Design and Synthesis of Collagen-binding Anti-microbial Proteins

Mona Ghannad

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the MSc Degree in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

Abstract

The Herpes simplex virus (HSV) is a virus that commonly infects the skin, and mucous membrane of the mouth, genitalia, and the eye. HSV-1 is the strain that is most commonly associated with corneal infections, and it is the most frequent cause of corneal blindness in North America [1]. Currently no cure is available, and many limitations are characterized by the currently available synthetic antiviral drugs, which suggest the need for other potential drug alternatives and delivery strategies. Anti-microbial peptides are naturally occurring peptides that are potent killers of a broad range of micro-organisms, including bacteria, fungi, and viruses [2]. AMPs are known to be a key component of the innate immune response at the human ocular surface. The human cathelicidin-derived AMP, LL-37, expressed in human corneal epithelial cells provides a wide range of protection against viral pathogens such as HSV-1 [3]. My thesis research addressed the design and recombinant production of hybrid AMP sequences containing LL-37 with the potential ability to form chemical or physical associations with a Collagen scaffold material, such as those used in current artificial cornea constructs to address the need for alternative anti-viral drugs. Three fusion proteins were tested, and compared for feasible design anti-microbial peptide expression and purification in *E. coli*. It was illustrated that the thioredoxin and SUMO fusion systems are good candidates for successful recombinant production of active designed peptides. The point-mutated LL-37 sequence was successfully expressed and purified using the thioredoxin fusion system. It was demonstrated that this modified LL-37 was effective against HSV-1 infection. The SUMO system was used to express the bio-functional LL-37 containing a collagen-binding sequence. Further work is required to address issues regarding recombinant AMP production, such as increasing enzymatic cleavage efficacy, and minimizing proteolytic degradation or modification.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACV	Acyclovir
AMP	Anti-microbial peptide
BSA	Bovine Serum Albumin
CBD	Collagen binding domain
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ELP	Elastase-like protein
FBS	Fetal bovine serum
FPLC	Fast performance liquid chromatography
HCEC	Human corneal epithelial cells
HCl	Hydrochloric acid
HPLC	High performance liquid chromatography
HSV	Herpes Simplex Virus
IMAC	Immobilized Metal Ion Affinity Chromatography
IMPACT	Intein Mediated Purification system with an Affinity Chitin-binding Tag
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilodalton
LL(C)-37	LL-37 with a cysteine mutation in the third amino acid position

MALDI	Matrix Assisted Laser Desorption Ionization
MALDI-TOF	Matrix Assisted Laser Desorption Ionization - Time of Flight
NaCl	Sodium Chloride
Ni-NTA	Nickel-nitrilotriacetic acid
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PNK	Polynucleotide kinase
PVDF	Polyvinylidene fluoride
pI	Isoelectric point
PMSF	Phenylmethylsulfonyl fluoride
RT	Room temperature
RU	Resonance unit
SDS	Sodium Dodecyl Sulfate
SPR	Surface Plasmon Resonance
SUMO	Small Ubiquitin-Like Modifier
TBS	Tris Buffered Saline
Ulp1	(Ubiquitin-like protein)-specific protease 1
UV	Ultraviolet

Contributors to Thesis

Bareiss, Bettina

Bettina was a graduate student in Dr. Griffith's lab. She performed the anti-HSV-1 protection assay.

Couture, Jean-Francois

Dr. Couture is a professor in the department of Biochemistry, Microbiology, and Immunology at the University of Ottawa. He provided the Smt3 (SUMO containing) vector, pHIS2 (modified pET-22b(+) based) expression vector, and non-commercial SUMO Protease I.

Dick, Scott

Dr. Dick is a researcher who did the cloning work for the intein fusion systems, pLET1, and the His-tag system. He also did the purification and Maldi analysis of the fusion expressed in the His-tag system, mentioned in Appendix B.

Li, Fengfu

Dr. Li is a researcher who coated the COOH-1 chip with collagen for the SPR instrument, mentioned in Appendix B.

Wan, Fan

Fan is a graduate student in Dr. Harden's lab. She performed the collagen-binding testing of the fusion expressed in the His-tag system using the SPR instrument, mentioned in Appendix B.

Acknowledgements

I would like to thank Dr. Harden and Dr. Griffith for their guidance and supervision, and for making this invaluable learning experience possible. I thank Dr. Couture, and Dr. Mah who provided insight into direction of my research.

I would also like to thank my colleagues from both labs for their support.

I would like to thank my family, specially my dear parents for all their support.

CHAPTER 1

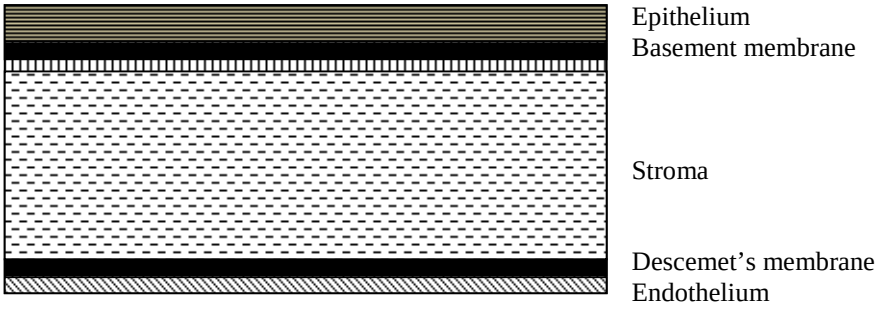
INTRODUCTION

1.1 Cornea

The human cornea is the outermost layer of the eye. The most important function of the cornea is to transmit and refract light onto the retina [4]. In addition to being transparent, the cornea also protects the intraocular contents. However, the tissue is not vascularized and does not contain lymphoid cells to nourish or protect it against infection. Consequently the cornea relies on the conjunctiva for much of its immune responses [5]. The conjunctiva is a major support tissue that covers the sclera and lines the inside of the eyelids, which also contributes to immune surveillance and helps to prevent the entrance of microbes into the eye.

The corneal tissue consists of a thick central transparent connective tissue (stroma) that is covered by epithelia on both sides, separated from the stroma by basement membranes (Figure 1-1) [5, 6]. The stroma comprises about 90% of the cornea's thickness and consists primarily of water and collagen [6]. Collagen provides the cornea with strength, elasticity, and form. The fibril-forming Type I collagen is the primary corneal collagen [4]. Collagen I consists of two types of polypeptide chain sequences that come together into the triple helix through disulphide bond formation. These collagen molecules then align along the helix axes and assemble into higher order polymers known as collagen fibrils. Collagen fibrils often aggregate on a higher order of structure and make up larger, cablelike bundles called collagen fibers [5]. Type I collagen also forms networks *in vitro*. The protein initially assembles into fibers that elongate and intermesh together to form a gel.

Figure 1-1: Schematic drawing of human corneal layers. Sketch showing the corneal stromal layer, bounding membranes layers, epithelium and endothelium. The stroma accounts for approximately 90% of the cornea's thickness. Separating the stroma from the epithelium is the basement membrane, while between the stroma and the endothelium is Descemet's membrane. Adapted from Ref [6].



1.2 Innate Immunity as a Component of Immune Defense at the Ocular Surface

The ocular surface is constantly exposed to the exterior world, which means that it is vulnerable to infection by the potentially harmful range of microorganisms found in the environment. In addition to the cornea's lack of vascularization, the moist mucosal surface places it at particular risk for infections. The ability of the outer ocular system to recognize and eliminate foreign pathogens is critical for the preservation of sight. Any damage to the cornea caused by injury or disease can cause visual impairment or blindness. While the cornea's tough and resilient structure serves as a mechanical barrier that prevents pathogens from reaching the rest of the eye, immunological defense mechanisms have also evolved to further protect the eye [5]. The cornea is protected by the mucosal immune system that incorporates both innate and adaptive response mechanisms present in the tissue and tear film [5].

Innate immunity is the first line of defense against corneal infection [5]. In addition to the physical barriers, such as the eyelids, there are several molecular and cellular components of the ocular innate immune system such as tears, epithelial cells, keratocytes, corneal nerves, the complement system, and interferons, which protect the cornea surface.

Small cationic Anti Microbial Peptides (AMPs), which have direct antimicrobial effect against a broad range of microorganisms, are a major component of the innate immune system at the human ocular surface [7]. The tear film is an essential functional element of immune defense in the ocular surface containing mucins, lyzozyme, lactoferrin, and several AMPs [5]. AMPs in the tear film, not only stop the invading organism if it breaches the outer defense, they also activate cellular processes like migration, proliferation, and cytokine production, which simultaneously signals the adaptive immune system to prepare for action. Apart from the

different protective antimicrobial substances in the tear fluid, the conjunctiva and cornea epithelia also produce an array of AMPs to kill invading organisms [7].

1.3 Anti-Microbial Peptides as part of Innate Anti-viral Agents

Anti-microbial peptides, as signified by their name, are naturally occurring peptides that are produced by animals and plants to fight infection [8]. In recent years, hundreds of microbicidal peptides have been identified and isolated in the Antimicrobial Sequences Database from plants and animals [9, 10]. Endogenous antimicrobial peptides are also known to be a major component of the innate immune response at the human ocular surface. As part of the innate immune system, these peptides are present constitutively, but are increased with inflammation and injury. AMPs are produced in part by various epithelial cells, including human corneal and conjunctival cells, and are potent killers of a broad range of micro-organisms, including bacteria, fungi, and viruses [2].

AMPs as a group are active against both Gram positive and Gram negative bacteria, some fungi, and some viruses, although the individual peptides differ in their ability to kill different pathogens [11]. The vast majority of AMPs are cationic and can be categorized based on their size, secondary and tertiary structures, and other characteristics such as the presence or absence of disulfide bridges. AMPs are either linear (with a potential to form amphipathic α-helical, β-sheet structures), or circular (ie. theta-defensins) [10].

Despite their significant diversity in structure, it is generally accepted that the AMPs exert their antimicrobial activity similarly through several common features, such as having a net positive charge and amphiphilicity. The major mode of action of AMPs is through adhering to and/or penetrating the pathogen's cell membrane, resulting in cell death. The precise

mechanisms of their action are poorly understood [9, 11]. However, it is clear that AMPs are able to initially bind to the anionic components on the surface lipid membranes of pathogens via their positive charge [2]. After this non-specific initial interaction occurs, the hydrophobic portion of the AMPs may integrate with and possibly penetrate the membrane, inducing either transient pores or a detergent-like solubilisation [2]. AMPs that act through a nonreceptor-mediated mechanism are able to act specifically on pathogens, because the inner membrane of most pathogens is composed of negatively charged phospholipids. In contrast, the phospholipids comprising the membrane of normal mammalian cells are asymmetrically distributed and maintain a neutral net charge. Thus, host cells are resistant to lysis, due to their different lipid composition. Moreover, due to the nonspecific mode of action of AMPs, there is far less bacterial resistance developing towards AMPs compared to conventional antibiotics. It should be noted that not all AMPs have the same mode of action, toxicity or the same targets, as different parameters such as their relative positive charge, play major roles in their biological activity [10].

1.3.1 LL37, a Model Ocular Defense Peptide for HSV Infection

One class of antimicrobial peptides, the cathelicidin-derived peptides, are known to be an important component of the mammalian innate immune system [12]. Cathelicidins consist of a highly conserved N-terminal region that is homologous to cathelin, a cysteine protease inhibitor, and a less conserved C-terminal antimicrobial region that differs among species [3]. The human cationic antimicrobial peptide of 18 kDa (hCAP 18), is the only cathelicidin identified in humans [3]. Various epithelial cells, including human corneal and conjunctival cells produce hCAP18 [3, 7]. Upon stimulation, the cathelin domain is cleaved proteolytically by serine proteases to release the mature COOH-terminal LL-37 [3].

LL-37 is a 37-residue, 4.7 kDa cationic anti-microbial peptide that begins with two Leucines. LL-37 forms a linear, amphipathic, α-helical peptide, since it is organized in such a way that it contains a hydrophobic side opposed to its cationic side [2]. LL-37 is the only cathelicidin-derived antimicrobial peptide found in humans. LL-37 is expressed in human corneal epithelial cells, and is shown to be upregulated by injury or exposure to invading pathogens [11]. LL-37 provides a wide range of protection against bacteria, fungi, and viral pathogens such as Pseudomonas and HSV-1 [3]. LL-37 has also been shown to stimulate angiogenesis, which is the growth of new blood vessels from pre-existing vessels, and also to promote epithelial healing in skin wounds [3]. The amino acid sequence of the native LL-37 found in humans is **LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTES**.

1.4 Impact of Herpes Simplex Virus on Vision

The Herpes simplex virus (HSV) type-1 and -2 can commonly infect the skin, mucous membrane of the mouth, genitalia, central nervous system, and the eye in humans. Although both HSV-1 and HSV-2 are able to enter the same cell types, HSV-1 is the strain that is believed to be responsible for the majority of corneal infections. It is the most frequent cause of corneal blindness in North America [13]. Approximately 400, 000 individuals in the United States have ocular herpes, with 50, 000 new and recurrent cases each year [14].

The virus is transmitted through direct contact of mucosal surfaces, secreted fluids, or abraded skin of an infected individual. Ocular herpes infection may be transmitted either directly via fluid spread or indirectly through neuronal spread from a location that is not ocular, such as the mouth mucosa.

The entry of HSV into ocular cells can occur through several pathways altered by several factors such as the individual's immune response. Virus replication is followed by its entry into various ocular structures [14]. Following initial exposure, a primary infection occurs, that can give rise to a range of clinical signs that include epithelial ulceration and corneal inflammation, which could develop into a more serious infection of the corneal epithelium and stroma called herpes keratitis (HSK). Following recovery from primary infection, the virus develops life-long latency in host neuronal cells and possibly the cornea, and is easily triggered for reactivation [14, 15]. Reactivation of latently infected neurons may be triggered by varied stimuli such as stress, fever, sunlight, or immune suppression, weeks or even years after the initial occurrence [15]. Recurrence of HSV-1 infection in the eye may eventually cause scarring, thinning, vascularisation of the cornea, and stromal keratitis, which can lead to loss of vision and possibly blindness [13].

1.4.1 Treatments Available and their Limitations

Although currently no cure is available, treatment with anti-viral drugs helps to stop the herpes virus from multiplying. There are a few synthetic antiviral drugs available for treating ocular herpes. Of these, acyclovir and its derivatives, administered both topically and orally, are most commonly used as a treatment for HSV epithelial keratititis [13]. Compared to the other drugs, acyclovir is more specific towards the target virus and has less toxicity to corneal cells in typically administered doses [15]. Even so, there are many limitations characterized by the currently available therapies with acyclovir. For instance, the absorption of orally administered acyclovir is rather slow, inconsistent, and incomplete with an oral bioavailability of less than 30% [16]. Moreover, acyclovir can cause nausea, diarrhea, rash, or headache, and like all oral

antivirals, long-term usage may suppress the immune system [16]. With topical treatment, the drug can be directly targeted to the site of action, and the reduction in circulating drug levels may in turn reduce adverse effects. However, for topical applications, the absorption of the drug is typically very slow and requires the assistance of a permeation enhancer [16]. For the case of ocular topical acyclovir delivery systems, only modest efficacy has been demonstrated due to a number of limitations, including the tendency for the drug to be washed away by tear fluid during blinking [16].

Lacking effective treatment, HSV induced corneal blindness may result due to heavy corneal damage and scarring. Treatment failure eventually leads to surgical interference. Currently corneal transplantation is the treatment of choice for herpes induced blindness as a result of corneal damage and scarring. However, donor corneas for transplantation are not readily available, and in case of autoimmune conditions or disease like HSV infection, this procedure has a low success rate (e.g. 22% success at 5 years for HSV compared to 73% success at 5 years for non-HSV grafts) [4]. One reason for the low success rate is that the stress of surgery induces HSV virus reactivation, which may lead to graft rejection.

1.5 Determinants and Motifs of Thesis Project

The limitations characterized by synthetic antiviral drug therapy suggest the need for other potential drug alternatives, as well as an improved delivery strategy. In the case of high risk ocular HSV-infected transplants, due to the shortage of donor corneas and the high risk of rejection, biomaterial-based corneal substitutes made primarily of collagen, are being investigated as a possible treatment alternative [17-19]. An advantage this engineered treatment

model may enable is the possibility that anti-viral drugs specific for HSV infection may also be incorporated within the implant to treat or suppress HSV-reactivation triggered by the surgery.

In recent years, AMPs have received growing interest, due to the demand for new drugs resulting from the emergence of multi-drug resistant pathogens [8]. Incorporating HSV specific AMPs into novel biomaterial based corneal substitutes may prove useful in suppressing HSV-1 reactivation triggered by the surgery, and hence reduce graft rejection.

The potent anti-microbial activity of LL-37 and its effectiveness against Pseudomonas and HSV-1, make LL-37 an important ocular surface defense peptide, and a strong candidate as an alternative anti-viral drug. Thus, the inclusion of LL-37 in biomaterial based corneal substitutes made primarily of collagen might therefore be of considerable value in the prevention or suppression of HSV infection as a cause of graft failure.

This thesis project focused on the development and biosynthesis of new hybrid AMP sequences with the potential ability to form chemical or physical associations with a Collagen scaffold material, such as those used in current artificial cornea constructs. In order to conduct basic scientific studies and clinical trials, large quantities of antimicrobial peptides are required. Isolation of AMPs from natural sources is typically a complex and inefficient approach that precludes obtaining peptides in large amounts. On the other hand, synthetic approaches have the potential for large scale production of native AMPs and their designer variants with altered sequences [20]. At present, the relatively high cost of chemical synthesis of AMPs is a limiting factor for large-scale production [20]. As an alternative, recombinant DNA technology enables cost-effective biological routes to AMP production. The relatively low cost and easy scale-up of the recombinant approach, as well as the relative ease of modification of peptide sequences, makes it a very attractive means for large-scale production of anti-microbial peptides [20].

However, there are often challenges with the biosynthesis and purification of AMPs that must be taken into consideration. For instance, a well chosen expression host and cloning strategy is often essential for obtaining good AMP yields.

The effective recombinantly produced anti-microbial peptide will need to be integrated in the biomaterials scaffolds, and possess sufficient flexibility to maintain its anti-infective functionality. Thus, a collagen-binding domain sequence, and a spacer sequence were selected for possible attachment and display of the AMP. Using standard molecular biology methods, these sequences can be expressed with the AMP, using a fusion tag (ie. thioredoxin) to enhance peptide expression and purification. Further background details concerning production and design of biofunctional AMP (LL-37) are discussed below.

1.5.1 Carrier proteins for fusion expression of antimicrobial peptides in *E. coli*

Several anti-microbial peptides have been successfully produced using standard molecular biology techniques, most commonly in *Escherichia coli* [20]. However, due to their antimicrobial characteristics, simply expressing an AMP sequence may not be a very effective strategy. Rather, anti-microbial peptides produced in *E. coli* are often expressed as fusion proteins, in order to mask or minimize the peptide's potential lethal effect towards the bacterial host and also to protect the peptides from proteolytic degradation [20].

The intein-mediated system is the third most commonly used carrier for recombinant production of antimicrobial peptides. According to the Recombinantly-produced Antimicrobial Peptides Database, nine AMPs were found using this system. It distinguishes itself from all other purification systems by its ability to purify in a single chromatographic step, a native recombinant protein without the use of a protease, thus being cost efficient. This is enabled by

utilizing the inducible self-cleavage activity of the protein termed intein, to sever the target protein from the affinity tag. Low-temperature induction is usually required to obtain soluble fusion.

Thioredoxin has been the most frequently used carrier protein according to the Recombinantly produced Antimicrobial Peptides Database, as more than 20% of all reported antimicrobial peptides fusions were expressed using thioredoxin [20]. Thioredoxin is highly soluble and acts as a chaperone, and may thus promote soluble expression of recombinant proteins/peptides in the *Escherichia coli* cytoplasm [20]. It is noteworthy to mention that low-temperature induction is not always required to obtain soluble thioredoxin fusion [20]. AMP peptides may be fused to thioredoxin through simple molecular biology techniques, and once the fusion is expressed the antimicrobial peptide may be released from the fusion protein.

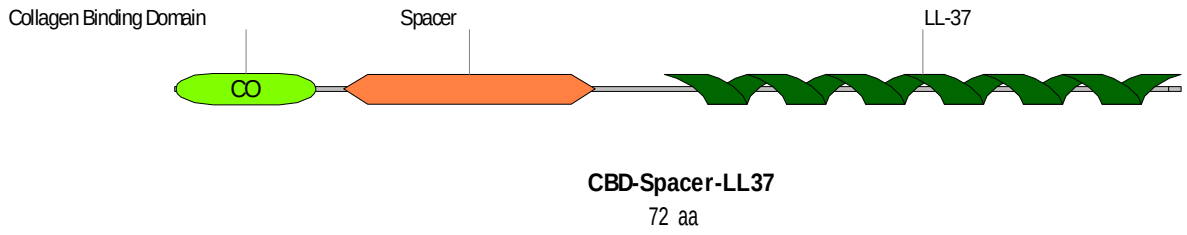
"Recently, SUMO has been used as a novel fusion carrier for the production of recombinant proteins [20]." The Sumo-based peptide expression system utilizes a small ubiquitin-like modifier (SUMO) that allows expression, purification, and generation of native proteins in *E.coli* [21]. SUMO has a hydrophobic core and a hydrophilic surface, thus making it highly soluble. As a fusion partner, SUMO may enhance protein expression levels, as well as the solubility of recombinantly expressed proteins by enhancing proper protein folding [21]. SUMO's small size (~100 amino acids, ~ 11 kDa), allows for a relatively high peptide-to-carrier ratio, similar to thioredoxin, which favours peptide yield and expression. In addition to incorporating the advantages of traditional fusion systems, the existence of SUMO Protease I, offers a unique advantage to the SUMO fusion system. The SUMO Protease I specifically recognizes the tertiary structure of the SUMO protein, and cleavage by SUMO protease I results in the production of target protein with native N-terminated fusion design [21].

LL-37 has previously been successfully expressed in *E. coli* using a thioredoxin fusion, as well as a GST fusion [20]. However, expression levels were rather low, with reported values of ~2 mg/L and ~0.3 mg/L for the thioredoxin and GST fusions, respectively [20]. When LL-37 expression yield is compared to expression yields of other similar antimicrobial peptides produced using the thioredoxin fusion system, it can be observed that LL-37 has quite low yields. For instance, β-Defensin 2, β-Defensin 2, and β-Defensin 4 have a reported yield of 346 mg/L, 689 mg/L, and 140 mg/L respectively when expressed with thioredoxin in *E. coli* [20]. In order to improve LL-37's expression yield other fusion proteins, expression strategies, and protocols may prove advantageous, due to the different characteristics they present. However, as discussed below, it is possible that low soluble expression yields of LL-37 are partially due to structural characteristics unique to LL-37, which may render any fusion containing the LL-37 sequence more difficult to produce recombinantly compared to other antimicrobial peptides.

1.5.2 Design Considerations of AMP-based Constructs

An effective anti-microbial peptide will need to be integrated into the Collagen-based biomaterials scaffolds used in artificial corneas in such a way as to preserve the antimicrobial activity of the AMP domain. To do so, my design includes a collagen-binding domain (CBD) sequence, separated from the AMP domain by a soluble, disordered peptide spacer sequence, as sketched in Figure 1-2. The focus of this work is on LL37 as a characteristic AMP sequence. Using standard molecular biology methods, the hybrid CBD-spacer-LL-37 chimera can be expressed in *E.coli*, using an appropriate fusion tag (several fusion systems were explored, as discussed below) to enhance peptide expression and purification.

Figure 1-2: Schematic drawing of the AMP-Spacer-CBD Construct.



CBD-Spacer-LL37

72 aa

1.5.3 Attachment of AMP to the Scaffold's Surface

The hybrid AMP construct was designed for facile attachment onto a collagen-based cornea graft, via a collagen-binding domain. Collagens are the most abundant proteins in mammals. Collagen-rich extracellular matrices are not only critically important for the biomechanical properties of tissues, but are also intimately involved in cell adhesion and migration during growth, differentiation, morphogenesis, and wound healing. As previously mentioned, Type I collagen is the predominant structural component in artificial corneas. Since protein-protein interactions are known to be very specific, a known targeted collagen-binding sequence was chosen for this purpose. The sequence that was adopted was a Type I collagen-binding sequence derived from von Willebrand's factor previously investigated by Andrades *et al*, [22]: WREPSFMALS.

The original collagen-binding sequence derived from von Willebrand's factor sequence contains a cysteine instead of the methionine. However, it was observed that the cysteine interfered with the disulfide bonds formation of adjacent peptide sequences. Since our antimicrobial peptide contains two disulfide bonds, this modification, which was noted to make no difference in collagen-binding activity, was retained.

1.5.4 Spacer between AMP and Collagen-Binding-Domain

A spacer between the CBD and the AMP domains was included in the design to provide a flexible junction between the CBD and AMP sequences that would potentially increase the bioavailability of the AMP domain when the CBD domain is immobilized on the collagen substrate. The spacer block was designed to be both flexible and resistant to proteolysis, and

may also enhance expression yield and recovery of the designed peptide. The following disordered sequence, [23] was adopted for our purposes: GSTSGSGKPGSGEGSTKG.

OBJECTIVES AND HYPOTHESIS**:**

This project focuses on the design and production of biofunctional AMP (LL-37), namely modified LL-37, and the CBD-spacer-LL-37 (WREPSFMALS- GSTSGSGKPGSGEGSTKG – LL-37) discussed above. My hypotheses are that (1) AMP-based peptides with the ability to form chemical (ie. disulfide formation) or physical association (ie. through Collagen-binding attributes) can be designed and produced recombinantly, and (2) that such AMP-based systems will retain their innate antimicrobial activity. The thesis work proceeded by pursuing three main objectives:

(1) Design of the biofunctional LL-37 sequences

(2) To determine which expression system is best suited for facile expression and purification of the modified LL-37 and the CBD-spacer- LL-37 construct. Three fusion systems were investigated before adopting thioredoxin and SUMO (small ubiquitin related modifier) as novel fusion partners for expression.

(3) To determine whether the purified construct has maintained its sequence integrity and thus potential activity.

The following chapters present the pursuit of these objectives. In Chapter 2, the materials and methods used are presented. Chapter 3 presents the expression, purification and testing results of the more successful thioredoxin fusion system for the point-mutated LL-37 with a Cys substitution allowing for possible chemical attachment to collagen substrates. Chapter 3 also presents preliminary results of the SUMO fusion system for expression and purification of CBD-spacer- LL-37. Chapter 4 provides a discussion of the work. Finally, some supplementary studies are described in several appendices.

CHAPTER 2

MATERIALS AND METHODS

Restriction enzymes and the Quick Ligation Kit were purchased from New England Biolabs. PCR primers were purchased from Integrated DNA Technologies. PCR reagents were purchased from Invitrogen. The PCR Purification Kit, QIA-quick gel extraction, and Plasmid extraction kit were purchased from Promega. Acrylamide, bisacrylamide, SDS, Tricine, and 10-20% Tris-Tricine precast gels were purchased from Fisher Scientific. *E. coli* DH10β (maintained in our laboratory) was used for subcloning and plasmid amplification. *E. coli* BL21 (maintained in our laboratory) was used as the expression host. Ni-NTA Agarose was purchased from Qiagen. The HiPrep 16/10 QFF column used for FPLC was purchased from GE Healthcare. The commercial SUMO Protease 1 was purchased from Life Sensors. The Smt3 (SUMO containing) vector, pHIS2 (modified pET-22b(+) based) expression vector, and non-commercial SUMO Protease I were generously offered by Dr. Couture. The SeeBlue® Plus2 Pre-Stained Standard was purchased from Invitrogen. The Lane Marker Reducing Sample buffer (5x), and GelCode Blue Stain Reagent were purchased from Thermo Scientific. QuantiChrom^{TM} Urea Assay Kit was purchased from BioAssay Systems. Monoclonal Mouse IgG anti-LL-37 antibody was purchased from Cell Sciences. Mouse IgG anti-His antibody was purchased from Novagen. Western kit and AP Conjugate Substrate development kit were purchased from Novagen. All chemicals, unless otherwise stated, were obtained from Sigma-Aldrich Canada Ltd or Fisher Scientific.

Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) was performed as follows for each amplification. The reaction mixture containing PCR template, Primers, dNTP, $MgSO_4$, 10x PCR pfu buffer, and DNA polymerase (pfu) was aliquoted into duplicate small PCR tubes with a total volume of 50 µl per tube. The PCR reactions were performed by running 35 cycles with a temperature profile of 94°C for 5min, 1 min at 63°C, and 2 min at 72°C using the mastercycler ep gradient machine from Eppendorf (Fisher).

2.1 Construction of SUMO-(CSG)-LL-37 expression vector

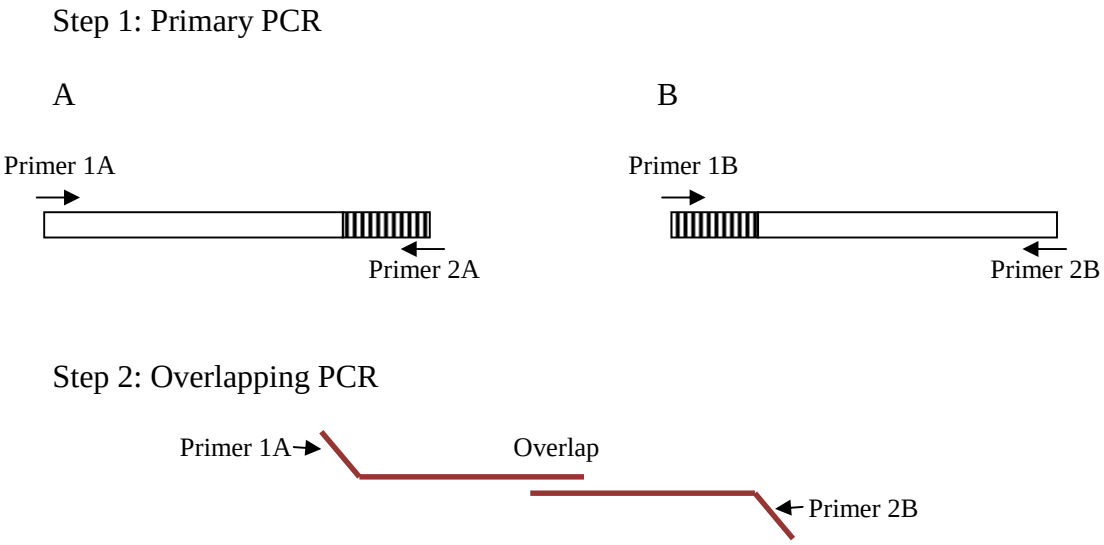
PCR primers were designed to amplify the SUMO and LL-37 sequences from their respective templates, Smt3 and pLET1 (PCR step #1). The full length SUMO-LL-37 sequence was generated progressively in a single reaction by overlap extension PCR (PCR step #2), and cloned in a modified pET-22b(+) based plasmid (which was obtained from Dr. Couture) with the NdeI and XhoI restriction sites. The general PCR steps are outlined in Figure 2-1.

PCR step #1 a- The PCR primers used to amplify the 6x His-SUMO sequence from the Smt3 vector (which was a generous gift from Dr. Couture) with an overlapping LL-37 containing sequence:

Forward primer 1: 5’ CAT ATG CAC CAT CAC CAT CAC CAT TCG 3’ (Underlined sequence indicates NdeI restriction site)

Reverse primer 1: 5’ ATC GCC CAG CAG **GCC GCT GCA** ATA CGT AGC ACC ACC AAT CTG TTC 3’(Underlined sequence indicates part of **CSG-**LL-37 sequence, seamlessly adjacent to the C-term of the SUMO sequence)

Figure 2-1: Outline of PCR steps for seamless fusion vector construction. In step 1, PCR primers are used to amplify the SUMO fusion with part of the sequence for the gene of interest. Additionally, the sequence for the gene of interest is amplified with part of the SUMO fusion sequence. In step 2, PCR primers are used to amplify the two Amplicons from step 1, with an overlapping sequence, to produce a seamless fusion construct.



PCR step #1 b- The PCR primers used to amplify the **CSG**-LL-37 sequence from the pLET1 vector (which was a generously constructed and offered by Dr. Dick) with an overlapping SUMO containing sequence:

Forward primer 2: 5' **TGC AGC GGC** CTG CTG GGC GAT TTT TTT CGC AAA AGC AAA GAA 3' (The underlined sequence indicates the modification of the Cys mutation back to Gly)

Reverse primer 2: 5' CTC GAG TTA TTA GCT TTC GGT GCG CGG CAC CAG GTT 3' (Underlined sequence indicates XhoI restriction site)

PCR step #2- The full length SUMO-(SCG)-LL-37 sequence was generated progressively in a single reaction by overlap extension PCR, using forward primer 1, and reverse primer 2, and the PCR purified products of PCRs #1a and #1b. The full length SUMO-(SCG)-LL-37 sequence was digested with the NdeI and XhoI restriction enzymes, and then ligated (using Quick Ligation Kit) into previously digested pHIS2 vector to create the corresponding expression vector SUMO-(SCG)-LL-37. Accuracy of the inserted cDNA was confirmed by automated DNA sequencing. The SUMO-(SCG)-LL-37 construct was transformed into the *E. coli* strain BL21 for protein expression.

2.2 Expression of Thioredoxin and SUMO fusion constructs

Protein expression was performed as following. A single transformed colony was grown in 10 ml LB media containing 100 mg/ml ampicillin at 37°C with shaking at 250 rpm. The cells were grown overnight and then the 10 ml culture was transferred into a 1L fresh LB medium containing 100µl/ml Amp. The culture was incubated in an air shaker at 37°C until the OD_{600} reached 0.8. Protein expression was induced with 1mM IPTG and cells were harvested after 4

hours of incubation at 37°C. All cells were harvested by centrifugation at 8000xg for 30 min at 4°C. For protein purification, three 1L cultures were grown and harvested.

2.3 Purification of Thioredoxin fusion construct, denaturing conditions

Bacterial pellet (~3 g wet cells from 3 L culture) was resuspended in 100 ml of denaturing purification buffer A (100 mM NaH_2PO_4, 10 mM Tris-Cl, 8 M urea, pH 8.0). The whole cell lysate was then centrifuged at 25,000 rpm for 30 min to separate soluble and insoluble portions. The supernatant was collected and applied to a column containing 10 mL Ni-NTA resin pre-equilibrated with buffer A. After washing the column with buffer A, at pH 8.0, fusion protein was washed with buffer A at pH 6.3 and 5.9. The fusion protein was eluted with buffer A at pH 4.5. The fusion protein fractions were pooled and dialyzed in 10mM Tris, 20mM NaCl, pH 8.0 buffer using dialysis tube with 1 kDa MWCO (molecular weight cut-off) to remove urea.

2.4 Purification of Thioredoxin and SUMO fusion constructs, native conditions

Bacterial pellet (~3 g wet cells from 2 L culture) was resuspended in 100 ml of native purification buffer A (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0), lysozyme to a final concentration of 200 µg/ml, and PMSF to a final concentration of 1mM. The whole cell lysate was then centrifuged at 25,000 rpm for 30 min to separate soluble and insoluble portions. The supernatant was collected and applied to a column containing 10 mL Ni-NTA resin pre-equilibrated with buffer A. After washing the column with buffer A containing 20 mM imidazole, fusion protein was washed with buffer A containing 50 mM and 100 mM imidazole. The fusion protein was eluted with buffer A containing 250 mM imidazole.

2.5 P-LL(C)-37 peptide release and purification

The fusion protein fractions were pooled and dialyzed in 10mM Tris, 20mM NaCl, pH 8.0 buffer using dialysis tube with 1 kDa MWCO (molecular weight cut-off) to remove imidazole, which is necessary for acid cleavage step. The LL(C)-37 containing fusion protein, was cleaved in 50% formic acid (v/v) at 50°C for 50 hours. After cleavage, the Pro residue was left at the N-terminus of LL(C)-37. This recombinant LL(C)-37 is hereafter called P-LL(C)-37. The pH of this sample at 0.5 was brought to 2.0 by adding 6M NaOH. The sample was dialyzed in dH_20 using dialysis tube with 1 kDa MWCO (molecular weight cut-off) to remove formic acid, and was subsequently lyophilized.

The His-tagged thioredoxin carrier and residual undigested fusion proteins were rapidly removed by using HiPrep 16/10 QFF anion exchange column using the UNICORN program software. In brief, the lyophilized mixture was dissolved in 1ml of dH_2O, and 2ml of buffer 1 (10mM Tris, 20 mM NaCl, pH 8.0). The sample was subject to FPLC using HiPrep 16/10 QFF column equilibrated with buffer 1, and fractions were collected at a flow rate of 3 ml/min using a linear gradient of increasing NaCl concentration to 1M. The P-LL(C)-37, which has a pI of 10.5, is eluted in 10mM Tris, 20mM NaCl, pH 8.0 buffer as flow-through in fraction 4. The thioredoxin-LL(C)-37 fusion and thioredoxin are retained longer with pI's of 7.9 and 5.3 respectively. Each sample fraction correlating with a peak was run on SDS-PAGE (15%) for analysis.

2.6 Mass spectrometry

To substantiate the identity of purified P-LL(C)-37 and CBD-Spacer-LL(C)-37, mass spectra was obtained using MALDI-TOF mass spectrometry at the National Research Center

(Ottawa, Ontario) and at the Center for Advanced Research in Environmental Genomics at the University of Ottawa.

2.7 Antiviral Assay

2.5 X 10^4 immortalized human corneal epithelial cells (HCEC) were grown to 80 % confluence in 12-well plates supplemented with Keratinocyte Serum-Free Medium (KSFM; Invitrogen, Burlington, Canada) in a humidified tissue culture incubator. HSV-1 viruses were added to the HCEC cultures at a ratio of 0.1 MOI (10^4 pfu HSV) to simulate viral re-infection (or re-activation), along with free ACV and recombinant P-LL(C)-37 at increasing concentrations and incubated for 24 hours. Cytotoxicity was tested using live/dead staining (Molecular Probes kit) and a MTT assay (Ngamwongsatit, *et al*. 2008).

2.8 Construction of Thioredoxin-SUMO-CBD-Spacer-LL(C)-37 (pLET1-Spacer-CBD) expression vector

The design was to insert PCR LL-37 containing construct into pET-32a(+) in this order: Thioredoxin-His tag –AspPro- Collagen binding domain- Spacer – LL-37 –Stop. The following retriction sites were designed between the modules. EcoRI-AspPro-Collagen binding Domain-SalI + SalI-Spacer-pmlI-NotI + NotI-LL37-Stop codon-XhoI.

CBD module design and cloning:

Annealing oligos were designed for the CBD module with the following sequence: N-term GAA TTC **GAC CCG** TGG CGC GAA CCG AGC TTC ATG GCG CTG AGC GTC GAC C-term. (Underlined sequences at the 5' indicate EcoRI restriction site and SalI restriction site at the 3'. **Asp/Pro** sequence was placed at the N-term, following the EcoRI restriction site,

and is indicated as bold.) The annealing oligos were designed so that following the annealing step, the CBD domain could be ligated into the EcoRI/SalI digested vector without further work, as shown in the following annealing design:

EcoRI **Asp/Pro** SalI

5’ AATTC**GACCCG** TGG CGC GAA CCG AGC TTC ATG GCG CTG AGC G 3’

3’ G**CTGGGC** ACC GCG CTT GGC TCG AAG TAC CGC GAC TCG CAG CT 5’

The CBD oligos were ordered with the following sequences:

5’ AATTC**GACCCG** TGG CGC GAA CCG AGC TTC ATG GCG CTG AGC G 3’

5’ TC GAC GCT CAG CGC CAT GAA GCT CGG TTC GCG CCA **CGG GTC** G 3’

The oligos were annealed using PNK reaction followed by heating to 95°C cooled to RT. The pET-32a(+) plasmid vector was digested with the EcoRI/SalI restriction sites, and the annealed CBD sequence with complementary restriction sites was cloned into the digested vector.

Spacer module design and cloning:

Annealing oligos were designed for the spacer module with the following sequence: N-term GTC GAC GGC AGC ACC AGC GGT AGC GGC AAA CCG GGT AGC GGC GAA GGT AGC ACC AAA GGC CAC GTG/GCG GCC GCA C-term. (Underlined sequence indicates SalI restriction site at the N-term and PmlI/NotI restriction site at the C-term.) The annealing oligos were designed so that following the annealing step, the spacer domain could be ligated into the SalI/NotI digested vector without further work, as shown in the following annealing design:

SalI NotI

5’ TC GAC GGC AGC ACC (...spacer...) ACC AAA GGC CAC GTG/GC 3’

3’ G CCG TCG TGG (... spacer...) TGG TTT CCG GTG CAC/CGC CGG 5’

The CBD oligos were ordered with the following sequences:

5’ TC GAC GGC AGC ACC AGC GGT AGC GGC AAA CCG GGT AGC GGC GAA GGT AGC ACC AAA GGC CAC GTG/GC 3’

5’ GGC CGC/CAC GTG GCC TTT GGT GCT ACC TTC GCC GCT ACC CGG TTT GCC GCT ACC GCT GGT GCT GCC G 3’

The oligos were annealed using PNK reaction followed by heating to 95°C cooled to RT. The pET-32a(+) plasmid vector containing the CBD sequence was digested with the SalI/NotI restriction sites, and the annealed spacer sequence with complementary restriction sites was cloned into the digested vector.

LL-37 PCR design and cloning:

To amplify LL-37 from the pLET1 construct by PCR, the following primers were designed.

Forward Primer: 5’ GCG GCC GCA CTG CTG **TGC** GAT TTT TTT CGC 3’ (Underlined sequence indicates NotI restriction site, and the **Cys** mutation on the third amino acid of LL-37 is indicated in bold.)

Reverse Primer: 5’ CTC GAG TTA TTA GCT TTC GGT GCG CGG CAC CAG GTT 5’ (Underlined sequence indicates XhoI restriction site.)

The Amplicon was digested with NotI/XhoI restrictions enzymes and purified using PCR clean up kit. The pET-32a(+) plasmid vector containing the CBD-Spacer sequence was digested with the NotI/XhoI restriction sites, and the LL(C)-37 sequence with complementary restriction sites was cloned into the digested vector.

2.9 Construction of SUMO-CBD-Spacer-LL(C)-37 expression vector

PCR primers were designed to amplify the SUMO and CBD-Spacer-LL(C)-37 sequences from their respective templates, Smt3 and pLET1-Spacer-CBD (PCR step #1). The full length SUMO- CBD-Spacer-LL(C)-37 sequence was generated progressively in a single reaction by overlap extension PCR (PCR step #2), and cloned in a modified pET-22b(+) based plasmid (which was obtained from Dr. Couture) with the NdeI and XhoI restriction sites. The general PCR steps are outlined in Figure 2-1.

PCR step #1 a- The PCR primers used to amplify the 6x His-SUMO sequence from the Smt3 vector (which was a generous gift from Dr. Couture) with an overlapping CBD containing sequence:

Forward primer 1: 5' CAT ATG CAC CAT CAC CAT CAC CAT TCG 3' (Underlined sequence indicates NdeI restriction site)

Reverse primer 1: 5' CAT GAA GCT CGG TTC GCG CCA ATA CGT AGC ACC ACC AAT CTG TTC 3' (Underlined sequence indicates part of CBD sequence, seamlessly adjacent to the C-term of the SUMO sequence)

PCR step #1 b- The PCR primers used to amplify the CBD-Spacer-LL(C)-37 sequence from the pLET1-CBD-Spacer vector with an overlapping SUMO containing sequence:

Forward primer 2: 5' TGG CGC GAA CCG AGC TTC ATG GCG CTG AGC GTC GAC 3'

Reverse primer 2: 5' CTC GAG TTA TTA GCT TTC GGT GCG CGG CAC CAG GTT 3' (Underlined sequence indicates XhoI restriction site)

PCR step #2- The full length SUMO-CBD-Spacer-LL(C)-37 sequence was generated progressively in a single reaction by overlap extension PCR, using forward primer 1, and reverse primer 2, and the PCR purified products of PCRs #1a and #1b. The full length SUMO-CBD-

Spacer-LL(C)-37sequence was digested with the NdeI and XhoI restriction enzymes, and then ligated (using Quick Ligation Kit) into previously digested pHIS2 vector to create the corresponding expression vector SUMO-CBD-Spacer-LL(C)-37. Accuracy of the inserted cDNA was confirmed by automated DNA sequencing. The SUMO-CBD-Spacer-LL(C)-37 construct was transformed into the *E. coli* strain BL21 for protein expression.

2.10 Cleavage of SUMO-CBD-Spacer-LL(C)-37 fusion construct

The fusion protein fractions were pooled and dialyzed in 10 mM Tris, 150 mM NaCl, 1mM DTT, pH 8.0 buffer using dialysis tube with 1 kDa MWCO (molecular weight cut-off) to remove imidazole, which inhibits SUMO Protease I cleavage. SUMO protease was added to the purified fusion protein, and the mixture was incubated in buffer conatinuing10mM Tris, 150 mM NaCl, 1mM DTT, pH 8.0 at 4°C, RT, or 30°C for 24, 6, and 2 h respectively.
The cleavage efficiency was checked by either SDS-PAGE (15%) or Tris-Tricine (10-20%) gradient gel analysis, stained with Coomassie Blue. Proteins were transferred to PVDF membrane at 15 V for 15 min, and stained with Coomassie Blue prior to amino acid analysis.

2.11 Intein Fusion Systems

An intein fusion partner was used to assess the expression and purification of modified LL-37. The results of the intein fusion systems are presented and discussed in Appendix A.

IMPACT Protein Purification System

The IMPACT (Intein Mediated Purification system with an Affinity Chitin-binding Tag) protein purification system was employed initially. AMP is fused to a self-cleavable intein tag in which a chitin binding domain allows affinity purification of the fusion precursor on a chitin

column. In the presence of thiols such as DTT, the intein undergoes specific self-cleavage which releases the target protein from the chitin-bound intein tag resulting in a single-column purification of the target protein. (Supplementary Fig. 1, [24])

Bioseparation Using Self-Cleaving Elastin-Like Polypeptide (ELP) Tags

Self-cleaving ELP tags consist of repeating pentapeptides of VPGXG (X = any amino acid), which are fused to a self-cleaving intein. Upon expression of this construct, it is expected that as the cell lysate solution is heated to 30–40 °C and the salt concentration is increased, the ELP-intein tag becomes insoluble, and precipitates (Supplementary Figure 3, [25]). As the His-tagged LL(C)-37 is self-cleaved from the intein, and the ELP-intein is precipitated out, it is expected that the His- tagged LL37 will be in the supernatant in solution.

CHAPTER 3

RESULTS

3.1 Modified LL-37 System

We focused on LL-37 sequence as a model AMP to design a recombinant expression system. To test express and purify the modified AMP LL-37, three different fusion partners were assessed. Following successful AMP purification in the thioredoxin fusion system, we opted to test the antiviral protection activity and the effectiveness of recombinant modified LL-37, expressed in bacteria in comparison to the synthetic LL-37 made on the peptide synthesizer. The LL-37 sequence that was ligated in the intein and thioredoxin fusion systems was modified to incorporate a cysteine mutation in the third amino acid position for possible future crosslinkage purposes onto collagen of biomaterials based corneal substitutes. The LL-37 sequence that was ligated in the SUMO fusion system was modified to incorporate a cysteine followed by serine and glycine at the N-terminus. The cysteine was to enable possible future crosslinkage onto collagen of biomaterials based corneal substitutes, and the serine and glycine were purposed to serve as a small spacer.

After inclusion of the modified LL-37 sequence in the expression vector, each of the candidate fusion systems (intein, thioredoxin, and SUMO) were further investigated for improved peptide expression and purification. The expression and purification of the intein, and SUMO fusion systems are mentioned in Appendix A, and the reason as to why they were deemed unsuitable for our purposes is discussed. The expression and purification of the more successful thioredoxin fusion system, followed by characterization and testing of the purified modified LL-37 is next presented. These results will help establish recombinant LL-37's

potential as antiviral drug to be used and incorporated in biomaterial based corneal substitutes made primarily of collagen.

3.1.1 Design and Cloning of Modified AMP LL-37 Fusion Systems: Intein, Thioredoxin and SUMO fusion tags.

3.1.1.1 Intein Fusion Systems

IMPACT Protein Purification System

We used the pTYB expression vectors, which allowed the fusion of the cleavable intein tag to be at either the C-terminus (pTYB2, C-terminal fusion) or the N-terminus (pTYB12, N-terminal fusion) of LL(C)-37. This flexibility in fusion protein construction was intended to maximize the probability of successful expression and purification. Both of these expression vectors were constructed by Dr. Dick, and were generously given to me to follow through with expression and purification.

Bioseparation Using Self-Cleaving Elastin-Like Polypeptide (ELP) Tags

Molecular biology was used to add a His-tag to the LL(C)-37. The His-tagged LL(C)-37 was fused to the gene encoding the self-cleaving ELP-intein tag, by cloning it into an elastomer vector (pEI). This expression vector was constructed by Dr. Dick, and was generously given to me to follow through with expression and purification.

3.1.1.2 Thioredoxin Fusion System

The construct for LL(C)-37 expression ligated into a modified pET-32a(+) plasmid, containing a His-tag for affinity purification and a formic acid cleavage site for peptide release,

referred to as pLET1, is depicted in the schematic drawing (Figure 3-1). The pLET1 expression vector was constructed by Dr. Dick, and was generously given to me to follow through with expression and purification.

In order to reduce the cost of the enzyme required for peptide cleavage, we opted to release the AMP by chemical cleavage using formic acid. Through our primer design, we created an Asp-Pro chemical cleavage site, between the His-tag and LL(C)-37. Following cleavage a Proline residue is left at the N terminus of LL(C)-37, a possible disadvantage of this method. This recombinant Proline-LL(C)-37 is hereinafter called P-LL(C)-37. The whole fusion protein may be separated from the whole cell extract using the His-tag, which has binding affinity to Ni-NTA column. This can be done either in denaturing conditions, or in native condition. After the AMP is cleaved from the thioredoxin fusion using acid digest, the AMP peptide will be separated from its carrier protein using chromatography technique such as ion exchange column or HPLC.

3.1.1.3 **SUMO Fusion System**

In collaboration with a group who generously provided us with their modified pET-22b(+) vector and the SUMO sequence, we designed the LL-37 sequence to be ligated in the vector and expressed as a SUMO fusion (Figure 3-2). The constructed SUMO fusion vector encoded the LL-37 sequence with a three amino acid linker (Cys-Ser-Gly) at the N-terminus. The cysteine was included to enable possible future chemical crosslinking onto the collagen matrix of corneal substitutes, and the serine and glycine were intended to serve as a flexible spacer. The SUMO-(CSG)-LL-37 fusion protein was constructed using a two step PCR design for seamless fusion. The correctness of the construct sequence was confirmed by DNA sequencing.

Figure 3-1: The DNA sequence encoding AMP human LL-37 into the pET-32a(+) plasmid, referred to as pLET1.

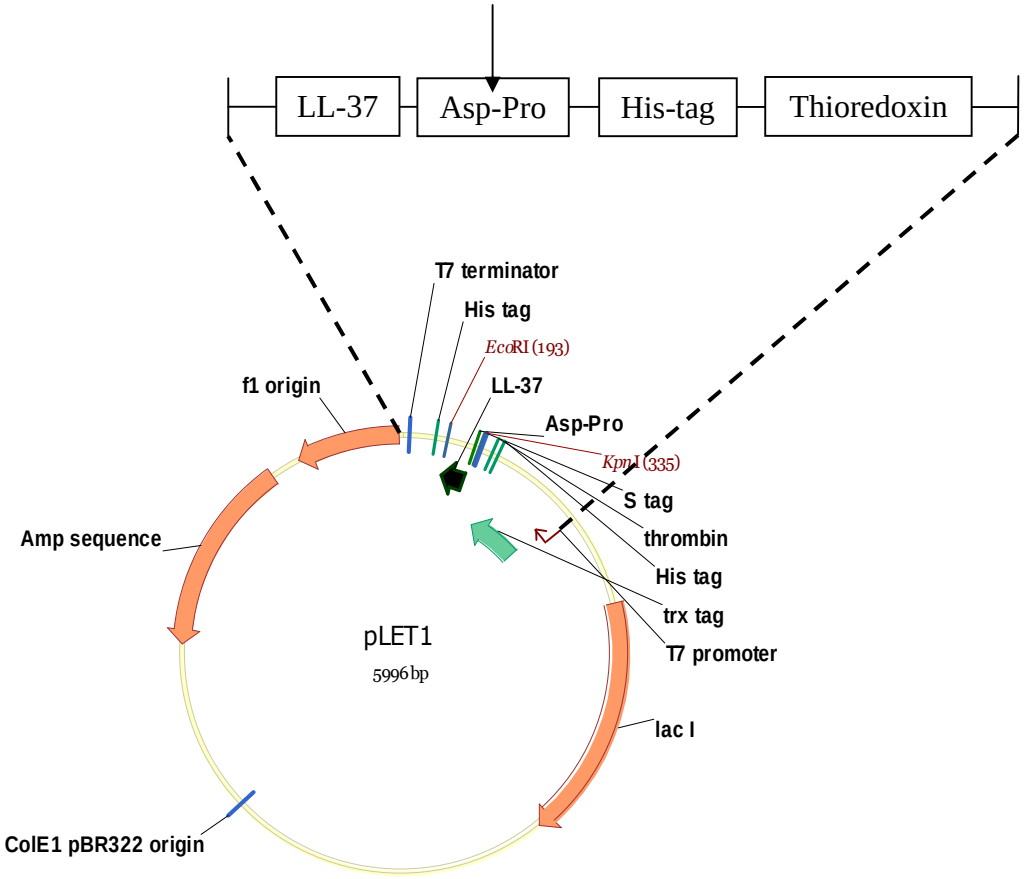
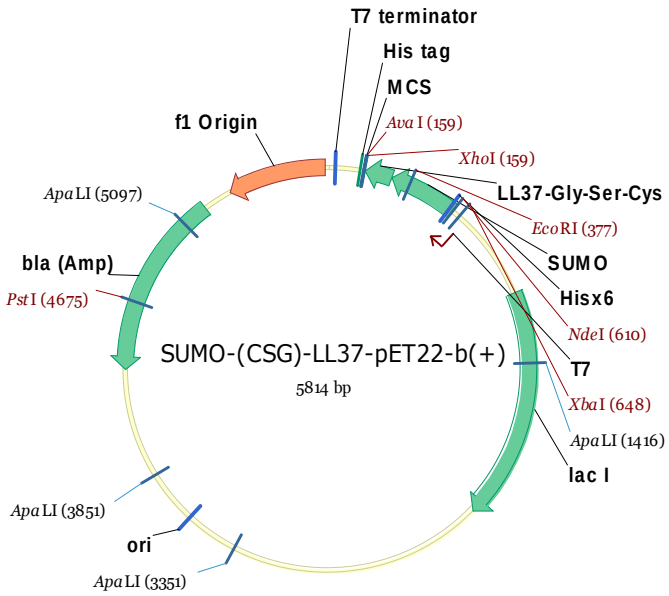


Figure 3-2: The DNA sequence encoding SUMO-(CSG)-LL-37 was ligated into a modified pET-22b(+) based plasmid.



The production of a native N-terminated fusion following SUMO Protease I cleavage is quite advantageous compared to the formic acid cleavage system, which leaves a Proline residue at the N-terminus. Due to the unique activity of SUMO Protease I, the enzyme never cleaves within the protein of interest and thus, the cleavage may be more controlled in comparison to a chemical cleavage. Once cleaved, the native protein is then able to be purified by removal of the SUMO fusion protein and SUMO Protease I by affinity chromatography on a nickel-chelating resin by binding of the His tag that is present on both fusion and protease. Thus, recombinant proteins expressed from the pET-based SUMO vector, and cleaved by SUMO Protease I should result in production of native protein.

Application of the SUMO fusion strategy to antimicrobial peptides have not been widely reported [20]. As with any enzymatic cleavage system, the high cost of enzyme are limiting. To reduce the cost, SUMO Protease I with a His tag for affinity purification may be produced in proteomic labs, reducing the large-scale purification cost substantially. However, it may be a challenge to obtain non-commercial protease with consistent activity. Thus, there might be limitations in both options of using either commercial or non-commercial protease that must be considered.

3.1.2 Expression and Purification of Modified AMP LL-37 Fusion using the more successful Thioredoxin Fusion System

Escherichia coli strain BL21 harboring the pLET1 vector was utilized to express the fusion protein containing LL(C)-37. The expression of the full chimera was induced in the presence of IPTG. The amount of the expressed fusion protein as judged from Coomassie Blue

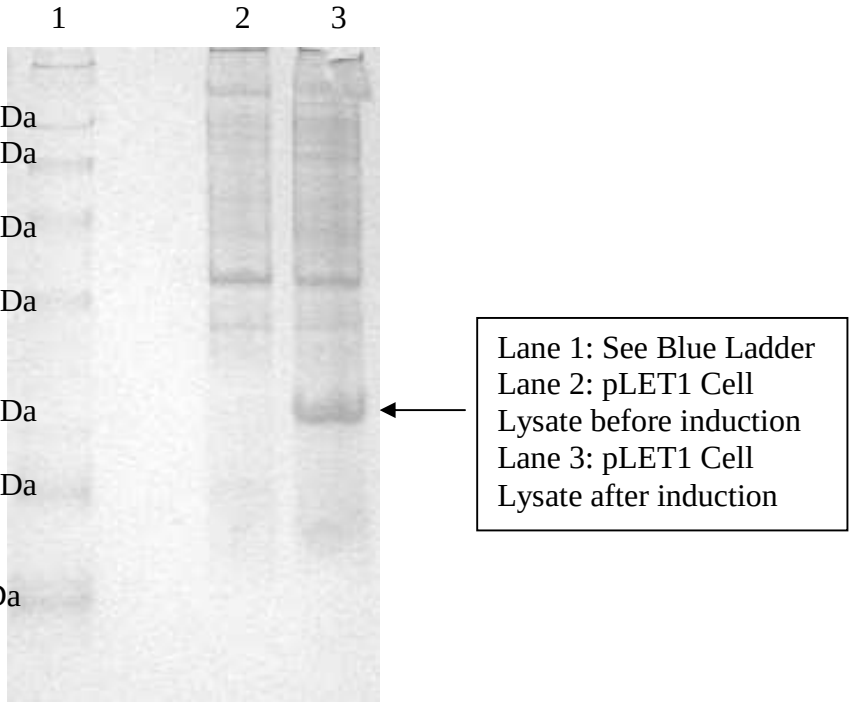
staining (Figure 3-3) was comparatively higher compared to our previously tried expression systems incorporating the intein fusion partner as mentioned in Appendix A.

The thioredoxin-LL(C)-37 fusion protein was purified from the whole cell free extract using a Ni-NTA nickel column equilibrated with binding buffer. This was initially carried out in denaturing conditions but due to the problems observed in our initial trial using denaturing conditions discussed below it was subsequently carried out in native conditions.

To purify the thioredoxin-LL(C)-37 fusion protein from the cell lysate, the pLET1 bacteria pellet was initially resuspended in pH 8.0 buffer containing urea. After centrifugation, the fusion protein was found primarily in the supernatant of whole cell lysate, and was loaded onto a previously equilibrated Ni-NTA nickel column. In the wash step, the pH of the urea buffer was dropped to 6.3 and 5.9, and the full fusion protein at the molecular weight of 23 kDa was expected to come off the column in the buffer with a low pH 4.5. However, it was observed that the LL(C)-37 seemed to be prematurely cleaved from the thioredoxin fusion, and thus appearing in the wash, as well as the elute fractions (Figure not shown). The elute fractions thus contained 3 bands, corresponding to the full fusion protein at 22kDa, the thioredoxin-His tag at 17kDa, and the Proline-LL(C)-37 at around 5kDa. It appeared that low pH buffer, and the high amount of urea encourage premature cleavage of the AMP from the carrier protein. Thus, much of the peptide was lost in the wash steps, as well as the elute fractions prior to cleavage. This significantly reduced the amount of full fusion protein collected for cleavage step. After performing acid cleavage, and running it through an anion exchange column to separate it from the carrier protein, there was very small trace of Proline-LL(C)-37, as it was barely visible on SDS-PAGE. Thus, we proceeded onto purifying the cell free extract under native conditions, to overcome such premature cleavage.

Figure 3-3: Expression of pLET1 in *E. coli*. *E. coli* grown in LB media was induced at 37°C for 4 hours. Expression of thioredoxin-LL(C)-37 fusion protein at 23kDa (indicated by an arrow) as followed by SDS-PAGE (15%) analysis.

(b)



Prior to purification in native conditions, the amino acid sequence of thioredoxin, and Asp-Pro-LL37 were both analyzed using the PROWL ProFound digestion computer program for protease activity (Figure 3-4). It was observed that it contains quite a number of Lysine residues, which might be prone to serine protease activity, such as trypsin. PMSF very effectively inhibits both trypsin and chymotrypsin, since they are both serine proteases.

To purify the thioredoxin-LL(C)-37 fusion protein from the cell lysate, the pLET1 bacteria pellet was resuspended in native purification buffer, containing low amounts of imidazole and lysis buffer, with 0.1mM PMSF. After centrifugation, the fusion protein was found primarily in the supernatant of whole cell lysate (Figure 3-5, lane 2), and was loaded onto a previously equilibrated Ni-NTA nickel column. The column was washed with slightly higher concentration of imidazole to reduce non-specific binding, and finally eluted with a very high concentration of imidazole. It can be observed that the majority of the unwanted proteins were removed by Ni-NTA affinity chromatography (Figure 3-5, lanes 3-7), and most of the fusion protein is eluted in the first collected fraction (Figure 3-5, lane 8).

The elution fractions were pooled together, and dialyzed to remove imidazole. This step was necessary prior to the 50% Formic acid digest carried out at 50˚C, to release the peptide from the fusion protein. Following cleavage, it appears that P-LL(C)-37 is being released from the fusion (Figure 3-6, circled band). In the presence of formic acid, the band intensity of the full length fusion protein slightly above 22 kDa was significantly reduced. Meanwhile, two predominant bands corresponding to the carrier protein (17 kDa) and P-LL(C)-37 (5 kDa) appeared, indicating a successful cleavage of the fusion (Figure 3-6). It should be noted that the smear that is visible on the SDS-PAGE is perhaps an indication of either degradation or nonspecific cleavage occurring during the cleavage process, as the fusion protein band was a

Figure 3-4: Trypsin cleavage sites of pLET1 fusion protein. Trypsin cleavage sites of thioredoxin, and Asp-Pro-LL(C)-37 sequence based on their amino acid sequence using a Protein analysis program, from http://prowl.rockefeller.edu/.

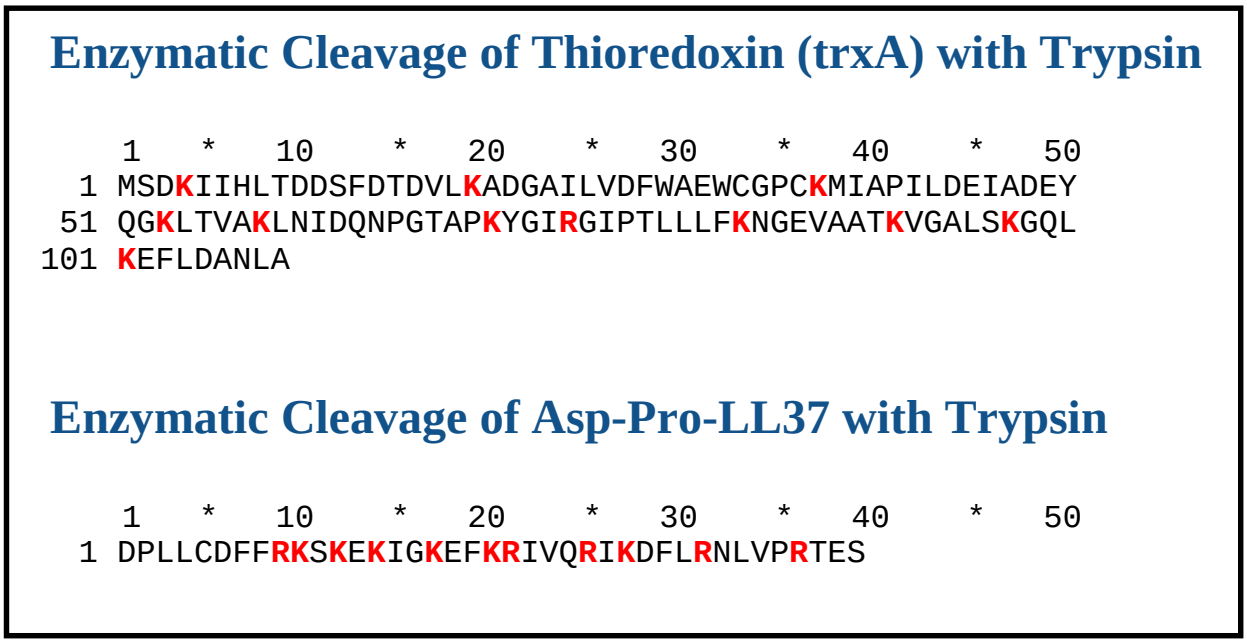


Figure 3-5: Purification of His-tagged LL-37 fusion protein on IMAC (10 mL) under native conditions, as followed by SDS-PAGE (15%) analysis. Lane 1, See Blue Ladder; Lane 2, cell free extract; Lane 3, column flow through; Lanes 4 and 5, column wash with 20mM imidazole; Lanes 6 and 7, column wash with 50 and 100mM imidazole respectively; Lanes 8-10, protein fractions eluted with 250mM imidazole from Qiagen Ni-NTA resin.

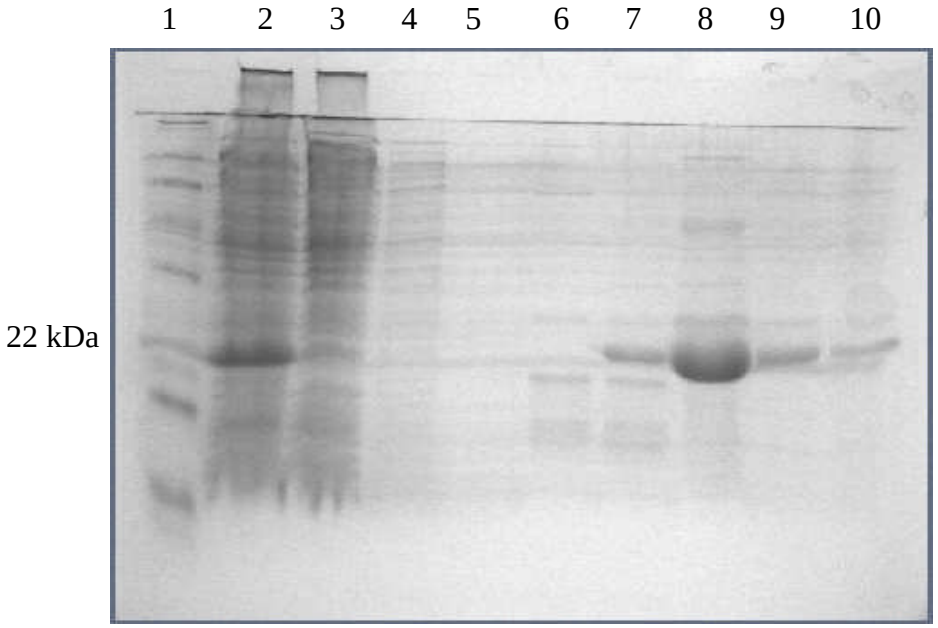
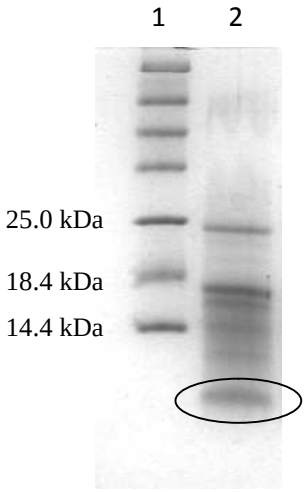


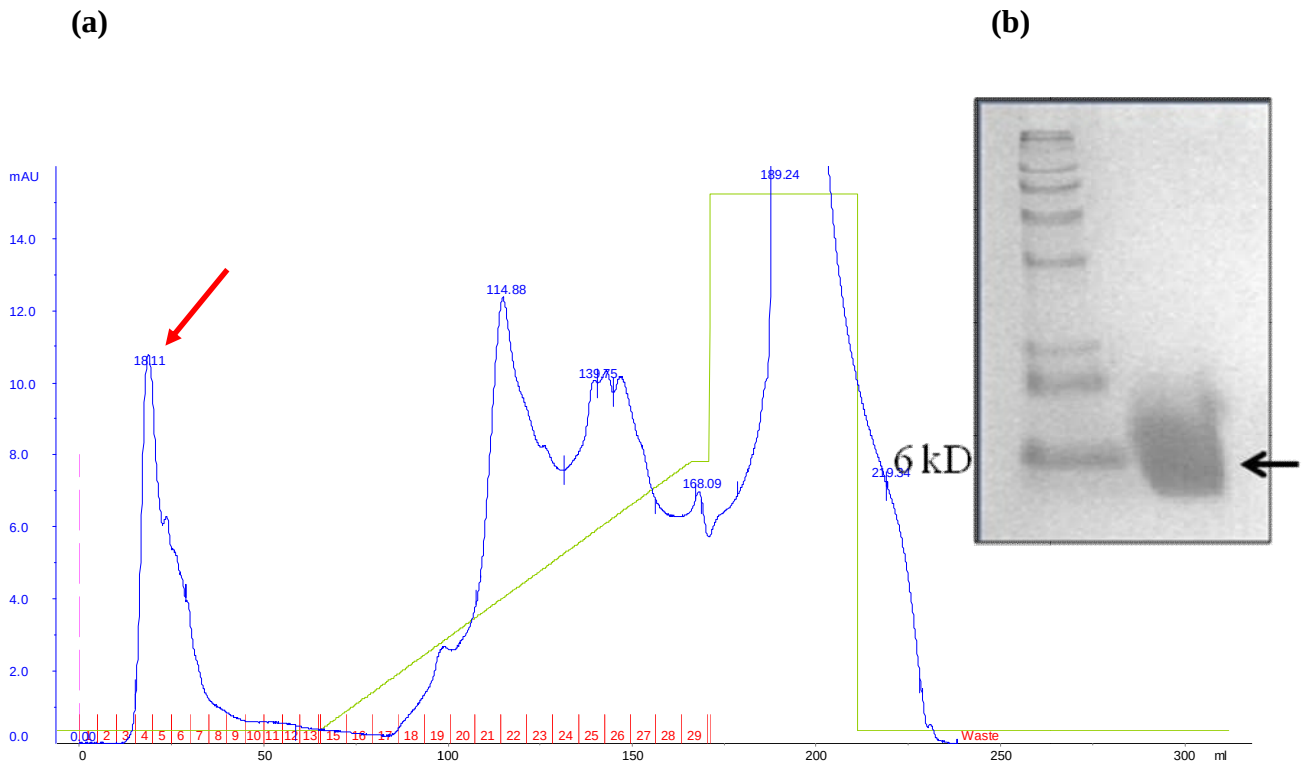
Figure 3-6: SDS-Page (15%) analysis of pLET1 cleavage. Lane 1, Ladder; Lane 2, P-LL(C)-37 (circled) was cleaved off the thioredoxin fusion, following a 50hr Formic acid (50%) cleavage.



single band prior to cleavage. This could be due to the harsh cleavage condition, although reducing the cleavage reaction time or the temperature still produced similar pattern, only with darker fusion band (reduced cleavage).

The reaction mixture was then lyophilized to remove formic acid. The His-tagged carrier and residual undigested fusion proteins as well as non-specifically cleaved proteins were then rapidly removed by FPLC, using an anion exchange column (Figure 3-7). The P-LL(C)-37, which has a pI of 10.5, is eluted in 10mM Tris, 20mM NaCl, pH 8.0 buffer as flow through (Figure 3-7 a, fraction 4), while the thioredoxin-LL(C)-37 fusion and thioredoxin are retained longer with pIs of 7.9 and 5.3 respectively. SDS-PAGE analysis of the fraction 4 shows a single band at around 5 kDa consistent with the molecular weight of P-LL(C)-37 (Figure 3-7 b).

Figure 3-7: Separation of P-LL(C)-37 from the carrier thioredoxin by FPLC. Following a 50hr Formic acid (50%) cleavage, the sample was loaded onto an anion exchanger column. (a) The P-LL(C)-37, which has a pI of 10.5, is eluted in 10mM Tris, 20mM NaCl, pH 8.0 buffer as flow through (fraction 4). SDS-PAGE (15%) analysis of the fraction 4 shows a single band at ~5 kDa (b) consistent with the molecular weight of P-LL(C)-37.



3.1.3 Characterization and Testing of Modified AMP LL-37

The recombinantly purified AMP was further analyzed and characterized to investigate its sequence integrity and potential anti-infective activity. P-LL(C)-37 has a calculated molecular weight of 4636. The mass determined by MALDI-TOF MS analysis is 4499, which agrees 97% with the calculated value (Figure 3-8). The discrepancy may be due to some modification of the peptide due to degradation or modification of the amino acid side-chains due to formic acid activity.

To test the recombinant P-LL(C)-37 anti-microbial and anti-HSV-1 activity, a viral protection assay was carried out by Mrs. Bareiss. Our recombinantly expressed P-LL(C)-37 was tested for its anti-viral activity against HSV-1 in HCEC cells, in comparison to the drug Acyclovir (Figure 3-9, Bareiss, B). P-LL-(C)-37 peptide was serially diluted and tested against 0.1 MOI HSV-1 on HCEC. The following comparison table (Table 3-1), summarizes the effective concentration of synthetic LL-37 and recombinant P-LL(C)-37 at which 90%, 75%, and 50% (used as standard) of the cells are living. The control panel was untreated cells set to 100% living cells, and the cells in the other treatments were then counted and the percent living cells were calculated. It can be observed that the results correlate, as the numbers are effectively comparable and similar in each column of percent living cells, for both synthetic LL-37 and recombinant modified LL-37. In summary, at a concentration of 1µg/ml both the synthetic and recombinant LL-37 showed no effect. At a concentration of 1µg/ml to 20µg/ml, both types of AMP LL-37 showed a concentration dependent protection against HSV. At a concentration greater than 20µg/ml, both types of AMP LL-37 protect against HSV infection. Although there are some limitations with exact measurment due to the method used in this experiment, these results still positively demonstrate our thesis objective.

Figure 3-8: Mass spectrum of purified P-LL(C)-37 on MALDI-TOF mass spectrometry.

The molecular weight determined for the recombinant protein was 4499, calculated mass is 4636.

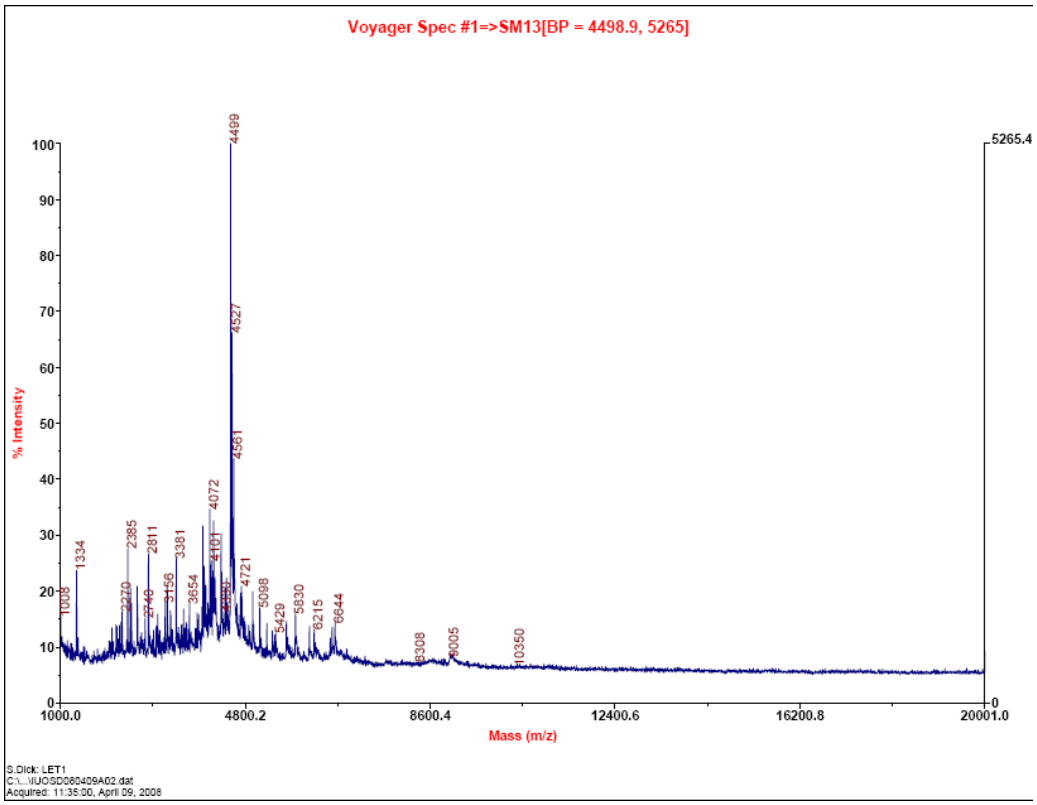
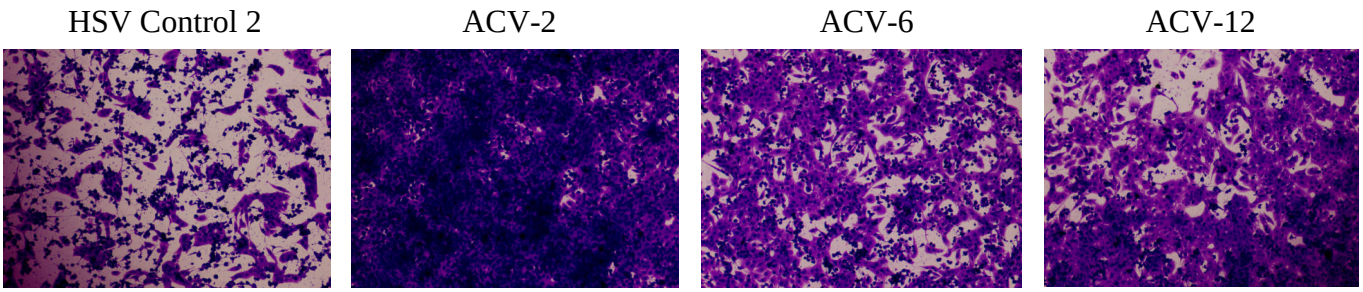


Figure 3-9: Antiviral activity of recombinant P-LL(C)-37 against HSV-1. (a) HCEC cells were infected with 0.1 MOI HSV-1. Virus protection is dose-dependent. When no drug is added (HSV Control 2), the cells are clearly dying. As a control the drug ACV was added to one row of cell culture plate from higher concentration (ACV-2) to lower concentration (ACV-12). As the concentration of the drug decreased, the level of protection also decreased.

(b) Recombinant P-LL(C)-37 was added to another row of cell culture plate, from highest concentration in LL37-2 to lowest concentration in LL37-12. At a concentration of 17µg/mL 90% of the cells are living, at 7µg/mL 75% of the cells are living, and at 4.8µg/mL only 50% of the cells are living, which means it has no effect against HSV-1 infection. The standard deviation was calculated to be ±2%. (In collaboration with Mrs. Bareiss, B)

(a)



(b)

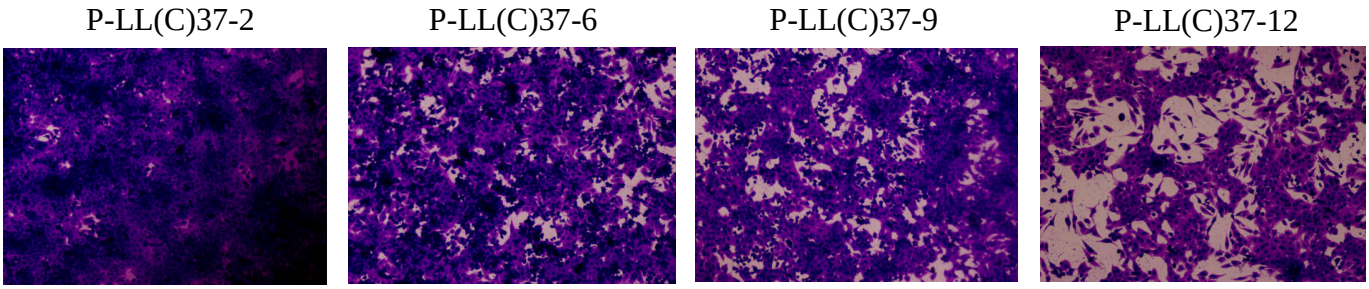


Table 3-1: Equivalent concentrations of recombinant P-LL(C)-37 in comparison to synthetic LL-37 and at which 90%, 75%, and 50% of the cells are living. Virus protection is concentration dependent for both Synthetic LL-37 and Recombinant P-LL(C)-37.

	Concentration at which 90% cells Living, 10% cells dead from HSV infection	**Concentration at which 75% cells Living, 25% cells dead from HSV infection**	**Concentration at which 50% cells Living, 50% cells dead from HSV infection (No effect)**
Synthetic LL-37	19.5 µg/mL	9.7 µg/mL	4.8 µg/mL
Recombinant P-LL(C)-37	17 µg/mL	7 µg/mL	3 µg/mL

3.2 CBD-Spacer-LL37 System

As previously mentioned, the effective recombinant anti-microbial peptide will need to be integrated in the biomaterials scaffolds, with sustainable length and flexibility. To achieve this goal, a collagen-binding domain and a spacer sequence were selected to be ligated at the N-terminus of LL-37 for possible attachment and proper form following peptide expression. Using standard molecular biology methods, the selected collagen-binding domain and spacer sequences were used to create our designed biofunctional AMP LL-37. This construct was designed in both thioredoxin and SUMO fusion systems for improved peptide expression and purification.

The SUMO fusion system was later considered as the better choice for expression and purification of CBD-Spacer-LL-37 construct. The limitations of the thioredoxin fusion system in the acid cleavage step caused us to be cautious about the possible variability and inconsistencies that might arise due to acid reduction of biofunctional side-chains. This is particularly important when functional domains such as the collagen-binding domain are introduced in our designed peptide. In addition, following acid cleavage a Proline residue is left at the N terminus of our biofunctional AMP, something which is avoidable in the SUMO fusion system with proper design.

3.2.1 Design and Cloning of the CBD-Spacer-LL37 Fusion Systems

3.2.1.1 Thioredoxin Fusion System

The expression system initially used to design our biofunctional AMP was chosen based on the preliminary purification data of P-LL(C)-37 (in pLET1 expression vector) to be the pET32a(+) vector. The primers were designed in such a way that the cloning was done in blocks (AMP; Spacer; CBD), with each of the 3 block (ie. CBD) having unique restriction sites at either

end in order to accommodate change in this modular ligation. Thus, the Spacer or CBD may be swapped with other spacer, or CBD sequences, whenever desired. To clone the spacer and CBD sequence in the pLET1, it was required to position the AMP LL-37 at the C-term, since the C-term of LL37 is the active part for AMP activity and should be free in order to retain its activity (Figure 3-10). The LL-37 sequence that was inserted in the thioredoxin fusion system incorporated the modified cysteine mutation in the third amino acid position, as the pLET1 construct was used as PCR template.

Collagen-binding domain is a functional domain, and any side-chain modification or degradation that modifies and affects the potential activity of the product is highly undesirable. Thus, even though, the proposed CBD-Spacer-LL37 was designed and cloned in the pET32-a(+) plasmid with thioredoxin fusion tag (Figure 3-10), another construct was designed for the SUMO fusion system.

3.2.1.2 **SUMO Fusion System**

The designed CBD-Spacer-LL(C)-37 sequence was cloned in the SUMO-tagged pET-based vector that we had obtained from the group who had generously produced us with the cloning and expression plasmids (Figure 3-11). The DNA sequence from the pLET1-Spacer-CBD construct, with a cysteine mutation in the third amino acid of LL-37, was used as template for the PCR amplification. Following a two step PCR, the CBD-Spacer-LL(C)-37 sequence was fused seamlessly to the SUMO fusion (Figure 3-11). The recombinant plasmid with correct reading frame was confirmed by DNA sequencing.

Figure 3-10: Vector map and Cloning of LL37 with Spacer and CBD in *E. coli*. The DNA sequence encoding human LL-37, Spacer, and CBD into the pET-32a(+) plasmid, referred to as pLET1- Spacer- CBD. Note that LL-37 is designed to have its C-term free, to not interfere with its AMP activity.

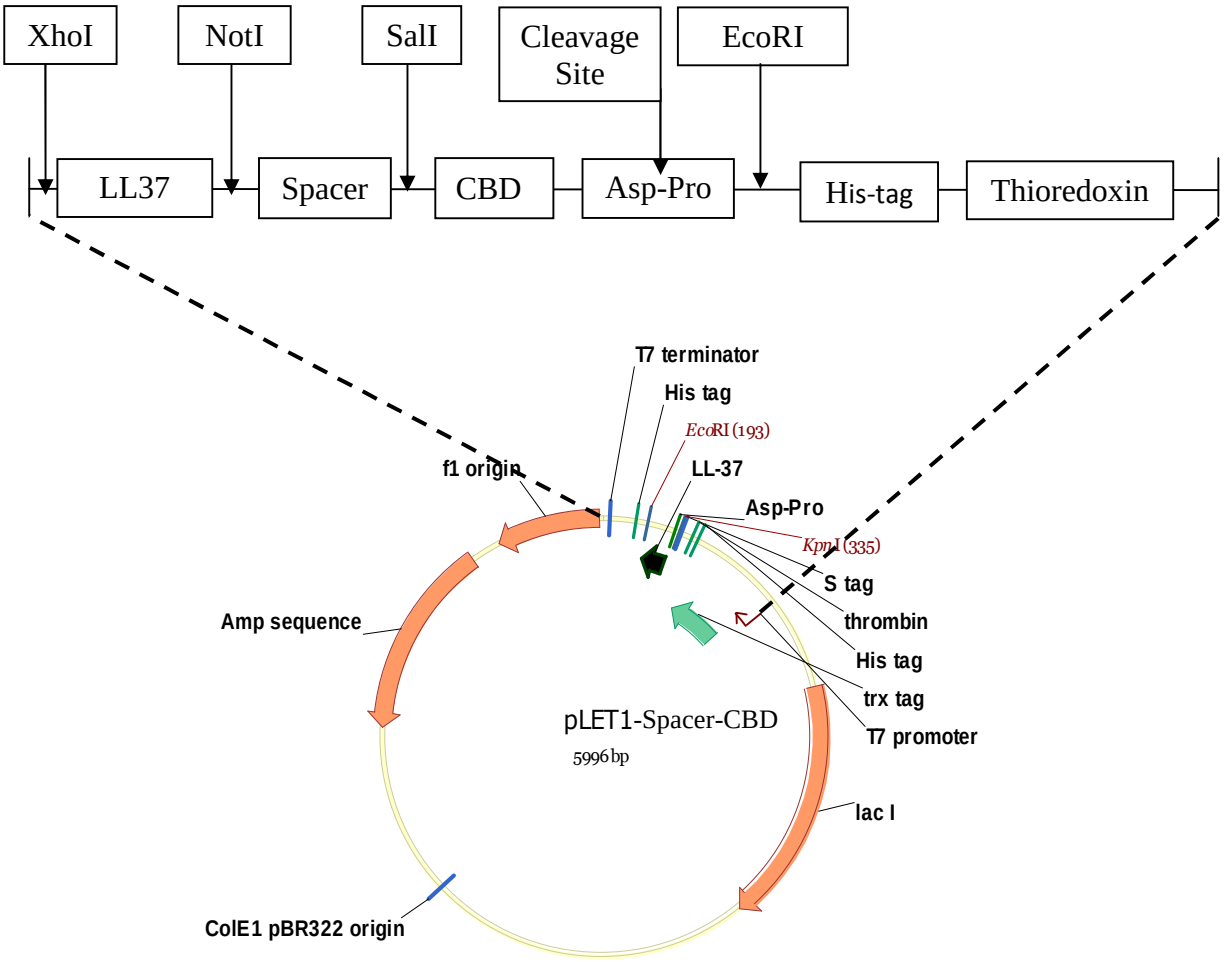
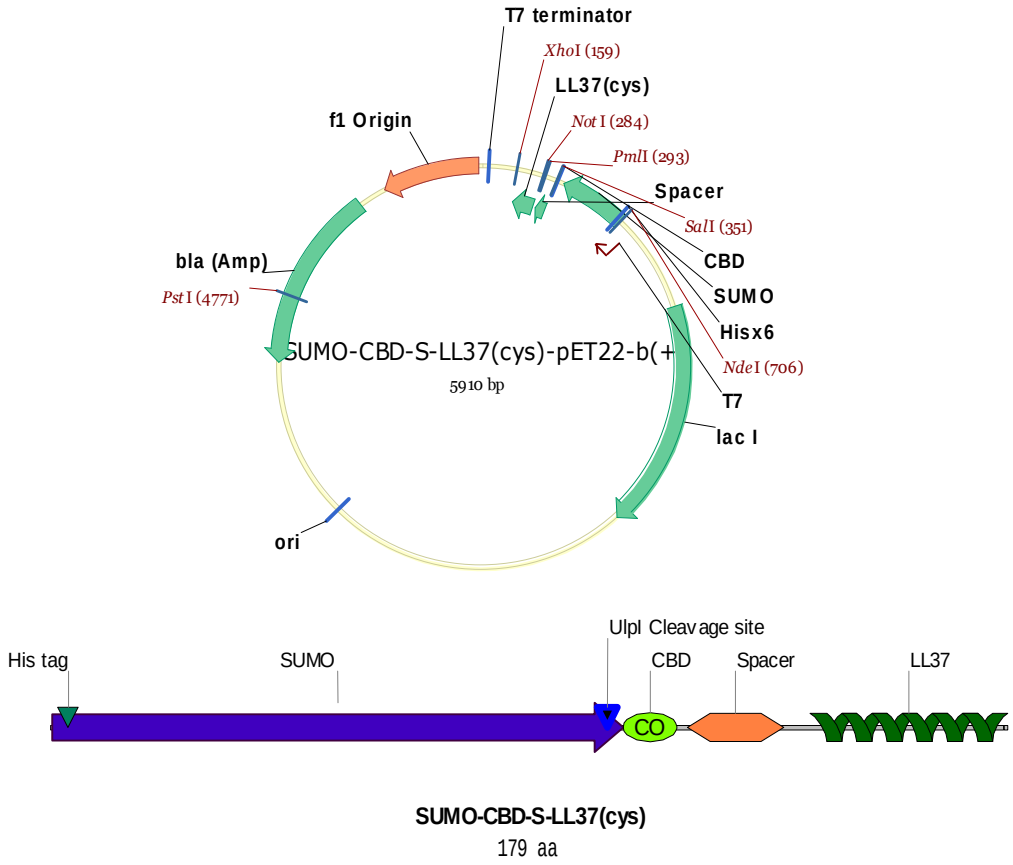


Figure 3-11: Vector map and schematic diagram of SUMO-CBD-Spacer-LL(C)-37 Fusion.

The DNA sequence encoding SUMO-CBD-Spacer-LL(C)-37, into the pET-22b(+) plasmid. A schematic diagram of the expected expressed chimera is depicted below



3.2.1.3 **His Tag System**

Additionally, due to the limitations of the enzymatic cleavage in AMP containing sequences as observed and presented in the purification section below, it was later considered to clone the His-tagged CBD-Spacer-LL(C)-37 construct in a pET –based vector without any fusion. The DNA sequence encoding the 6xHis-CBD-Spacer-LL(C)-37 was ligated in a pET21d vector by Dr. Dick, with XbaI and XhoI restriction sites using the appropriate primers designed by me. The recombinant plasmid with correct reading frame was confirmed by DNA sequencing. It was to be investigated if the full correct chimera containing the AMP sequence may be readily expressed in bacteria, with the addition of the extra 34 amino acid at the N-term contained in the 6xHis-CBD-Spacer sequence. Swapping longer CBD and spacer sequences may be incorporated in the future design as they might enhance the chances of expression and purification. The limited data available from the purification and characterization of the 6xHis-CBD-Spacer-LL(C)-37 fusion construct performed by Dr. Dick and Ms. Wan is presented in Appendix B.

3.2.2 Expression and purification of CBD-Spacer-LL37 using the SUMO-based peptide expression system and the Ni-NTA purification system.

The SUMO expression system was used to express the designed LL-37 anti-microbial peptide with the CBD-Spacer sequence. The SUMO expression system was chosen mainly for the advantages it offered in producing a biofunctional fusion protein with intact functional domains. The addition of the CBD-Spacer sequence also offers the advantage of a slight increase in molecular weight following cleavage and thus an increase chance in its visibility on SDS-PAGE analysis, and more facile purification. As we had previously experienced difficulty with

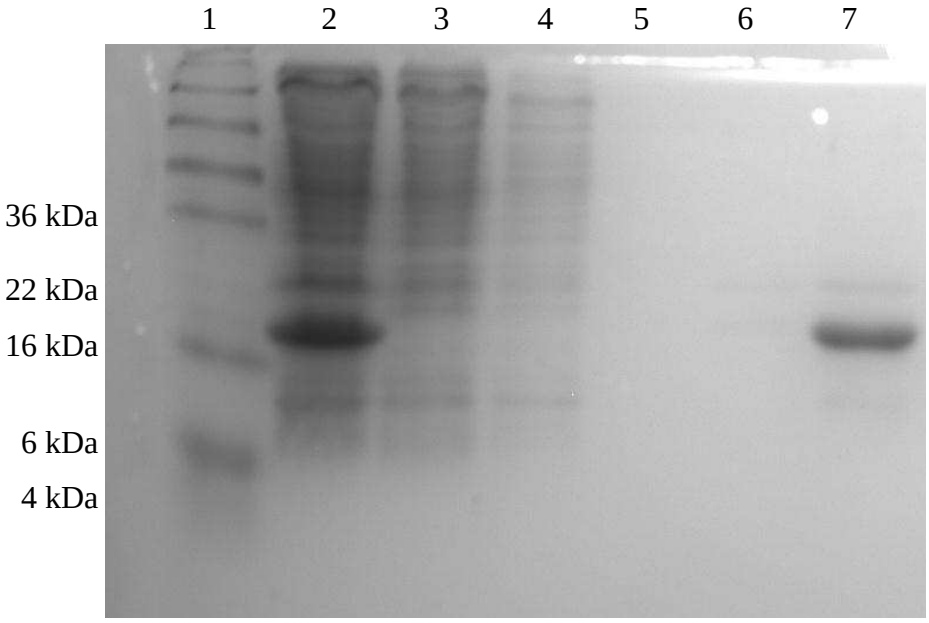
the cleavage and purification of modified LL-37 using the SUMO system (Appendix A), we initially focused on optimization of cleavage conditions.

Escherichia coli strain BL21harboring the SUMO- CBD-Spacer-LL(C)-37 vector was utilized to express the fusion protein. The expression of the full chimera was induced in the presence of IPTG. The molecular weight of the SUMO-His tag is approximately 13 kDa, and the molecular weight of CBD-Spacer-LL(C)-37 is approximately 7.5 kDa. The expected weight of the expressed fusion protein SUMO- CBD-Spacer-LL(C)-37 is approximately 21 kDa. Following induction, a 21 kDa band corresponding to SUMO- CBD-Spacer-LL(C)-37 chimera is visible, as evident based on the relative migration on SDS-PAGE analysis. The SUMO- CBD-Spacer-LL(C)-37 bacteria pellet was resuspended in native purification buffer, containing low amounts of imidazole, PMSF, and lysis buffer. After centrifugation and filteration, the fusion protein was found primarily in the supernatant of whole cell lysate (Figure 3-12, lane 2).

The SUMO- CBD-Spacer-LL(C)-37 fusion was purified from the supernatant of cell lysate using a Ni-NTA nickel column equilibrated with binding buffer. After binding, the Ni-NTA nickel column was washed with slightly higher concentration of imidazole to reduce non-specific binding, and finally eluted with a very high concentration of imidazole. It can be observed that the majority of the unwanted proteins were removed by Ni-NTA affinity chromatography (Figure 3-12, lanes 3-6), and most of the fusion protein is eluted in the first collected fraction (Figure 3-12, lane 7).

As the eluted fusion protein is in buffer containing 250 mM imidazole, there is a need to exchange the buffer for efficient SUMO Protease 1 cleavage, as imidazole concentrations higher than 150 mM can adversely affect the activity of the protease. The isolated SUMO- CBD-Spacer-LL(C)-37 fusion was subsequently dialysed using a 500 Da Molecular Weight Cut Off

Figure 3-12: Purification of His-tagged SUMO-CBD-Spacer-LL(C)-37 fusion protein on IMAC (10 mL), as followed by SDS-PAGE (15%) analysis. Lane 1, See Blue Ladder; Lane 2, cell free extract; Lane 3, column flow through; Lanes 4-6, column wash with 20mM imidazole (40 mL each); Lane 7, protein fraction eluted with 250mM imidazole from Qiagen Ni-NTA resin.

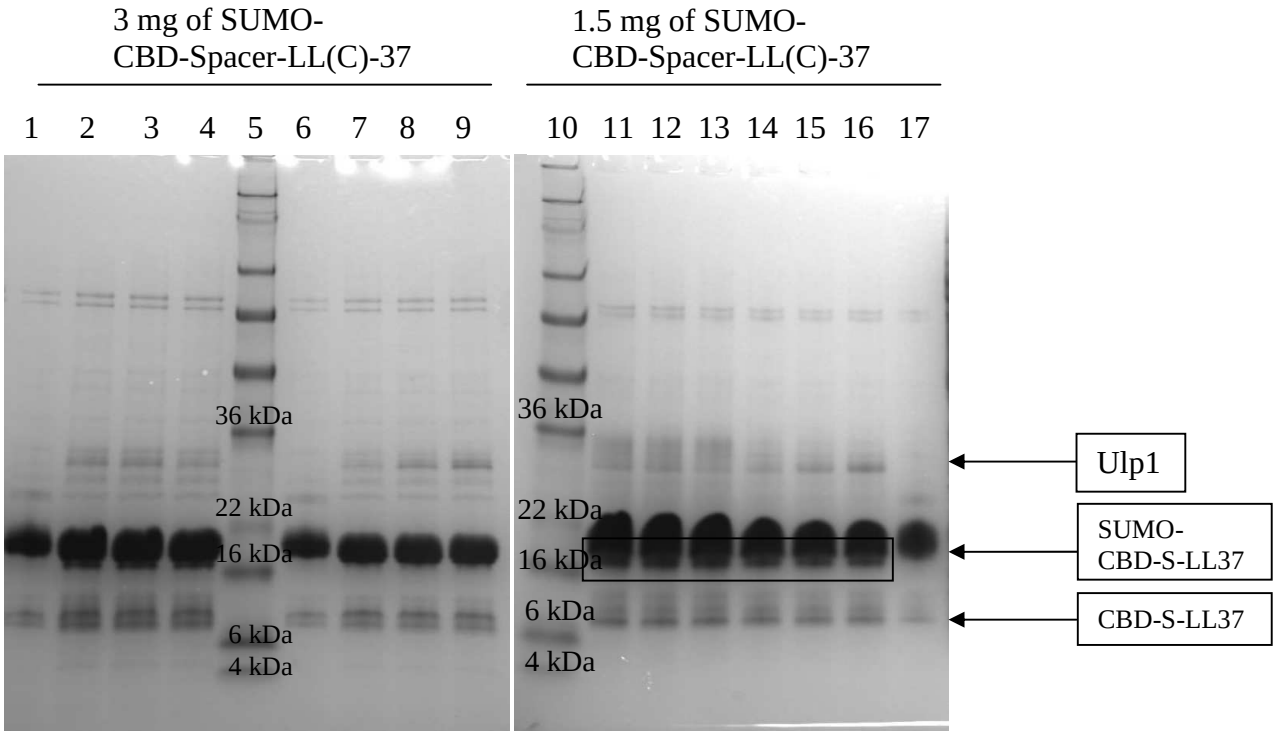


(MWCO) dialysis tube, to a compatible SUMO Protease I cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 1mM DTT, pH 8.0) prior to SUMO Protease I cleavage.

To have a cost effective system, we initially collaborated with the group of J.-F. Couture, who generously produced us with the SUMO Protease I synthesized in their lab. The SUMO Protease I may have a variable range of activity depending on the method of production and various other factors. In particular, it was not possible to obtain the quantified specific protease activity with regards to any one of the enzyme batches we obtained. In addition, the protease had never been tested out with an AMP sequence. As their cleavage requirements differed from ours, following their exact cleavage protocol could only be used merely as a guideline for us. Thus, we systematically proceeded to change the cleavage protocol to improve our results. Through extrapolation of what should work based on our collaborator's experience and data, we could expect to successfully cleave a relatively larger amount of fusion with the addition of a small amount of non-commercial enzyme. However, as following that guideline didn't produce successful cleavage for us (Appendix A), we cleaved 1 mL aliquots (3mg/ml) of our fusion, varying certain parameters as shown below.

Previously, as a precaution to avoid any change in our fusion protein of interest, the cleavage was carried out at 4°C overnight. However, to determine if the cleavage efficiency in our case was affected by temperature, we cleaved 1mL (3mg/ml) of the SUMO-CBD-Spacer-LL(C)-37 fusion with 50 µl of non-commercial SUMO Protease I at three different conditions as described following (Figure 3-13). The cleavage was set up at 4°C overnight, at room temperature for 6 hours, and at 30°C for 2 hours (Figure 3-13, lanes 2-4). Following the enzymatic cleavage and analyzing the sample on 10-20% Tris-Tricine gradient gel, it is clear that the fusion band is

Figure 3-13: Ulp1 cleavage of His-tagged SUMO-CBD-Spacer-LL(C)-37 fusion protein at different temperatures and with varying enzyme amount, as followed by Tris-Tricine Gradient (10-20%) analysis. Lane 1, SUMO-CBD-Spacer-LL(C)-37 (3mg) uncleaved fusion; Lanes 2-4, SUMO-CBD-Spacer-LL(C)-37 (3mg) cleaved at 4°C overnight, room temperature for 6 hrs, and 30°C for 2 hrs respectively; Lane 5, See Blue Ladder; Lane 6, SUMO-CBD-Spacer-LL(C)-37 (3mg) uncleaved fusion; Lane 7-9, SUMO-CBD-Spacer-LL(C)-37 (3mg) cleaved with 50 µl, 100 µl, and 150 µl of Ulp1 respectively. Lane 10, See Blue Ladder; Lanes 11-13, SUMO-CBD-Spacer-LL(C)-37 (1.5 mg) cleaved at 4°C overnight, room temperature for 6 hrs, and 30°C for 2 hrs respectively; Lanes 14-16, SUMO-CBD-Spacer-LL(C)-37 (1.5 mg) cleaved with 50 µl, 100 µl, and 150 µl of Ulp1 respectively; Lane 17, SUMO-CBD-Spacer-LL(C)-37 (1.5 mg) uncleaved fusion.



slightly shifted lower and we may be getting some cleavage with the addition of the SUMO Protease I, although the cleaved bands are not very strong (Figure 3-13, lanes 1-2). It also seems that the cleavage is not affected by different temperature. To determine if the cleavage efficacy may be improved with the addition of more enzyme, 1 mL (3mg/ml) of the SUMO-CBD-Spacer-LL(C)-37 fusion was cleaved with 50 µl, 100 µl, and 150 µl of enzyme respectively at 4°C overnight (Figure 3-13, lanes 7-9). The cleavage efficacy does not seem improved within this range (Figure 3-13). The same cleavage plan with varying temperatures, as well as enzyme amounts was set up with 500 µl (3mg/ml) of the SUMO-CBD-Spacer-LL(C)-37 fusion (Figure 3-13, lanes 10-17). Although half the amount of fusion was cleaved with 50 µl, 100 µl, and 150 µl of enzyme respectively at 4°C overnight, only a very small fraction of the fusion was cleaved and the cleavage efficiency remained very low. There appears to be some spontaneous cleavage in the fusion without the addition of any protease. This may be an indication of degradation. A Western analysis with Anti-LL-37 and Anti-His antibody further indicated that the small band just above the 6 kDa corresponds to the CBD-Spacer-LL(C)-37 peptide, and the very thick band just below 22 kDa corresponds to the SUMO-carrier fusion as expected (Figure not shown).

According to the guidelines we obtained from our collaborator, using 150 µl of enzyme to cleave 1.5 mg of fusion should result in very efficient cleavage. However, the efficiency of SUMO-CBD-Spacer-LL(C)-37 cleavage with SUMO Protease I was quite low. We hypothesized that the cysteine in the third position of LL-37 of our sequence may be forming a disulfide linkage with the cysteine of SUMO Protease I, thereby undermining its activity. Thus, an experiment was set up with 500 µl (3mg/ml) of SUMO-CBD-Spacer-LL(C)-37 fusion with 50 µl of SUMO Protease I, and an increasing amount of DTT concentration up to the maximum limit

allowed for our cleavage (Figure 3-14). However, there was no difference observed in cleavage efficacy.

LL-37 aggregates in solution, and this may hinder the structure and limit the enzyme from accessing the active site for efficient cleavage [20]. Thus, an experiment was set up with 500 µl (3mg/ml) of SUMO-CBD-Spacer-LL(C)-37 fusion with 50 µl of SUMO Protease I, in varying solution conditions (ie. pH, detergent concentration). The experiment was set up in conditions that would discourage peptide aggregation, while maintaining compatible cleavage conditions with SUMO Protease I (Figure 3-15). However, there was no improvement in cleavage efficacy was observed. It should be noted that breaking up the alphahelical peptide aggregation is very difficult. In future, ideas such as introduction of sequence modifications in the design peptide may be considered to reduce peptide aggregation while preserving antimicrobial activity.

To increase the efficiency of the cleavage, an experiment was set up with 1000 µl (3mg/ml) of SUMO-CBD-Spacer-LL(C)-37 fusion and 500 µl of SUMO Protease I, in order to increase the enzyme's chances of reaching the fusion's active site (Figure 3-16). This was initially presumed to be a very large ratio of protease-to-substrate, according to all guidelines. However, the result was surprisingly successful. The cleaved CBD-Spacer-LL(C)-37 construct is very visibly clear standing a little above the 6 kDa marker. The ratio of cleaved product to fusion looks reasonable, and the cleavage seems efficient. Adding substantially more protease seems to have improved our results, and cleavage efficiency. Unfortunately, we ran out of this batch of SUMO Protease I that had been generously provided over the course of the year. We were provided with a new batch of protease that unexpectedly formed a precipitate once thawed and was inactive. We had not anticipated this issue, and unfortunately we were unable to repeat the

Figure 3-14: Ulp1 cleavage of His-tagged SUMO-CBD-Spacer-LL(C)-37 fusion protein with increasing amount of DTT, as followed by Tris-Tricine Gradient (10-20%) analysis. Lane 1, See Blue Ladder; Lane 2, SUMO-CBD-Spacer-LL(C)-37 fusion uncleaved; Lanes 3-8, SUMO-CBD-Spacer-LL(C)-37 cleaved in 0, 1 mM, 2 mM, 5 mM DTT, 10 mM DTT, and 20 mM DTT respectively.

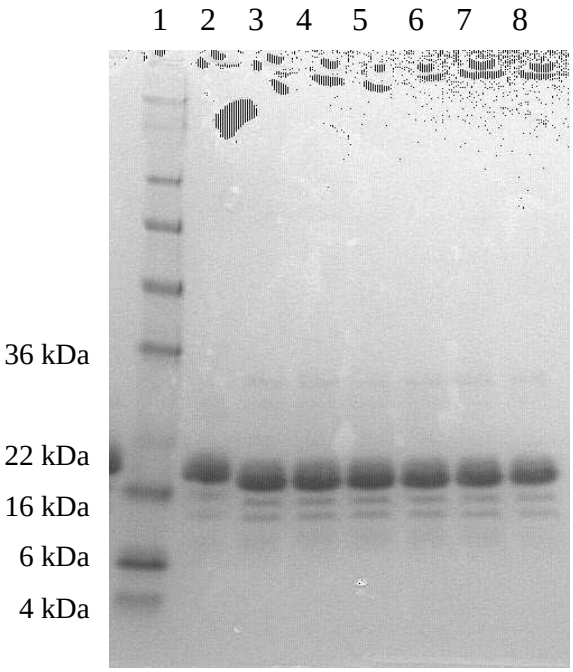


Figure 3-15: Ulp1 cleavage of His-tagged SUMO-CBD-Spacer-LL(C)-37 fusion protein in conditions that discourage peptide aggregation, as followed by Tris-Tricine Gradient (10-20%) analysis. Lane 1, SUMO-CBD-Spacer-LL(C)-37 fusion uncleaved; Lane 2, cleaved in 1 mM DTT; Lane 3, cleaved at pH 5.5 - 1 mM DTT; Lane 4, cleaved at pH 9.5 - 1 mM DTT; Lane 5, cleaved in 2 M Urea - 1 mM DTT; Lane 6, cleaved in 15% Triton-X - 1 mM DTT.

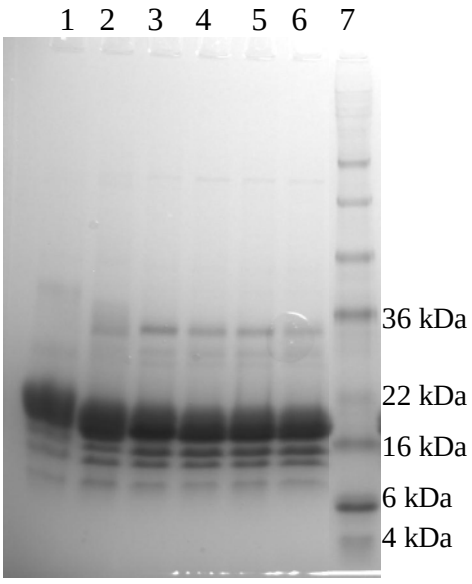
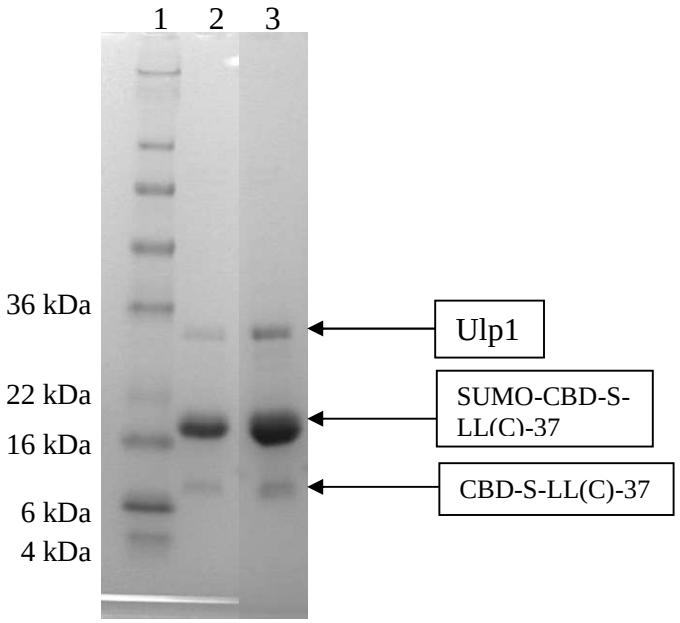


Figure 3-16: Ulp1 cleavage of His-tagged SUMO-CBD-Spacer-LL(C)-37 fusion protein, as followed by Tris-Tricine Gradient (10-20%) analysis. Lane 1, See Blue Ladder; Lanes 2-3, SUMO-CBD-Spacer-LL(C)-37 fusion cleaved at 4°C overnight with 10 µl and 5 µl of sample loaded in each lane respectively.

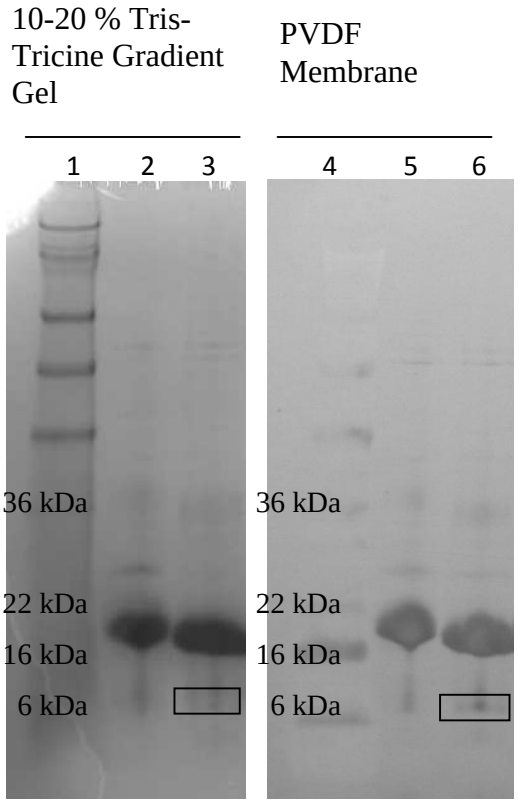


experiment with the affordable enzyme to carry out the rest of our purification. We opted to obtain a small amount of the commercially available SUMO Protease I for further verification and characterization of our previously observed successful cleavage results with the non-commercial protease, as evident based on the Tris-Tricine Gel analysis.

A small amount (200 Units) of the SUMO Protease I was purchased with a cleavage activity of 10-100 µg of substrate per Unit of enzyme. As with previous experience based on our SUMO-CBD-Spacer-LL(C)-37 fusion protein, it was assumed that 1 Unit of enzyme cleaves 10 µg of fusion. A cleavage was set up with 500 µl (3mg/ml) of SUMO-CBD-Spacer-LL(C)-37 fusion with 150 Units of SUMO-Protease I (Figure 3-17). However, as compared to the cleavage observed in Figure 3-16, it appears that the cleavage is very inefficient. The sample was transferred to PVDF membrane, following Tris-Tricine gel analysis.

The band in the enclosed box was thought to correspond to the cleaved CBD-Spacer-LL(C)-37 construct. The enclosed box with a very small amount of Coomassie stained sample was excised from the PVDF membrane and sent for amino acid analysis. However, as the quantity was very small our results were not considered valid and the presumed identity of the sample could not be verified. As the amount of commercial protease that could be obtained was limited, we were not able to obtain enough cleaved material, for further characterization and testing.

Figure 3-17: Cleavage of His-tagged SUMO-CBD-Spacer-LL(C)-37 fusion protein using commercial SUMO Protease I, as followed by Tris-Tricine Gradient (10-20%) and PVDF analysis also stained with Coomassie Blue. Lane 1, See Blue Ladder; Lanes 2, SUMO-CBD-Spacer-LL(C)-37 fusion uncleaved; Lane 3 SUMO-CBD-Spacer-LL(C)-37 fusion cleaved at 4°C overnight. The samples were transferred to PVDF membrane, as seen in lanes 4-6. Lane 4, See Blue Ladder; Lanes 5, SUMO-CBD-Spacer-LL(C)-37 fusion uncleaved; Lane 6 SUMO-CBD-Spacer-LL(C)-37 fusion cleaved at 4°C overnight.



CHAPTER 4

DISCUSSION

Utilizing AMPs with specific activity towards HSV-1 virus may enhance the potential treatments that are currently being investigated for patients with ocular herpes. The incorporation of AMP LL-37 in biomaterial-based corneal substitutes made primarily of collagen may aid as a treatment alternative and be of considerable value in the prevention or suppression of HSV infection as a cause of graft failure.

In Chapter 3, I described a novel design of biofunctional LL-37-containing sequences with the potential ability to form chemical or physical associations with a Collagen scaffold material. Several fusion systems were explored for the development and biosynthesis of the designed sequences, and it was determined that the thioredoxin fusion expression system is best suited for facile expression and purification of the modified LL-37 with a cysteine substitution allowing for possible chemical attachment to collagen substrates through disulfide formation. The recombinant purified modified LL-37 was shown to inhibit HSV-1 reactivation in HCEC cells, with comparable effective correlation in comparison to the drug Acyclovir. The SUMO fusion expression system was determined as best suited candidate for expression and purification of the CBD-spacer- LL-37 construct. Results indicate that these novel sequences may be effectively produced in bacteria and retain their potential activity. However, these selected expression systems need to be further investigated. A critical important consideration in successful biosynthesis and purification of AMPs is a well chosen expression system and purification strategy.

Aside from the toxicity issue and other technical challenges, the unique structural characteristics of AMPs also raise several challenges and add another obstacle to recombinant production. The lack of a carrier specifically designed for peptides may be a limiting factor for successful fusion expression [20]. The current systems available, such as most of the ones used in this project, are all for the general purpose of recombinant protein production, with the exception of tandem repeats. Even though enzymatic cleavage is usually efficient at releasing target protein from its fusion carrier, in the case of AMPs it is found to be less efficient than chemical cleavage. For example, when the human antimicrobial peptide LL-37 is expressed as thioredoxin and GST fusions, it cannot be efficiently released by thrombin or factor Xa cleavage [26]. A possible reason for this is that certain AMPs such as LL-37 tend to form oligomers, which block the cleavage site. [20]

The intein system that requires no auxiliary enzyme or chemicals reagents for carrier removal, offers a great advantage to target protein purification. However, its uninduced self-cleavage activity limits its value for antimicrobial peptide production as premature cleavage of the fusion protein is an intrinsic problem associated with the intein system, which causes loss in the yield of target protein. [20, 27] It is upon induction with thiol reagents or pH and temperature shift, that the target protein is released via intein mediated self-cleavage. In order for the intein-mediated cleavage to be effective, the fusion protein should be in soluble, correctly folded form. Growth temperature optimization is sometimes needed to get soluble expression. However, it has been noted that even at lower protein expression temperatures, the average yield of non-antimicrobial peptides is still much higher than that of antimicrobial peptides. The lower yield of antimicrobial peptides, such as LL37, could be due to uncontrolled autocleavage, which releases a small amount of peptide that negatively influences the cell's protein-making machinery. For

proteins other than antimicrobial peptides, the loss due to *in vivo* cleavage can probably be offset by the gain from the relatively simple purification of this approach as long as the uncontrolled cleavage is kept to a low level. However, for AMPs this could be a serious problem, since even trace amounts of released peptide may be fatal to the host. [20, 27] The very low expression levels observed in the pTYB expression systems, whether C-terminally fused, or N-terminally fused corresponds with this reasoning. In conclusion this fusion system was identified as a poor choice for LL-37 containing peptide expression and purification.

Most commonly used carrier proteins such as thioredoxin, intein, and SUMO promote proper folding and enhance solubility of the fused peptides. However, certain carriers such as the ELP tag system are specially designed to promote inclusion-body formation. For short peptides without disulfide linkages, refolding is generally not required to restore activity, and insoluble expression may have the advantage of facile purification. There is evidence to suggest that increasing salt concentrations causes the peptide to form increasing amounts of secondary structures, to favour oligomeric assemblies where hydrophobic faces are hidden, and eventually precipitate. In addition, LL-37 upon oligomerizing is likely to precipitate out of solution and thus be significantly less present in the supernatant fraction. However, since protein expression levels in tandem design are usually not proportional to the degree of multimerziation, it does not necessarily assure an improved peptide yield [20, 28, 29]. In addition to low expression levels, this method of purification did not produce a single identifiable band for LL-37. Even though a few numbers of modifications were done, such as changing the amount of salt concentration in buffers, and altering the time and temperature of the heat cycles, the same multiple banding pattern were observed in the final purification step. This method of purification was clearly not ideal for purifying a LL-37 containing peptide. Since many antimicrobial peptides have been

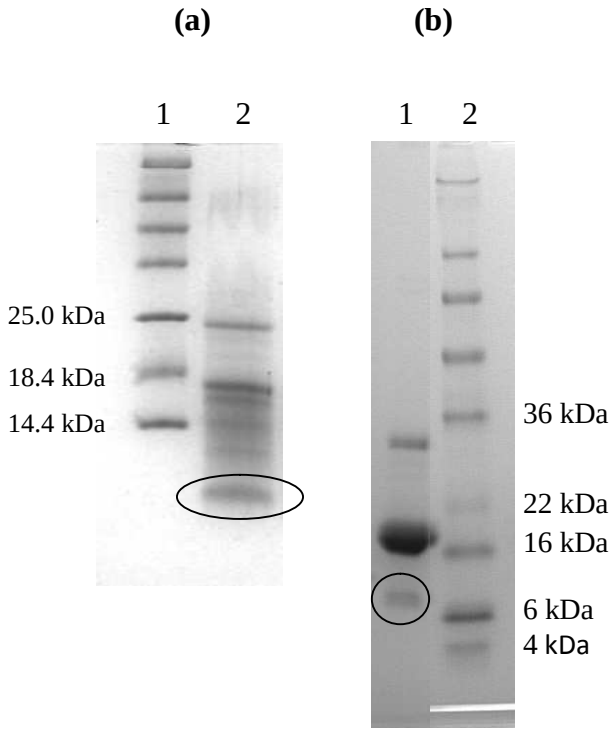
successfully expressed as soluble fusions, forcing fusion proteins into inclusion bodies is not essential in protecting the bacterial host from the toxicity of AMPs [20].

To increase the efficiency of releasing the modified LL-37 from the thioredoxin fusion carrier, we designed a chemical cleavage site in the pET-232 series vector to accommodate a formic acid cleavage site. Although the purified recombinant modified LL37 produced in the thioredoxin system showed a concentration dependent protection against HSV-1 virus, using formic acid cleavage was rendered unreliable for the possible side chain modification risks it imposed in the production of the biofunctional peptide containing the Collagen binding functional domain. It is important to have a system in which we can produce intact designed peptides, for fully reproducible and consistent data that includes correct mass spec data. The advantage of using thioredoxin fusion tag is the better level of expression of the modified AMP LL-37, and the disadvantage was that following cleavage a Proline residue is left at the N-terminus of the targeted peptide.

The SUMO fusion system, like thioredoxin, promotes proper folding and soluble expression of the fused proteins. Since SUMO has a hydrophobic core and a hydrophilic surface, it is highly soluble. In addition, due to the small size of the SUMO protein, there is a higher stoichiometric mass ratio of the target protein. Since the size of AMPs is usually small, will help increase the net yield. Furthermore, SUMO protease is highly efficient and specific. The mechanism of cleavage also allows for the production of target proteins with native N-terminus when the fusion is properly designed. Due to such characteristics, the SUMO fusion system offers an attractive alternative for recombinant production of antimicrobial peptides, although application of the SUMO fusion strategy to antimicrobial peptides has not been widely reported. When comparing the formic acid cleavage and SUMO protease cleavage of their corresponding

fusions, and the release of the target LL-37 containing peptides, it is clear that the protease fusion is a much more elegant cleavage and better suited for AMP sequences that contain functional domain (Figure 4-1). The challenges we experienced in terms of cleavage efficiency and peptide yield may certainly be attributed to the steric hindrance imposed by LL-37 oligomerization, and provide an explanation as to how an increase of SUMO protease facilitated this. Although this may explain and support the inefficiency of the cleavage, further optimization and alternative methods need to be considered to improve cleavage efficiency and peptide release in order to ultimately enhance recombinant expression of designed peptides containing AMP sequences.

Figure 4-1: Comparison of pLET1 formic acid cleavage vs. SUMO-CBD-S-LL37 Ulp1 enzymatic cleavage, as followed by gel analysis. (a) Lane 1, See Blue Ladder; Lane 2, P-LL-37 (circled) was cleaved off the thioredoxin fusion, following a 50hr Formic acid (50%) cleavage. **(b)** Lane 1, CBD-Spacer-LL(C)-37 (circled) was cleaved off the SUMO fusion an overnight Ulp1 enzymatic cleavage at 4°C; Lane 2, See Blue Ladder. Cleavage with Ulp1 does not produce a smear.



APPENDICES

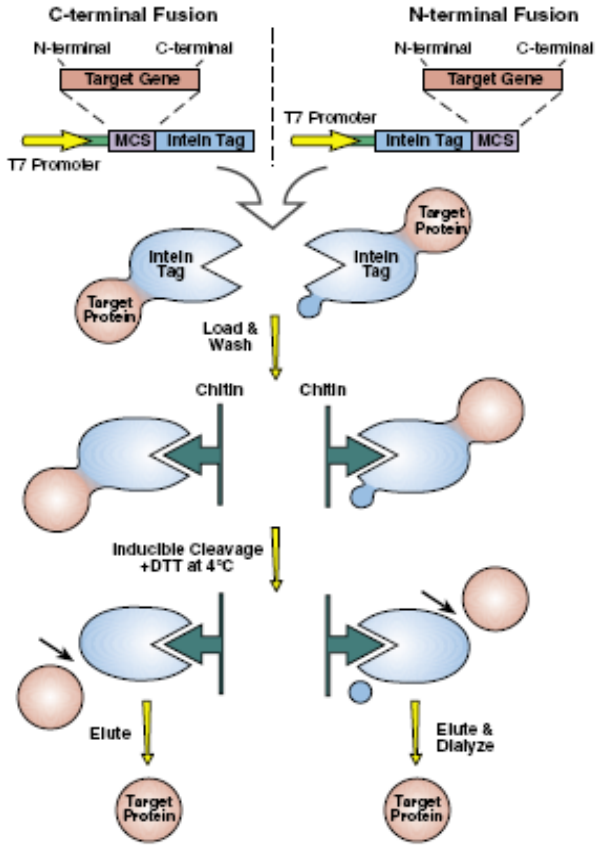
Appendix A

1.1 Expression and Purification of Modified LL-37 using the IMPACT purification system.

The anti-microbial peptide LL(C)-37 was cloned into the pTYB2 (C-terminal fusion) and pTYB12 (N-terminal fusion) expression vectors (Supplementary Fig 1). *Escherichia coli* strain BL21 harbouring the pTYB2 or pTYB12 vectors, was utilized to express the intein fusion protein containing LL(C)-37. The expression of the full chimera was induced in the presence of IPTG. Lower induction temperatures (ie. 15°C) were found necessary to improve the intein-tag partnered to hard-to-express peptides expression yields. Following fusion expression, the cell lysate was then purified over the IMPACT protein purification system, and the protein samples were analyzed by SDS-PAGE (15%) analysis (Supplementary Fig.2, A and Fig. 2, B).

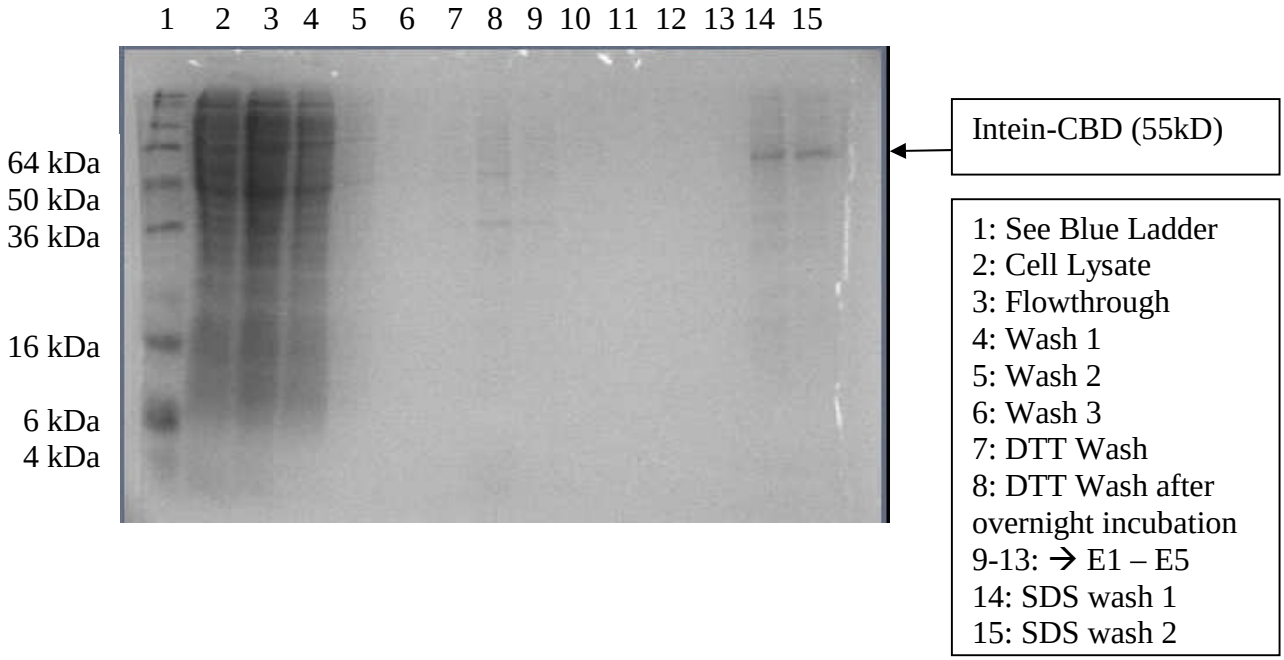
The amount of the expressed fusion protein as judged from Coomassie Blue staining was relatively low in this system, when the induced cell lysate was compared to the uninduced cell lysate (figure not shown). Most of the fusion protein was found in the soluble supernatant of the cell lysate. The LL(C)-37 containing fusion protein was purified from the supernatant of cell lysates using a chitin bead column (New England Bio Labs) equilibrated with binding buffer. The 4.5 kDa LL(C)-37 was expected to be eluted in the DTT washes following self cleavage. However, in both constructs, no lower band indicative of the peptide was visible (Supplementary Fig.2, A and Fig. 2, B). The Intein-CBD was to be eluted in the SDS washes at 55kDa. However, since the cleaved or uncleaved chimera proteins were relatively close in molecular weight, it was not possible to conclude whether or not the full fusion protein containing the LL(C)-37 peptide had been cleaved (Supplementary Fig.2, A and Fig. 2, B). It may be an issue of visualization of

Supplemental Figure 1: A schematic illustration of the IMPACT System.

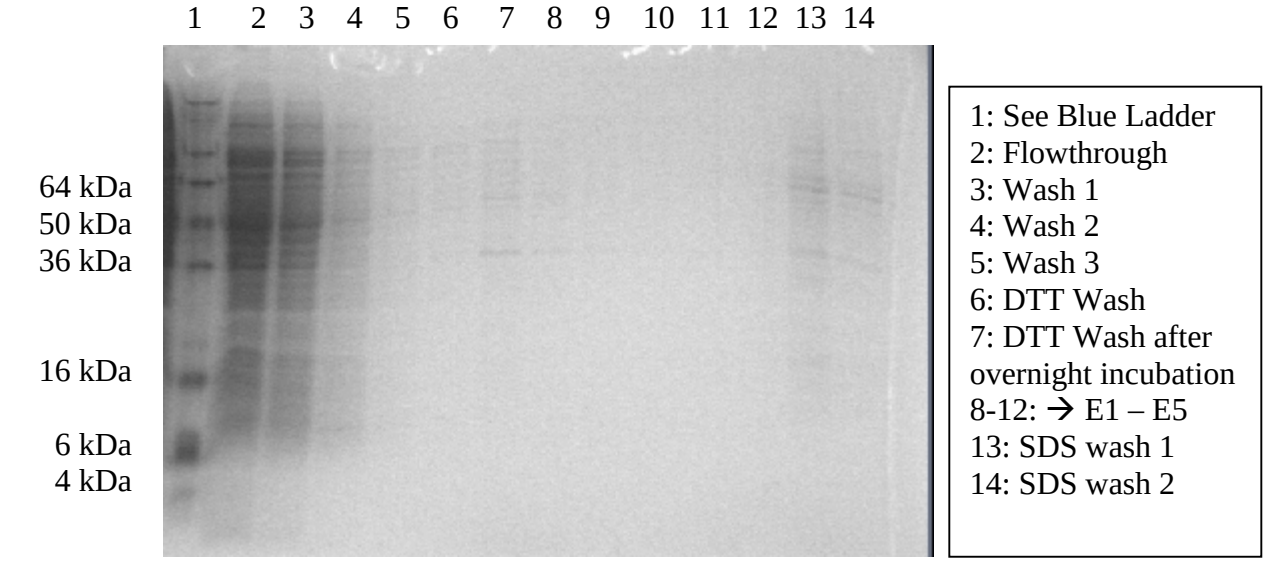


Reference: Figure obtained from New England BioLabs Instruction Manual. Catalog # #E6900S [24].

Supplemental Figure 2, A: Purification of LL(C)-37 on a chitin bead column (10 mL), using the pTYB2 (C-terminal fusion) expression system, followed by SDS-PAGE (15%) analysis.



Supplemental Figure 2, B: Purification of LL(C)-37 on a chitin bead column (10 mL), using the pTYB12 (N-terminal fusion) expression system, followed by SDS-PAGE (15%) analysis.



the small cationic peptide on SDS-page due to low expression levels that leads to low amounts of peptide, or perhaps the tendency of the small cationic peptide to interact with the sides of the column and get easily lost during the purification process.

In order to better visualize cleaved vs. uncleaved chimera, and distinguish between the two possibilities for the bands observed in the SDS washes, molecular biology was used to add a His-tag to the LL(C)-37 in the PTYB2 (C-terminal fusion) expression vector. The full chimera was expressed and purified similarly (Figure not shown). However, no bands corresponding to the His-tagged LL(C)-37 (expected at around 7kDa), following thiol treatment, were visible in the elution fractions. Additionally by means of doing a Western, using anti-His antibody, we were not able to distinguish between cleaved vs. uncleaved intein-carrying chimera as we had hoped. The bands in the SDS washes (following elution that should contain the cleaved intein tag without any His-tag) were observed in the Western, which would suggest that the full chimera was not cleaved. However, as other bands not specific for our full chimera, were also observed non-specifically, it was merely inconclusive due to the non-specific binding of the antibody.

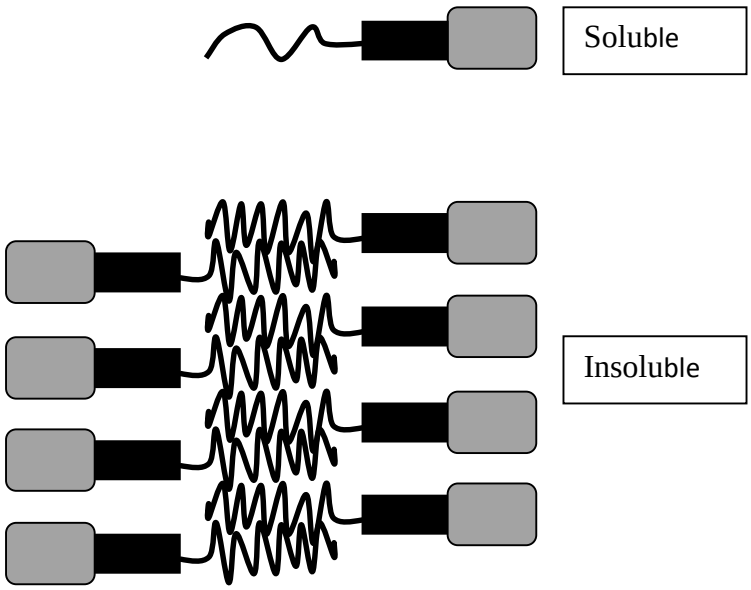
1.2 Expression and Bioseparation of modified LL(C)-37 using self-cleaving elastin-like polypeptide (ELP) tag system

His-tagged LL(C)-37 was fused to the gene encoding the self-cleaving ELP-intein tag, by cloning it into an elastomer vector (pEI). *Escherichia coli* strain BL21 harboring this vector was utilized to express this fusion protein containing LL(C)-37. The expression of the full chimera was induced in the presence of IPTG. Lower induction temperatures (ie. 15°C) were also necessary to improve the intein-tag partnered to hard-to-express peptides expression yields. The expression level in this system was relatively low, as was the case with the IMPACT expression system. In an attempt to overcome that problem we preceded with a bioreactor run, where

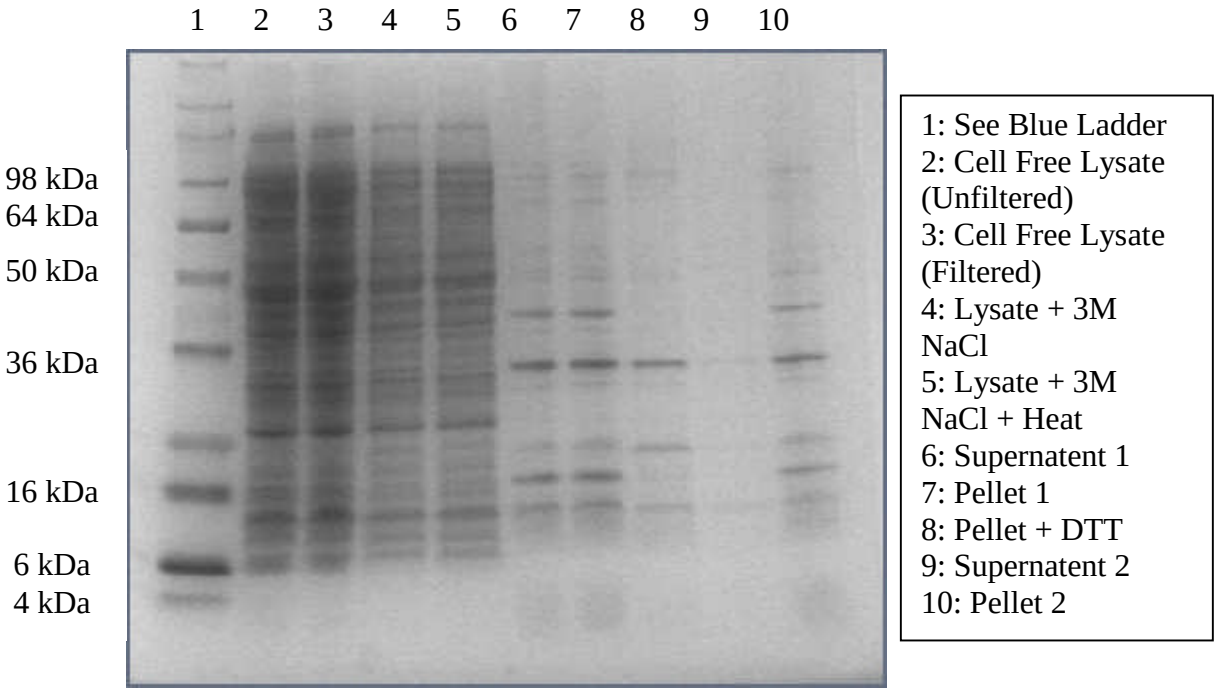
expression levels are typically much higher compared to a shake flask. This approach also has the advantage of producing more bacteria by weight, when the volume of starting media is held constant. To express the full chimera, a 5L medium containing 500 µg/ml Amp was inoculated with 100ml freshly grown culture. The culture was incubated in an air shaker at 37°C until the OD_{600} reaches 10, and IPTG induced at 15˚C overnight. The expression level following the larger scale bioreactor run was improved, when the induced cell lysate was compared to uninduced cell lysate (Figure not shown). Over 400g of pellet was harvested in total. One pellet of 36g was used for this experiment. An equal volume of 3M NaCl was added to the cell lysate. The resulting sample was heated at 30°C for 10 min, and centrifuged for 10 min. The pellet was dissolved in 5 ml of pH 6.0 buffer, containing 50mM DTT, incubated at room temperature over-night. The next day, an equal volume of 3M NaCl was added to the samples, and the heat and centrifuge cycle was repeated.

It was expected that only the His-tagged LL(C)-37 would be soluble and thus be present in the supernatant, while the ELP-intein tag would be in the insoluble precipitated pellet (Supplementary Fig.3). The protein samples were analyzed by SDS-PAGE (15%) analysis (Supplementary Fig.4). The ELP tag has an expected molecular weight of approximately 66kDa, the His-tagged LL(C)-37 is has an expected molecular weight of approximately 7kDa, and the ELP fusion has an expected molecular weight of approximately 72kDa.

Supplemental Figure 3: A schematic illustration of the ELP-intein tag System.



Supplemental Figure 4: Purification of LL(C)-37 on a chitin bead column (10 mL), using the ELP-intein tag system, from a 5L starting culture, followed by SDS-PAGE (15%) analysis.



Unexpectedly, following purification steps there are still many bands appearing in each lane of the late stages of the purification process (Supplementary Fig.4, lanes 6-10). It is rather difficult to distinguish, which band if any is the His-tagged LL(C)-37 that we expected to see in the second supernatant (Supplementary Fig.4, lane 9). There is no band at the expected molecular weight of 7 kDa visible in lane 9, directly corresponding to His-tagged LL(C)-37. It is plausible that His-tagged LL(C)-37 may have precipitated out due to self-association into an insoluble precipitate after being heated in the presence of salt or perhaps formed higher oligomers due to high salt concentration. Since LL-37 has a net charge of +6 at physiological pH, it is possible for the peptide to run somewhat differently, but a difference of this magnitude is not expected.

Furthermore, a Western was done using Anti-His antibody, to distinguish LL(C)-37 containing band (Figure not shown). However, multiple bands at different molecular weights were observed, which suggests either or a combination of the following: degradation of the fusion, oligomerization of cleaved His-tagged LL(C)-37, or unspecific-binding of antibody. One clear band that always appears in our purification is the bands at 36 kDa, which might indicate the LL(C)-37 is oligomerizing, or perhaps degradation of the full fusion.

As it is common for anti-His antibody to give false positive binding, a Western was also done using anti-LL37-Antibody (Figure not shown). However, still the problem of observing multiple unspecific bands within the purification lanes persisted, that interestingly corresponded with the anti-His-Antibody Western bands.

It was evident that although expression level was within reasonable range in larger scale expression, this purification method was not successful in producing an analyzable single band peptide.

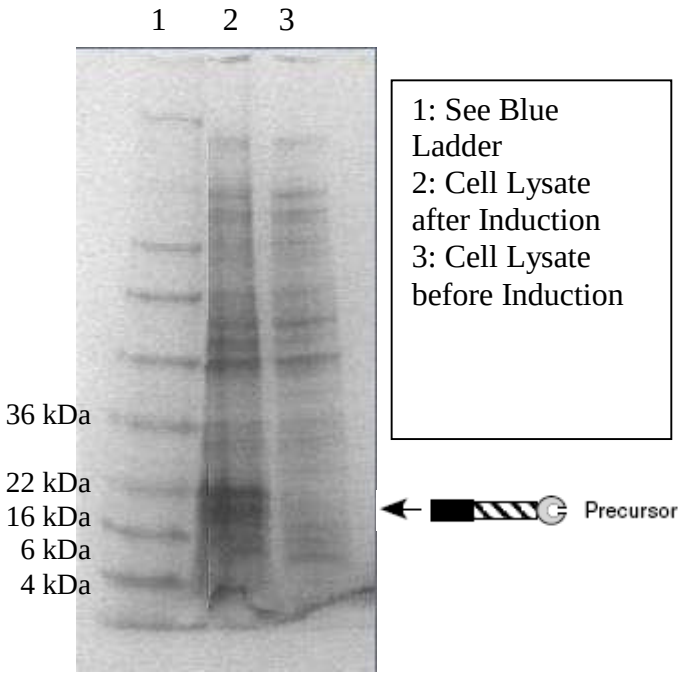
Additionally, we tried to purify the fusion using several other separation methods, such as using a nickel column, which has an affinity to bind the His tag. The 6xHis tag should interact with the Ni-NTA resin. After the wash step, the His-tagged LL(C)-37 is expected to lose affinity and come off the column with a two step elution in buffers with a low pH of 5.9 and 4.5. The protein samples were analyzed on SDS-PAGE (15%) (Figure not shown). However, although they were fewer bands in the elution fractions from the Ni-NTA column compared to previous observations, there were still multiple bands observed, and we were not able to distinguish any clear band corresponding to the LL(C)-37 molecular weight. Westerns using Anti-His Antibody, and Anti-LL37 Antibody were also performed, but unfortunately the same problem of having multiple unspecific bands visible in each lane was observed. It could thus be concluded that non-specific cleavage and degradation were perhaps reasons for multiple bands, and the intended purification steps of this system did not result in a purified analyzable peptide, and specifically one that corresponded with His-tagged LL(C)-37.

1.3 Expression and purification of modified LL-37 using SUMO System

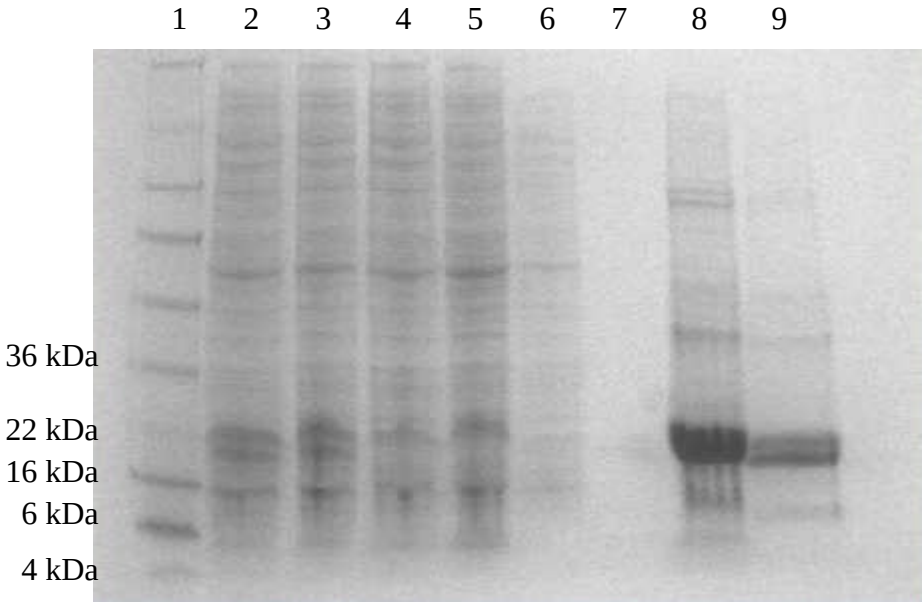
The (CSG)-LL-37 sequence was fused seamlessly to the SUMO fusion via a two step PCR, and ligated into the pET-based vector. *Escherichia coli* strain BL21 harboring the SUMO-(CSG)-LL-37 expression vector, was utilized to express the SUMO fusion protein containing (CSG)-LL-37. The expression of the full chimera was induced in the presence of IPTG. The SUMO-(CSG)-LL-37 fusion was expressed with good expression levels as judged from Coomassie Blue staining (Supplementary Fig.5). The calculated molecular weight of the SUMO-His tag is approximately 13 kDa, and the molecular weight of (CSG)-LL-37 is approximately 5 kDa. The expected weight of the expressed fusion protein SUMO-(CSG)-LL-37 is approximately 18 kDa. Following induction, an 18 kDa band corresponding to SUMO-(CSG)-LL-37 chimera was visible, following SDS-Page (15%) analysis (Supplementary Fig.5, lane 2). The band is visibly absent in the control that is cell lysate mixture right before induction (Supplementary Fig.5, lane 3).

The SUMO-(CSG)-LL-37 bacteria pellet was resuspended in native purification buffer, containing low amounts of imidazole, PMSF, and lysis buffer. After centrifugation and filtration, the fusion protein was found primarily in the supernatant of whole cell lysate (Supplementary Fig.6, lane 2-3). The SUMO-(CSG)-LL-37 fusion protein was purified from the supernatant of cell lysate using a Ni-NTA nickel column equilibrated with binding buffer. After binding, the Ni-NTA nickel column was washed with slightly higher concentration of imidazole to reduce non-specific binding, and finally eluted with a very high concentration of imidazole. It can be observed that the majority of the unwanted proteins were removed by Ni-NTA affinity chromatography (Supplementary Fig.6, lanes 4-7), and most of the fusion protein is eluted in the first collected fraction (Supplementary Fig.6, lane 8).

Supplemental Figure 5: Expression of SUMO-(CSG)-LL-37 in *E. coli*. Expression of SUMO-(CSG)-LL-37 fusion protein at 18kDa (indicated by an arrow) induced with IPTG for 4 hours, at 37°C, as followed by SDS-PAGE (15%) analysis. Lane 1, See Blue Ladder; Lane 2, cell free extract after induction; Lane 3, cell free extract before induction.



Supplemental Figure 6: Purification of His-tagged SUMO-(CSG)-LL-37 fusion protein on IMAC (10 mL), as followed by SDS-PAGE (15%) analysis. Lane 1, See Blue Ladder; Lane 2, cell free extract; Lane 3, cell free extract filtered; Lanes 4-5, column flow through (50 mL each); Lanes 6-7, column wash with 20mM imidazole (40 mL each); Lanes 8-9, protein fractions eluted with 250mM imidazole from Qiagen Ni-NTA resin.

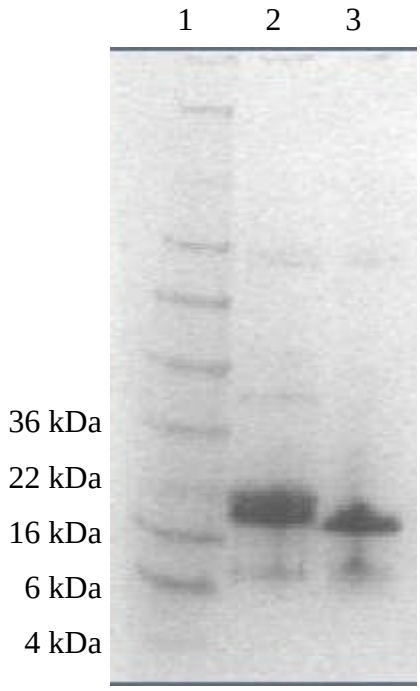


As the eluted fusion protein is in buffer containing 250 mM imidazole, there is a need to exchange the buffer for efficient SUMO Protease 1 cleavage, since imidazole concentrations higher than 150 mM can adversely affect the activity of the protease. The isolated SUMO-(CSG)-LL-37 fusion protein was subsequently dialyzed using a 500 Da Molecular Weight Cut Off (MWCO) dialysis tube, to a compatible SUMO Protease I cleavage buffer (20 mM Tris-HCl, 150 mM NaCl, 1mM DTT, pH 8.0) prior to SUMO Protease I cleavage.

To have a cost effective system, we worked in collaborated with a group who generously produced us with a non-commercial enzyme. However, we were not provided with the quantified specific protease activity corresponding to our batch, and the SUMO Protease I activity ranges greatly based on its preparation methodology and condition. To overcome this problem, we followed our collaborators guideline as a starting point. However, it became clear that we required some deviations and modifications to establish an efficient cleavage reaction according to our specific fusion characteristics.

As a precaution to avoid any change in our fusion of interest, the cleavage was carried out at 4°C overnight. Following the enzymatic cleavage and analyzing the sample on 15% SDS-PAGE, it appears that the fusion band is slightly shifted lower (Supplementary Fig.7, lanes 1-2). However, this shift does not necessarily indicate a cleavage according to its molecular weight apparent on the SDS-PAGE, and could merely indicate the result of a buffer exchange, as the released (CSG)-LL-37 is not visualized on the SDS-PAGE stained with Coomassie Blue (Supplementary Fig.7). As we had previously experienced difficulty visualizing the release of the small LL-37 peptide, and a 5 kDa may only produce a small shift in the SUMO-fusion's running on SDS-PAGE, we need to verify if the cleavage was successful. Since the SUMO Protease I was a generous offer, we were not well equipped with setting up a positive control.

Supplemental Figure 7: Cleavage of His-tagged SUMO-(CSG)-LL-37 fusion with SUMO Protease I, as followed by SDS-PAGE (15%) analysis. Lane 1, See Blue Ladder; Lane 2, SUMO-(CSG)-LL-37elute uncleaved; Lane 3, SUMO-(CSG)-LL-37 elute cleaved;



Without a positive control or a quantified protease activity, we could only continue to further explore this system, as these tools are essential in order to continue the project.

To further analyze the SUMO expression system we tried to redo the experiment and quantify the amount of fusion protein we are cleaving, rather than estimate based on SDS-PAGE analysis. The fusion was over-expressed as before, and 20g of the bacteria pellet containing the expressed fusion was purified accordingly. The results were a replication of what we had obtained in our first purification. The eluted fusion fractions were pooled into 150 mL of cleavage compatible buffer. Based on the UV absorbance reading in accordance to the amino acid sequence of SUMO-(CSG)-LL-37, we were able to quantify that we obtained approximately 800mg of SUMO-(CSG)-LL-37 fusion protein in our elute. Following cleavage, of course with the assumption of complete cleavage, based on the (CSG)-LL-37 to SUMO-(CSG)-LL-37 sequence ratio, we should consequently have approximately 200 mg of (CSG)-LL-37 in 150 ml of cleaved protein mixture, which is 1.25mg/ml of pure (CSG)-LL-37. A cleavage reaction was set up again with multiple aliquots of the sample, ranging from 10-100 mLs, with the addition of 50 µl of SUMO Protease I for each reaction. However, the same result was experienced and no band corresponding with the cleaved (CSG)-LL-37 was observed. Although at an approximate concentration of 1.25mg/ml or perhaps even 1/10 of that amount, the cleaved (CSG)-LL-37 is expected to be visualized on SDS-PAGE analysis. This is a clear indication of incomplete and insufficient cleavage activity.

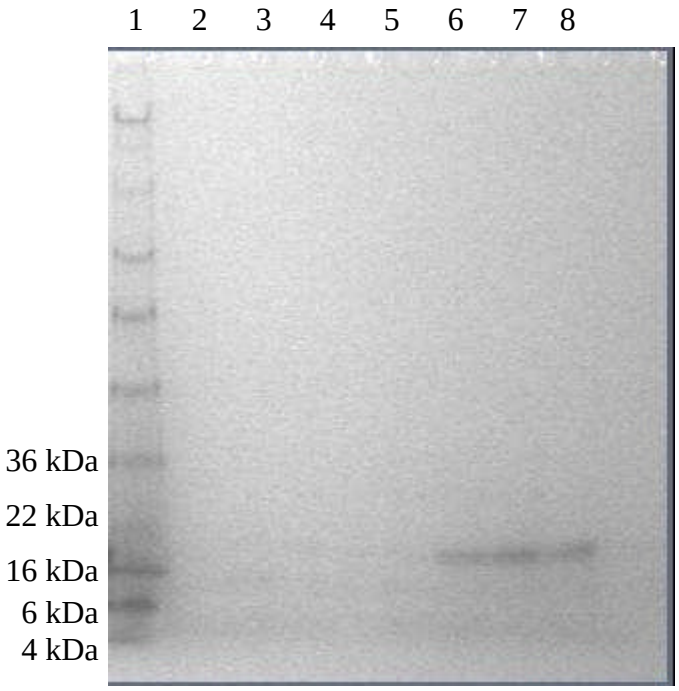
In accordance with the purification design following completed fusion cleavage, the His-tagged SUMO carrier and SUMO Protease I, as well as residual undigested fusion proteins were then to be removed by an IMAC column, carried out in denaturing conditions with a high Urea concentration. It was to be expected that the flow-through fraction would contain the cleaved and

non-His-tagged (CSG)-LL-37, and the elute fractions were to contain the His-tagged SUMO fusion as well as the SUMO Protease I and any undigested fusion protein. We carried out with this purifications step to further explore our system to pinpoint the problem, as the SDS-PAGE analysis results of the previous step indicated that we might have encountered a problem that we had not anticipated. We run our sample through an IMAC column, followed by an SDS-PAGE analysis of the collected fractions. It was depicted that there are no (CSG)-LL-37 visible in the flow through fractions as desired as perhaps due to insufficient fusion cleavage. The elute fractions do contain a band that corresponds to approximately 18 kDa (Supplementary Fig.8).

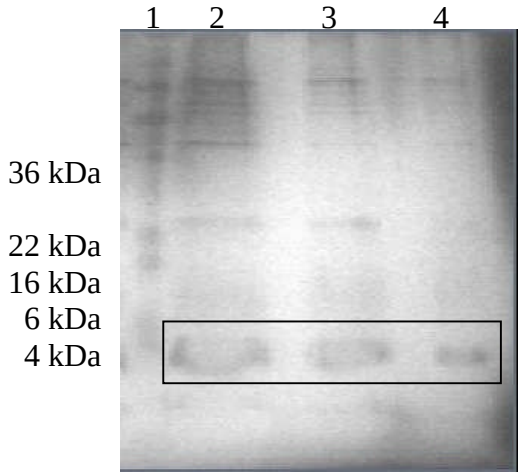
The flow through fraction that was to contain the cleaved (CSG)-LL-37 following the completion of SUMO Protease I cleavage, was dialyzed using 500 Da MWCO Dialysis tube in dH_2O, to eliminate Urea and any other trace salts. It was then lyophilized and a white powder was obtained. The dried white powder was re-suspended in small amount of dH2O at a very high concentration (70mg in 300µg dH2O), and the sample was run on SDS-PAGE (Supplementary Fig.9).

Peptide bands at a low molecular weight corresponding with the expected weight of (CSG)-LL-37 were observed with Silver Stain (Supplementary Fig.9). However, submission of 20 mg, and 150 mg of the sample for Maldi and Amino Acid Analysis respectively were inconclusive (date not shown). As Urea was prominent in the purification procedure, a Urea and BCA Assay was performed to verify that the sample was in fact protein and not Urea salt before sample submission. Although it was confirmed that this did in fact contain amino acids, the small single (CSG)-LL-37 peptide was not identified. This might indicate that the sample we obtained was degraded fusion fragments.

Supplemental Figure 8: Purification of His-tagged SUMO-(CSG)-LL-37 fusion on IMAC, as followed by SDS-PAGE (15%) analysis. Lane 1, See Blue Ladder; Lanes 2-5, Flow through lanes empty although expected to contain cleaved (CSG)-LL-37; Lanes 6-8, wash and elute containing SUMO fusion (with 4M Urea from Qiagen Ni-NTA resin).



Supplemental Figure 9: Silver staining of (CSG)-LL-37. Lane 1, See Blue Ladder; Lane 2, 10µl of (CSG)-LL-37; Lane 3, 5µl of (CSG)-LL-37; Lane 4, 3µl of (CSG)-LL-37.



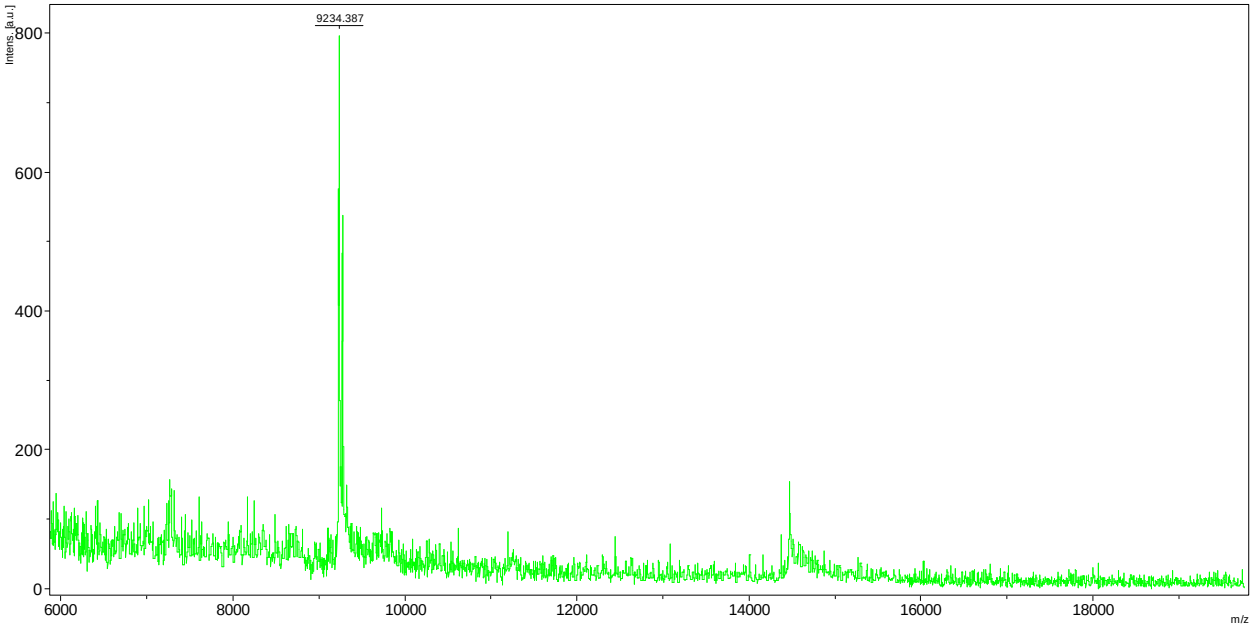
Appendix B

1.4 Characterization and testing of expressed and purified 6xHis-CBD-Spacer-LL37 using a pET-based expression system.

The DNA sequence encoding the 6xHis-CBD-Spacer-LL(C)-37 was ligated in a pET21d vector and expressed in *E. coli*. The protein was easily observed to be expressed following induction, and was purified using HPLC. The MALDI is approximately 400 Daltons higher than the expected weight (Supplementary Fig. 10).This could be due to “signal” or “flag” sequences, in which amino acids that should not have been expressed or that should have been removed afterwards were expressed or not removed following expression. It could also be that the protein has a contaminant attached, which is increasing the expected weight molecular weight. However, as not lower than the expected weight, it is thought that our protein of interest was intact. The sample was subsequently tested for collagen binding activity using SPR. The red channel was coated with EDC for 6.5min and then collagen for 40min (Supplementary Fig. 11a). Solutions were prepared by Dr. Li, and the rest of the experimental work was carried out by Ms. Wan. Approximately 0.12ug/cm^2 of collagen was coated onto the red channel of the chip by converting the unit. 0.05mg/ml of 6xHis-CBD-Spacer-LL(C)-37 was passed through collagen-coated chip (Supplementary Fig. 11b). In order to compare the difference of 2 channels when LL-37 collagen-binding-2 passed through the collagen-coated chip, the response was zeroed. The difference of response between red and blue channels after LL37 collagen-binding-2 passed through the sensor is around 100RU. Since red channel was coated with collagen and blue channel was not, this result shows that the CBD of the LL-37 construct binds to collagen. However, this is preliminary and further work is required to further prove that this is not merely non-specific binding. In addition, further work is required

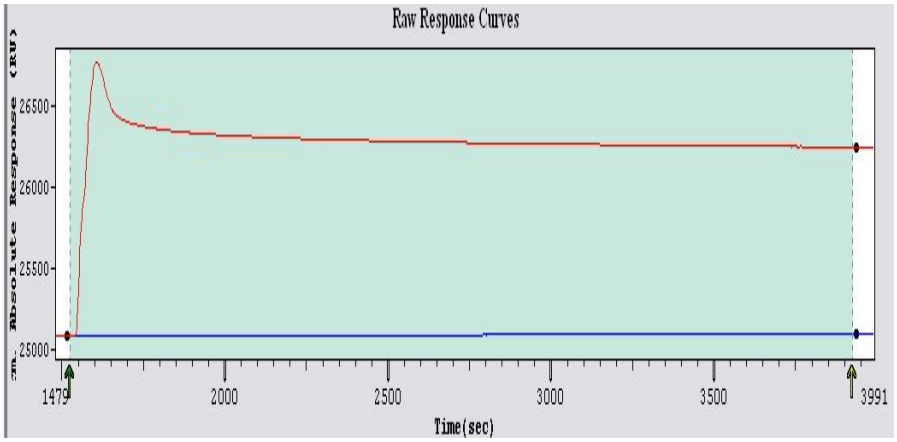
to demonstrate that the expressed and purified peptide obtained as a single peak shown in MALDI analysis, is in fact the desired peptide containing the correct sequence. The MALDI discrepancy should be further analyzed by amino acid analysis, to show it contains the selected amino acid sequence. Only then can the additional weight be attributed to other plausible factors according to the collective results, and discussed upon. Swapping longer CBD and spacer sequences may be incorporated in the future design as they might enhance and better the chances of expression and purification.

Supplemental Figure 10: Mass spectrum of the purified 6xHis-CBD-Spacer-LL(C)-37 on MALDI-TOF mass spectrometry. The molecular weight for the recombinant xHis-CBD-Spacer-LL(C)-37 was 9234, calculated mass is 8882. Dr. Dick, S.

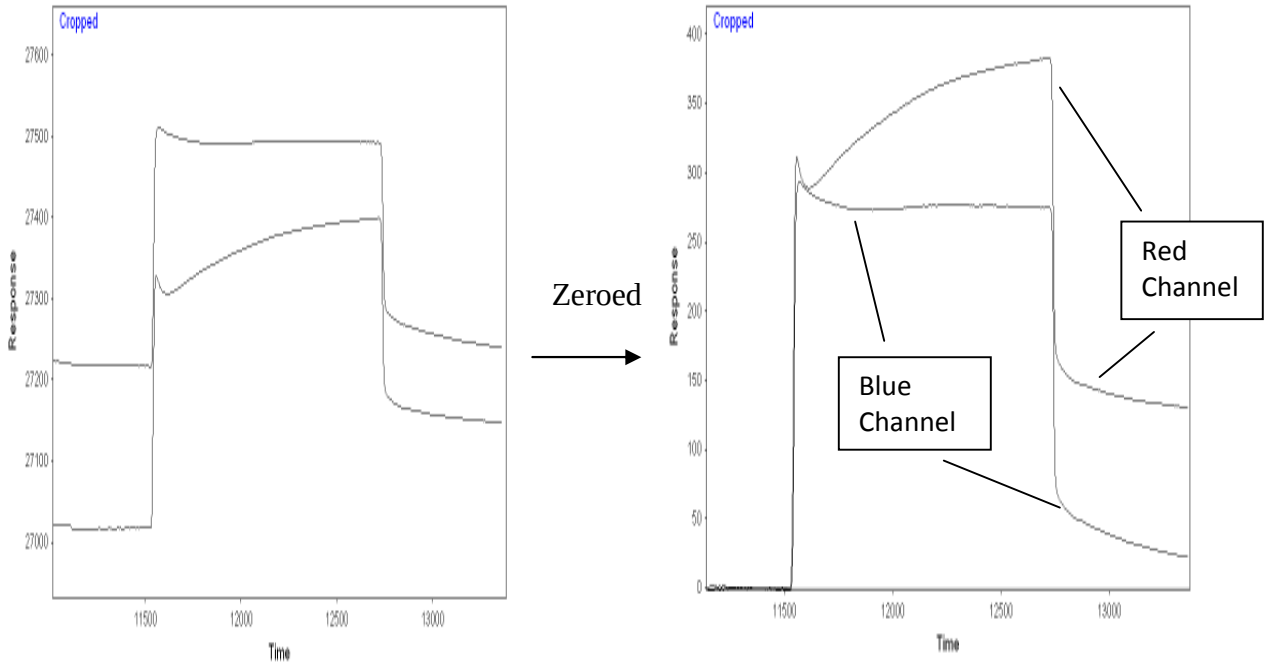


Supplemental Figure 11: Collagen binding testing of 6xHis-CBD-Spacer-LL(C)-37 using SPR. Wan, F.

(a) Coating COOH-1 chip inside of SPR instrument by injection with porcine type I collagen



(b) 0.05mg/ml of 6xHis-CBD-Spacer-LL(C)-37 passed through collagen-coated chip



References

1. Hicks, C.R., et al., *Outcomes of implantation of an artificial cornea, AlphaCor: effects of prior ocular herpes simplex infection.* Cornea, 2002. **21**(7): p. 685-90.
2. Izadpanah, A. and R.L. Gallo, *Antimicrobial peptides.* J Am Acad Dermatol, 2005. **52**(3 Pt 1): p. 381-90; quiz 391-2.
3. Gordon, Y.J., et al., *Human cathelicidin (LL-37), a multifunctional peptide, is expressed by ocular surface epithelia and has potent antibacterial and antiviral activity.* Curr Eye Res, 2005. **30**(5): p. 385-94.
4. Griffith, M., et al., *Artificial human corneas: scaffolds for transplantation and host regeneration.* Cornea, 2002. **21**(7 Suppl): p. S54-61.
5. Knop, E. and N. Knop, *Anatomy and immunology of the ocular surface.* Chem Immunol Allergy, 2007. **92**: p. 36-49.
6. Michelacci, Y.M., *Collagens and proteoglycans of the corneal extracellular matrix.* Braz J Med Biol Res, 2003. **36**(8): p. 1037-46.
7. Garreis, F., M. Gottschalt, and F.P. Paulsen, *Antimicrobial peptides as a major part of the innate immune defense at the ocular surface.* Dev Ophthalmol. **45**: p. 16-22.
8. Skerlavaj, B., et al., *Biological characterization of two novel cathelicidin-derived peptides and identification of structural requirements for their antimicrobial and cell lytic activities.* J Biol Chem, 1996. **271**(45): p. 28375-81.
9. Gallo, R.L., et al., *Identification of CRAMP, a cathelin-related antimicrobial peptide expressed in the embryonic and adult mouse.* J Biol Chem, 1997. **272**(20): p. 13088-93.
10. Shai, Y., *Mode of action of membrane active antimicrobial peptides.* Biopolymers, 2002. **66**(4): p. 236-48.
11. McDermott, A.M., *Defensins and other antimicrobial peptides at the ocular surface.* Ocul Surf, 2004. **2**(4): p. 229-47.
12. Durr, U.H., U.S. Sudheendra, and A. Ramamoorthy, *LL-37, the only human member of the cathelicidin family of antimicrobial peptides.* Biochim Biophys Acta, 2006. **1758**(9): p. 1408-25.
13. Toma, H.S., et al., *Ocular HSV-1 latency, reactivation and recurrent disease.* Semin Ophthalmol, 2008. **23**(4): p. 249-73.
14. Farooq, A.V., T. Valyi-Nagy, and D. Shukla, *Mediators and mechanisms of herpes simplex virus entry into ocular cells.* Curr Eye Res. **35**(6): p. 445-50.
15. Kleymann, G., *Novel agents and strategies to treat herpes simplex virus infections.* Expert Opin Investig Drugs, 2003. **12**(2): p. 165-83.
16. Cortesi, R. and E. Esposito, *Acyclovir delivery systems.* Expert Opin Drug Deliv, 2008. **5**(11): p. 1217-30.
17. Fagerholm, P., et al., *A biosynthetic alternative to human donor tissue for inducing corneal regeneration: 24-month follow-up of a phase 1 clinical study.* Sci Transl Med. **2**(46): p. 46ra61.
18. Liu, W., et al., *Recombinant human collagen for tissue engineered corneal substitutes.* Biomaterials, 2008. **29**(9): p. 1147-58.
19. Merrett, K., et al., *Tissue-engineered recombinant human collagen-based corneal substitutes for implantation: performance of type I versus type III collagen.* Invest Ophthalmol Vis Sci, 2008. **49**(9): p. 3887-94.

20. Li, Y., *Carrier proteins for fusion expression of antimicrobial peptides in Escherichia coli.* Biotechnol Appl Biochem, 2009. **54**(1): p. 1-9.
21. Technologies, I.L. *ChampionTM pET SUMO Protein Expression System*. For high-level expression and enhanced solubility of recombinant proeins in E. coli and cleavage of native protein 2004 [cited.
22. Andrades, J.A., et al., *A recombinant human TGF-beta1 fusion protein with collagen-binding domain promotes migration, growth, and differentiation of bone marrow mesenchymal cells.* Exp Cell Res, 1999. **250**(2): p. 485-98.
23. Whitlow, M., et al., *An improved linker for single-chain Fv with reduced aggregation and enhanced proteolytic stability.* Protein Eng, 1993. **6**(8): p. 989-95.
24. Biolabs, N.E., *Protein Purification System Now Featuring Fusion to C- or N- Terminus of the Target Protein*. 2006.
25. Banki, M.R., L. Feng, and D.W. Wood, *Simple bioseparations using self-cleaving elastin-like polypeptide tags.* Nat Methods, 2005. **2**(9): p. 659-61.
26. Piers, K.L., M.H. Brown, and R.E. Hancock, *Recombinant DNA procedures for producing small antimicrobial cationic peptides in bacteria.* Gene, 1993. **134**(1): p. 7-13.
27. Wood, W.J., L. Huang, and J.A. Ellman, *Synthesis of a diverse library of mechanism-based cysteine protease inhibitors.* J Comb Chem, 2003. **5**(6): p. 869-80.
28. Lee, J.H., et al., *Enhanced expression of tandem multimers of the antimicrobial peptide buforin II in Escherichia coli by the DEAD-box protein and trxB mutant.* Appl Microbiol Biotechnol, 2002. **58**(6): p. 790-6.
29. Wang, F., et al., *Fusion expression of human beta-defensin-2 from multiple joined genes in Escherichia coli.* Prep Biochem Biotechnol, 2004. **34**(3): p. 215-25.