [PubMed: 18093091]
15. Han R, Caswell CC, Lukomska E et al. (2006) Binding of the low-density lipoprotein by streptococcal collagen-like protein Scl1 of *Streptococcus pyogenes*. *Mol Microbiol* 61 (2):351–367. 10.1111/j.1365-2958.2006.05237.x [PubMed: 16856940]
16. Pålman LI, Marx PF, Mörgelin M et al. (2007) Thrombin-activatable fibrinolysis inhibitor binds to *Streptococcus pyogenes* by interacting with collagen-like proteins A and B. *J Biol Chem* 282(34):24873–24881. 10.1074/jbc.M610015200 [PubMed: 17553807]
17. Reuter M, Caswell CC, Lukomski S et al. (2010) Binding of the human complement regulators CFHR1 and factor H by streptococcal collagen-like protein 1 (Scl1) via their conserved C termini allows control of the complement cascade at multiple levels. *J Biol Chem* 285(49):38473–38485. 10.1074/jbc.M110.143727 [PubMed: 20855886]
18. Squeglia F, Bachert B, De Simone A et al. (2014) 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* 289 (8):5122–5133. 10.1074/jbc.M113.523597 [PubMed: 24356966]
19. McNitt DH, Choi SJ, Allen JL et al. (2019) Adaptation of the group A *Streptococcus* adhesin Scl1 to bind fibronectin type III repeats within wound-associated extracellular matrix: implications for cancer therapy. *Mol Microbiol* 112(3):800–819. 10.1111/mmi.14317 [PubMed: 31145503]
20. McNitt DH, Choi SJ, Keene DR et al. (2018) Surface-exposed loops and an acidic patch in the Scl1 protein of group A *Streptococcus* enable Scl1 binding to wound-associated fibronectin. *J Biol Chem* 293(20):7796–7810. 10.1074/jbc.RA118.002250 [PubMed: 29615492]
21. Golser AV, Scheibel T (2018) Routes towards novel collagen-like biomaterials. *Fibers* 6 (2):21 10.3390/fib6020021

22. Han R, Zwiefka A, Caswell CC et al. (2006) Assessment of prokaryotic collagen-like sequences derived from streptococcal Scl1 and Scl2 proteins as a source of recombinant GXY polymers. *Appl Microbiol Biotechnol* 72 (1):109–115. 10.1007/s00253-006-0387-5 [PubMed: 16552563]
23. Peng YY, Howell L, Stoichevska V et al (2012) Towards scalable production of a collagen-like protein from *Streptococcus pyogenes* for biomedical applications. *Microb Cell Factories* 11:146–146. 10.1186/1475-2859-11-146
24. Peng YY, Stoichevska V, Madsen S et al. (2014) A simple cost-effective methodology for large-scale purification of recombinant non-animal collagens. *Appl Microbiol Biotechnol* 98 (4):1807–1815. 10.1007/s00253-013-5475-8 [PubMed: 24402415]
25. Cosgriff-Hernandez E, Hahn MS, Russell B et al. (2010) Bioactive hydrogels based on designer collagens. *Acta Biomater* 6 (10):3969–3977. 10.1016/j.actbio.2010.05.002 [PubMed: 20466083]
26. Yoshizumi A, Fletcher JM, Yu Z et al. (2011) Designed coiled coils promote folding of a recombinant bacterial collagen. *J Biol Chem* 286(20):17512–17520. 10.1074/jbc.M110.217364 [PubMed: 21454493]
27. Peng YY, Stoichevska V, Schacht K et al. (2014) Engineering multiple biological functional motifs into a blank collagen-like protein template from *Streptococcus pyogenes*. *J Biomed Mater Res A* 102(7):2189–2196. 10.1002/jbm.a.34898 [PubMed: 23913780]
28. Yu Z, An B, Ramshaw JA et al. (2014) Bacterial collagen-like proteins that form triple-helical structures. *J Struct Biol* 186(3):451–461. 10.1016/j.jsb.2014.01.003 [PubMed: 24434612]
29. Yu Z, Visse R, Inouye M et al. (2012) Defining requirements for collagenase cleavage in collagen type III using a bacterial collagen system. *J Biol Chem* 287(27):22988–22997. 10.1074/jbc.M112.348979 [PubMed: 22573319]
30. Seo N, Russell BH, Rivera JJ et al. (2010) An engineered alpha1 integrin-binding collagenous sequence. *J Biol Chem* 285 (40):31046–31054. 10.1074/jbc.M110.151357 [PubMed: 20675378]
31. Bronk JK, Russell BH, Rivera JJ et al. (2014) A multifunctional streptococcal collagen-mimetic protein coating prevents bacterial adhesion and promotes osteoid formation on titanium. *Acta Biomater* 10(7):3354–3362. 10.1016/j.actbio.2014.04.005 [PubMed: 24732634]
32. An B, DesRochers TM, Qin G et al. (2013) The influence of specific binding of collagen-silk chimeras to silk biomaterials on hMSC behavior. *Biomaterials* 34(2):402–412. 10.1016/j.biomaterials.2012.09.085 [PubMed: 23088839]
33. Parmar PA, St-Pierre JP, Chow LW et al. (2016) Harnessing the versatility of bacterial collagen to improve the chondrogenic potential of porous collagen scaffolds. *Adv Healthc Mater* 5(13):1656–1666. 10.1002/adhm.201600136 [PubMed: 27219220]
34. Peng YY, Yoshizumi A, Danon SJ et al. (2010) A *Streptococcus pyogenes* derived collagen-like protein as a non-cytotoxic and non-immunogenic cross-linkable biomaterial. *Biomaterials* 31(10):2755–2761. 10.1016/j.biomaterials.2009.12.040 [PubMed: 20056274]
35. Browning MB, Dempsey D, Guiza V et al. (2012) Multilayer vascular grafts based on collagen-mimetic proteins. *Acta Biomater* 8 (3):1010–1021. 10.1016/j.actbio.2011.11.015 [PubMed: 22142564]
36. Cereceres S, Touchet T, Browning MB et al. (2015) Chronic wound dressings based on collagen-mimetic proteins. *Adv Wound Care (New Rochelle)* 4(8):444–456. 10.1089/wound.2014.0614 [PubMed: 26244101]
37. Maier T, Drapal N, Thanbichler M et al. (1998) Strep-tag II affinity purification: An approach to study intermediates of Metalloenzyme biosynthesis. *Anal Biochem* 259(1):68–73. 10.1006/abio.1998.2649 [PubMed: 9606145]
38. Greenfield NJ (2007) Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* 1:2876 10.1038/nprot.2006.202
39. Brodsky-Doyle B, Leonard KR, Reid KB (1976) Circular-dichroism and electron-microscopy studies of human subcomponent C1q before and after limited proteolysis by pepsin. *Biochem J* 159(2):279–286. 10.1042/bj1590279 [PubMed: 187173]
40. Sherratt MJ, Graham HK, Kielty CM et al. (2000) ECM macromolecules: rotary shadowing and scanning transmission electron microscopy In: Streuli CH, Grant ME (eds) *Extracellular Matrix Protocols*. Humana Press, Totowa, NJ, pp 119–132. 10.1385/1-59259-063-2:119

Fig. 1. Screening of *E. coli* colonies for the production of rScl proteins. Example of 12% SDS-PAGE representing the small-scale induction of rScl expression. Three different *E. coli* colonies (lanes 1–2, 3–4, and 5–6), expressing a potential rScl construct, were grown at 37 °C and protein production was compared. Both noninduced (lanes 1, 3, 5) and induced (lanes 2, 4, 6) samples were analyzed on a gel. The induced protein band of expected 31-kDa molecular weight is marked

Fig. 2.

Impact of incubation temperature on rScl expression. In a 10-mL induction protocol, the same *E. coli* colony was grown either at 30 °C (lanes 1–3) or 37 °C (lanes 4–6). Bacterial cells were treated with buffer P and samples were analyzed by 12% SDS–PAGE. Lanes 1 and 4 are noninduced samples; lanes 2 and 5 were loaded with supernatant fractions and lanes 3 and 6 with cell debris fractions. The induced protein band of 23-kDa molecular weight is shown

Fig. 3.

Comparison of protein extraction methods using buffer P and cell lytic buffer. 10-mL cultures of two different *E. coli* colonies (colony 1, lanes 1–6; colony 2, lanes 7–8) were grown at 37 °C, and cells, noninduced or induced with AHTC, were processed using the buffer P (lanes 3–4, 7–8) or cell lytic B (lanes 1–2, 5–6) methods (*see* Note 6). Samples were analyzed by 12% SDS–PAGE. Samples in lanes 1 and 2 show noninduced controls; lanes 3–8 show samples induced for rScl expression. Lanes 1, 3, 5, and 7 show supernatant fraction, and lanes 2, 4, 6 and 8 show cell debris fractions. Lane 9 shows a previously purified rScl preparation

⁶. We recommend that three 10-mL samples are tested at the same time: (1) noninduced, (2) induced and processed with buffer P, and (3) induced and processed with cell lytic B.

Fig. 4. 4–20% SDS–PAGE of two distinct elution patterns represented by rScI constructs 1 and 2, following affinity chromatography. Supernatants from a 1-L induction protocol were loaded onto Strep-Tactin sepharose, washed, and eluted in 500- μ L fractions. (i) rScI construct 1: (lane 1) flow through, (lane 2) wash, (lanes 3–8) elutions 1–6, (lane 9) purified rScI marker. (ii) rScI construct 2: (lane 1) flow through, (lanes 2–6) elutions 1–5, (lane 7) example of avidin pellet fraction showing the presence of rScI protein. Left lane, protein size marker [kDa]

Fig. 5. Circular dichroism (CD) spectra of an rSc11 protein derived from an M1 strain. Wavelength scan showing the collagen triple helix present at 25 °C (solid line—folded conformation) and lacking at 50 °C (dashed line—unfolded conformation); dotted vertical line shows 220_{nm}-maximum molar ellipticity, a characteristics of collagen triple helix. (This research was originally published in the *Journal of Biological Chemistry*. Xu Y, Keene DR, Bujnicki JM, Höök M, Lukomski S. Streptococcal Sc11 and Sc12 proteins form collagen-like triple helices. *J Biol Chem*. 2002; 277:27312-8

Fig. 6. Electron microscopy images of the rotary shadowed samples of rScI1 and rScI2 constructs. Purified rScI constructs maintain shared “lollipop-like” two-domain organization. Bars: 50 nm