Přírodovědecká fakulta Jihočeská univerzita v Českých Budějovicích

Faculty of Science University of South Bohemia in České Budějovice

Expression and purification of adhesive recombinant proteins, sericin 2 and salivary gland secretion 3

Bachelor's thesis

Trang Thanh Vu

Supervisor: Žurovec Michal, prof. RNDr. CSc.

České Budějovice, Czech Republic

2019

Vu, T., 2019: Expression and purification of adhesive recombinant proteins, sericin-2 and salivary gland glue 3. Bc. Thesis, in English. – 39p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation: Recombinant proteins derived from *Bombyx mori Ser-2* gene and *Drosophila melanogaster Sgs3* gene were expressed in bacterial expression systems, purified and partially characterised. The identity of the expressed recombinant proteins was verified by mass spectroscopy. The recombinant proteins were tested for their ability to coat hydrophobic surfaces and sequentially serve as a substrate for the attachment of cells in tissue culture. The quality of the recombinant protein surface coating was examined by scanning electron microscopy. The results show that further optimisation for the purification, solubilization and refolding of these recombinant proteins is needed to incorporate their potential as biomaterials in the future.

I hereby declare that I have worked on my bachelor's thesis independently and used only the sources listed in the bibliography. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full / in shortened form resulting from deletion of indicated parts to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defence in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

In Ceske Budejovice on 11.12.2019 ....

Trang Thanh Vu

Acknowledgment

I would like to thank my supervisor prof. Michal Žurovec and everyone working in his laboratory in especially Yu-Hsien Lin, M.Sc and Brož Václav, Ph.D. for teaching me all the procedures and knowledge behind them, as well as discussing the results with me. Thank you for providing the proteins I was able to work with all the additional information. I had a lot of fun working and talking with everyone, and I was glad I was always welcome to ask the many questions I had.

I would like to give thanks to Tereza Konikova together with Hana Sehadova who helped with Electron microscopy.

I would also like to thank my friends and family for encouraging me and supporting me through this journey.

Table of Contents

1. INTRODUCTION 1

1.1. Bioadhesives 1
1.2. Applications 2
1.3. Adhesive proteins from insects 2
1.4. Production of recombinant proteins 4

2. AIM OF THE THESIS 6

3. MATERIALS 7

3.1. Recombinant protein sequences 7
3.2. Expression system 7
3.3. Growth Media and Chemicals 8

4. METHODS 9

4.1. Plasmid transformation of component bacterial cells by heat shock 9
4.2. Expression of Recombinant Proteins 9
4.3. Purification of Recombinant Protein from Cytosol 10
4.4. Purification of Recombinant Protein from Inclusion Bodies 11
4.5. Dialysis 11
4.6. Lyophilization and final treatment 12
4.7. Protein detection 12
4.8. Determination of recombinant protein concentrations 13
4.9. Characterization of recombinant proteins 14

5. RESULTS 16

5.1. Pilot experiments for expression of various recombinant proteins 16
5.2. Expression and purification of recombinant proteins 20
5.3. Cell Culture Tests 27
5.4. SEM analysis 32

6. DISCUSSION 35

6.1. Protein expression and purification 35
6.2. Properties of Ter1 and Sgs3 36
6.3. Application of adhesive recombinant proteins 37

7. CONCLUSION 39

7.1. Optimising the expression of adhesive recombinant proteins in E. coli 39
7.2. Purification optimisation of recombinant Bombyx sericin 2 (Ter1) and recombinant Drosophila Salivary gland secretion 3 (Sgs3) 39
7.3. Preliminary characterization of isolated recombinant proteins 39

8. REFERENCES 40

9. APPENDIX 45

9.1. Appendix 1 – Additional Materials 45
9.2. Appendix 2 – Additional results 49

Abbreviations:

AA	Acrylamide
APS	Ammoniumperoxodisulphate
BSA	Bovine serum albumin
CBB	Coomassie brilliant blue
EB	Elution Buffer
IBB	Inclusion Body Buffer
IPTG	isopropyl-β-D-thiogalactopyranosid
LB	Luria-Bertani medium
MW	Molecular Weight
RT	Room temperature
SD	Standard deviation
SDS – PAGE	Sodium dodecyl sulphate – polyacrylamide gel electrophoresis
SEM	Scanning Electron Microscope
SEM	Standard Error of Mean
Ser-2	sericin 2
WB	Washing Buffer

1. Introduction

Adhesives are substances capable of holding at least two surfaces together in a strong and permanent manner. Sealants are substances capable of attaching to at least two surfaces, thereby, filling the space between them to provide a barrier or protective coating (Petrie 2007). Adhesives hold substrates together under the desired end-use conditions. More than 30 chemical families of polymers and resins are used to formulate adhesives and sealants. With the increasing emphasis on environmental sustainability and biocompatibility, there has been a great interest in the development of biological adhesives. Biological adhesives are natural polymeric substances predominantly composed of proteins and carbohydrates. Apart from being biocompatible and less harmful to the environment, adhesive proteins pose other advantageous properties in the development of new materials in various areas.

1.1. Bioadhesives

The incorporation of biological products such as proteins for the assembly of materials and other uses present diverse advantages due to their naturally versatile properties as well as customisable physical, chemical and functional properties (Greer et al., 2009). Biomaterials with adhesive properties, known as bioadhesives, are produced by more than 100 natural organisms for purposes such as defence, attachment and construction (Richter et al., 2018). Examples of such organisms include common insects such as silkworms (*Bombyx mori)* and fruit flies (*Drosophila melanogaster)*, as well as moths and butterflies (*Lepidoptera*) (Yonemura et al., 2006). In seas, organisms like marine mussels (*Mytilus edulis*) also produce adhesive proteins (Kord Forooshani et al., 2017).

The use of adhesive biomaterials has become increasingly popular for its universal applications in cosmetics, food and medicine as some exhibit antibacterial and tensile characteristics that allow the attachment on practically any substrates (Richter et al. 2018). This includes the application as tissue adhesives to aid wound closure and healing, drug delivery and implantation of medical devices as well as tissue engineering and dental applications (van Byern and Grunwald, 2010).

In contrast to synthetic adhesives which often require the use of hazardous chemicals and contribute to the production of toxic waste, bioadhesives tend to raise less concerns regarding the environment, health and economic aspects (Richter et al., 2018). Hence, natural

bioadhesives pose to be commercially attractive due to their biodegradability and biocompatibility working at a variety of conditions (Richter et al., 2018).

To optimize the characteristics and expand the fields of applications, the artificial production of these biomaterials has been of large interest. This is particularly the case for the expression of recombinant adhesive proteins in bacteria (Weisman et al., 2009). However, it has been proven to be difficult to express the natural adhesive protein sequences of insects in bacterial cells due to their large size and repetitive sequences and must therefore be modified into recombinant genes for transgenic production (Weisman et al., 2009).

1.2. Applications

Fibrin glue, consisting of human fibrinogen, is an example of a topical biological tissue adhesive intended to adhere surgical meshes and imitate natural tissue sealing (Brennan, 1991). Its replication of the final stages of coagulation prove to be biocompatible by being completely absorbed during wound healing, thus preventing excessive blood loss and allowing effective blood control during surgical procedures (Brennan, 1991).

Sericin, the glue protein produced by *Bombyx mori*, has also shown to be particularly favourable for the facilitation of wound healing without causing allergic reactions. The properties of sericin in dressing or cream form allows more rapid healing while decreasing inflammatory effects (Aramwit et al., 2011). Sericin in form of a gel film used for wound healing was demonstrated to be flexible, non-toxic and less damaging on regenerated tissue cells when taken off (Teramoto et al., 2008).

Biological adhesives are also particularly useful for drug delivery purposes by attaching the drug delivery system to a biological location. A common use is the attachment of drugs to the mucus coat that exhibits the advantage of enhancing the bioavailability, being well-compliant with patients when orally introduced and other possibilities (Donnelly et al., 2011). Unlike other drug delivery systems, mucosal routes allow the avoidance of first-pass effect and pre-systematic removal in the digestive tract while being deemed as safe (Bruschi et al., 2005).

1.3. Adhesive proteins from insects

Many insects need to produce adhesive structures to facilitate their survival. These are widely employed for the attachment to surfaces, resisting external forces from detachment as well as restraining their predators, and mating partners. Another aspect in which the adhesive mechanisms are naturally used is to defend themselves and their eggs

(Betz, 2010). Adhesive proteins can also be found in the silk of *Bombyx mori (B. mori)* or in the salivary gland secretion of *Drosophila melanogaster (D. melanogaster)*.

1.3.1. Sericin

While the silk produced from *B. mori* has been used as a biomaterial in the textile industry for decades, there are various other components of silk such as antibacterial or adhesive proteins with antioxidant activities, or good biocompatibility that are attractive for commercial and industrial purposes (Kunz et al., 2016). Other than exhibiting adhesive properties, sericin is nowadays also found for the use of dietary antioxidants by being able to neutralize free radicals in the human body (Aramwit et al., 2011). Sericin has also been applied in the cosmetic industry due to its moisturizing abilities stemming from the high serine and glycine content. Its collagen enhancing capabilities have shown to also boost elasticity and prevent wrinkle development (Aramwit et al., 2011). However, a large part of sericin is discarded in the sericulture, making the use of them environmentally more sustainable for the silk production industry and economically beneficial in countries where silk is processed, like China and India (Kunz et al., 2016).

The silk produced by *B. mori* is composed of two major components, fibroin and sericin (Aramwit et al., 2011). Sericins are a group of polypeptides with a significantly high number of serine amino acids. There are three main types of sericin proteins found in the silk of *B. mori* cocoons (Takasu et al, 2002). Soluble sericins, being very hydrophilic proteins, act as glue by enveloping the fibroin fibre in adhesive layers to form an insoluble silk cocoon from the silk fibres (Kundu et al., 2008). The high serine ratio (33.2 – 39 mole%) of the sericin polypeptide compositions allows the formation of biodegradable materials through cross-linkage and copolymerization of the side chains (Elzoghby et al., 2015; Takasu et al, 2002).

B. mori larvae produce large quantities of sericin towards the end of the fifth larval instar to aid the construction of firm cocoons for larval metamorphosis to adults (Kunz et al., 2016). Sericins are made of three genes, which produce several glycoproteins due to the alternative splicing (Kunz et al., 2016).

The major genes that account for the synthesis of sericin are *Ser1, Ser2* and *Ser3*. Although the *Ser2* gene, discovered by Michaille et al. (1990), is similar to *Ser1* in composition due to its size, it has been demonstrated to be more complex than other silk proteins encoding genes as it contains 13 exons. The products resulting from the crude expression of Ser2 are suggested to be responsible for the adhesive coating of the silk filaments required for firm attachment of the cocoon (Kludkiewicz et al., 2009). It has been furthermore

suggested that the higher number of charged amino acid residues are correlated to the adhesive properties of the protein (Kludkiewicz et al., 2009).

1.3.2. Salivary gland secretion

During the third instar of *D. melanogaster*, salivary gland secretion (*Sgs*) genes are activated and immediately deactivated once the puparium has formed. Proteins expressed are from the larval salivary glands and are necessary for *Drosophila* puparium adhesion to surfaces to undergo metamorphosis (Da Lage et al., 2018). The secretion and efficacy of the glue during metamorphosis is important for the development of *D. melanogaster* as it is particularly vulnerable during this stage (Da Lage et al., 2018). There are 7 genes grouped into the *Sgs* gene family with known sequences that all start with a signal peptide (Da Lage et al., 2018). Their nomenclature is based on the visualisation of the gene products on an electrophoresis gel which are as follows: *Sgs1* (band 1), *Sgs3* (band 3), *Sgs4* (band 4), *Sgs5* (band 5), *Sgs7*, and *Sgs8* and *Eig71Ee* (Korge, 1975).

Sgs3, being one of the larger genes, exhibits to be more variable in its products due to repetitive sequences (Da Lage et al., 2018). The expression of this gene has been identified to be regulated by 3 separate regions, two of which (proximal element and distal element) serve for the tissue and developmental stage specific expressions (Lehmann, 1995). There are laboratory studies focusing on the gene expression and adhesive properties of protein products. However, there has been little research conducted on recombinant *Sgs* genes.

1.4. Production of recombinant proteins

Many insect proteins exhibit unique structural or adhesive properties, but it is difficult to obtain them in sufficient quantity and quality for any applications. Biotechnological advances allow their production through genetic engineering using heterologous expression.

There are many expression systems available for the production of recombinant proteins where host organisms include bacteria, yeast or filamentous fungi (Rosano et al., 2014). Nonetheless, the most commonly employed expression system for recombinant proteins is in *Escherichia coli (E. coli)* as genetic engineering tools are increasingly designed for this host. Depending on the target protein and its nature, a bacterial expression system may be limited in the production of proteins that require a eukaryotic post-translational modification. In such case, the protein would need to be expressed in a eukaryotic host (Rosano et al., 2014; Ferrer-Mirralles et al, 2015).

However, over the past decade a diversity of proteins from higher organisms such as humans were successfully expressed in *E. coli* if the molecular mass is expected to be below around 60 kDa (Gräslund et al., 2008). In this host, proteins can also be expressed in insoluble

form through the formation of inclusion bodies of *E. coli*. Other advantages that *E. coli* as an expression host organism exhibits are fast growth kinetics allowing the exploration of protein expression in an inexpensive and time-effective manner (Rosano et al., 2014). It has also been noted that the *E. coli* system should be used primarily and be extensively explored before alternative systems are used (Gräslund et al., 2008).

The general process for such a system to produce recombinant proteins starts with gene isolation, the process of selecting and isolating the target gene. The insert is cloned into a selected expression vector and then transferred into another organism, known as the expression host. In the new organism, the gene is expected to function as it does in the original organism to produce the desired product such as an active protein. The success and of the gene transfer and the yield of the protein expression are, however, heavily dependent on the expression organism, as well as the conditions in which it is cultured in (Ferrer-Miralles et al., 2015).

For prokaryotic host cells, plasmids which are small circular DNA molecules are commonly used as expression vectors due to their versatility and simple structure that facilitates gene alterations (Griffiths et al., 1999). Currently, plasmids can be customized to be composed of a combination of numerous cloning sites, promoter selection markers, and other genetically engineered sites to suit the target gene and host (Rosano et al., 2014). For high-yield production of proteins, a transcriptional promoter with strong abilities to control the gene expression is essential (Sorenson and Mortensen, 2005). To clone and express recombinant proteins in *E. coli,* the pET system (Novagen, Inc.) with the T7 promoter is the most frequently used for recombinant proteins due to its diversity in applications (Sorensen and Mortensen, 2005).

The *E. coli* strain, BL21, carries λDE3 lysogen which encodes for T7 RNA polymerase that is required for the inducible expression of recombinant proteins in the pET expression system. This is the most suitable strain for large-scale protein productions (Gräslund et al, 2008). The transcription of T7 polymerase is induced by the binding of IPTG to the *lacUV5* promoter. This is followed by the transcription of the target gene from the T7 promoter. *E. coli* RNA polymerase cannot detect the T7 promoter and thus cannot initiate the transcription, resulting in only minimal background expression (Sorensen and Mortensen, 2005).

2. *Aim of the thesis*

- Optimising the expression of two insect adhesive recombinant proteins in *Escherichia coli BL21*
- Optimization of purification of recombinant *Bombyx sericin 2* (Ter1) and recombinant *Drosophila Salivary gland secretion 3* (Sgs3)
- Preliminary characterization of isolated recombinant proteins

3. *Materials*

3.1. *Recombinant protein sequences*

The sequences used for the expression of adhesive proteins were derived from the natural proteins. To optimize the sequence for the expression in bacteria, the sequences were synthesized *de novo* by Gene Universal Inc. (Newark, DE, USA) and inserted into pET-15b expression vectors with NcoI-XhoI as cloning sites. The amino acid sequences are shown in Table 1.

Table 1: Sequences of recombinant proteins, Ter1 and Sgs3

Gene Name	Organism Source	Molecular Weight, Residues	Sequence
Ter1	*Bombyx mori*	27.8 kDa, 247 residues	MGKDEEYSEQ NSSNKSFNDG DASADYQTKS KKVEKNSARD KKEKEKTDTR NSDGTYKTSE REKEQSSRVN QSKGSNSRDS SESDKSGRKV NKETETYSDK DAQTSESERT QSKEKKNTAP KNKGKKGTST ETDGVTKNAS KQKEKVPKDG SKSSTNDSEG KQKNKDQSKG QKNNQDGQDS STNENSKKTD DNVAKKEEPN NQKREQKGKT RCGSRKTESS KAKEDRSKKS TTDKDQRDDK KHHHHHH
Sgs3	*Drosophila melanogaster*	17.9kDa, 168 residues	MAPPTQQSTT QPPCTTSKPT TPKQTTTQLP CTTPTTTKAT TTKPTTTKAT TTKATTTKPT TTKQTTTQLP CTTPTTTKQT TTQLPCTTPT TTKPTTTKPT TTKPTTTKPT TTKPTTTKPT TTKPTTTKPT TTKPTTTKPT TTKPTTTKPT TTKPTTTKPT TTHHHHHH

3.2. *Expression system*

Table 2: Construct

Name	Description	Source	Catalogue Nr.
pET15b	Expression Vector	Novagen	69661-3
One Shot™ BL21 Star™ (DE3) Chemically Competent *E. coli*	*E. coli,* Component Cell	Invitrogen	C6010-03

The pET-15b vector system (Novagen, Inc.) that carries a N-terminal His•Tag sequence and contains the T7 promoter needed for the expression of proteins in *E. coli* was kindly provided by the molecular biology laboratory of the Biology Centre CAS, Entomology institute/University of South Bohemia to be used for this research.

3.3. Growth Media and Chemicals

All media were autoclaved or prepared under a sterile procedure.

Table 3: Luria-Bertani (LB) medium for 1 L

Substance	Amount
Trypsin	10 g
NaCl	10 g
Yeast extract	5 g
diH2O	
pH	7.0

Table 4: Luria-Bertani (LB)-Agarose medium for 1 L

Substance	Amount
Trypsin	10 g
NaCl	10 g
Yeast extract	5 g
Agar	15 g
diH2O	
pH	7.0

Table 5: Cl8+ Medium for Drosophila tissue culture (CM) for 250 mL. The pH of the mixture is adjusted to 6.5 by HCl (conc) and the whole mix is filter-sterilized.

Substance	Content
Shields and Sang M3 (Sigma-Aldrich)	9.85 g
$KHCO_3$	0.125 g
Drosophila extract	2.5 %
BSA (10%)	2 %
Penicilline/ Streptomycin	1 %
Insuline	125 IU^{-1}
diH_2O	
pH	6.5

Due to the scope of the materials, all other materials used, and solution compositions are listed in Section 8.1 Appendix 1 – Additional Materials.

4. Methods

4.1. Plasmid transformation of component bacterial cells by heat shock

The expression vector pET15b with desired expression constructs was transformed into *E. coli* cell strains, BL21 (DE3). The recombinant plasmid was mixed gently with thawed cells at a ratio of 5 – 25 ng of DNA and 50 µL of component cells. The mixture was incubated for 30 minutes on ice. Then a heat shock of 45 seconds at 42°C was applied, after which the mixture was immediately transferred on to ice again. To the transformation mixture, 400 µL of LB medium conditioned at 37 °C was added. The tube was incubated on a shaker for 60 minutes, 220 rpm at 37°C. Followed by centrifugation for 1 min, 2500 rpm at RT. 800 µL of the supernatant was removed from the tube and the pellet was suspended in the supernatant. 100 µL of these prepared cells were applied to an LB agar plate treated with ampicillin (final concentration 100 µg/mL) and scraped for a specific selection of transformants by means of a glass stick. The plate was incubated for 16 hours at 37°C, bottom-up. Using a sterile inoculation loop, one isolated colony was selected and streaked on a new LB agar plate using the streak-plate method. The new plate was incubated, bottom-up, for 16 hours at 37°C. The whole process was carried out under sterile conditions. The new plate was sealed with parafilm and stored at 4°C or immediately used for further analysis by gel electrophoresis.

4.2. Expression of Recombinant Proteins

4.2.1. Pilot experiment for the expression of recombinant proteins

The expression vector with the appropriate insert was transformed into component BL21 cells (according to section 4.1.) from which one colony was selected and inoculated in 3 ml of LB media. This medium was supplemented with ampicillin (final concentration of 0.1 mg/mL). The bacterial culture was incubated at 37°C on a shaker at 200 rpm overnight (at least 16 hours). After 16 hours, 100 µL from the overnight culture with the desired protein sequence was inoculated to 15 mL of sterile LB media. The culture was grown with ampicillin again (final concentration of 0.1 mg/mL).

The 15 mL bacterial culture was incubated at 37°C on a shaker at 200 rpm until the log phase (OD_{600} = 0.6 – 0.8) was reached. The optical density was measured by taking 300 µL of the culture every 30 minutes. A sample of 1 ml was collected and stored as a control (uninduced) sample at 4°C. Protein expression was induced by adding IPTG (final concentration of 1 mM). The bacterial culture was further incubated for 4 hours on a shaker at

200 rpm at 37°C. At each hourly interval a sample of 1 mL was collected and stored at 4°C for further analysis.

The collected samples were centrifuged for 1 minute, 1000 rpm at 4°C. The supernatant was removed from the sample and the pellet was resuspended in B-PER^{TM} Pierce^{TM} Complete Reagent (5 mL reagent/g of biomass). The suspension was pipetted until homogenous and incubated for 15 minutes at RT with gentle rocking. The lysate was centrifuged at 16000 × *g* for 20 minutes to separate the soluble proteins from the insoluble ones. The samples were prepared for further analysis.

4.2.2. Expression of proteins in larger quantities

The expression vector with the corresponding adhesive protein sequence was transformed into competent *E. coli* BL21 cells (according to section 4.1.). An LB media of 250 mL, supplemented with ampicillin (final concentration: 0.1 mg/mL), was inoculated with 1.5 mL of overnight culture (prepared according to section 4.2.1). After the bacterial cultures reached log phase ($OD_{600} = 0.6 - 0.8$), protein expression was induced by IPTG at a final concentration of 1 mM, followed by a 4-hour incubation on a shaker at 200 rpm at 37°C. The bacterial culture was centrifuged for 10 min at 3000 rpm in 4°C. The supernatant was removed, and the pellet was frozen at -80°C or immediately used to prepare the bacterial lysate.

4.3. Purification of Recombinant Protein from Cytosol

4.3.1. Preparation of Bacterial lysate

The bacterial pellet (prepared according to section 4.2.2) was resuspended in 20 mL of NLB lysis buffer, 2 ml of Lysozyme and 50 µL of RNAse. This mixture was resuspended until homogeneous, followed by an incubation of 30 minutes on ice with gentle rocking. The lysate was sonicated at 10 seconds for 10 pulses (with intervals of 10 seconds between each pulse), with 40% of the sonication performance. This was followed by centrifugation in Eppendorf tubes of 1.5 mL for 15 min, at 3000 rpm in 4 °C. From the supernatant, 20 µL of the sample was collected. The remaining pellets were frozen for inclusion body purification. The supernatant was frozen at -20°C or immediately used for purification of the soluble recombinant protein by affinity chromatography.

4.3.2. Affinity chromatography

The sequence of the recombinant proteins contained a tag of six-histidine molecules that allow the affinity purification using Nickel-Agarose. A 2 ml Ni-NTA agarose column was prepared exactly according to the Ni-NTA Purification System kit protocol (Invitrogen, 10 mL falcon tube). The Ni-NTA agarose was washed with 2 mL of distilled H_2O. Once the water was removed, 20 mL of the supernatant (prepared in section 4.3.1.) was added to the sealed

column, followed by an incubation on the rotator for 1 hour at 4°C. This allowed the proteins to bind to the resin. Unbound protein buffer was collected by gravity gradient through the column filter and re-applied to the column 2 more times to ensure increased binding of the desired proteins to the stationary phase. The unbound protein buffer was then removed, and a sample of 1 mL from this solution was stored at -20°C for further analysis. The column was then washed successively with NWB. Fractions of 1 mL were continuously collected. The purity of the sample was monitored by measuring the absorbance of the washing fractions until a local minimum absorbance was reached (A_{280} = 0.1 -0.01). This was followed by the release of bound proteins which was performed by applying 4 mL of NEB to the column and 2 mL of NEB for the re-elution. All fractions were stored for further analysis.

4.4. Purification of Recombinant Protein from Inclusion Bodies

4.4.1. Preparation of Bacterial Lysate

The bacterial pellet (prepared in section 4.3.1) was resuspended in 40 mL of IBB 1 and 4 mL of lysozyme. This was followed by an incubation of 45 minutes on ice, on a shaker until homogenous. The lysate was treated with protease inhibitor (1:100) and sonicated at 30 seconds × 5 pulses, with 60% sonication power (with intervals of 30 seconds between each pulse). The lysate was subsequently treated with 40 µL of RNAse, 20 µL of $MgCl_2$ (2M) and 10 µL of $CaCl_2$ (15mM), followed by a 60-minutes incubation at RT. The bacterial lysate was centrifuged for 15 min, 6000 rpm at 8°C. Then, the supernatant was removed, and the pellet resuspended in 2 mL of IBB 2 and sonicated at 10 seconds × 3 pulses with 10 seconds of rest between each pulse, vortexed and incubated at RT for 10 min. Then the inclusion bodies were centrifuged at full speed for 15 minutes at 8°C. The supernatant was removed and the washing process using IBB 2 was repeated a total of 2 times. After the supernatant was removed, the pellet was resuspended in 50 mL of IBB 3 and centrifuged at maximum rpm for 15 minutes, and the washing process using IBB 3 buffer was repeated 4 times in total. The supernatant was then removed, and the pellet was dissolved in 10 mL of 8 M urea. All fractions were stored for further analysis.

4.5. Dialysis

The elution samples were pipetted into the dialysis membrane (Servapor 29 mm, 12000-14000 MWCO) and dialysed against 1.5 L of dialysis solution which depended on the sample. A 1 x PBS solution was used for proteins purified from cell cytosol. Proteins purified

from inclusion bodies were dialysed against 10 mM NH_4HCO_3.The dialysis solution was stirred and exchanged every 12 hours for a total of 3 times and incubated at 4°C.

4.6. Lyophilization and final treatment

Dialysed samples were frozen at -80 ° C. Afterwards, the samples were lyophilized and stored at -20 °C. Prior to further analysis, the proteins that were purified from the cell cytosol, were dissolved in distilled H_2O. The proteins purified from the inclusion bodies were dissolved in concentrated formic acid. The amount of formic acid was dependent on the desired sample concentration.

4.7. Protein detection

4.7.1. Vertical Sodium-dodecyl-sulphate-Polyacrylamide gel electrophoresis (SDS-PAGE)

A protein marker and protein samples from various stages of protein expression process were applied to a prepared 10% polyacrylamide gel (see section 8.1 Appendix 1 – Additional Materials) for the detection of the recombinant proteins.

For sample preparation, 40 µL of the sample and 10 µL sample buffer (6X) were combined and denatured for 5 minutes at 96°C. The samples were then centrifuged at maximum speed for 5 mins at 4 °C. 20 µL of a sample was pipetted to a well of the gel.

The visualization of the proteins allowed the determination of its successful expression, molecular size (kDa) and concentration (densitometry). Gel electrophoresis was conducted in Running Buffer at 6 V/cm for 30 minutes, after which the voltage was increased 10 V/cm and run for 2 hours (or until the front reaches the end of the gel). The gel was stained with a CBB solution for at least 60 minutes (or overnight) on a shaker. The gel was washed with a destaining solution until the bands were desirably visible. Finally, the gel was photographed.

4.7.2. Western blot

To further analyse the expression of proteins, western blotting was performed by a gel transfer to the nylon membrane and protein bands were detected using antibody against 6x His. Proteins were separated by SDS-PAGE (according to section 4.7.1). The proteins from the unstained gel were transferred to a PVDF membrane in 1 × Electrotransfer buffer at 50 volts for at least 200 minutes.

After the transfer, the membrane was incubated for 2 hours in 1% BSA solution of PBS-tween at 4°C. The gel was stained with CBB while the membrane was washed 3 times for 15 minutes in a PBS-tween solution. The washing was followed, by an incubation of 1

hour with 2% BSA in PBS-tween and the primary anti-His antibody (1:1000). The membrane was then washed 3 times for 5 minutes with PBS-tween. The membrane was then incubated with secondary antibody (1:5000; anti-mouse conjugated with peroxidase) in a 1% BSA in PBS-tween solution for 1 hour, followed by 3 washes of 5 minutes in 1 x PBS solution. The membrane was prepared for imaging using the Pierce ECL Plus Western Blotting Substrate protocol and chemiluminescence signals were detected using the Luminescent Image Analyzer LAS 3000. The membrane was then stained for 5 minutes in Ponceau S staining solution (3% TCA, 0.5g/L Ponceau powder).

4.7.3. Sequential analysis

To verify the expressed proteins from the gels, peptide identification by tandem mass spectrometry was performed by Mgr Peter Konik at the Faculty of Science of the University of South Bohemia in a specialized laboratory for proteomic analysis. Protein separation was made using a polyacrylamide gel, the protein bands were stained by CBB and excised with a razor. The Q-Tof Premier (Waters) mass spectrometer connected to Nano Acquity UPLC (Waters) was used to analyse the protein samples. Data was processed by PLGS 2.3 (Waters). For the peptide identification, the UniProt and NCBI databases were used, containing the collection of *Bombyx mori* and *Drosophila melanogaster* sequences.

4.8. Determination of recombinant protein concentrations

4.8.1. Nanodrop

Using NanoDrop^{TM}, the concentrations of the purified proteins (elution samples) were measured by performing the A280 Absorbance Protein measurement. Desjardins et al. (2009) described this method as suitable for the quantification of purified proteins with extremely small volumes that contain Tryptophan, Tyrosine, Phenylalanine residues or Cysteine-Cysteine di-sulfide bonds and additionally exhibit absorbance at 280 nm.

To determine the protein concentration more precisely, the extinction coefficient was entered. The molar extinction coefficient, a measure of a sample's molar absorptivity (Anthis et al., 2013), for the ultraviolet absorption at 280 nm was determined through the amino acid sequence of the protein using the ProtParam tool of the Expasy Bioinformatics Research Portal (Gasteiger et al., 2005).

The approximate concentrations were used to determine which fractions should be combined together in order to dialyze the samples when the elution fractions were divided into a larger number.

4.8.2. Densitometry

Vertical electrophoresis on 10% polyacrylamide gel for Sodium-dodecyl-sulphate-Polyacrylamide gel electrophoresis (SDS-PAGE) was performed with the purified proteins. After the staining of the gel with the Coomassie blue protein dye, the gel was photographed. Recombinant protein concentrations were determined from this image. The software, Image Quant ™, was used to quantify the size of specific bands that characterize the observed proteins. Bovine serum albumin (BSA) of known concentration was used as a standard.

4.9. Characterization of recombinant proteins

4.9.1. Cell culture experiments

The purified proteins (obtained from section 5.6.), were the effects of protein coating on plastic surfaces of regular polystyrene plates tested on *Drosophila* cells. The Cl.8+ cells require special treatment of polystyrene surfaces with gas plasm or coated with polylysine for their growth. Without the attachment of the cells, apoptosis occurs. The Cl.8+ cells originate from the third instar wing imaginal disc cell line, CME W1 (Peel, 1992). The cells were cultivated in Shields and Sang M3 (Sigma-Aldrich) Insect medium. The cells were diluted 1:3 during passaging.

Each polystyrene petri dish of a diameter of 55 mm was labelled with 6 circles (8 mm diameter) so that 5 μL of each recombinant protein (Ter1 and Sgs3 of 1 mg/mL in water) and the negative control (*BSA* of 0.5 mg/mL concentration) could each be plated twice. The protein solutions were allowed to dry in a sterile box. The coordinates for photographing the cells were inscribed on the outside of the petri dish. There were 3 locations plotted for each circle.

The Cl.8+ cells were diluted to a concentration of 5×10^5 cells/mL in CM. A 5 mL solution of medium was added per plate with a 10 μL anti-fungal solution (Fungin™). The medium already contained penicillin and streptomycin. The cell culture was incubated for 3 days under standard conditions. Every 24 hours of incubation, photographic evidence of the cultures was taken under a light microscope (Leica Dmi8). The photos were taken at the same, well-marked point of the culture dish. The number of cells were manually counted from the photo and recorded. The cell viability was assessed according to their morphology. Healthy cells must be elongated and contain pseudopodia. Cell clusters were not counted in the results.

4.9.2. Scanning Electron Microscopy

Samples of adhesive proteins were spread over nonporous glass or plastic surfaces that were coated with gold. The samples were then characterised by scanning electron microscopy using JEOL JSM-6300 scanning electron microscope at the Laboratory of Electron Microscopy of the Institute of Parasitology, ASCR.

5. *Results*

5.1. *Pilot experiments for expression of various recombinant proteins*

After the plasmids (prepared in section 4.1.) were plated and incubated, a single recombinant *E. coli* colony was selected and inoculated in LB media for culturing. The expression of the proteins was verified by inducing different recombinant proteins in the BL21 Competent *E. coli* cells and analysis by separation on SDS-PAGE. Following the procedure in section 4.2.1, the samples were not purified prior to the analysis. The same procedure was used to analyse the expression at different induction times (1 – 4 hours). The expression of different proteins was compared to determine which are feasible for further analysis as not all of them were expected to show meaningful results.

The protein marker was applied first (lane M). For each recombinant protein, the samples were applied in the following order: insoluble bacterial lysate (pellet) with induced protein expression after 4 hours, bacterial lysate prior induction of protein expression and soluble bacterial lysate (supernatant) with induced protein expression. The results of the protein expression kinetics are presented in Figure 1 and 2. Based on the pilot experiments, we selected Ter1 and Sgs3 for further work.

Figure 1 demonstrates the protein separation from the soluble components of the bacterial lysate containing different recombinant Sericin proteins (Ser1, Ter1, Ser3) of which the natural variants are originally found in the silk from *Bombyx mori*. Ter1 which is a recombinant protein derived from *sericin 2*, was elected for further analysis. In Figure 1, it can be seen that lane 11 showed Ter1 with the most prominent band from the supernatant of the induced bacterial lysate. This is a significant result when comparing it to lane 8 which represented the protein expression in the uninduced Ter1 sample. This pattern was not observed for Ser1 samples in contrast. The bands for Ser1 samples are not distinctively different between the uninduced and induced samples, making it unclear whether the desired protein was expressed. Ter1 thus became a preferable candidate for further analysis.

A distinctive band at around 40 to 50 kDa can be found in the soluble phase of the induced Ter1 sample. The exact mass is not quite clear due to the separation of the protein marker. Although the theoretical molecular weight of Ter1 (27.7 kDa) is lower, this band shows strong correspondence to the natural sericin protein.

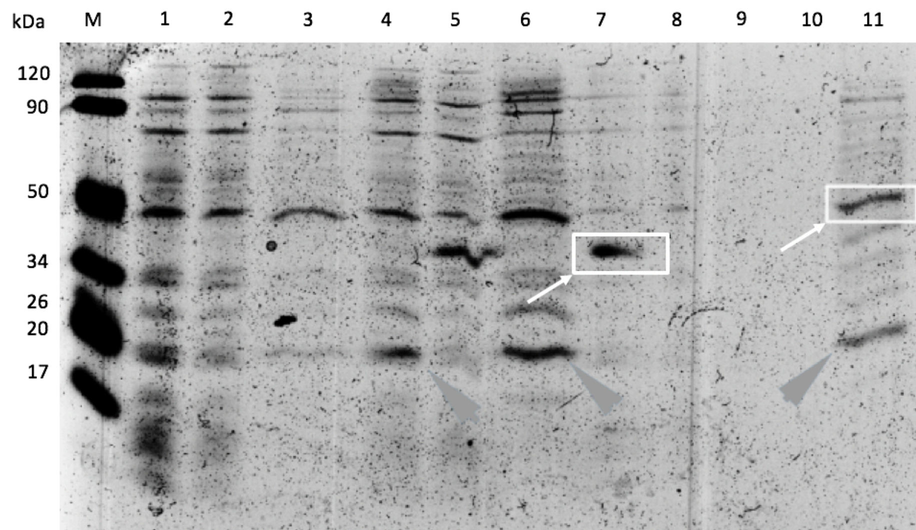


Figure 1: SDS -PAGE analysis of three different crude recombinant sericin proteins (Ser1, Ser3, Ter1) after bacterial expression.
Lane M: Protein marker (Pierce #26612)
Lane 1: Ser1 (15.5 kDa) induced expression in pellet
Lane 2: Ser1 (15.5 kDa) uninduced expression in pellet
Lane 3: Ser1 (15.5 kDa) induced expression in bacterial lysate
Lane 4: Ser3 (16.8 kDa) induced expression in pellet
Lane 5: Ser3 (16.8 kDa) uninduced expression in pellet
Lane 6: Ser3 (16.8 kDa) induced expression in bacterial lysate
Lane 7: Ter1 (27.7 kDa) induced expression in pellet
Lane 8: Ter1 (27.7 kDa) uninduced expression in pellet
Lane 11: Ter1 (27.7 kDa) induced expression in bacterial lysate
The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of Ter1 samples are indicated by an arrow and a white frame.

According to Takasu et al. (2007) it is often observed for sericin proteins to move slower and to show a higher estimated molecular mass by SDS-PAGE than their expected molecular size due to their structural composition. Depending on the gel, the observed mass could be twice as much as the expected. This verified the expression of Ter1.

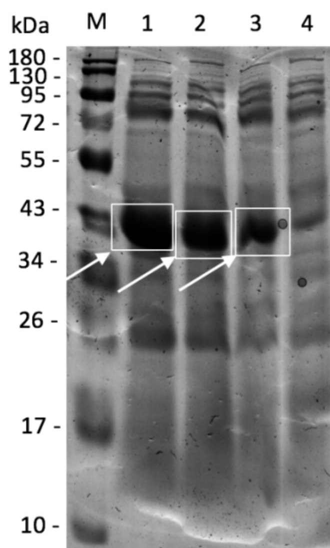


Figure 2: SDS-PAGE of different induction times (1-4 hours) for the expression of recombinant protein Ter1 (27.7 kDa). Induction using IPTG was performed after the log phase (OD600 = 0.6-0.8) was reached.
Lane M: protein marker (Thermo #26616)
Lane 1: 4 hours of induction
Lane 2: 3 hours of induction
Lane 3: 2 hours of induction
Lane 4: 1 hour of induction
The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of Ter1 is indicated by an arrow and a white frame.

Following the procedures of 4.2.1. for the micro-scalar protein expression, different times for the induction period were tested on Ter1 protein production to investigate optimal conditions. The recommended time for protein expression is 3-4 hours (Novagen, Inc.). This is also evident in Figure 2 where the intensity of the bands increases as the induction period increases. An induction of 4 hours showed the highest intensity, and thus was used as the optimum induction period for large-scale protein expression. Additionally, the intensive bands highlighted as Ter1 in Figure 2, are the same molecular mass as was seen in Figure 1. Thus, the success of the protein's expression was reconfirmed.

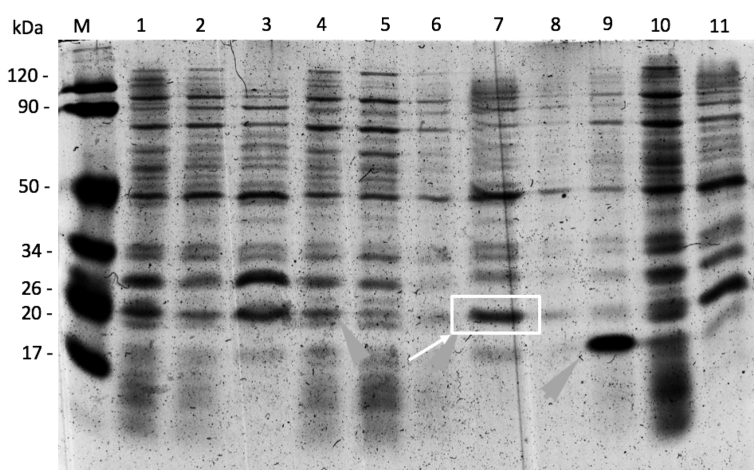


Figure 3: SDS-PAGE of various crude recombinant adhesive proteins (Cados, Sgs3, Barnacle*). All three proteins originate from different organisms.*

Lane M: Protein marker (Pierce #26612)
Lane 1: Cados (14.5 kDa) induced expression in pellet
Lane 2: Cados (14.5 kDa) uninduced expression in pellet
Lane 3: Cados (14.5 kDa) induced expression in bacterial lysate
Lane 4: Sgs3 (17.9 kDa) induced expression in pellet
Lane 5: Sgs3 (17.9 kDa) uninduced expression in pellet
Lane 7: Sgs3 (17.9 kDa) induced expression in bacterial lysate
Lane 9: Barnac (17.8 kDa) induced expression in pellet
Lane 10: Barnac (17.8 kDa) uninduced expression in pellet
Lane 11: Barnac (17.8 kDa) induced expression in bacterial lysate
The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of Sgs3 samples are indicated by an arrow and a white frame.

Several control recombinant adhesive proteins (Cados, Sgs3, Barnacle) which are natively produced by *Rhyacophila obliterate*, *Drosophila* and *Megabalanus rosa* respectively, were also cultured and expressed. For Sgs3 samples with an expected molecular mass of 17.9 kDa, a distinct band was determined in the induced cell cytosol sample that cannot be found in the uninduced sample at around ~20 kDa. Due to the correspondence to the expected molecular, we considered the expression of Sgs3 as successful. As the observed expressed protein was heavier than expected, the presence of Sgs3 needed further validation by performing western blotting for the detection by anti-His antibody. The proteins were detected

using luminescence (secondary antibody was conjugated with peroxidase). Finally, the membrane was also stained with Ponceau S.

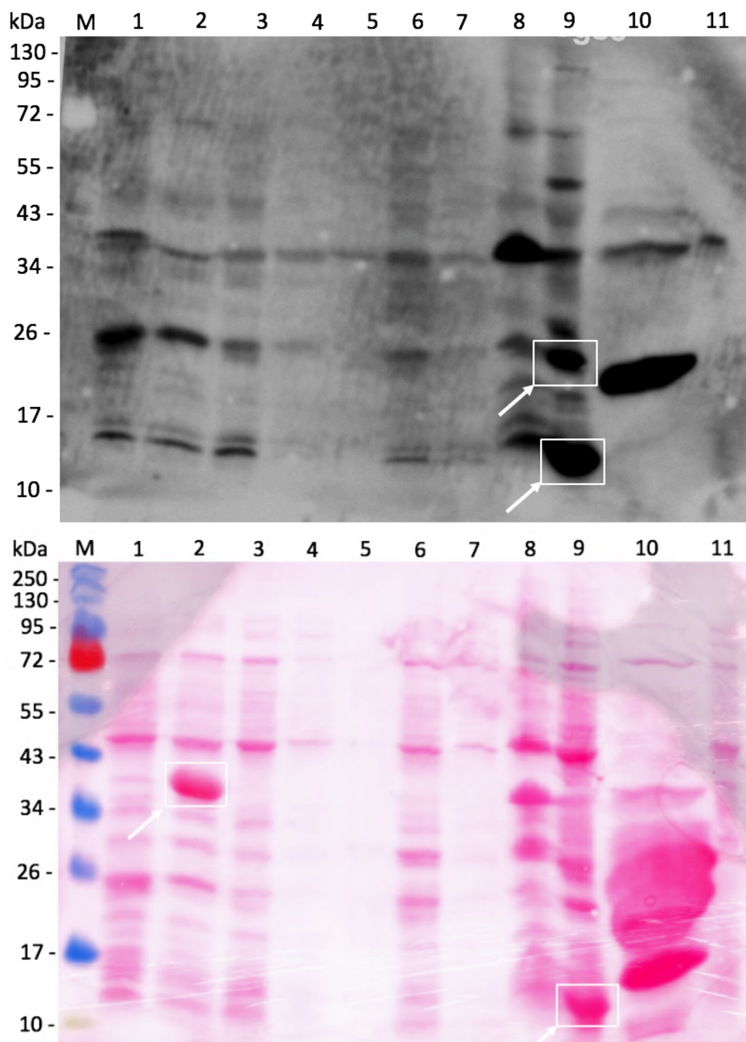


Figure 4: Visualization of various crude recombinant proteins after 4 hours of induced expression using western blot technique. (A) Imaging of membrane through chemiluminescence using Luminescent Image Analyzer LAS 3000 (B) Imaging of membrane using Ponceau S staining.
Lane M: Protein marker (Thermo #26616)
Lane 1: Seh1 bacterial lysate as control
Lane 2: Ter1 (27.7 kDa) bacterial lysate
Lane 3: Ser1 (15.5 kDa) bacterial lysate
Lane 4: Ser1 (15.5 kDa) uninduced pellet
Lane 5: Ser1 (15.5 kDa) uninduced pellet
Lane 6: CadOS (14.5 kDa) bacterial lysate
Lane 7: CadOS (14.5 kDa) uninduced pellet
Lane 8: Ser3 (16.8 kDa) bacterial lysate
Lane 9: Sgs3 (17.9 kDa) bacterial lysate
Lane 10: Seh1 (10.9 kDa) purified protein as control
Lane 11: His-tag positive control
The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of selected proteins are indicated by an arrow and a white frame.

The strongly visible band corresponding to the induced supernatant sample of Sgs3 at around 20 kDa that was identified in Figure 3 (Lane 7) is also visible on the western blot (see Figure 4(A), Lane 9). However, the band at ~20 kDa was not very strongly visible when stained with Ponceau S (see Figure 4(B), Lane 9). Figure 4 also showed an intensive band at around 17 kDa for the expression of Sgs3 which corresponds more closely to the expected molecular mass of 17.9 kDa. This band is additionally more prominently visible on the Ponceau S stained membrane (see Figure 4 (B, Lane 9)).

While the expression of the recombinant protein Ter1 is not clearly visible when the luminescence detection method was used, a strong band can be seen in Figure 4 (B) for the expression of soluble *Ter1* at around 40 kDa (Lane 2). This is representative of Ter1 protein production observed in Figures 1 and 2. As a positive control, His-Tag was applied (see Lane 11 in Figure 4, (A) and (B)) to furthermore distinguish samples from the control sample. As the expression of both Ter1 and Sgs3 did not overlap with the control sample, the production of the desired proteins was reaffirmed.

5.2. Expression and purification of recombinant proteins

5.2.1. Purification of recombinant proteins

Ter1 and Sgs3 proteins were expressed in larger quantities (according to section 4.2.2.). This was followed by further purification by affinity chromatography. The purification of the proteins was visualized by vertical electrophoresis.

5.2.1.1. Purification of Ter1

Figure 6, (A) and (B), both show that a significant production of Ter1 was collected during the elution phase with a molecular weight of around 40 kDa as it was observed in the initial analysis and its western blot (see section 5.1., Figure 4 (B)). It must be noted, however, that there was a loss of the desired proteins before the elution process as intensive bands with the expected molecular weight of Ter1 were detected in the fractions not initially bound to the chromatography column, the flow- through fractions (see Figure 6 (A), Lane 1 – 3). Following these results, the method was optimised to reduce loss in the flow-through fractions for further research by re- collecting and re-applying the flow-through to the column two more times before discarding the solution and continuing with washing of the resin. The uninduced sample in lane 9 of Figure 6 (B) is not visible due to difficulties encountered while loading the sample on the gel. The uninduced sample was highly viscous due the lack of denaturing and purification steps. This made it challenging to compare the induced expression of desired proteins to the natively expressed proteins of *E. coli*.

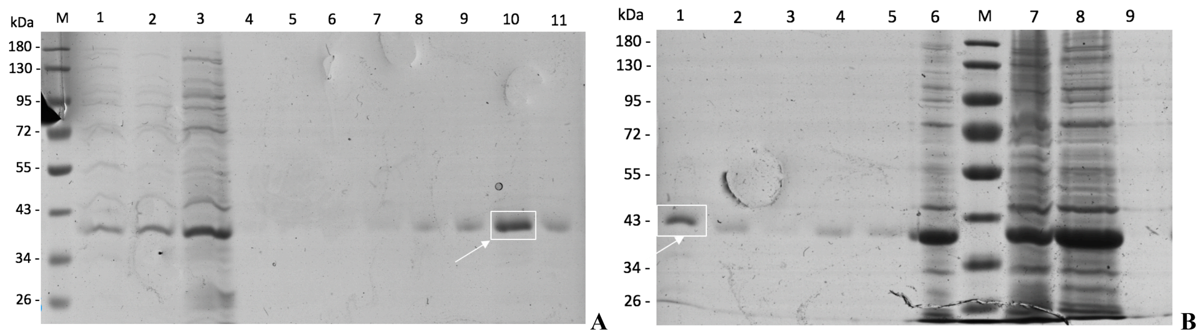


Figure 6: (A) SDS-PAGE analysis of purification fractions of Ter1 (27.7 kDa) by affinity chromatography (Left)
Lane M: Protein marker (Thermo #26616)
Lane 1, 2, 3: Flow-through Fractions
Lane 4, 5, 6, 7, 8: Washing Fractions
Lane 9, 10, 11: Elution Fractions
The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of selected proteins are found at around 43 kDa. This is strongly visible in Lane 10, indicated by a white arrow and white frame.

Figure 6: (B) SDS-PAGE analysis of purification fractions of Ter1 (27.7 kDa) by affinity chromatography (Right)
Lane M: Protein marker (Thermo #26616)
Lane 1, 2, 3: Elution fractions
Lane 4, 5: Re-elution fractions
Lane 6: Lysed Supernatant prior to purification
Lane 7, 8: Induced *Ter1* bacterial culture prior to lysis
Lane 9: Uninduced *Ter1* bacterial culture
Due to difficulties encountered during the loading of the gel with the uninduced sample, Lane 9 appears to be void of sample. The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of the selected proteins are indicated by an arrow and a white frame.

The most intensive band during elution was found in lane 10 of Figure 6 (A), meaning that in this fraction the highest concentration of Ter1 was obtained. The correspondence of this band to the desired protein, Ter1 (recombinant *sericin 2*), was later confirmed using mass spectrometry (see section 5.2.2.). The purification of Ter1 from endogenous bacterial proteins was shown especially successful in contrast to the crude extracts such as the bacterial lysate (Figure 6 (B), Lane 6) which contained a variety of different proteins that are no longer present in the purified fractions.

5.2.1.2. Purification of Sgs3

From Figure 7, showing the Sgs3 purification, it can be seen that the elution fractions had two distinct bands present of size ~26 kDa and ~17 kDa both of which were also previously observed on the western blotted membrane (see Figure 4 (A)). The heavier band only appeared during elution as a desired recombinant protein is expected during affinity chromatography purification. Yet, as it was observed for the purification of Ter1 in section 5.2.1.1., desired recombinant proteins that were used in this analysis may be subject to poor binding of this Ni-Agarose column. This was also supported by the molecular weight of the heavier band which was found in the uninduced sample but not in the fractions of proteins that did not bind to the chromatography column (Flow-through fraction). Additionally, the band found at ~17 kDa had a molecular weight close to the expected. This size also corresponded

to the protein indicated as Sgs3 on the Ponceau S stained western blotted membrane (see Figure 4 (B)).

The heavier band was likely a contaminant bacterial protein that was able to bind effectively to the chromatography column due to the possible presence of Histidines residues This band also corresponds to the protein with a weight of ~26kDa that was seen in the uninduced sample (Figure 7, Lane 1). This band was distinctively seen with this size on a different SDS-PAGE analysis (see Figure 8, Lane 1).

For further analysis, it was suggested to analyse the elution fractions against a sample of bacterial lysate prior to the application on the chromatography column as well. It was also suggested to explore whether the poor binding of the recombinant proteins to the column continues to occur with different purification procedures.

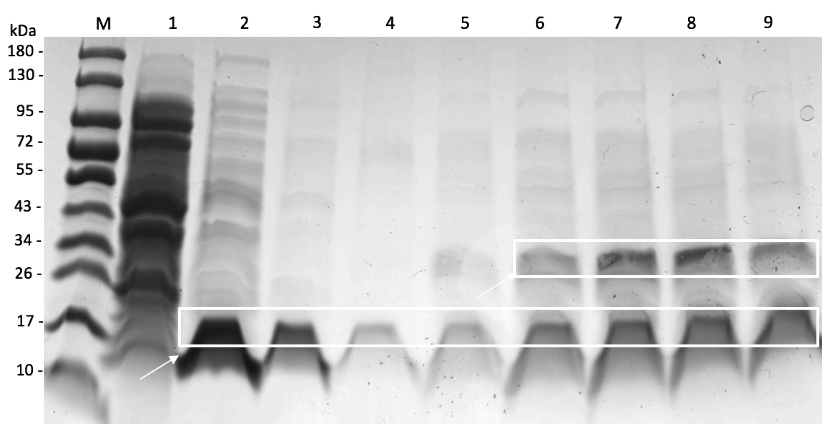


Figure 7: SDS-PAGE (Gradient gel 4~20%, Bio-Rad #4561096) analysis of purification of Sgs3 (17.9 kDa) by affinity chromatography.
Lane M: Protein marker (Thermo #26619)
Lane 1: Uninduced Sgs3 bacterial culture
Lane 2: Flow-through fraction
Lane 3, 4: Washing fractions
Lane 5, 6, 7, 8, 9: Elution fractions
Although there are two bands visible during elution, the heavier band was not visible in the flow-through fraction. However, the lower band corresponds to the expected molecular weight and is not expressed in the uninduced sample. The expression of Sgs3 as the lower band was later verified through mass spectrometry. The numbers next to the protein marker (M) indicate the protein weight in kDa. Expression of the selected proteins are indicated by an arrow and a white frame.

5.2.2. Protein verification using proteomic analysis

The presence of the recombinant proteins was further verified by performing tryptic peptide analysis on a mass spectrophotometer according to section 4.7.3.

For the verification of Ter1, the distinct band from the elution fraction (Figure 6 (A), Lane 10) was cut out from the gel for further proteomic analysis.

Table 6: Mass spectrometry results of Ter1 that show the overlapping sequences to databases

Database	Organism	Accession Number	Sequence	Peak Molecular Weight / Da
UniProt	*Bombyx mori*	D2WL76	(K)SFNDGDASADYQTK(S)	1518,633
			(K)ETETYSDKDAQTSESER(T)	1975,825
			(K)DEEYSEQNSSNK(S)	1429,576

The results from the mass spectrometry verified the expression and analysis of the correct recombinant protein (see Table 6). The results had a clear correspondence to sequences of the protein *sericin 2* produced by *Bombyx mori.*

For Sgs3, the two distinct bands from the analysis seen in Figure 7 in the elution fraction (Lane 8) were both included for further analysis. The heavier band (~26 kDa) gave no distinct results that clearly matched the native Sgs3 protein sequence from *Drosophila melanogaster*. This band contained various proteins such as keratin and ribosomal proteins, which were considered as contaminants or proteins natively produced by *E. coli.* The band with the lower molecular weight (~17 kDa) in contrast exhibited strong correlation to the desired protein sequence across different databases (see Table 7). This molecular weight is also comparable to the theoretical molecular weight of recombinant Sgs3. Hence, the successful expression of recombinant Sgs3 protein was verified.

Table 7: Mass spectrometry results of Sgs3 that show the overlapping sequences to the database

Database	Organism	Accession Number	Sequence	Peak Molecular Weight / Da
NCBI	*Drosophila melanogaster*	AAF50056.1	(K)QTTTQLPCTTPTTTK(Q)	1621,810
		AAF59342.1	(K)FVQEFAR(D)	896,4619
			(R)SSTQTDLQMDR(N)	1281,582
			(K)DSPVTDSDQFLSRSSTQTDLQMDR(N)	2745,213
UniProt	*Drosophila melanogaster*	P02840	(K)QTTTQLPCTTPTTTK(A)	1621,810
			(L)PCTTPTTTK(A)	949,4655

5.2.3. Further testing of recombinant Sgs3 protein expression

5.2.3.1. Protein expression at lower temperatures

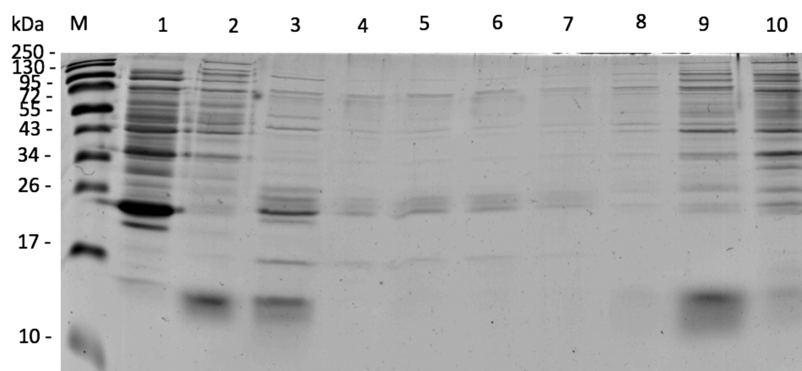


Figure 8: SDS-PAGE analysis of the purification of Sgs3 (17.9 kDa) by affinity chromatography at 25°C.
Lane M: Protein marker (Thermo #26616)
Lane 1: Uninduced sample of bacterial culture grown at 37°C
Lane 2: Flow through fraction of bacterial culture grown 37°C
Lane 3: Elution Fraction of bacterial culture grown at 37°C
Lane 4, 5, 6, 7: Elution fraction of bacterial culture grown at 25°C
Lane 8, 9: Washing fractions of bacterial culture grown at 25°C
Lane 10: Flow through fractions of bacterial culture grown at 25°C
The condition during the growth of the bacteria and induction was reduced in attempt to minimize the unspecific bacterial protein yield that remained bound during elution. However, the yield of all proteins was reduced and much of the desired protein was lost during flow-through. It can be clearly seen that the heavier protein found in the elution fraction corresponds to a bacterially expressed protein that is also found in the uninduced sample.

In an attempt to reduce the undesired proteins produced during the protein expression of Sgs3 which were visible during the elution (see Figure 7, Lane 6 – 9; Figure 9, Lane 1), the temperature at which the cells were cultured at was lowered to 25°C (RT). Allowing lower temperatures during cell growth and induction, in many cases improved the yield of the soluble product by reducing protein aggregation and protein degradation that cause the accumulation of insoluble inclusion bodies (Sorensen et al., 2005).

As can be seen in Figure 8, the yield of the desired protein significantly reduced when a lower temperature for cell growth and induction was used without significantly reducing the yield of the endogenous bacterial protein. This is especially clear when comparing Lane 3 and 4, which present the elution fractions of a cell culture grown at 37°C and one grown at 25°C. Hence, it was determined that 37°C is the optimal temperature for cell cultures and induction. This is also in accordance to Mühlmann et al. (2017) where it was reported that 37°C is an optimal temperature to achieve high yields when inducing with an IPTG concentration of 1mM as it was done for this procedure.

It was further noted that the intensity of the bands representing the flow-through fractions at both temperatures, was practically identical. Sgs3 was similarly not detected in the elution fraction (see Figure 8, Lane 4 – 7). As the desired recombinant protein was largely lost before the washing procedure had been performed, similar to Ter1, reducing the temperature

for bacterial cell culture growth simultaneously reduced the affinity of the protein to the column.

Figure 8 also showed more clearly that the recombinant protein Sgs3 was slightly smaller in size than expected, with an observed molecular weight of ~15kDa.

5.2.3.2. Inclusion body purification analysis of Sgs3

In attempt to obtain the pure recombinant Sgs3 protein we used procedures for the isolation from the inclusion bodies, which were supposed to contain the desired recombinant proteins in denatured state. The procedure with more washing steps was performed as described in section 4.4. The samples were analysed and evaluated using vertical electrophoresis.

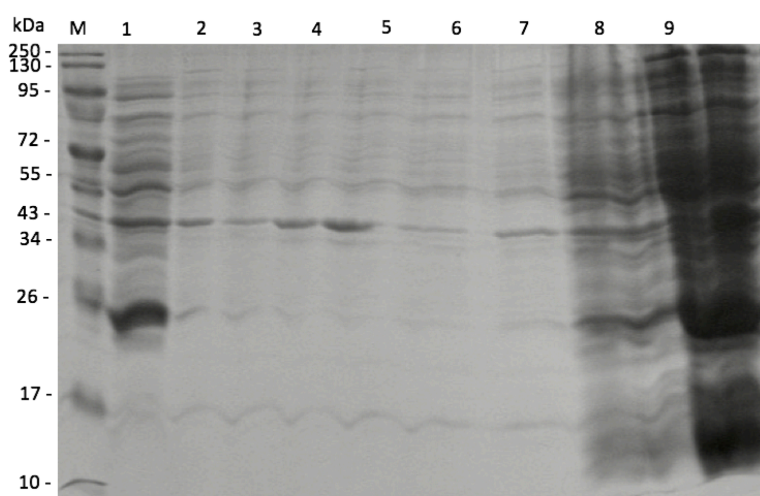


Figure 9: SDS-PAGE analysis of inclusion body purification of Sgs3 (17.9 kDa) by affinity chromatography.
Lane M: Protein marker (Thermo #26616)
Lane 1: Uninduced Sgs3 sample of bacterial culture
Lane 2: Soluble Purified Sgs3
Lane 3, 4, 5: Washing Fraction using IBB3
Lane 6, 7: Washing Fraction using IBB2
Lane 8: Supernatant prior to IBB2
Lane 9: resuspended pellet after sonification
The proteins purified from inclusion bodies gave a purer product as the bacterially expressed proteins were removed although the yield is overall lower than purified proteins from the cell cytosol

In Figure 9, there are some light bands observed during the washes at the expected molecular mass (~17 kDa) which are not observed in the uninduced sample. The yield of expressed proteins appeared overall lower in the inclusion bodies than in the soluble phase. Nonetheless, the yield of the presumably bacterial protein with the size of ~26 kDa was significantly lower than of the desired protein meaning that the inclusion body purification of Sgs3 produced purer recombinant proteins which are needed for further testing.

5.2.4. Densitometry

Densitometry was performed to determine the concentration of the expressed proteins using ImageQuant™. The standards used for the densitometric analysis were BSA of the following concentrations: 1.5, 0.75, 0.375, 0.1875, 0,09375 mg/mL. The data units are Optical Density (OD) and the background subtraction method was Global.

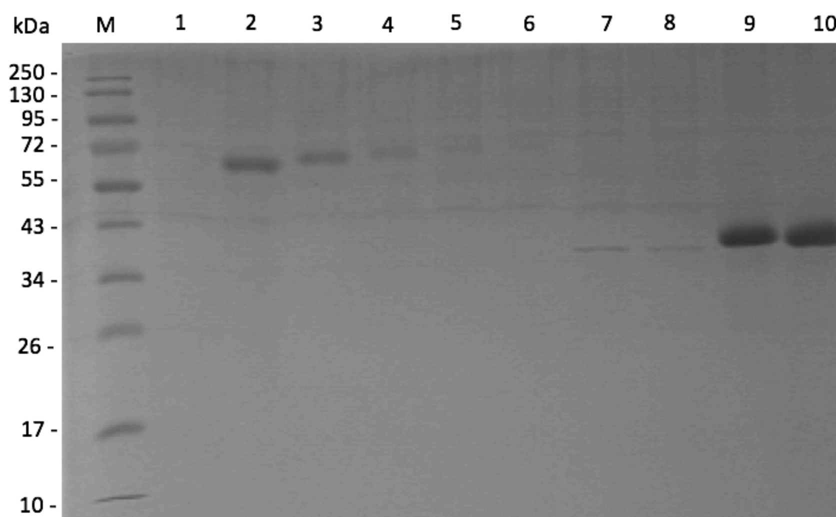


Figure 10: SDS-PAGE of *BSA* (66.5 kDa) standard solutions, Sgs3 (17.9 kDa) and Ter1 (27.7 kDa) for densiometric analysis.
Conc = (Vol – 2E+06) ÷ 576219
R^2 = 0.9789
Background subtraction method: Global
Data units: Optical Density (OD)

Lane M: Protein marker (Thermo #26616)
Lane 1: Blank sample
Lane 2, 3, 4, 5, 6: Standard *BSA* (66.5 kDa) solutions
Lane 7, 8: Purified Sgs3 (17.9 kDa) proteins from inclusion bodies
Lane 9, 10: Purified Ter1 (27.7 kDa) proteins from bacterial lysate

The calculated mean concentration obtained for Ter1 using the gel shown in Figure 10, was 3.1 mg/mL with an R-squared value of 0.9789. Unfortunately, the bands for Sgs3 were predominantly not very clear and the proteins with the desired molecular weight at around 17 kDa were not visible, making the densitometry analysis on the Sgs3 sample not possible. This may have been due to the poor solubility of isolated recombinant proteins making it difficult to load the proteins purified inclusion bodies on the gel. Only light bands are visible on the gel but not of the correct size. Therefore, the expressed Sgs3 was likely lost during the dialysis procedure. Further analysis suggested, that the protein was not properly dissolved. However complete solubilisation could be easily achieved by using formic acid as a solvent.

5.3. Cell Culture Tests

The coating on hydrophobic surfaces is required for *Drosophila* cells such as *Cl.8+* which must be attached to a surface to grow. Cell culture tests were conducted to according to section 4.9.1. Cell cultures were cultivated on polystyrene petri dishes that were coated with the respective adhesive recombinant protein. The growth of cell cultures was monitored over 3 days.

BSA, a protein without any known adhesive properties, was used as a negative control variable for comparison. Each test was conducted twice for increased reliability. Only healthy cells were counted according to their morphology (extended shape, presence of pseudopodia).

It should be noted that Ter1 did not fully dissolve in the distilled water.

5.3.1. Growth of *Drosophila* Cl.8+ cells on Ter1 coating

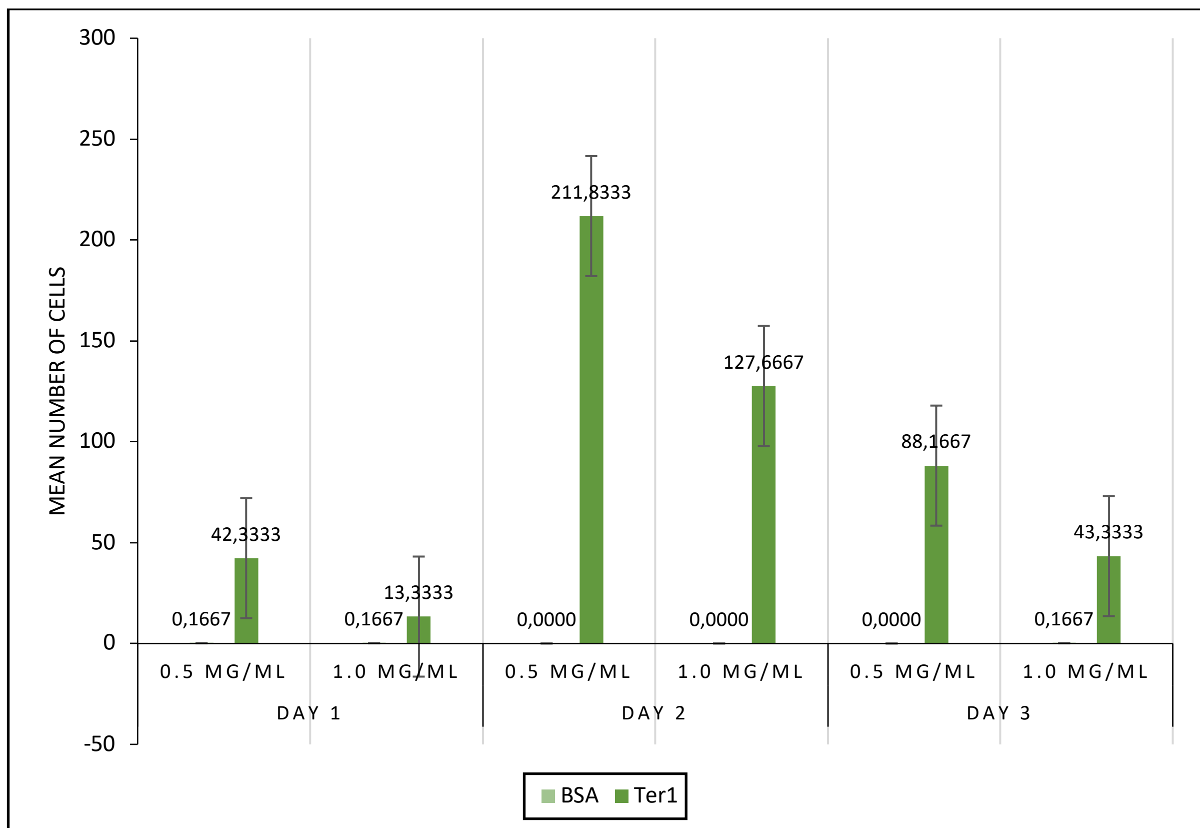


Figure 11: Cell culture mean number of cells for Ter1 in the first trial. On petri dishes, circles with a diameter of 8mm (0.50 cm^2) were coated with the purified proteins dissolved in water. BSA was used as a negative control variable. There were 6 (3 of each circle) samples taken to calculate the mean number of surviving cells. The specific standard deviation and standard error can be found in section 9.2, Appendix 2 – Additional results.

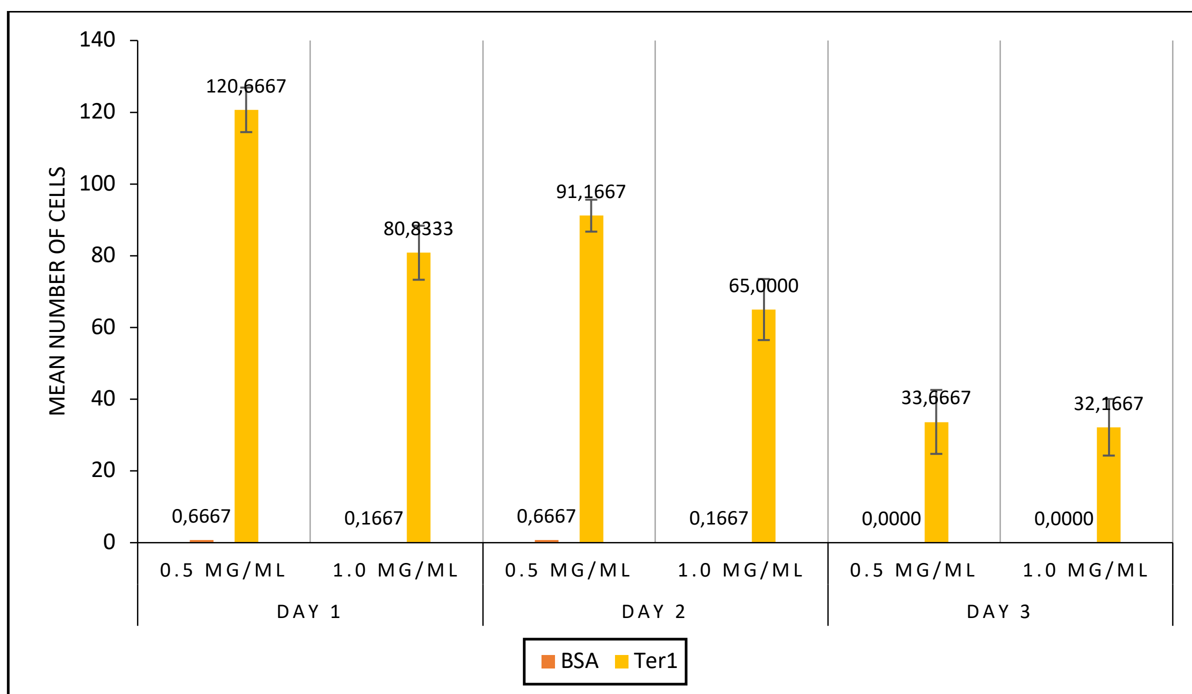


Figure 12: Cell culture mean number of cells for Ter1 in the second trial. On petri dishes, circles with a diameter of 8mm (0.50 cm^2) were coated with the purified proteins dissolved in water. BSA was used as a negative control variable. There were 6 (3 of each circle) samples taken to calculate the mean number of surviving cells. The specific standard deviation and standard error can be found in section 9.2, Appendix 2 – Additional results.

As shown in Figure 11 and 12 which represent the *Drosophila* cell growth on the purified Ter1 protein coating, it was observed that the number of cells generally decreased as the number of days increased for both recombinant protein concentrations. In Figure 12, there was a sharp decrease visible between Day 2 and 3. The cell count reduced by at least half for both concentrations. Although this pattern was also generally observed in the first batch (see Figure 11), it must be noted that the standard error and standard deviation (see section 9.2, Appendix 2 – Additional results, Table 12) were much larger for the concentration of 1.0 mg/mL during the first trial.

The significant decrease in cell growth resulted from the growing number of clusters that were formed (cells in clusters were not counted) due to the dense population, strongly suggesting that the coating of our recombinant proteins may not be stable and the cells detached. Moreover, there was a larger cell culture growth detected when a lower protein concentration (0.5 mg/mL) was used on average in both trials, on all days. Figure 12 illustrates a significant increase in cell growth from Day 1 to Day 2 for the concentration 0.5 mg/mL. This indicates the ability of the cell population to grow with the aid of Ter1 as an adhesive protein coating if the space is not limited. It must be emphasized that the average number of cells in the first trial for Ter1 (1.0 mg/mL) was significantly lower than in the second trial (see Figure 11).

5.3.2. Growth of Drosophila cells on Sgs3 coating

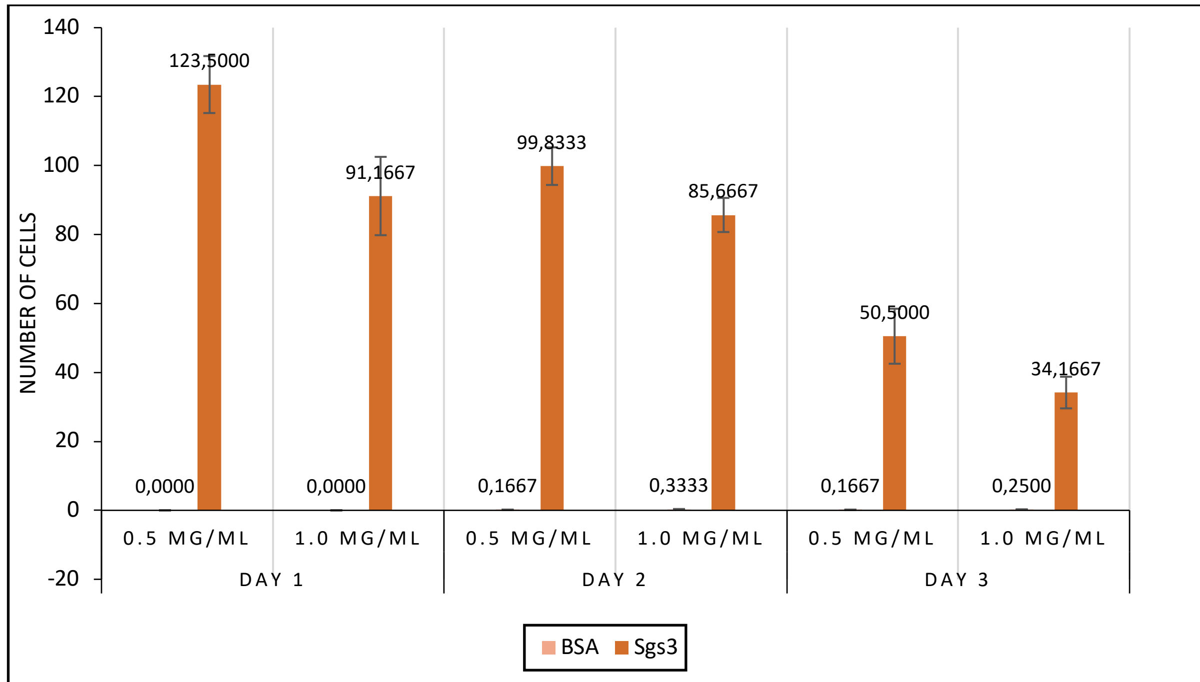


Figure 13: Cell culture mean number of cells for Sgs3 in the first trial. On petri dishes, circles with a diameter of 8mm (0.50 cm^2) were coated with the purified recombinant proteins dissolved in water. BSA was used as a negative control variable. There were 6 (3 of each circle) samples taken to calculate the mean number of surviving cells. The specific standard deviation and standard error can be found in section 9.2, Appendix 2 – Additional results.

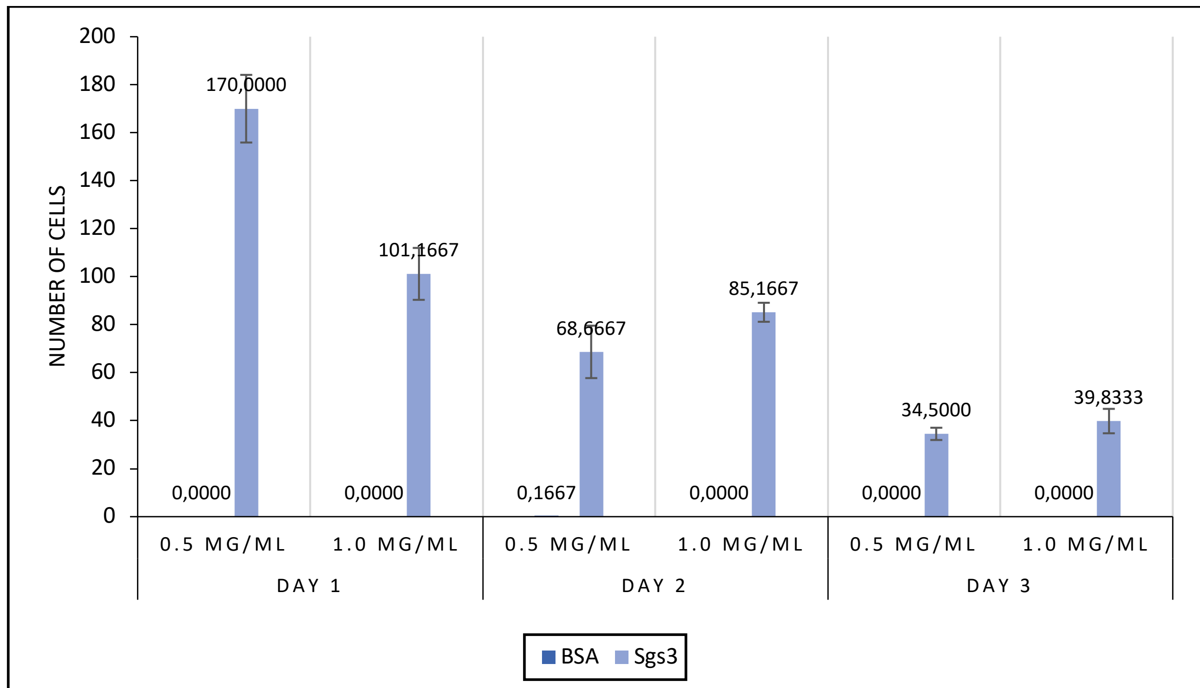
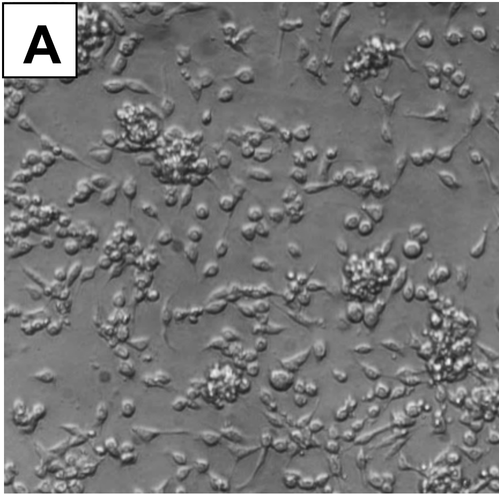


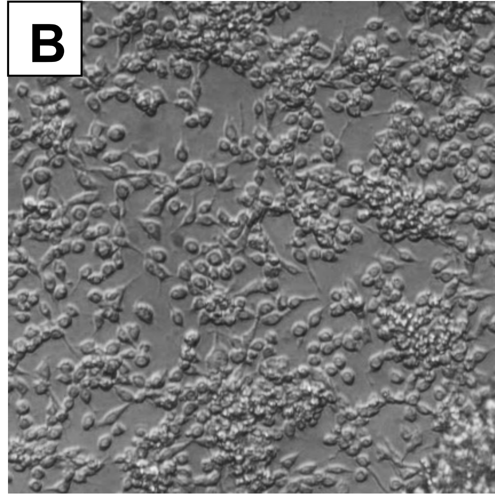
Figure 14: Cell culture mean number of cells for Sgs3 in the second trial. On petri dishes, circles with a diameter of 8mm (0.50 cm^2) were coated with the purified proteins dissolved in water. BSA was used as a negative control variable. There were 6 (3 of each circle) samples taken to calculate the mean number of surviving cells. The specific standard deviation and standard error can be found in section 9.2, Appendix 2 – Additional results.

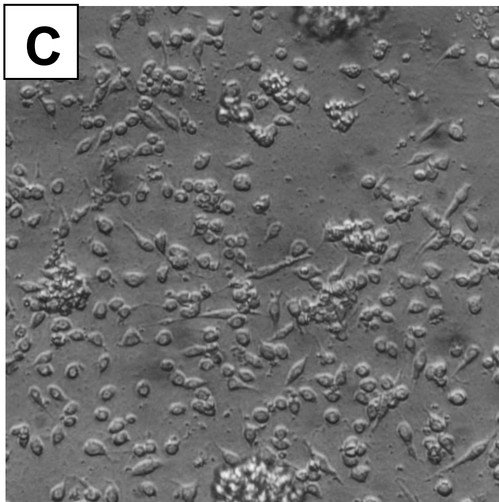
Overall, the growth of Cl.8+ cells on Sgs3 surface coating presented similar growth patterns to Ter1. However, Figure 13 and 14, show that cell growth for areas with Sgs3 protein coating of 1.0 mg/mL decreased at a slower rate than in areas of protein coating with lower

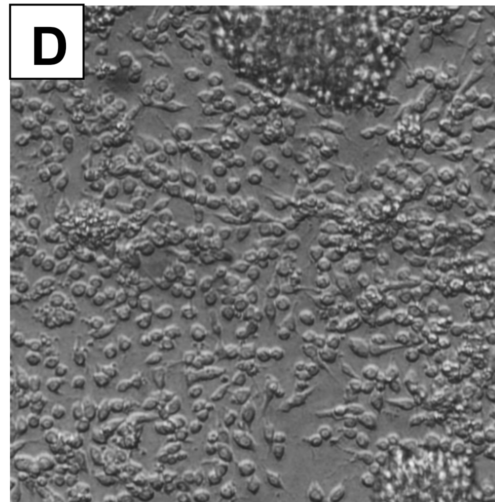
concentration. In Figure 14, the cell count significantly decreases by 60% between Day 1 and Day 2 for the 0.5 mg/mL coating. This decline was steep enough for the more highly concentrated sample to surpass the cell count of the former after 48 hours. The protein coating with higher concentration only experienced a decrease by 26%. With increasing time, the population multiplied under preferable conditions, yet as the space remained limited the population became too dense and cells started to cluster causing a large number of the cells to die. Therefore, if the sample started with a large population, it was more subject to clustering due to the cell overgrowth and poor adherence of the proteins to the cells causing the population to reduce significantly over time. This phenomenon can be visually observed in Figures 15.

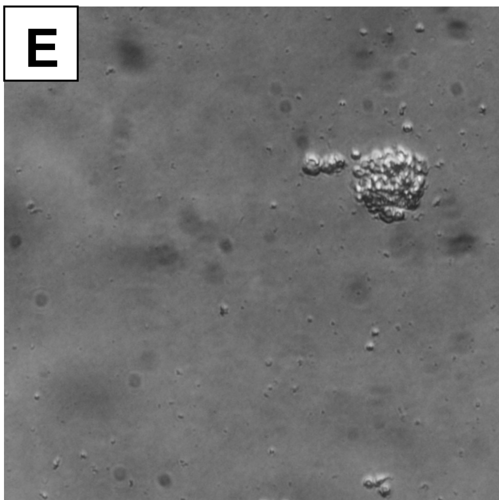
However, for the cell culture tests using our recombinant proteins, it can be observed that they support cell growth. Some adhesive properties were evident as there was a significant cell growth in all cases especially in comparison to the negative control (BSA) which nearly showed no cells on all days for all trials. The minimum number of cells adhering on BSA coated surfaces is an important control variable for comparison as it presents the amount of background 'noise' (Hemphries, 2009). Taken together, these results show that recombinant Ter1 and Sgs3 proteins*,* are able to coat the surface of the polystyrene and support grow of adherent cells to some extent.











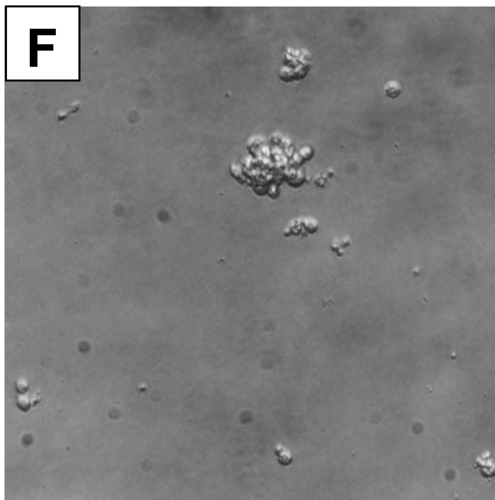


Figure 15: Cl8+ cell adhesive test with (A) 0.5 mg/ml of *Sgs3* after 1 day, (B) Cl8+ cell adhesive test with 0.5 mg/ml of *Sgs3* after 2 days at same location as A, (C) 0.5 mg/ml of Ter1 after 1 day, (D) Cl8+ cell adhesive test with 0.5 mg/ml of Ter1 after 2 days at same location as C, (E) 0.5 mg/ml of BSA after 1 day, (F) 0.5 mg/ml of BSA after 2 days at same location as E. Petri dishes were coated with the recombinant protein before the cells were applied. Only healthy cells were counted for the analysis. Clusters were not considered Although the increase in cell numbers increased visibly, more clusters also visibly formed

5.4. SEM analysis

For further characterization of adhesive proteins, we compared our Ter1 protein with a closely related control protein, BmS2 – GL2*,* kindly received from Tereza Konikova (Tereza Staskova 2012, Master thesis, Faculty of Sciences JCU) for comparison. The BmS2 – GL2 is a hybrid protein between Ter1 and a mussel glue (*Mytilus edulis* prot. Q25460). The Ter1 gene sequence encoding the recombinant BmS2 – GL2 can be found in section 9.2. Appendix 2 – Additional results. The control protein, BmS2 – GL2, was dissolved in concentrated formic acids and directly spread on the plastic or glass surface.

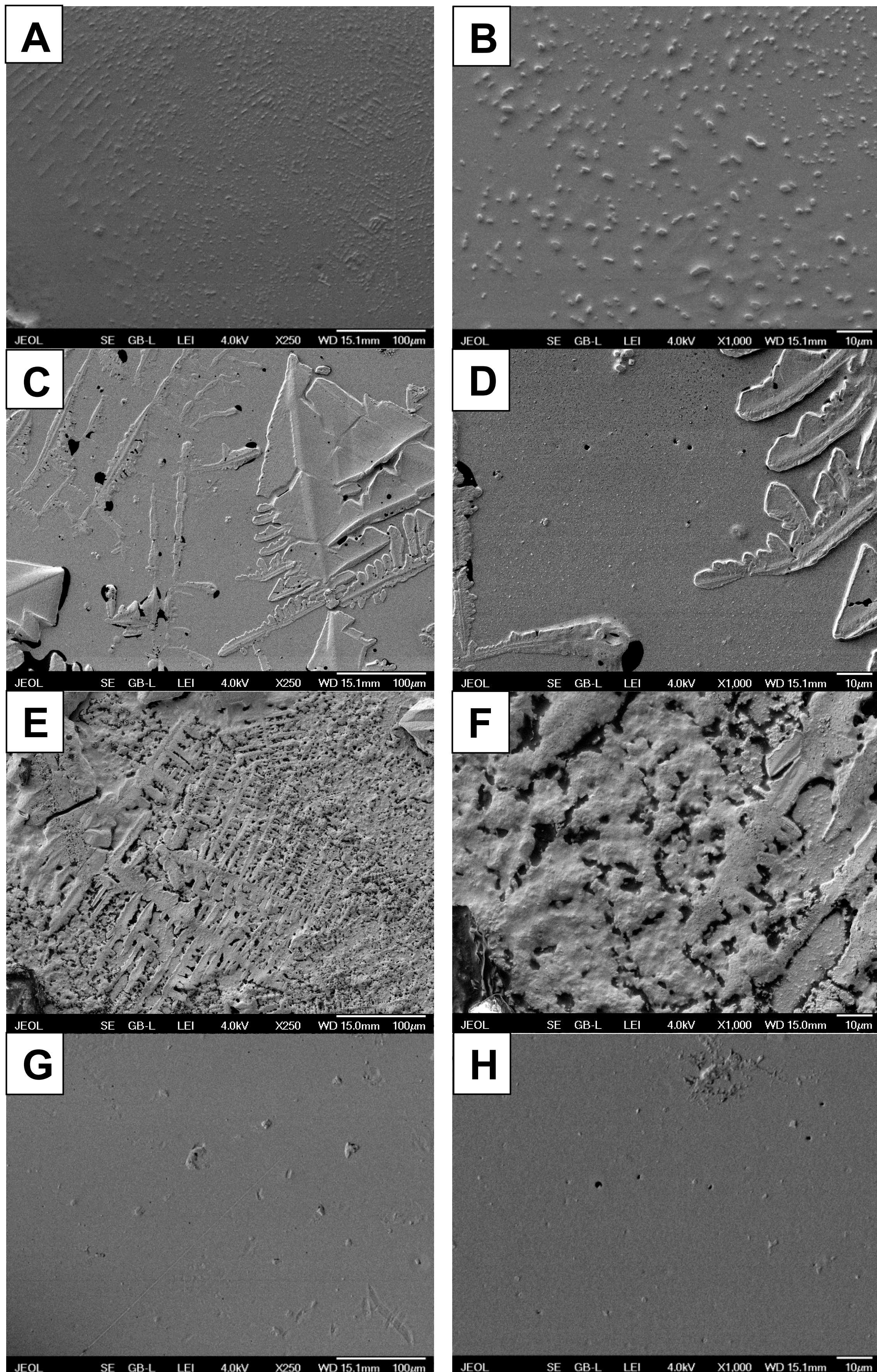


Figure 16: Analysis of physical structure of recombinant proteins, Ter1 and *BmS2 -GL2,* on plastic surfaces (petri dishes) by Scanning Electron Microscopy. SEM of (A) control variable *PBS* 250, (B) control variable *PBS* 1000x, (C) Ter1 layer 250x, (D) Ter1 layer 1000x, (E) BmS2-GL2 layer 250x, (F) BmS2-GL2 layer 1000x, (G) Commercial sericin (Sigma #S5201) 250x, (H) Commercial sericin (Sigma #S5201)

Through the analysis of the proteins, the spreading efficiency of the protein was determined. For the control samples (PBS and commercial sericin) seen in Figure 16, A – B and G – H respectively, some small salt crystals without protein layer are generally visible. The PBS sample, however, showed the formation of small bumps. This is still strongly in contrast to the related recombinant protein, BmS2-GL2*,* which formed a high number of clusters and, thus, a highly porous cement (see Figure 16, E and F). Ter1, while still not forming smooth cement surfaces like the control variables, appeared structurally less damaged where only a few larger clusters formed (see Figure 16, C and D).

Figure 17 also clearly presented the same pattern when the proteins were applied to a glass surface. Ter1 was observed to have a more even cement and nonporous spreading on a glass surface than a plastic surface when comparing Figure 16 (B and C) and Figure 17 (A and B). The highly porous surface and structural damage were still present when BmS2 -GL2 was applied to a glass surface instead of a plastic one.

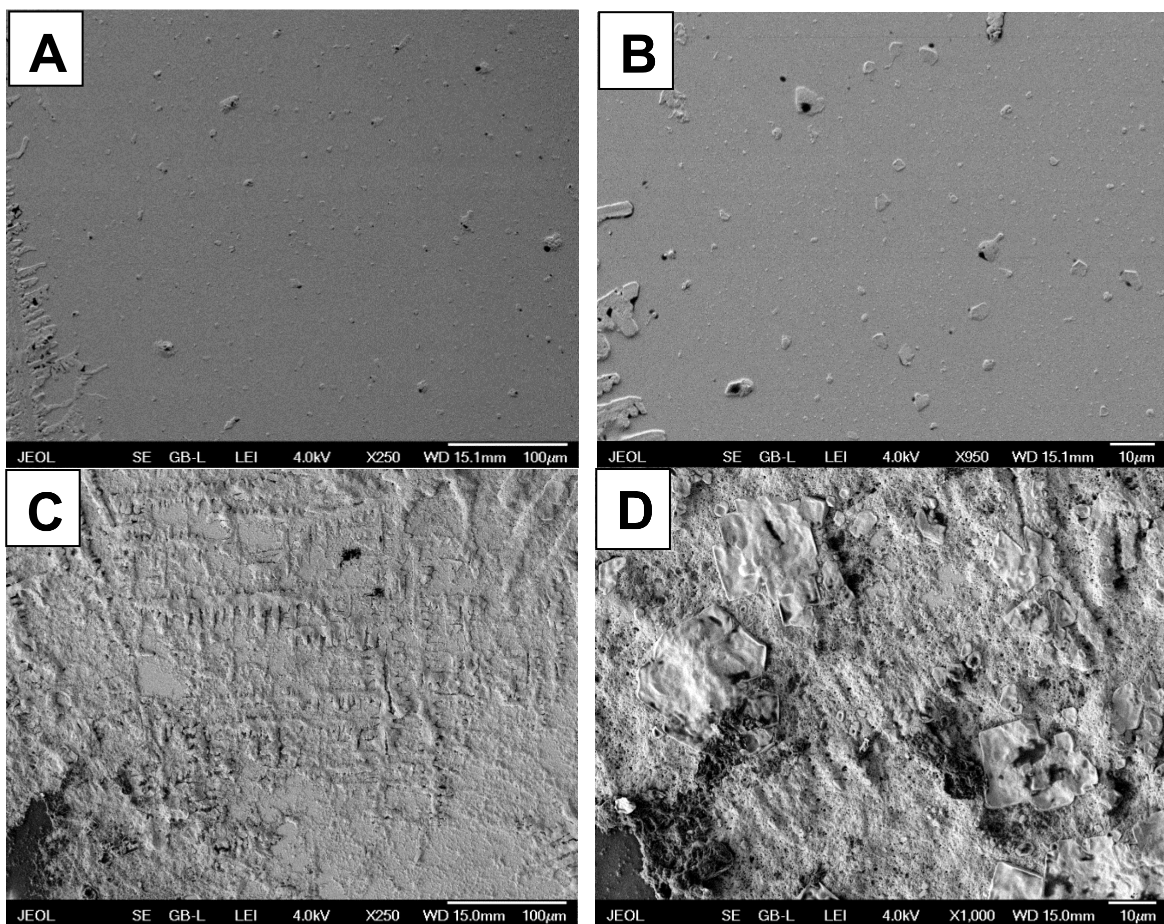


Figure 17: Analysis of physical structure of recombinant proteins, Ter1 and BmS2-GL2, on glass by Scanning Electron Microscopy. SEM of (A) Ter1 layer 250x, (B) Ter1 layer 1000x, (C) BmS2-GL2 layer 250x, (D) BmS2-GL2 layer 1000x.

6. Discussion

6.1. Protein expression and purification

While target proteins obtained from native hosts often present challenges, recombinant protein production in prokaryotic expression systems offers various advantages that allow a high-yielding and cost-effective approach with simple procedures (Sorensen et al, 2005). For small scale productions of recombinant Ter1 protein, it was possible to obtain a sufficient amount of yield in standard LB growth media. Ter1 was observed to be much heavier than its expected theoretical molecular weight. This corresponds to the electrophoretic mobility of studied serine-rich proteins where sericins were migrating at double the size than expected (Huang et al, 2003).

While the expression and purification of Ter1 in soluble form presented fewer challenges, Sgs3 was difficult to purify in its soluble form and the techniques had to be optimised. Although the methods for recombinant expression and purification are simple, factors such as the culture to flask volume ratio, IPTG concentration, pH, temperature, agitation and induction period must often be tested and optimised to pursue the most efficient methodology due to recombinant proteins being diverse and thus may each present different challenge during the expression and purification in practice (Collins et al., 2013; Rosano and Ceccarelli, 2014). This was evident through the analysis of Sgs3 where contaminating proteins were eluted along the low amount of desired protein during the purification although the same methods as with Ter1 were used. Gräslund et al. (2008) suggests that it is relatively common for an endogenous *E. coli* protein to elute during purification if the recombinant protein expression is limited as it occurred in our case. They thus suggest performing mass spectrometry for protein identification as it was done in section 5.2.2.

For the proteomic analysis, it should be highlighted that the proteins are identified according to the predicted tryptic peptides (trypsin exclusively cleaves C-terminal to arginine and lysine residues) and their molecular weights that are found in either a public domain database or a custom-made database. Thus, if the result only returns a single aligning sequence, it should not be immediately considered as a direct identification or verification of the protein due a possible random error occurring (Rohrbough et al, 2006). A minimum of two corresponding sequences would be necessary for a true protein verification. Rohrbough et al. (2006) describes the increasing number of corresponding sequences to the protein as an increasing confidence in protein identification.

The elimination of inclusion bodies is difficult under the stress conditions applied for the overexpression of proteins. Such conditions cause proteins that are generally difficult to express to form insoluble aggregates (Hwang et al, 2014). Recombinant proteins are densely packed in the inclusion bodies (Ferrer-Miralles et al., 2015) and due to this, the extraction of proteins from inclusion bodies was preferable for Sgs3 as the purification from bacterial lysate did not yield the soluble target protein, isolated from endogenous bacterial contaminants. The purification of the proteins from inclusion bodies, however, requires harsh dissolving procedures and denaturation due their lacking solubility. Although this is an essential step for the purification from inclusion bodies, it likely caused the misfolding during the recovery of proteins (Hwang et al, 2014). Due to these procedures, it is often found that that the yield of active proteins, recovered from inclusion bodies is very low although the proteins in inclusion bodies may be pure (Yamaguchi et al., 2014) - which was also evident for the analysis of our overexpressed Sgs3.

For further research, different cleavage methods and refolding buffer systems to maximize the recovery of inclusion bodies by optimising protein concentration and physicochemical conditions are recommended to explore as a significant yield can be prospectively acquired from inclusion bodies (Basu et al., 2011). More various conditions for the growth of the bacterial cultures to optimize the yield are also suggested. Currently, the demand for resources and time needed to produce a sufficient amount of our recombinant proteins is still relatively high.

6.2. Properties of Ter1 and Sgs3

Efficient cell adhesion onto extracellular surfaces in tissue engineering is a crucial topic (Ciofani et al. 2013). The investigation of the coating capacity and ability to support cell-adhesion of our recombinant proteins showed that proteins formed coatings on polystyrene, strongly decreasing the hydrophobicity of the surface. The protein coating allowed cell attachment and growth which suggested that recombinant proteins showed some adhesive properties. It was not assessed whether the quality of protein coating was also in correlation to the death of the cells as Humphries (2009) describes that different substrates can also cause changes in cell morphology. The significant reduction of healthy cells over time was assumed to be mainly caused by the limited cell adherence properties of the recombinant proteins. Vancha et al. (2004) reported that in the absence of coating, the cells adhere to the surface poorly and strongly adhere to each other instead resulting in the formation of clusters. The weak adherence thus caused a significant reduction of healthy cells through medium changes as it was similarly observed for Ter1 and Sgs3 coating due to uneven cell distribution (Vancha

et al., 2014). In addition, overgrowth may be another contributing factor to the tendency of cluster formation. It is therefore suggested to conduct more cell culture experiments with a lower concentration of the cells in order to observe a more gradual growth pattern and allow a closer perspective on the survival of cells as well as determine the minimum concentration needed for cell attachment to occur.

This would also reduce the impact of random errors where the number of observed cells was significantly low. This may have been caused by the movement of the protein coating by external forces such as relocating the petri dish before it dried in the designated area. This resulted in a significantly low number of growing cells in some of the analysed areas. The insufficient or improper coating of the protein is a common problem that is encountered when measuring cell adhesion (Humphries, 2009).

An important property of surface coatings is the stability of such coating (Winandy et al. 2018). Due to low solubility and denatured character of our isolated recombinant proteins, their strength of adhesion and coating durability was very poor while the natural proteins are considered as very strong adhesives.

Solubilization is a critical procedure needed for the purification of proteins from inclusion bodies. 70% formic acid is a very strong denaturing solvent, which can formylate serine and threonine residues and cause permanent protein damage (Hwang et al., 2014). Our Ter1 protein was completely insoluble in water and thus had to be exclusively dissolved in formic acid for the SEM analysis. The use of formic acid, however, resulted in some clustering. For the highly related protein, BmS2-GL2, it caused additional damage as the formic acid drastically degraded the structure of the protein, as evident from the SEM figures. For further use of our recombinant adhesives it will be necessary to replace formic acid with some milder detergent.

Another challenge that was encountered was the over-time continuously decreasing amount of the lyophilized Ter1 sample stored at -20°C as it became completely anhydrous within a short period of time. As it finally became insoluble in water, it was dissolved in formic acid. Evidently the recombinant proteins are sensitive and challenging to handle.

6.3. Application of adhesive recombinant proteins

Silk materials possesses great potential for the application as biomedical coating material because of the non-toxic character, the good biocompatibility and the low body immune response (Bokner et al 2016). Although most of the data is required to be repeated for further verification, there is a strong prospective potential for such proteins to be applied

in various fields. Until present, it has been proven to be difficult to artificially synthesise adhesive proteins without losing their natural properties at a larger scale.

Studies on recombinant spider silk-like proteins expressed in *E. coli* for the potential incorporation in biomaterial showed similar problems, (Bowen et al., 2018; Teulé et al., 2011). These proteins revealed that there is a certain size limit for bacterial expression, and a larger construct produced a lower yield despite of the strong T7 promoter employed for gene expression (Teulé et al., 2009). However, while the plasmids with *Ter1* and *Sgs3* constructs contain ampicillin resistance genes, Teulé et al. (2007) replaced this gene with the kanamycin resistance gene which resulted in higher yield of the target proteins due to longer gene expression. In comparison to the yield of our recombinant Ser2 silk protein, the yield of the spider silk-like protein reaches up to 5 times as much (Teulé et al., 2009). The published results with the recombinant synthesis of spider proteins show that the artificially produced proteins can retain their advantageous natural properties and can prospectively be engineered this way in order to yield large quantities and be cross-linked into new biomaterials (Scheibel, 2004). Our Ter1 and Sgs3 production, still needs further optimisation and analyses of their properties to resemble the native proteins.

Other recombinant silk materials studied in other laboratories include honeybee silk proteins, which unlike those of silkworms and spiders are easier to recombinantly produce due to smaller and less repetitive sequences (Campbell et al., 2014). Sutherland et al. (2019) reported the purification of honeybee silk proteins from inclusion bodies to have a purity of over 99% which is exceedingly high due to the formers ability to readily self-assemble into the α-helices structure.

Further research is needed for optimising the key steps of our recombinant protein isolation and keeping their structure in order to mass produce the recombinant proteins in heterologous expression systems. In the future, this would allow further characterisation of physical properties such as strength, toughness and extensibility as well as nontoxic and antibacterial behaviour. However, both Ter1 and Sgs3 have shown the potential to become bioadhesives useful for surface coating and the support of adherent cell growth.

7. Conclusion

7.1. Optimising the expression of adhesive recombinant proteins in E. coli

- Within this work the recombinant proteins Ter1 and Sgs3 were expressed in bacterial expression systems in *E. coli* BL21 cells.
- For small-scale analysis, different proteins were expressed and analysed before choosing Ter1 and Sgs3 as preferable candidates.
- Different conditions were tested for the growth of bacterial cultures to find the optimal expression conditions of the proteins.

7.2. Purification optimisation of recombinant Bombyx sericin 2 (Ter1) and recombinant Drosophila Salivary gland secretion 3 (Sgs3)

- Ter1 was purified under native conditions from bacterial lysate by affinity chromatography. The purification of Sgs3 from bacterial lysate was contaminated by endogenous bacterial proteins. Sgs3 was following purified from bacterial inclusion bodies under denaturing conditions, which produced slightly better yields. Purified recombinant proteins were dialysed and lyophilised
- Problems for Sgs3 purification on a larger scale remained. Methods thus require further optimisation for inclusion body solubilization and protein refolding.

7.3. Preliminary characterization of isolated recombinant proteins

- Tissue cell culture tests were conducted by coating polystyrene surfaces with the recombinant proteins and applying Cl.8+ cells which require attachment for cell growth.
- The cell growth was monitored over three days for 2 different concentrations of coated proteins used. The quantitive tests show that recombinant Ter1 and Sgs3 at these stages were significantly different from control BSA. However, the coating showed relatively low stability.
- Recombinant Ter1 and a control (highly related protein, BmS2 – GL2) protein coating was visualized using SEM. Ter1 showed crystal-like structures and both displayed signs of structural damage due to the formic acid solvent.
- Although more tests should be conducted on different surfaces, both recombinant sericin and salivary gland secretion proteins have the potential to be applied as adhesive biomaterials in the future when the properties of the protein are less affected by isolation and solubilization procedures.

8. *References*

Anthis, N.J., Clore, G.M., 2013. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm: Sequence-Specific Protein Concentration at 205 nm. Protein Science 22, 851–858. https://doi.org/10.1002/pro.2253

Aramwit, P., Siritientong, T., Srichana, T., 2012. Potential applications of silk sericin, a natural protein from textile industry by-products. Waste Manag Res 30, 217–224. https://doi.org/10.1177/0734242X11404733

Basu, A., Li, X., Leong, S.S.J., 2011. Refolding of proteins from inclusion bodies: rational design and recipes. Appl Microbiol Biotechnol 92, 241–251. https://doi.org/10.1007/s00253-011-3513-y

Betz, O., 2010. Adhesive Exocrine Glands in Insects: Morphology, Ultrastructure, and Adhesive Secretion, in: von Byern, J., Grunwald, I. (Eds.), Biological Adhesive Systems. Springer Vienna, Vienna, pp. 111–152. https://doi.org/10.1007/978-3-7091-0286-2_8

Bowen, C.H., Dai, B., Sargent, C.J., Bai, W., Ladiwala, P., Feng, H., Huang, W., Kaplan, D.L., Galazka, J.M., Zhang, F., 2018. Recombinant Spidroins Fully Replicate Primary Mechanical Properties of Natural Spider Silk. Biomacromolecules 19, 3853–3860. https://doi.org/10.1021/acs.biomac.8b00980

Brennan, M., 1991. Fibrin glue. Blood Reviews 5, 240–244. https://doi.org/10.1016/0268-960X(91)90015-5

Bruschi, M.L., de Freitas, O., 2005. Oral Bioadhesive Drug Delivery Systems. Drug Development and Industrial Pharmacy 31, 293–310. https://doi.org/10.1081/DDC-52073

Campbell, P.M., Trueman, H.E., Zhang, Q., Kojima, K., Kameda, T., Sutherland, T.D., 2014. Cross-linking in the silks of bees, ants and hornets. Insect Biochemistry and Molecular Biology 48, 40–50. https://doi.org/10.1016/j.ibmb.2014.02.009

Ciofani, G., Genchi, G.G., Liakos, I., Athanassiou, A., Mattoli, V., Bandiera, A., 2013. Human recombinant elastin-like protein coatings for muscle cell proliferation and differentiation. Acta Biomaterialia 9, 5111–5121. https://doi.org/10.1016/j.actbio.2012.10.016

Collins, T., Azevedo-Silva, J., da Costa, A., Branca, F., Machado, R., Casal, M., 2013. Batch production of a silk-elastin-like protein in E. coli BL21(DE3): key parameters for optimisation. Microb Cell Fact 12, 21. https://doi.org/10.1186/1475-2859-12-21

Da Lage, J.-L., Thomas, G.W.C., Bonneau, M., Courtier-Orgogozo, V., 2019. Evolution of salivary glue genes in Drosophila species. BMC Evol Biol 19, 36. https://doi.org/10.1186/s12862-019-1364-9

Desjardins, P., Hansen, J.B., Allen, M., 2009. Microvolume Protein Concentration Determination Using the NanoDrop 2000c Spectrophotometer. JoVE 1610. https://doi.org/10.3791/1610

Donnelly, R., Shaikh, R., Raj Singh, T., Garland, M., Woolfson, Ad., 2011. Mucoadhesive drug delivery systems. J Pharm Bioall Sci 3, 89. https://doi.org/10.4103/0975-7406.76478

Elzoghby, A.O., Elgohary, M.M., Kamel, N.M., 2015. Implications of Protein- and Peptide-Based Nanoparticles as Potential Vehicles for Anticancer Drugs, in: Advances in Protein Chemistry and Structural Biology. Elsevier, pp. 169–221. https://doi.org/10.1016/bs.apcsb.2014.12.002

Ferrer-Miralles, N., Saccardo, P., Corchero, J.L., Xu, Z., García-Fruitós, E., 2015. General Introduction: Recombinant Protein Production and Purification of Insoluble Proteins, in: García-Fruitós, E. (Ed.), Insoluble Proteins. Springer New York, New York, NY, pp. 1–24. https://doi.org/10.1007/978-1-4939-2205-5_1

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607

Gräsland, S., Nordlund, P., Weigelt, J., et al., 2008. Protein production and purification. Nat Methods 5, 135–146. https://doi.org/10.1038/nmeth.f.202

Greer, A.M., Huang, Z., Oriakhi, A., Lu, Y., Lou, J., Matthews, K.S., Bondos, S.E., 2009. The Drosophila Transcription Factor Ultrabithorax Self-Assembles into Protein-Based Biomaterials with Multiple Morphologies. Biomacromolecules 10, 829–837. https://doi.org/10.1021/bm801315v

Griffiths, A.J., Gelbart, W., Miller, J., Lewontin, R.C., 1999. Modern Genetic Analysis, 2nd Print. ed. W.H. Freeman, New York.

Huang, J., Valluzzi, R., Bini, E., Vernaglia, B., Kaplan, D.L., 2003. Cloning, Expression, and Assembly of Sericin-like Protein. J. Biol. Chem. 278, 46117–46123. https://doi.org/10.1074/jbc.M307792200

Humphries, M.J., 2009. Cell Adhesion Assays, in: Even-Ram, S., Artym, V. (Eds.), Extracellular Matrix Protocols. Humana Press, Totowa, NJ, pp. 203–210. https://doi.org/10.1007/978-1-59745-413-1_14

Hwang, P.M., Pan, J.S., Sykes, B.D., 2014. Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in *Escherichia coli*. FEBS Letters 588, 247–252. https://doi.org/10.1016/j.febslet.2013.09.028

Kludkiewicz, B., Takasu, Y., Fedic, R., Tamura, T., Sehnal, F., Zurovec, M., 2009. Structure and expression of the silk adhesive protein Ser2 in Bombyx mori. Insect Biochemistry and Molecular Biology 39, 938–946. https://doi.org/10.1016/j.ibmb.2009.11.005

Kord Forooshani, P., Lee, B.P., 2017. Recent approaches in desi42ahn42hesiveshesive materials inspired by mussel adhesive protein. J. Polym. Sci. Part A: Polym. Chem. 55, 9–33. https://doi.org/10.1002/pola.28368

Korge, G., 1975. Chromosome puff activity and protein synthesis in larval salivary glands of Drosophila melanogaster. Proceedings of the National Academy of Sciences 72, 4550–4554. https://doi.org/10.1073/pnas.72.11.4550

Kundu, S.C., Dash, B.C., Dash, R., Kaplan, D.L., 2008. Natural protective glue protein, sericin bioengineered by silkworms: Potential for biomedical and biotechnological applications. Progress in Polymer Science 33, 998–1012. https://doi.org/10.1016/j.progpolymsci.2008.08.002

Kunz, R.I., Brancalhão, R.M.C., Ribeiro, L. de F.C., Natali, M.R.M., 2016. Silkworm Sericin: Properties and Biomedical Applications. BioMed Research International 2016, 1–19. https://doi.org/10.1155/2016/8175701

Lehmann, M., 1996. Drosophila Sgs genes: Stage and tissue specificity of hormone responsiveness. Bioessays 18, 47–54. https://doi.org/10.1002/bies.950180110

Michaille, J.-J., Garel, A., Prudhomme, J.-C., 1990. Cloning and characterization of the highly polymorphic Ser2 gene of Bombyx mori. Gene 86, 177–184. https://doi.org/10.1016/0378-1119(90)90277-X

Mühlmann, M., Forsten, E., Noack, S., Büchs, J., 2017. Optimizing recombinant protein expression via automated induction profiling in microtiter plates at different temperatures. Microb Cell Fact 16, 220. https://doi.org/10.1186/s12934-017-0832-4

Novagen, Inc. pET system manual. United States and Canada. T[B0]55 8th Edition 02[/9]9. 8th ed. 1991. Available from: http:/, n.d.

Peel, D.J., Milner, M.J., 1992. The expression of PS integrins in Drosophila melanogaster imaginal disc cell lines. Roux's Arch Dev Biol 201, 120–123. https://doi.org/10.1007/BF00420423

Novagen, Inc. pET system manual. United States and Canada. TB055 8th Edition 02/99. 8th ed. 1991. Available from: http://research.fhcrc.org/content/dam/stripe/hahn/methods/biochem/pet.pdf

Petrie, E.M., 2007. Handbook of adhesives and sealants, 2nd ed. ed, McGraw-Hill handbooks. McGraw-Hill, New York.

Richter, K., Grunwald, I., von Byern, J., 2018. Bioadhesives, in: da Silva, L.F.M., Öchsner, A., Adams, R.D. (Eds.), Handbook of Adhesion Technology. Springer International Publishing, Cham, pp. 1597–1640. https://doi.org/10.1007/978-3-319-55411-2_53

Rohrbough, J.G., Breci, L., Merchant, N., Miller, S., Haynes, P.A., 2006. Verification of single-peptide protein identifications by the application of complementary database search algorithms. J Biomol Tech 17, 327–332.

Rosano, G.L., Ceccarelli, E.A., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. Front. Microbiol. 5. https://doi.org/10.3389/fmicb.2014.00172

Scheibel, T., 2004. Spider silks: recombinant synthesis, assembly, spinning, and engineering of synthetic proteins. Microb Cell Fact 3, 14. https://doi.org/10.1186/1475-2859-3-14

Sørensen, H., Mortensen, K., Peter, H., 2005. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. Microb Cell Fact 4, 1. https://doi.org/10.1186/1475-2859-4-1

Sørensen, H.P., Mortensen, K.K., 2005. Advanced genetic strategies for recombinant protein expression in Escherichia coli. Journal of Biotechnology 115, 113–128. https://doi.org/10.1016/j.jbiotec.2004.08.004

Sutherland, T.D., Vashi, A.V., Kardia, E., Sriskantha, A., Rapson, T.D., Hall, R.N., Werkmeister, J.A., 2019. Biocompatibility and immunogenic response to recombinant honeybee silk material. J. Biomed. Mater. Res jbm.a.36695 https://doi.org/10.1002/jbm.a.36695

Takasu, Y., Yamada, H., Tamura, T., Sezutsu, H., Mita, K., Tsubouchi, K., 2007. Identification and characterization of a novel sericin gene expressed in the anterior middle silk gland of the silkworm Bombyx mori. Insect Biochemistry and Molecular Biology 37, 1234–1240. https://doi.org/10.1016/j.ibmb.2007.07.009

Takasu, Y., Yamada, H., Tsubouchi, K., 2002. Isolation of Three Main Sericin Components from the Cocoon of the Silkworm, *Bombyx mori*. Bioscience, Biotechnology, and Biochemistry 66, 2715–2718. https://doi.org/10.1271/bbb.66.2715

Teramoto, H., Kameda, T., Tamada, Y., 2008. Preparation of Gel Film from *Bombyx mori* Silk Sericin and Its Characterization as a Wound Dressing. Bioscience, Biotechnology, and Biochemistry 72, 3189–3196. https://doi.org/10.1271/bbb.80375

Teulé, F., Cooper, A.R., Furin, W.A., Bittencourt, D., Rech, E.L., Brooks, A., Lewis, R.V., 2009. A protocol for the production of recombinant spider silk-like proteins for artificial fiber spinning. Nat Protoc 4, 341–355. https://doi.org/10.1038/nprot.2008.250

Teulé, F., Furin, W.A., Cooper, A.R., Duncan, J.R., Lewis, R.V., 2007. Modifications of spider silk sequences in an attempt to control the mechanical properties of the synthetic fibers. J Mater Sci 42, 8974–8985. https://doi.org/10.1007/s10853-007-1642-6

Von Byern, J., Grunwald, I. (Eds.), 2010. Biological adhesion systems: from nature to technical and medical application. Springer, Wien ; New York ; London.

Vancha, A.R., Govindaraju, S., Parsa, K.V., Jasti, M., González-García, M., Ballestero, R.P., 2004. Use of polyethyleneimine polymer in cell culture as attachment factor and lipofection enhancer. BMC Biotechnol 4, 23. https://doi.org/10.1186/1472-6750-4-23

Weisman, S., Haritos, V.S., Church, J.S., Huson, M.G., Mudie, S.T., Rodgers, A.J.W., Dumsday, G.J., Sutherland, T.D., 2010. Honeybee silk: Recombinant protein production, assembly and fiber spinning. Biomaterials 31, 2695–2700. https://doi.org/10.1016/j.biomaterials.2009.12.021

Winandy, L., Hilpert, F., Schlebusch, O., Fischer, R., 2018. Comparative analysis of surface coating properties of five hydrophobins from Aspergillus nidulans and Trichoderma reseei. Sci Rep 8, 12033. https://doi.org/10.1038/s41598-018-29749-0

Yamaguchi, H., Miyazaki, M., 2014. Refolding Techniques for Recovering Biologically Active Recombinant Proteins from Inclusion Bodies. Biomolecules 4, 235–251. https://doi.org/10.3390/biom4010235

Yonemura, N., Sehnal, F., Mita, K., Tamura, T., 2006. Protein Composition of Silk Filaments Spun under Water by Caddisfly Larvae. Biomacromolecules 7, 3370–3378. https://doi.org/10.1021/bm060663u

9. Appendix

9.1. Appendix 1 – Additional Materials

9.1.1. Buffers and Solutions

The liquid medium used to cultivate the *E. coli* bacteria is Luria-Bertani (LB) medium. All the media were sterilized or prepared under sterile conditions

Table 2: Composition of Native Lysis Buffer (NLB). The buffer is adjusted to pH 8, by addition of HCl (conc.).

Substance	Concentration / mM
Na_2HPO_4	25
NaCl	250
Lysozyme	0.1 %
diH_2O	
pH	8.0

Table 3: Composition of Native Binding Buffer (NBB). The buffer is adjusted to pH 8, by addition of HCl (conc.).

Substance	Concentration / mM
Na_2HPO_4	25
NaCl	250
diH_2O	
pH	8.0

Table 4: Composition of Native Washing Buffer (NWB). The buffer is adjusted to pH 8, by addition of HCl (conc.).

Substance	Concentration / mM
Na_2HPO_4	25
NaCl	250
Imidazole	20
diH_2O	
pH	8.0

Table 5: Composition of Native Elution Buffer (NEB). The buffer is adjusted to pH 8, by addition of HCl (conc.).

Substance	Concentration / mM
Na_2HPO_4	25
NaCl	250
Imidazole	250
diH_2O	
pH	8.0

Table 6: Composition of Inclusion Body Buffer 1 (IBB 1). The buffer is adjusted to pH 7 by NaOH (1M).

Substance	Concentration / mM
NaCl	100
TRIS/ HCl	100
EDTA	1
diH_2O	
pH	7.0

Table 7: Composition of Inclusion Body Buffer 2 (IBB 2). The buffer is adjusted to pH 7 by HCl conc.

Substance	Content
NaCl	1.5 M
Triton X-100	2 %
EDTA	60 mM
diH_2O	
pH	7.0

Table 8: Composition of Inclusion Body Buffer 3 (IBB 3). The buffer is adjusted to pH 7 by NaOH (1M).

Substance	Concentration / mM
NaCl	100
TRIS/ HCl	100
EDTA	20
diH_2O	
pH	7.0

Table 10: Composition of a 10% separation Gel for a single gel for SDS-PAGE.

Substance	Volume / µL
33% Acrylamide/Bis-acryamide (29:1)	2290
4x Tris-HCl (pH 8.8)	1875
TEMED	5.6
10 % Ammoniumperoxodisulphate	56
diH_2O	3125

Table 11: Composition of a 5% stacking gel for a single gel for SDS-PAGE.

Substance	Volume/ µL
AA (33%)	525
4x Tris-HCl (pH 6.8)	940
TEMED	5.6
APS (10%)	56
diH_2O	2250

Table 12: Composition of the Running Buffer (RB). The percentage of SDS in the mixture is 1%

Substance	Concentration / M
Tris Base	0.25
Glycine	1.92
Sodium dodecyl-sulphate (SDS)	
diH_2O	

Table 7: Composition of the 10x concentrated Electrotransfer Buffer for 1 L

Substance	Mass / g
Tris (hydroxymethyl)-aminomethane (Tris)	30.3
Glycine	144
diH_2O	

Table 7: Composition of Destaining Solution for SDS Page for 1 L

Substance	Volume / mL
Methanol	250
Acetic Acid	100
diH_2O	

Table 8: Composition of one liter of the 1x concentrated ETB.

Substance	Volume/ mL
ETB 10x	100
Methanol	200
diH_2O	700

Table 9: Composition of the 10x concentrated Phosphate-buffered salt solution (PBS) for 1 L3

Substance	Mass/ g
NaCl	80
KCl	2
Na_2HPO_4	14.4
KH_2PO_4	2.4
diH_2O	
pH	7.4

Table 10: Composition of PBS-Tween® (PBS-T). PBS 1x is gained from dilution of the PBS 10x with diH_2O.

Substance	Percentage/ %
PBS 1x	99.5
Tween®	0.05

9.1.2. Protein Purification Kit

Table 11: Protein purification kit

Name	Description	Source / Catalogue Nr.
Pierce™ Centrifugation Columns	Affinity chromatography columns, 10 mL	Pierce #89898
Ni-NTA Purification System	Purification of recombinant proteins	Invitrogen #Invitrogen K95001

9.1.3. SDS-Page analysis materials

Table 12: SDS-PAGE analysis materials

Name	Description	Source / Catalogue Nr.
PageRuler™ Plus Prestained Protein Ladder	Protein Marker, 10 – 250 kDa	Thermo #26619
Pierce™ Prestained Protein MW Marker	Protein Marker, 20 – 120 kDa	Pierce #26612
PageRuler™ Plus Prestained Protein Ladder	Protein Marker, 10 – 180 kDa	Thermo #26616
4–20% Mini-PROTEAN® TGX™ Precast Protein Gels, 15-well, 15 µl #4561096	4–20% precast polyacrylamide gel	Bio-Rad #4561096

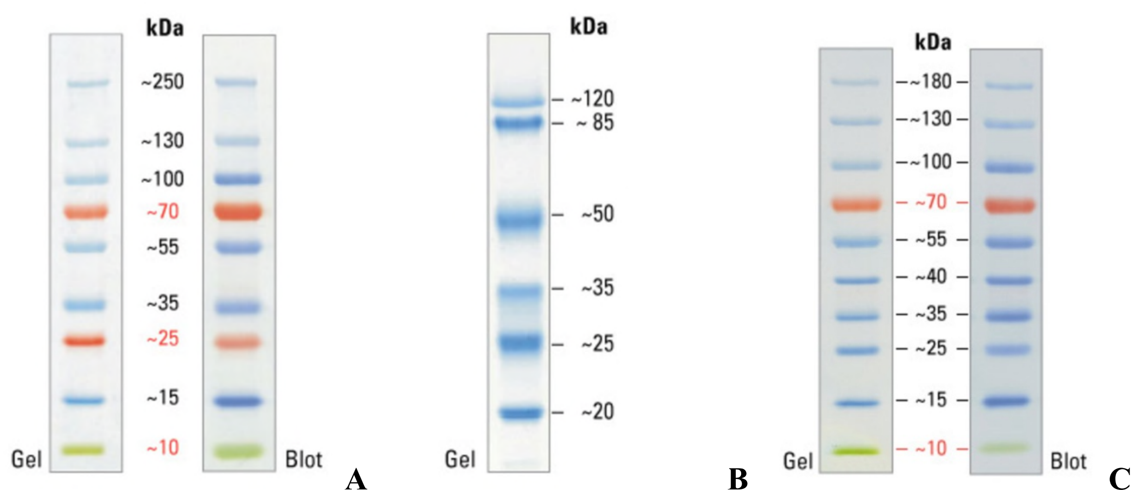


Figure 1: (A) PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa, Thermo #26619 (Thermo Scientific, 2019). (B) Pierce™ Prestained Protein MW Marker, 20 to 120 kDa, Pierce #26612 (Thermo Scientific, 2019) (C) PageRuler™ Plus Prestained Protein Ladder, 10 to 180 kDa, Thermo #26616 (Thermo Scientific, 2019)

9.1.4. Western Blot analysis materials

Table 13: Western Blot analysis materials

Name	Description	Source / Catalogue Nr.
Pierce™ ECL Plus Western Blotting Substrate	Western blot imaging	Pierce #32132
Monoclonal Anti-polyHistidine antibody produced in mouse	Primary antibody, 1:1000 dilution	Sigma #H1029
Anti-Mouse IgG (whole molecule)–Peroxidase antibody produced in goat	Secondary antibody, 1:5000	Sigma #A4416
Nitrocellulose Membrane, 45µm	Western blot	BioRad #162-0115

9.1.5. Other kits and materials

Table 14: Other kits and materials

Name	Description	Source / Catalogue Nr.
B-PER	Lysis buffer	Thermo #78248
SERVAPOR® dialysis tubing	Dialysis membrane	SERVA #44146
Petri plates, round, diam. 55mm, height 14,2mm, aseptic, 3 vents	Petri dishes	Merci #331 999 000 060

9.1.6. Instrumentation

Table 15: Instrumentation

Name	Description	Source
Centrifuge 5804R	Centrifuge	Eppendorf
Ultrasonis homogenizer 4710	Sonicator	Cole Palmer
Freeze dryer ALPHA 1-2 / LD p lus	Lyophilize	Christ
DMI 8 inverted microscope	Light microscope	Leica
Luminescent Image Analyzer LAS 3000	Chemiluminence detector	Fujifilm
Q-Tof Premier	Mass spectrometer	Waters
Nano Acquity UPLC	Liquid chromatography	Waters

Table 16: Software

Name	Description	Source
GS-900™ Calibrated Densitometer	Densitometry	BioRad
ImageQuant™ TL	Densitometry	GE Healthcare Life Sciences

9.2. *Appendix 2 – Additional results*

9.2.1. Densitometry analysis

Table 12: Concentration determination of recombinant protein, *Ter1* and *Sgs3*, using densitometric analysis

Sample	Type	Concentration / mg mL^{-1}
BSA1	Standard	1.5
BSA2	Standard	0.75
BSA3	Standard	0.375
BSA4	Standard	0.1875
BSA5	Standard	0.09375
Ter1 – 1	Protein	3.002357
Ter1 – 2	Protein	3.261900

9.2.2. Cell culture analysis

Table 13: Cell culture analysis of Ter1 for the first trial

Day	Concentration	Mean	± SEM	SD
Day 1	0.5 mg/ml	42,3333	4,0469	9,9130
	1.0 mg/ml	13,3333	3,1798	5,5076
Day 2	0.5 mg/ml	211,8333	17,3713	42,5508
	1.0 mg/ml	127,6667	2,2485	11,8462
Day 3	0.5 mg/ml	88,1667	10,2710	25,1588
	1.0 mg/ml	43,3333	9,4634	23,1805

Table 14: Cell culture analysis of Ter1 for the second trial

Day	Concentration	Mean	± SEM	SD
Day 1	0.5 mg/ml	120,6667	6,2057	15,2009
	1.0 mg/ml	80,8333	7,5340	18,4544
Day 2	0.5 mg/ml	91,1667	4,4528	10,9072
	1.0 mg/ml	65,0000	8,5206	20,8710
Day 3	0.5 mg/ml	33,6667	8,9132	21,8327
	1.0 mg/ml	32,1667	7,8885	19,3227

Table 15: Cell culture analysis of Sgs3 for the first trial

Day	Concentration	Mean	± SEM	SD
Day 1	0.5 mg/ml	123,5000	8,2573	20,2262
	1.0 mg/ml	91,1667	11,3590	27,8239
Day 2	0.5 mg/ml	99,8333	5,4554	13,3629
	1.0 mg/ml	85,6667	4,9441	12,1106
Day 3	0.5 mg/ml	50,5000	7,9739	19,5320
	1.0 mg/ml	34,1667	4,5856	11,2324

Table 16: Cell culture analysis of Sgs3 for the first trial

Day	Concentration	Mean	± SEM	SD
Day 1	0.5 mg/ml	170,0000	14,0973	34,5311
	1.0 mg/ml	101,1667	10,8149	26,4909
Day 2	0.5 mg/ml	68,6667	10,9260	26,7632
	1.0 mg/ml	85,1667	3,9616	9,7040
Day 3	0.5 mg/ml	34,5000	2,5528	6,2530
	1.0 mg/ml	39,8333	5,0887	12,4646

9.2.3. SEM analysis

9.2.3.1. Synthetic sequence encoding the recombinant protein BmS2 – GL2

The recombinant BmS2-GL2 protein sequence comprises 246 amino acids and encodes a protein of 30.2 kDa.

```
MEFDYEKANY RSPSHRDYEK ANYRSPSHRD YEKANYRSSH RGSEFDYEKA NYRSPSHRDY
EKANYRSPSH RDYEANYRSP SHRDYEKANY RSPSHRDYEK ANYRSPSHRD YEANYRSPSH
RDYEKANYRS PSHRDYEKAN YRSPSHRDYE ANYRSPSHRD YEKANYRSPS HRDYEKANYR
SPSHRDYEAN YRSPSHRDYE KANYRSPSHR DYEKANYRSP SHRDYEANYR SPSHRGSRGS HHHHHH
```