BACTERIAL EXPRESSION OF SELF-ASSEMBLING PEPTIDE HYDROGELATORS

by

Cem Sonmez

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Biochemistry

Spring 2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

BACTERIAL EXPRESSION OF SELF-ASSEMBLING

PEPTIDE HYDROGELATORS

by

Cem Sonmez

Approved: _____

Murray Johnston, Ph.D.
Chair of the Department of Chemistry and Biochemistry

Approved: _____

George H. Watson, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____

James G. Richards, Ph.D.
Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Joel P. Schneider, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Brian J. Bahnson, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Sharon Rozovsky, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Darrin J. Pochan, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

First of all, I would like to thank to my advisor Dr. Joel Schneider for such a graduate experience opportunity with invaluable supervision, support and encouragement throughout my graduate career at the UD and at the National Cancer Institute. I would also express my gratitude to my committee Dr. Sharon Rozovsky, Dr. Brian Bahnson and Dr. Darrin Pochan for their guidance and scientific discussions.

Throughout my graduate studies, there have been many people in the Schneider Lab that I would like to thank for their assistance, knowledge, guidance, support and patience. It was great to work with such colleagues that helped me developing my scientific skills; not only Ronak Rughani, Radhika Nagarkar, Monica Branco, Chomdao Sinthuvanich, Mike Giano, Dan Smith and Kate Nagy that joined the lab before me, but also Salome Veiga, Iris Geisler, Scott Media and Yuji Yamada. I want to mention Sameer Sathaye from the Pochan Lab, as well.

I would like to express my thanks to colleagues at NCI that I had a chance to meet and discuss scientifically. Many thanks are given to Molecular Targets Lab where I conducted my experiments in my first years and Barry O`keefe for welcoming me to his lab which include but not limited to Lauren Krumpe, Carrie Saucedo, Shilpa Kurian and Koreen Ramessar. Also Jacek Lubkowski from Macromolecular Crystallography Lab and Dominic Esposito from Protein Expression Lab were very

helpful and supportive by not only providing materials but also sharing their time and knowledge.

I would like to thank the faculty and staff of the Department of Chemistry and Biochemistry at UDEL and the Chemical Biology Laboratory at NCI for their assistance. I would like to mention the people that helped me with the revision of my dissertation, as well. Many thanks are given to Scott and Kate from my lab, Koreen from MTL and Jim Kelley from CBL.

Many years here at NCI have been shared with great friends from NCI, as well. Much gratitude is expressed to Erkan Kiris, Humeyra Taskent and Efe Sezgin for the wonderful times spent together with fun, joy and happiness. I got great moral support by their friendship. I would also like to acknowledge and sincerely thank many people from NCI`s international postdoctoral community that I had great time together out of the lab, as well.

Without their love, motivation, support and faith, I would not have the life skills that has traveled me to the end of my dissertation so I want to dedicate this work to my family, my mom, dad, sister, brother and wife; Nuran Sonmez, Raci Sonmez, Berna Sonmez, Ege Sonmez and Deniz Tavli Sonmez. I want to especially express my appreciation to my wife, Deniz, for her endless support, love and patience during my graduate career and to my daughter, Ela, for always smiling back at me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

For tissue regeneration and drug delivery applications, various architectures are explored to serve as biomaterial tools. Via *de novo* design, functional peptide hydrogel materials have been developed as scaffolds for biomedical applications. The objective of this study is to investigate bacterial expression as an alternative method to chemical synthesis for the recombinant production of self-assembling peptides that can form rigid hydrogels under physiological conditions. The Schneider and Pochan Labs have designed and characterized a 20 amino acid β-hairpin forming amphiphilic peptide containing a D-residue in its turn region (MAX1). As a result, this peptide must be prepared chemically. Peptide engineering, using the sequence of MAX1 as a template, afforded a small family of peptides for expression (EX peptides) that have different turn sequences consisting of natural amino acids and amenable to bacterial expression. Each sequence was initially chemically synthesized to quickly assess the material properties of its corresponding gel. One model peptide EX1, was chosen to start the bacterial expression studies. DNA constructs facilitating the expression of EX1 were designed in such that the peptide could be expressed with different fusion partners and subsequently cleaved by enzymatic or chemical means to afford the free peptide. Optimization studies were performed to increase the yield of pure peptide that ultimately allowed 50 mg of pure peptide to be harvested from one liter of culture, providing an alternate means to produce this hydrogel-forming peptide. Recombinant production of other self-assembling hairpins with different turn sequences was also

successful using this optimized protocol. The studies demonstrate that new β-hairpin self-assembling peptides that are amenable to bacterial production and form rigid hydrogels at physiological conditions can be designed and produced by fermentation in good yield at significantly reduced cost when compared to chemical synthesis.

Chapter 1

INTRODUCTION

1.1 Peptide Materials in Tissue Engineering and Drug Delivery

In nature, proteins are the main building blocks of living organisms. The vast diversity of life is a result of the variety in the amino acid sequence and composition and the consequent 3-dimensional structures which dictate the biological functions of proteins. The influence of protein structure on biological function was substantially impacted by Pauling and Corey`s papers in 1951 detailing α-helix and β-sheet protein secondary structures [1, 2].

Since this early work, mimicking and altering natural protein conformations have become a corner-stone of modern pharmaceutical and biotechnology research endeavors. The biotechnology field explores these varying architectures to develop tools for tissue engineering and drug delivery applications. Biomaterials are promising candidates for such therapies and combine various approaches at the level of design, structure and function. The structural elements of proteins, particularly peptides can serve as an inspiration for developing biomimetic molecules. Peptides are short chains of amino acids that are chemically accessible (through solid phase peptide synthesis),

functionally diverse (cell binding, proteolytic susceptibility, self-assembly) and amenable to *de novo* design [3]. Peptide-based biomaterials have been fabricated by a bottom-up approach in which individual components build up supramolecular self-assemblies [4]. When rationally designed, peptide self-assembly results in a diverse range of architectures like fibers, tapes, sheets and tubes. Although weak non-covalent bonds are utilized for individual intermolecular interactions, supramolecular stability can be achieved into the self-assembled state [5]. Furthermore, various functionalities can be incorporated into the material by including design elements at the level of the monomer. As a result, peptide-based hydrogels are promising materials for tissue engineering as they offer tunable tissue-like characteristics that allow cells to adhere, grow, proliferate, migrate or differentiate [6].

1.2 Overview of Peptide-based Hydrogel Systems

One particularly interesting application of peptide-based hydrogels is as extracellular matrix (ECM) mimetics in tissue engineering. The ECM supports cells and allows diffusion of molecules in a highly hydrated environment by forming 3 dimensional porous networks. Some peptide-based hydrogels mimic the natural ECM through the formation of physically rigid porous networks allowing nutrient, drug and waste diffusion [7]. As a result, these hydrogels have been used as scaffolds to offer mechanical support for encapsulated cells and as a depot for the release of loaded drugs [8]. Extracellular matrix derived peptides and proteins can be immobilized into

the material scaffolds to facilitate cell binding, division, differentiation and movement [9].

Distinct nanostructures with desired mechanical properties as well as the presentation of biochemical cues can be generated by using hydrogel-forming self-assembling peptides. Peptides can be tailored to self-assemble and form gels in response to various environmental cues such as temperature, pH and ionic strength [10]. This level of control over gelation allows one to change the properties (e.g. stiffness, gelation kinetics) of hydrogels for desired applications. By using non-covalent electrostatic interactions, hydrophobic effects and hydrogen bonding between the components of physical hydrogels, injectable gels can be generated for minimally invasive applications [11].

1.2.1 Peptide hydrogels formed from α-helix and β-sheet motifs

α-helix and β-sheet secondary structures are the main components in the construction of self-assembling peptides that form hydrogels. α-helices are characterized by a backbone spiral conformation in which every 3.6 residues make one α-helical turn extending about 5.4 Å. Helices are stabilized by hydrogen bonding between the carbonyl oxygen of i^{th} residues and the amine hydrogen of $(i+4)^{th}$ residues. Peptides can also adopt a β-sheet conformation, which is defined by an

extended structure where hydrogen bonds are formed between neighboring strands that are either parallel or antiparallel. By using the characteristics of these two secondary structural elements, many different peptide-based hydrogel systems have been reported [10, 12-19].

Peptides having alternating hydrophobic and hydrophilic amino acid residues tend to form β-sheet structure in which one face is hydrophobic and the other face is hydrophilic. For example, Brack et. al. showed that poly(Val-Lys), poly(Glu-Ala), poly(Tyr-Glu), poly(Lys-Phe), poly(Lys-Leu) formed β-sheet structure in a pH and salt dependent manner [20]. In another study, the repetitive sequence $(AEAEAKAK)_2$ (EAK16) containing alternating hydrophobic and hydrophilic residues was identified in the yeast protein zuotin and characterized by Zhang et. al. [21]. Complementary ionic interactions between negatively charged glutamate residues and the positively charged lysine residues facilitated self-assembly resulting in the formation of a mechanically stable macroscopic membrane that displayed enhanced proteolytic stability against various proteases [21]. A similar peptide sequence, $Ac\text{-}(RADA)_4\text{–}NH_2$ (RADA16-I), containing alternating arginine and aspartate residues in the hydrophilic face was designed. RADA16-I was able to disassemble mechanically by sonication and reassemble upon cessation of sonication. When dissolved in water and sonicated, RADA16-I forms ordered nanofiber scaffolds and it was shown that β-sheet content of the fibrils does not change with repetitive sonication steps [22]. Caplan et.

al. studied sequence variations for similar peptide sequences with alternating hydrophobic and charged residues whose charges are also alternating on the hydrophilic face. Salt concentrations required to form gels for the peptide sequences Ac-$(FKFE)_n$–NH_2 were compared for n=2,3 and 4. Ac-$(FKFE)_3$–NH_2 required the lowest salt concentration due to an increase in intermolecular attraction which outweighed a decrease in the chain entropy as the peptide assembled [23]. In the same study, variation of hydrophobic residues on the sequence Ac-$(XKXE)_3$–NH_2, where X is phenylalanine, isoleucine or valine, was investigated and it was shown that the more hydrophobic peptides require less salt in order to assemble.

Aggeli et. al. studied the membrane protein IsK with a single transmembrane domain from which shorter peptides that adopt β-sheet conformation were generated and characterized [24]. The results of these experiments lead to the *de novo* design of the 11-residue peptide P_{11}-2 (Ac-QQRFQWQFEQQ-NH_2) that forms self-supporting gels at 15 mg/mL concentration in water. P_{11}-2 is a glutamine rich sequence with alternating polar and aromatic residues that forms β-sheet structure. Self-assembly of the peptide was promoted by the methylene groups on the glutamine side chains that provided hydrophobic interactions along with phenylalanine and tryptophan residues, which also could form π-π interactions. Arginine and glutamic acid residues form ionic interactions that favored antiparallel alignment of the neighboring strands during assembly [25]. P_{11}-2 displays a concentration dependent higher order structure formation in which β-sheet monomers self-assemble and form tapes over a critical

concentration and as the concentration increases, ribbons are formed from paired tapes, then ultimately fibrils and fibers are formed [26, 27]. In other work Collier at. al. designed Q11 (Ac-QQKFQFQFEQQ- NH_2) based on the P_{11}-2 sequence and appended a 17-residue peptide epitope from chicken egg ovalbumin N-terminal to P_{11}-2 [28]. As the peptide assembles into fibrils, epitopes are displayed along the long-axis of the fibrils. The assemblies were assessed as vaccines. The authors showed that high antibody titers were achieved in mice studies without the use of adjuvants. To afford more rigid hydrogels, Q11 was modified for enabling native chemical ligation between the N-terminal cysteine and C-terminal thioester moiety. Such modification afforded gels having five times greater storage moduli [29]. In a separate study, an RGD motif was appended to Q11, improving cell adhesion and spreading of HUVECs [30]; this peptide may find use in cell-based applications.

In other work, hydrogel formation of the dipeptide Fmoc-Leu-Asp was shown by Vegners et. al. where gelation was observed upon cooling of a <1wt% aqueous peptide solution [31]. In the same study, the gel was employed as a carrier for antigen presentation. Inherently non-antigenic adamantanamine derivatives were incorporated into these gels during the cooling process and the gels were injected into rabbits without adjuvants to induce antigen presentation. High titers of specific antibodies were obtained with the use of this hydrogel. For tissue engineering applications, Jayawarna et. al. studied gelation of Fmoc-dipeptides by dissolving them in water and then increasing the pH to deprotonate the C-terminal carboxylic acid and solubilize the

peptide. Next, the pH was decreased slowly until a gel was formed. It was shown that different Fmoc-dipeptides gelled at different values of pH; Fmoc-Gly-Gly, Fmoc-Ala-Gly, Fmoc-Ala-Ala, Fmoc-Leu-Gly, Fmoc-Phe-Gly gelled below pH 4 whereas Fmoc-Phe-Phe gelled below pH 8 [32]. Gelation pH could then be tuned to physiological pH by combining these dipeptides to form self-supporting fibrous gels used for chondrocyte seeding and proliferation. Xu et. al. appended naphthalene to one of these Fmoc-dipeptides that self-assembled to form super helical nanofiber structures [33]. At low (0.07 wt%) concentrations, gel formation occurred with the resulting material shown to be biocompatible. When the naphthyl moiety was placed on the C-terminus of Phe-Phe dipeptides, gelation occured in a narrow pH range (pH 5-6). These gels showed a selective inhibition of cancer cell (HeLa) growth over non-cancerous cells [34]. The possible reasons were speculated due to the poor permeability of non-cancerous cells and/or interaction of essential serum proteins in the medium of HeLa cells to these dipeptides and subsequently inhibit cell growth.

Inducing amphiphilicity through the addition of an alkyl moiety to short peptides is an alternative way to produce peptides capable of gelling. Such molecules, called peptide amphiphiles (PA), combine hydrophobic alkyl chains and hydrophilic peptides to produce molecules that assemble into 3 dimensional networks of fibrils having high-aspect-ratios. PAs are typically 4 component systems; an alkyl chain, a short hydrophobic peptide sequence, a segment of charged amino acids (to help solubility and direct self-assembly) and a bioactive moiety to interact with cells and

proteins [35]. Alkyl chain length and the nature of the amino acids in each segment can be varied to control the mechanical properties, morphologies, self-assembly characteristics and the ultimate function of the PA [36]. Specifically, different resulting morphologies can be generated by changing the interaction of residues on the peptide portion of the PA. An alkyl chain followed by a VVEE peptide sequence results in cylindrical fiber morphology, whereas alternating valine and glutamate residues (VEVE) result in a bilayer-like structure that ultimately forms nanobelts that are wide and flat structures [37]. Encapsulated cells in PA matrices were viable and proliferate for weeks without cell-cycle arrest indicating biocompatibility. In addition, the hydrogelation kinetics of the PA gels were tuned by incorporating larger hydrophobic residues together with the existing hydrophilic residues (SLSLGGG vs. AAAAGGG). These amphiphiles can be injected as liquids in vivo where they gel very fast. Epitopes like IKVAV for neural attachment, migration and outgrowth, and RGDS for cell adhesion were also appended to PA gels to enhance their cellular activity [35].

Peptide quaternary structure can also be used for material formation. Coiled-coils, where two or more α-helices form a self-assembled supercoil, were designed according to the heptad repeat $(abcdefg)_n$ of coiled-coils. Positions a and d on the same face of one helix forms a hydrophobic strip that drives peptide folding via coiled-coil formation. Positions e and g interact (e.g. charge-charge interactions) at the coiled-coil interface and positions b, c and f are solvent exposed [38]. The stability and the

multimerization state of coiled coils can be controlled by varying the residues within the heptad repeat. In these systems, oligomeric α-helices wrap around each other to bury the hydrophobic surfaces efficiently affording coiled-coils that can be very long. For example, fibrin and keratin are natural examples of coiled coil structures. Gel formation by using coiled coils can be achieved in various ways. Linking coiled coil forming α-helices to polymers results in peptide-polymer-peptide triblock macromolecules that ultimately self-assemble into extended structures forming a competent hydrogel [39]. A similar strategy is to cap the ends of unstructured hydrophilic sequences with coiled-coil forming peptides to promote gelation [40]. By rational peptide design, temperature responsiveness, swelling and viscoelastic behaviors can be modulated [11]. Networks composed entirely of coiled coils can also be generated by designing sticky ended coiled coils, a strategy made popular by the Woolfson Group. Here, heptad extension at the C- and N-termini of distinct peptides drives oligomerization and fibril formation [38].

The ECM structural protein elastin, found in many elastic organs like lungs, skin and arteries, has inspired the development of elastin-based materials [41-44]. Polymers compose of repeating units of the native pentapeptide repeat VPGXG, where X is any amino acid except proline, showed tunable physical properties by incorporating various amino acids in the X position [45]. Elastin-like polypeptides (ELP) present temperature dependent sol-gel phase transitions when they are heated over a transition temperature, which is defined by peptide hydrophobic content,

polymer size and concentration. Random coil conformation at low temperatures gives way to the formation of β-turn spiral conformations of the backbone where the hydrophobic residues are buried as the temperature increases. At the inverse phase transition temperature, the spiral folded peptides go through coacervation observed by the turbidity of the solution due to aggregation [46, 47]. The inverse temperature transition behavior can be adjusted between 0-100^{o}C by genetically engineering the sequence so that desired characteristics can be tailored that enable drug delivery and tissue engineering applications [48, 49].

1.2.2 β-hairpin forming peptide hydrogels

The rational design of peptides by the Schneider and Pochan laboratories have produced a family of β-hairpin peptides capable of folding and self-assemble upon addition of an environmental trigger to form hydrogels [50, 51]. The *de novo* designed parent peptide MAX1 has the sequence of VKVKVKVK-V^{D}P^{L}PT-KVKVKVKV-$CONH_2$, in which two strands containing alternating hydrophobic valine and hydrophilic lysine residues flank a tetrapeptide sequence -V^{D}PPT- designed to adopt type II` β-turn. At low pH and ionic strength the lysine side chains are protonated resulting in intrastrand charge repulsion keeping the peptide in an unfolded state. Upon addition of a folding trigger (e.g. pH change, salt, temperature) MAX1 folds and self-assembles into a fibrillar network (Figure 1.1). The transition from random coil to β-hairpin conformation can be triggered by either addition of salt which screens the

positive charges or increasing the pH which deprotonates some of the lysine residues [52, 53]. In addition, increasing the temperature promotes the desolvation of hydrophobic residues and subsequent hydrophobic collapse. Once such an external stimulus is introduced, the peptide folds into an amphiphilic β-hairpin molecule characterized by a lysine-rich hydrophilic face and a valine-rich hydrophobic face. The folded hairpin structure subsequently facially associates via hydrophobic collapse of the valine-rich face resulting in a bilayer-like structure. Lateral association of the hairpins also occurs via intermolecular hydrogen bonds and side chain hydrophobic contacts formed between neighboring hairpin molecules, resulting in the formation of fibrils. The smooth hydrophobic surface of the hairpin is susceptible to mispacking during fibril growth. The sites along a growing fibril where mispacking occurs are nucleation sites where new fibrils can begin to grow. Nascent fibrils can grow in a new direction in 3 dimensions. Overall, this results in the formation of branching points resulting in interfibril cross-links [54]. Ultimately, mechanically rigid hydrogels are formed via the contribution of these interfibril cross-links as well as simple fibrillar entanglements.

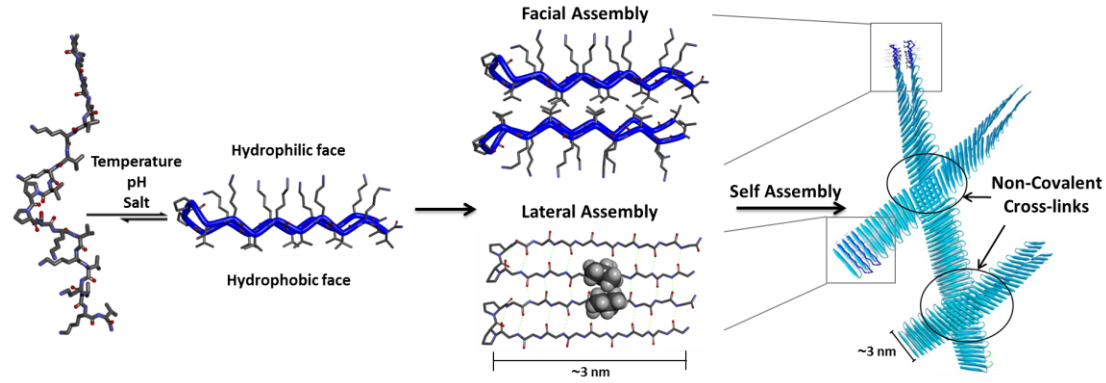


Figure 1.1 Folding and self-assembling mechanism of MAX1.

Rajagopal et. al. showed that the pH responsiveness of β-hairpin peptides can be tuned by decreasing the overall charge state of the peptide, leading to folding and self-assembly of the peptides triggered at lower pH values as compared to MAX1 [53]. It was also found that when the net positive charge of the peptide is decreased, the side chains are screened more effectively resulting in more peptide being incorporated into the hydrogel network, leading to a greater number of cross-links formed creating a more rigid gel [55, 56]. Based on this principle, MAX8 was designed through the substitution of Lysine 15 of MAX1 with a glutamic acid. This substitution decreased the net charge of the peptide from +9 to +7, leading to more rigid hydrogels (<40 Pa vs. 550 Pa for 0.5wt% gels after 1 h at 37^{o}C) at physiologically relevant pH and ionic strength conditions compared to MAX1 hydrogels [55].

As a result of the non-covalent nature of the gel cross-links, β-hairpin hydrogels show shear-thin and recovery capabilities. Specifically, these hydrogels go through transition from a visco-elastic solid to liquid-like under shear stress allowing them to act as an injectable scaffold or depot. During syringe delivery, the β-hairpin hydrogel flows like a viscous solution but upon the cessation of shear, it recovers its bulk properties at the site of delivery. It was proposed that the physical network fractures into domains during the application of shear and the gel flows as detached fibrils between the gel domains. Upon cessation of shear stress, the gel domains relax at the boundaries and interpenetrate into each other to percolate the whole volume again [57]. This shear-thinning and recovery property allows for the delivery of viable encapsulated cells homogenously distributed in the gel, with material properties similar to that before syringe delivery [58].

The syringe delivery of biocompatible hydrogels for cell encapsulation has been extensively studied in the lab. For example, gelation of MAX1 triggered by the addition of DMEM cell culture media resulted in rigid hydrogels whose surfaces presented cytocompatibility when model fibroblast cells (NIH 3T3) were tested [59]. In addition, MAX8 gels have been prepared containing a homogenous distribution of cells, which could shear-thin and recover without compromising cell viability [55, 57, 58].

1.3 Overview of Recombinant Peptide Production

While the chemical synthesis has been optimized for many peptides produced in the Schneider Lab, there is the opportunity for a more cost effective method of production through bacterial expression. Commonly, solid phase peptide synthesis is employed for small scale batches. While this method is rapid, it is limited by cost and yields during scale-up. Recombinant production is an alternative method that employs a host organism`s machinery for peptide synthesis. Antimicrobial peptides represent a class of peptides whose recombinant production has been investigated intensively. *E. coli* is the most employed organism among other host organisms like yeast, fungi, insect and plant for the production of antimicrobial peptides. There are many hurdles to overcome when engineering a system for efficient peptide production. One hurdle is the possible rapid degradation of the translated peptide by the host organism. To circumvent bacterial degradation pathways peptides are typically expressed as a fusion with a larger partner protein. Often times a tag (e.g. hexaHistidine/Histag) for purification is included in the design of the fusion protein. To free the peptide from its fusion partner following expression, a cleavage site is also incorporated in the sequence. In the case of antimicrobial peptide expression in bacteria, the peptides are generally fused to protein fusion partners that can mask the antimicrobial activity of the peptide or promote inclusion body formation in the cell to sequester the active peptide. Another hurdle is that highly repetitive sequences of target genes may result in poor expression levels and susceptibility to undesired mutations [60, 61]. Below,

precedent is given for the expression of antimicrobial and self-assembling peptides with corresponding yields to serve as a theoretical basis for the work presented in this thesis. The work published to date serves to validate the hypothesis that expression systems can be designed that allow efficient production of self-assembling peptides as well as peptides whose native function is to kill the very organism in which they are being produced.

ABP-CM4 is a 35-residue antimicrobial peptide from the cecropin family identified in insects and when secreted from *P. pastoris,* a yield of 15 mg/L was achieved [62]. A Thioredoxin (Trx) fusion of the ABP-CM4 peptide expressed in *E. coli* yielded 1.2 mg/L [63]. However, when expressed as a Glutathione S-transferase (GST) fusion in *E.coli*, the peptide could not be recovered. A chitin-binding domain (CBD) and intein fusion yielded 2.1 mg/L [64]. Although the yields were low, the work nicely demonstrated that bacterial expression is possible for this class of peptides.

Trx-Histag fusion of another cecropin (Mdmcec) expressed in *E. coli* resulted in a yield of 11.2 mg/L [65]. In other work, cathelicidins derived from the LL-37 peptide, which is a 37-residue peptide with two N-terminal leucines, were expressed in *E. coli* as a Trx fusion and the yields of the peptides were as follows; LL-23 5.4 mg/L, LL-29 2.1 mg/L, KR-20 0.7 mg/L, SK-29 1.7 mg/L [66]. The 22-residue peptide piscidin-1, identified in fish (hyrid-striped bass), was expressed in *E. coli* as a Histag-

ubiquitin fusion yielding 1.5 mg/L peptide [67]. A tandem repeat expression strategy for indolicidin derivative where 15 repeats of short anionic peptides with methionine at their N- and C-termini were interspersed between a cationic peptide fused to cellulose binding protein A ($CBP_{33\text{-}123}$), yielded 100 mg/L peptide functionalized with a C-terminal homoserine lactone moiety [68]. Another example of tandem expression is the expression of histonin repeats (2-32) fused to PurF protein. An endogenous furin cleavage site (RLKR) at the C-terminus was employed for cleavage and ultimately a yield of 167 mg/L peptide was achieved [69]. Table 1.1 lists the peptides and their corresponding yields. Generally, the yields are less than 10 mg/L with a few exceptions of relatively higher yields.

Table 1.1 Antimicrobial peptides expressed in bacteria. Peptide sequences, host organisms, fusion partner proteins and yields are shown.

Peptide	Host	Partner	Yield
ABP-CM4 MEDFQEDREGRSKHQRRYRQGWSSCRC RRSSCYY	*P. pastoris*	-	15 mg/L
ABP-CM4	*E. coli*	Trx	1.2 mg/L
ABP-CM4	*E. coli*	CBD-intein	2.1 mg/L
Mdmcec MKKIGKKIERVGQHTRDATIQTIGVAQQ AANVAATLKG	*E. coli*	Trx-Histag	11.2 mg/L
LL-23 LLGDFFRKSKEKIGKEFKRIVQR	*E. coli*	Trx	5.4 mg/L
LL-29 LLGDFFRKSKEKIGKEFKRIVQRIKDFLR	*E. coli*	Trx	2.1 mg/L
KR-20 KRIVQRIKDFLRNLVPRTES	*E. coli*	Trx	0.7 mg/L
SK-29 KRIVQRIKDFLRNLVPRTES	*E. coli*	Trx	1.7 mg/L
Piscidin-1 FFHHIFRGIVHVGKTIHRLVTG	*E. coli*	Histag-Ubiquitin	1.5 mg/L
Indolicidin* ILRWPWWPWRRK**M**AEAEPEAEPI**M**	*E. coli*	CBD derivative	100 mg/L
Histonin RAGLQFPVGKLLKKLLKRLKR	*E. coli*	PurF	167 mg/L

*Indolicidin sequence is underlined and anionic spacer sequence is between methionines shown in bold.

Pertinent to the work described in this thesis is the production of self-assembling peptides. The yields of recombinantly produced self-assembling peptides in the literature are summarized in Table 1.2. Tandem repeats (3) of $A\beta_{11-26}$ were expressed as a ketosteroid isomerase (KSI) fusion in *E.coli* and following CNBr cleavage, a yield of 10 mg/L was achieved [70]. RAD16 was expressed as tandem repeats fused to CBD in *R. eutropha* with an additional glutamate at the C-terminus of

the repeats for endoproteinase-Glu cleavage and 10.1 mg of peptide was recovered from a liter of culture [71]. SA2 and SA7 peptides were expressed by using small ubiquitin-related modifier (SUMO) protein as a fusion partner and cleavage was performed by SUMO protease. The overall yield of the peptides was 6 mg/L [72]. The 21-residue peptide from lac repressor called mLac21 was produced in *E.coli* as a KSI fusion and cleavage at the cysteine residues by using 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP) afforded the peptide with a final yield of 2.9 mg/L [73]. The P_{11}-2 peptide was fused to Trx with a tobacco etch virus (TEV) protease cleavage site N-terminal to P_{11}-2 and the ultimate yield was 13 mg/L [74]. Much higher yields around 90 mg/L were achieved by the McPherson laboratory for the P_{11}-4 peptide when 3 repeats of the peptide were fused to KSI and auto-induction media was employed [75]. CNBr cleavage and purification afforded a 90 mg/L yield, which was further improved to 203 mg/L when a modified version of auto-induction media was used [76]. In another study, the P_{11}-4 peptide and the complementary P_{11}-13 and P_{11}-14 peptides which are able to form hydrogel only when combined together, were expressed as SUMO fusion proteins and the yields were 35.3 mg/mL, 51.5 mg/mL and 17.8 mg/mL, respectively [77]. Although the yields of recombinant production of most self-assembling peptides are around 10 mg/L, the results of McPherson are outstanding and demonstrate that careful modification and optimization of the expression system can lead to excellent yields.

Table 1.2 Self-assembling peptides expressed in bacteria. Peptide sequences, host organisms, fusion partner proteins and yields are shown.

Peptide	Host	Partner	Yield
$A\beta_{11-26}$ $(EVHHQKLVFFAEDVGM)_3$	*E. coli*	KSI-Histag	10 mg/L
RAD16 $(RADARADARADARADAE)_4$	*R. eutropha*	CBD	10.1 mg/L
SA2 AAVVLLLWEE	*E. coli*	SUMO	6 mg/L
SA7 AAVVLLLWEEEEEEE	*E. coli*	SUMO	6 mg/L
mLac21 CKQLADSLMQLARQVSRLESA	*E. coli*	KSI	2.9 mg/L
P_{11}-2 QQRFQWQFEQQ	*E. coli*	Trx	13 mg/L
P_{11}-4 $(QQRFEWEFEQQM)_3$	*E. coli*	KSI-Histag	90 mg/L
P_{11}-4 QQRFEWEFEQQ	*E. coli*	KSI	203 mg/L
P_{11}-4 QQRFEWEFEQQ	*E. coli*	SUMO	35.3 mg/L
P_{11}-13 EQEFEWEFEQE	*E. coli*	SUMO	51.5 mg/L
P_{11}-14 (K) QQKFKWKFKQQ	*E. coli*	SUMO	17.8 mg/L

1.4 Research Objective and Organization of Thesis

The primary focus of this thesis is the investigation of bacterial expression of β-hairpin peptides as an alternative to production by chemical synthesis. Because each step that defines the methodology for bacterial expression can influence the overall yield, bacterial peptide production is challenging and requires careful consideration. For example, fusion protein and cleavage method selection, cloning strategy and

culturing method selection and downstream processing all need to be optimized to produce efficient yields.

The comprehensively studied β-hairpin peptide MAX1 contains a D-amino acid residue in its turn region and an amide functionality at its C-terminus which are not found in natural peptides, hence MAX1 cannot be produced in bacteria. Therefore, the first objective is to design novel β-hairpin sequences amenable for bacterial expression. However, these same peptides must also be capable a well-controlled self-assembly leading to the formation of rigid hydrogels at physiological conditions (pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C). The strategy followed here is to make changes to the MAX1 sequence affording new sequences amenable to bacterial production. Chemical synthesis of the new sequences will first be performed to quickly assess their folding and assembly properties and the material properties of their corresponding gels. The MAX1 derivatives which retained the ability to afford self-supporting hydrogels at physiological conditions will be selected for bacterial expression. In Chapter 2, the design and characterization of novel self-assembling peptides amenable to bacterial expression is presented. The new designs are first prepared by chemical synthesis as outlined in Chapter 2.

After selection of candidate peptides capable of efficient gel formation, my second objective was to establish an initial bacterial expression strategy as outlined in Chapter 3. This includes DNA construct design for alternative routes of peptide

expression and selection of expression and downstream processing methods. One model peptide, EX1, was chosen for the initial pilot expression studies. Pilot expression studies were performed to determine which strategies were successful. Scale-up and the established preliminary downstream processing methods afforded recombinantly produced peptide. These pilot studies, using the peptide EX1 are presented in Chapter 3.

In these pilot experiments, the overall yield was low; therefore my third objective was method refinement to improve the yield of pure recombinant β-hairpin peptide. If successful, this would result in a low-cost alternative to chemical synthesis. As will be shown in Chapter 4, optimization resulted in an increase in harvested biomass, as well as enhanced the recovery of the peptide. Solubility characteristics of the β-hairpin peptide were exploited during the downstream processing, which allowed extra purification steps to be eliminated. Once the final method was established, it was employed for two additional sequences, namely, EX2 and EX3 which were designed in Chapter 2. The yields of EX2 and EX3 were 31, and 15 mg/L which was slightly lower compared to 50 mg/L for EX1. Overall the work outlined in Chapters 2, 3 and 4 indicate that gel-forming β-hairpin peptides can be produced in bacteria.

Lastly in Chapter 5, problems encountered during the chemical synthesis of the peptides, namely succinimide formation and lysine deletion in different target peptides are discussed together with their consequences on the material properties of their corresponding gels. Synthetic methods that prevent succinimide formation are presented. In separate work, the consequences of lysine deletions specific positions within the hairpin on the elastic properties of their corresponding gels are also presented.

REFERENCES

1. Pauling, L. and R.B. Corey, *Atomic Coordinates and Structure Factors for 2 Helical Configurations of Polypeptide Chains.* Proceedings of the National Academy of Sciences of the United States of America, 1951. **37**(5): p. 235-240.

2. Pauling, L. and R.B. Corey, *The Pleated Sheet, a New Layer Configuration of Polypeptide Chains.* Proceedings of the National Academy of Sciences of the United States of America, 1951. **37**(5): p. 251-256.

3. Collier, J.H. and T. Segura, *Evolving the use of peptides as components of biomaterials.* Biomaterials, 2011. **32**(18): p. 4198-4204.

4. Drexler, K.E., *Molecular Engineering - an Approach to the Development of General Capabilities for Molecular Manipulation.* Proceedings of the National Academy of Sciences of the United States of America-Physical Sciences, 1981. **78**(9): p. 5275-5278.

5. Zhang, S.G., *Fabrication of novel biomaterials through molecular self-assembly.* Nature Biotechnology, 2003. **21**(10): p. 1171-1178.

6. Discher, D.E., P. Janmey, and Y.L. Wang, *Tissue cells feel and respond to the stiffness of their substrate.* Science, 2005. **310**(5751): p. 1139-1143.

7. Collier, J.H., et al., *Multi-component extracellular matrices based on peptide self-assembly.* Chemical Society Reviews, 2010. **39**(9): p. 3413-3424.

8. Yan, C.Q. and D.J. Pochan, *Rheological properties of peptide-based hydrogels for biomedical and other applications.* Chemical Society Reviews, 2010. **39**(9): p. 3528-3540.

9. Hirano, Y. and D.J. Mooney, *Peptide and protein presenting materials for tissue engineering.* Advanced Materials, 2004. **16**(1): p. 17-25.

10. Lowik, D.W.P.M., et al., *Stimulus responsive peptide based materials.* Chemical Society Reviews, 2010. **39**(9): p. 3394-3412.

11. Mart, R.J., et al., *Peptide-based stimuli-responsive biomaterials.* Soft Matter, 2006. **2**(10): p. 822-835.

12. Altunbas, A. and D.J. Pochan, *Peptide-Based and Polypeptide-Based Hydrogels for Drug Delivery and Tissue Engineering.* Peptide-Based Materials, 2012. **310**: p. 135-167.

13. Bowerman, C.J. and B.L. Nilsson, *Self-assembly of amphipathic beta-sheet peptides: insights and applications.* Biopolymers, 2012. **98**(3): p. 169-84.

14. Tsutsumi, H. and H. Mihara, *Soft materials based on designed self-assembling peptides: from design to application.* Mol Biosyst, 2013. **9**(4): p. 609-17.

15. Apostolovic, B., M. Danial, and H.A. Klok, *Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials.* Chem Soc Rev, 2010. **39**(9): p. 3541-75.

16. Maude, S., E. Ingham, and A. Aggeli, *Biomimetic self-assembling peptides as scaffolds for soft tissue engineering.* Nanomedicine, 2013. **8**(5): p. 823-847.

17. MacEwan, S.R. and A. Chilkoti, *Elastin-Like Polypeptides: Biomedical Applications of Tunable Biopolymers.* Biopolymers, 2010. **94**(1): p. 60-77.

18. Cui, H., M.J. Webber, and S.I. Stupp, *Self-assembly of peptide amphiphiles: from molecules to nanostructures to biomaterials.* Biopolymers, 2010. **94**(1): p. 1-18.

19. Woolfson, D.N. and Z.N. Mahmoud, *More than just bare scaffolds: towards multi-component and decorated fibrous biomaterials.* Chem Soc Rev, 2010. **39**(9): p. 3464-79.

20. Brack, A. and L.E. Orgel, *Beta-Structures of Alternating Polypeptides and Their Possible Prebiotic Significance.* Nature, 1975. **256**(5516): p. 383-387.

21. Zhang, S.G., et al., *Spontaneous Assembly of a Self-Complementary Oligopeptide to Form a Stable Macroscopic Membrane.* Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(8): p. 3334-3338.

22. Yokoi, H., T. Kinoshita, and S.G. Zhang, *Dynamic reassembly of peptide RADA16 nanofiber scaffold.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(24): p. 8414-8419.

23. Caplan, M.R., et al., *Control of self-assembling oligopeptide matrix formation through systematic variation of amino acid sequence.* Biomaterials, 2002. **23**(1): p. 219-227.

24. Aggeli, A., et al., *Peptides modeled on the transmembrane region of the slow voltage-gated IsK potassium channel: Structural characterization of peptide assemblies in the beta-strand conformation.* Biochemistry, 1996. **35**(50): p. 16213-16221.

25. Aggeli, A., et al., *Responsive gels formed by the spontaneous self-assembly of peptides into polymeric beta-sheet tapes.* Nature, 1997. **386**(6622): p. 259-262.

26. Davies, R.P.W., et al., *Self-assembling beta-sheet tape forming peptides.* Supramolecular Chemistry, 2006. **18**(5): p. 435-443.

27. Aggeli, A., et al., *Hierarchical self-assembly of chiral rod-like molecules as a model for peptide beta-sheet tapes, ribbons, fibrils, and fibers.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(21): p. 11857-11862.

28. Rudra, J.S., et al., *A self-assembling peptide acting as an immune adjuvant.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(2): p. 622-627.

29. Jung, J.P., et al., *Modulating the mechanical properties of self-assembled peptide hydrogels via native chemical ligation.* Biomaterials, 2008. **29**(13): p. 2143-2151.

30. Jung, J.P., et al., *Co-assembling peptides as defined matrices for endothelial cells.* Biomaterials, 2009. **30**(12): p. 2400-10.

31. Vegners, R., et al., *Use of a gel-forming dipeptide derivative as a carrier for antigen presentation.* J Pept Sci, 1995. **1**(6): p. 371-8.

32. Jayawarna, V., et al., *Nanostructured hydrogels for three-dimensional cell culture through self-assembly of fluorenylmethoxycarbonyl-dipeptides.* Advanced Materials, 2006. **18**(5): p. 611-+.

33. Yang, Z.M., et al., *Conjugates of naphthalene and dipeptides produce molecular hydrogelators with high efficiency of hydrogelation and superhelical nanofibers.* Journal of Materials Chemistry, 2007. **17**(9): p. 850-854.

34. Kuang, Y., Y. Gao, and B. Xu, *Supramolecular hydrogelators of N-terminated dipeptides selectively inhibit cancer cells.* Chemical Communications, 2011. **47**(47): p. 12625-12627.

35. Niece, K.L., et al., *Modification of gelation kinetics in bioactive peptide amphiphiles.* Biomaterials, 2008. **29**(34): p. 4501-4509.

36. Hartgerink, J.D., E. Beniash, and S.I. Stupp, *Peptide-amphiphile nanofibers: A versatile scaffold for the preparation of self-assembling materials.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(8): p. 5133-5138.

37. Cui, H., et al., *Self-assembly of giant peptide nanobelts.* Nano Lett, 2009. **9**(3): p. 945-51.

38. Woolfson, D.N., *Building Fibrous Biomaterials From alpha-Helical and Collagen-Like Coiled-Coil Peptides.* Biopolymers, 2010. **94**(1): p. 118-127.

39. Kopecek, J. and J.Y. Yang, *Peptide-directed self-assembly of hydrogels.* Acta Biomaterialia, 2009. **5**(3): p. 805-816.

40. Shen, W., et al., *Tuning the erosion rate of artificial protein hydrogels through control of network topology.* Nature Materials, 2006. **5**(2): p. 153-158.

41. Faury, G., *Function-structure relationship of elastic arteries in evolution: from microfibrils to elastin and elastic fibres.* Pathologie Biologie, 2001. **49**(4): p. 310-325.

42. Pasquali-Ronchetti, I. and M. Baccarani-Contri, *Elastic fiber during development and aging.* Microsc Res Tech, 1997. **38**(4): p. 428-35.

43. Keeley, F.W., C.M. Bellingham, and K.A. Woodhouse, *Elastin as a self-organizing biomaterial: use of recombinantly expressed human elastin polypeptides as a model for investigations of structure and self-assembly of elastin.* Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 2002. **357**(1418): p. 185-189.

44. Bellingham, C.M., et al., *Recombinant human elastin polypeptides self-assemble into biomaterials with elastin-like properties.* Biopolymers, 2003. **70**(4): p. 445-455.

45. Urry, D.W., *Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers.* Journal of Physical Chemistry B, 1997. **101**(51): p. 11007-11028.

46. Mackay, J.A. and A. Chilkoti, *Temperature sensitive peptides: engineering hyperthermia-directed therapeutics.* Int J Hyperthermia, 2008. **24**(6): p. 483-95.

47. Meyer, D.E. and A. Chilkoti, *Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides.* Biomacromolecules, 2004. **5**(3): p. 846-51.

48. Arias, F.J., et al., *Tailored recombinant elastin-like polymers for advanced biomedical and nano(bio)technological applications.* Biotechnology Letters, 2006. **28**(10): p. 687-695.

49. Saxena, R. and M.J. Nanjan, *Elastin-like polypeptides and their applications in anticancer drug delivery systems: a review.* Drug Deliv, 2013.

50. Schneider, J.P., et al., *Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide.* Journal of the American Chemical Society, 2002. **124**(50): p. 15030-15037.

51. Pochan, D.J., et al., *Thermally reversible hydrogels via intramolecular folding and consequent self-assembly of a de Novo designed peptide.* Journal of the American Chemical Society, 2003. **125**(39): p. 11802-11803.

52. Ozbas, B., et al., *Salt-triggered peptide folding and consequent self-assembly into hydrogels with tunable modulus.* Macromolecules, 2004. **37**(19): p. 7331-7337.

53. Rajagopal, K., et al., *Tuning the pH Responsiveness of beta-Hairpin Peptide Folding, Self-Assembly, and Hydrogel Material Formation.* Biomacromolecules, 2009. **10**(9): p. 2619-2625.

54. Yucel, T., et al., *Direct observation of early-time hydrogelation in beta-hairpin peptide self-assembly.* Macromolecules, 2008. **41**(15): p. 5763-5772.

55. Haines-Butterick, L., et al., *Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(19): p. 7791-7796.

56. Veerman, C., et al., *Gelation kinetics of beta-hairpin peptide hydrogel networks.* Macromolecules, 2006. **39**(19): p. 6608-6614.

57. Yan, C., et al., *Injectable solid hydrogel: mechanism of shear-thinning and immediate recovery of injectable beta-hairpin peptide hydrogels.* Soft Matter, 2010. **6**(20): p. 5143-5156.

58. Yan, C., et al., *Injectable Solid Peptide Hydrogel as a Cell Carrier: Effects of Shear Flow on Hydrogels and Cell Payload.* Langmuir, 2012. **28**(14): p. 6076-6087.

59. Kretsinger, J.K., et al., *Cytocompatibility of self-assembled ss-hairpin peptide hydrogel surfaces.* Biomaterials, 2005. **26**(25): p. 5177-5186.

60. Sallach, R.E., V.P. Conticello, and E.L. Chaikof, *Expression of a Recombinant Elastin-Like Protein in Pichia pastoris.* Biotechnology Progress, 2009. **25**(6): p. 1810-1818.

61. Bzymek, M. and S.T. Lovett, *Instability of repetitive DNA sequences: The role of replication in multiple mechanisms.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(15): p. 8319-8325.

62. Zhang, J., et al., *Expression and characterization of antimicrobial peptide ABP-CM4 in methylotrophic yeast Pichia pastoris.* Process Biochemistry, 2006. **41**(2): p. 251-256.

63. Li, B.C., et al., *Expression in Escherichia coli and purification of bioactive antibacterial peptide ABP-CM4 from the Chinese silk worm, Bombyx mori.* Biotechnology Letters, 2007. **29**(7): p. 1031-1036.

64. Chen, Y.Q., et al., *Expression of a cytotoxic cationic antibacterial peptide in Escherichia coli using two fusion partners.* Protein Expression and Purification, 2008. **57**(2): p. 303-311.

65. Xu, X.X., et al., *Expression and purification of a recombinant antibacterial peptide, cecropin, from Escherichia coli.* Protein Expression and Purification, 2007. **53**(2): p. 293-301.

66. Li, Y.F., X. Li, and G.S. Wang, *On-resin cleavage of bacterially expressed fusion proteins for purification of active recombinant peptides SK-29, KR-20, LL-29, and LL-23 from human sweat or skin.* Protein Expression and Purification, 2007. **55**(2): p. 395-405.

67. Moon, W.J., et al., *Recombinant expression, isotope labeling, refolding, and purification of an antimicrobial peptide, piscidin.* Protein Expression and Purification, 2007. **51**(2): p. 141-146.

68. Metlitskaia, L., et al., *Recombinant antimicrobial peptides efficiently produced using novel cloning and purification processes.* Biotechnology and Applied Biochemistry, 2004. **39**: p. 339-345.

69. Kim, J.M., et al., *High-level expression of an antimicrobial peptide histonin as a natural form by multimerization and furin-mediated cleavage.* Applied Microbiology and Biotechnology, 2008. **78**(1): p. 123-130.

70. Sharpe, S., W.M. Yau, and R. Tycko, *Expression and purification of a recombinant peptide from the Alzheimer's beta-amyloid protein for solid-state NMR.* Protein Expr Purif, 2005. **42**(1): p. 200-10.

71. Reed, D.C., et al., *Production and purification of self-assembling peptides in Ralstonia eutropha.* Protein Expr Purif, 2006. **46**(2): p. 179-88.

72. van Hell, A.J., et al., *Self-assembly of recombinant amphiphilic oligopeptides into vesicles.* Biomacromolecules, 2007. **8**(9): p. 2753-61.

73. Morreale, G., et al., *Bioprocess-centered molecular design (BMD) for the efficient production of an interfacially active peptide.* Biotechnol Bioeng, 2004. **87**(7): p. 912-23.

74. Hartmann, B.M., et al., *The Chromatography-Free Release, Isolation and Purification of Recombinant Peptide for Fibril Self-Assembly.* Biotechnology and Bioengineering, 2009. **104**(5): p. 973-985.

75. Riley, J.M., et al., *Bioproduction and Characterization of a pH Responsive Self-Assembling Peptide.* Biotechnology and Bioengineering, 2009. **103**(2): p. 241-251.

76. Kyle, S., et al., *Recombinant self-assembling peptides as biomaterials for tissue engineering.* Biomaterials, 2010. **31**(36): p. 9395-9405.

77. Prakash, A., et al., *Recombinant production of self-assembling beta-structured peptides using SUMO as a fusion partner.* Microb Cell Fact, 2012. **11**: p. 92.

Chapter 2

DESIGN OF SELF-ASSEMBLING PEPTIDES THAT CAN BE PREPARED BY FERMENTATION

2.1 Introduction

The extensively studied beta-hairpin peptide MAX1 contains a D-residue in its turn region and C-terminal amide which are functionalities not found in natural peptides. In bacteria, D-amino acids are synthesized by racemization, epimerization or stereospecific amination of α-ketoacids and they are mostly found in biomolecules such as peptidoglycan, teichoic acids and poly-γ-glutamate in extracellular or periplasmic space [1]. D-amino acids are not used in translation but can be incorporated into peptides for antibiotic synthesis by non-ribosomal peptide synthetases (NRPS). Bacteria produce nonribosomal peptides in the secondary metabolism by employing large multimodular NRPSs [2, 3]. During this process, not only 20 standard amino acids are used but also non-standard amino acids including methylated, hydroxylated or D-amino acids as well as lipids and carbohydrates can be incorporated to the peptides that can be in linear, branched, cyclic or poly-cyclic forms [4, 5]. The modular composition of NRPSs includes discrete enzymes each of which facilitates the addition of a specific monomer unit [6]. During recombinant production, bacterial ribosomal synthesis machinery is employed and MAX1 cannot be expressed

by this machinery due to having a D-proline in its sequence. Therefore, MAX1 must be prepared chemically.

The sequence of MAX1 was used as a design template and, the sequence changes to be performed on MAX1 to afford peptides amenable to bacterial expression and able to afford self-supporting hydrogels at physiological conditions (pH 7.4, 50 mM BTP, 150 mM NaCl, 37°C), are depicted in Figure 2.1. First, turn regions containing L-amino acids were incorporated into the MAX1 template. Next, the C-terminal carboxyl functionality found in natural peptides was incorporated, replacing the C-terminal amide of MAX1. Further change by substitution of lysine at position 15 on MAX1 to glutamic acid to better facilitate folding and self-assembly leading to stiffer hydrogels was implemented. New sequences were chemically synthesized to quickly assess their folding, assembly and material forming properties before initiating bacterial expression studies.

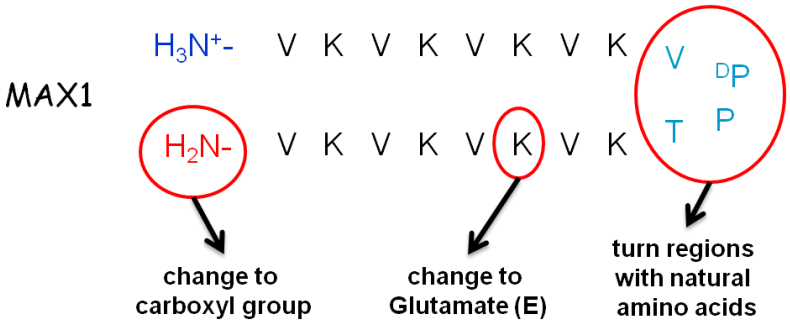


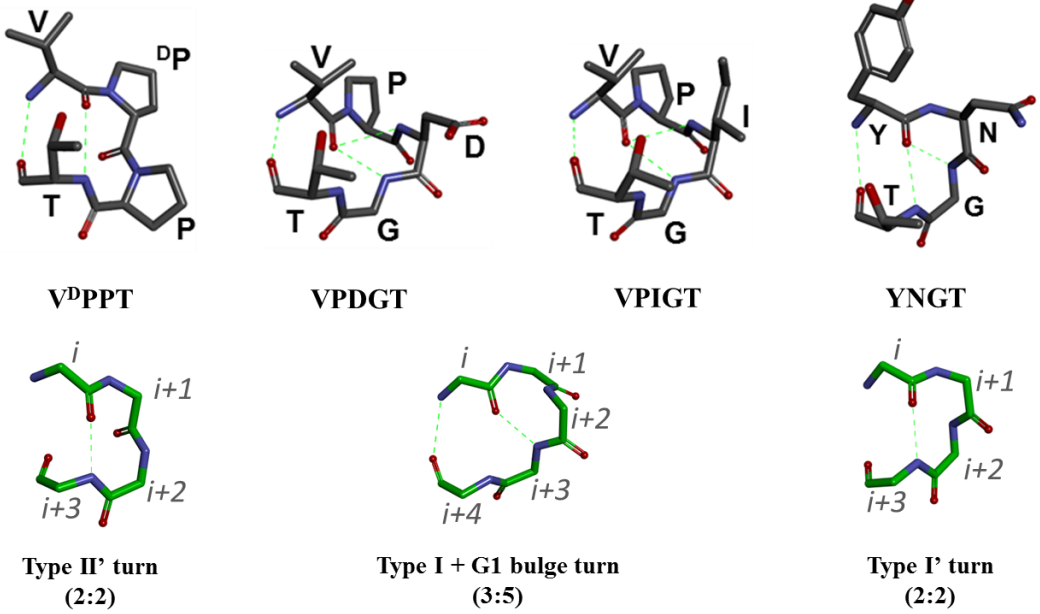
Figure 2.1 Peptide engineering on MAX1. Changes were performed on the turn region, C-terminus and lysine at position 15.

Natural L-amino acid containing turn structures used to replace the unnatural D-amino acid containing MAX1 turn structure are shown in Figure 2.2A with their putative turn types and hydrogen bonding patterns. The classification of turn types are mainly due to the backbone dihedral angles (Φ and Ψ) of turn residues at the *i+1* and *i+2* positions of the turn [7-11]. The new peptide sequences compared with the MAX1 sequence are shown in Figure 2.2B. The *i*+1 position of the MAX1 turn $V^{D}PPT$ has a D-Proline whose dihedral angles are restricted due to the covalent constraints of pyrrolidine ring formation and any L-amino acid in the *i*+2 position. These constrains strongly favor the formation of a stable type II` turn [12-15]. Figure 2.2A shows the stabilizing hydrogen bonding between the amide hydrogen and carbonyl oxygen of Valine and Threonine at the *i* and *i*+3 positions of the type II` turn, respectively. Previously, a former Schneider Lab member, Karthikan Rajagopal studied natural L-amino acid containing turn sequences where VPDGT and YNGT sequences were used to replace the $V^{D}PPT$ turn of MAX1. The resulting peptides were characterized at alkaline and high temperature conditions (pH 9 and 60^{o}C) with respect to their folding and self-assembly [16]. The turn sequence VPDGT was adapted from modification of an isolated native ubiquitin N-terminal sequence that forms a stable monomeric β-hairpin that was reported to adopt a type I + G1 bulged turn (NPDGT) [17]. The PDG sequence contained within this 5-residue turn was also shown to strongly promote a type I + G1 bulged turn in other reports [18-21]. Figure 2.2A shows the stabilizing hydrogen bonding formed between the amide hydrogen of Valine at position *i* and

carbonyl oxygen of Threonine at position *i*+4 as well as carbonyl oxygen of Valine at position *i* and amide hydrogen of Glycine at position *i*+3.

The YNGT turn sequence contains an Asn-Gly dipeptide at its central position and facilitates the formation of a type I` β-turn [22-25]. The hydrogen bonding pattern between Tyrosine at position *i* and Threonine at position *i*+3 is also shown in Figure 2.2A. For the third turn sequence VPIGT, the aspartate of VPDGT was replaced with isoleucine making the turn region more hydrophobic compared to the negatively charged hydrophilic aspartate containing sequence. The VPIGT turn is also expected to have a similar hydrogen bonding pattern to that of VPDGT that should stabilize the turn.

(A)



$V^{D}PPT$ | **VPDGT** | **VPIGT** | **YNGT**

Type II' turn (2:2) | **Type I + G1 bulge turn (3:5)** | **Type I' turn (2:2)**

(B)

MAX1	**H_3N^+ - VKVKVKVK $V^{D}PPT$ KVKVKVKV - NH_2**
MAX1(VPDGT)	**H_3N^+ - VKVKVKVK VPDGT KVKVKVKV - NH_2**
MAX1(VPIGT)	**H_3N^+ - VKVKVKVK VPIGT KVKVKVKV - NH_2**
MAX1(YNGT)	**H_3N^+ - VKVKVKVK YNGT KVKVKVKV - NH_2**
MAX1(VPDGT)-C	**H_3N^+ - VKVKVKVK VPDGT KVKVKVKV - COO^-**
EX1	**H_3N^+ - VKVKVKVK VPDGT KVEVKVKV - COO^-**
EX2	**H_3N^+ - VKVKVKVK VPIGT KVEVKVKV - COO^-**
EX3	**H_3N^+ - VKVKVKVK YNGT KVEVKVKV - COO^-**

Figure 2.2 Structures of MAX1 turn variants and peptide sequences. (A) Turn structures, turn backbone structures with residue numbers and turn types are shown in addition to turn region hydrogen bonding patterns. (B) Peptide sequences generated in this study.

Ultimately, a peptide amenable to bacterial expression must contain all L-residues and a free carboxy-terminus. The design strategy employed in this chapter is

to first replace the turn sequence of MAX1 with turns containing all L-residues. Initially, the replacement of the turn sequence was done strictly in the context of MAX1; that is, in the first generation of sequences, the C-terminal amide of MAX1 was left intact to assess the effect of altering the turn structure only. After studying these peptides (MAX1(VPDGT); MAX1(VPIGT); and MAX1(YNGT), Figure 2.2B), the effect of changing the carboxy amide to a free carboxylate was assessed on the MAX1(VPDGT) peptide alone. After this, I also studied the effect of changing the lysine at position 16, to a glutamate resulting in peptide EX1 based on the rational given below.

MAX1 was designed to form a rigid hydrogel in a pH dependent manner in which alkaline solution conditions are used as a folding trigger [26]. At near neutral pH (pH 7.4) and low ionic strength, the lysines of MAX1 (pK_a ~10.5) are protonated keeping the peptide in its unfolded state due to side chain charge repulsion. However, increasing the ionic strength of the solution using NaCl (150 mM) [27] can function as a folding trigger by screening the positive charges. In fact, increasing the concentration of NaCl beyond 150 mM results in an increase in the rate of folding, self-assembly and fibril network evolution. An alternative way to increase the folding kinetics is to design peptides with less positive net charge, which needs to be screened in order to initiate folding. Previously, it was shown that modification of MAX1`s net charge by substituting lysine with a glutamic acid lead to faster rates of folding and assembly, and hydrogels with enhanced mechanical rigidity at a specific pH [28]. In

fact, this substitution enabled hydrogelation to take place at pH 7.4 and 150 mM NaCl, conditions allowing physiological applications of the gel [29].

Finally, this lysine-to-glutamic acid substitution was also made in the context of the other two turn types resulting in peptides EX2 and EX3, Figure 2.2. For all the peptides, folding and assembly was assessed by circular dichroism following the evolution of β-sheet structure. Oscillatory rheology was used to investigate the mechanical properties of their resulting gels.

2.2 Varying the Turn Sequence in the Context of MAX1

The turn sequences consisting of L-amino acids (VPDGT, VPIGT and YNGT) were incorporated into sequence of MAX1, ultimately to derive sequences amenable to bacterial expression. These peptides were initially chemically synthesized to quickly assess their propensity to fold and assemble. The rheological properties of their corresponding gels at physiologically relevant conditions (50 mM BTP, 150 mM NaCl, pH 7.4, 37^{o}C) was also studied.

2.2.1 Characterization of MAX1 turn variants at physiological buffer conditions

The transition from random coil structure to beta-sheet structure was investigated via CD spectroscopy using dilute concentrations of peptides (150 μM).

MAX1(VPDGT), MAX1(VPIGT) and MAX1(YNGT) with C-terminal amidation were first studied to assess their potential to fold at physiologically relevant pH and NaCl concentrations. Figure 2.3A shows the mean residue ellipticity of the three peptides in addition to comparison with that of MAX1 at 216 nm, which is the characteristic wavelength indicative of β-sheet formation, as a function of temperature. The peptide concentration used in panel A was 150 μM. At this low concentration, if the sequence permits, the peptide will fold and self-assemble into β-sheet rich soluble fibrils, but not gel. As can be seen in the Figure, all peptides exist in the random coil conformation and don't exhibit any structural transformation into beta sheet conformation even at higher temperatures. The slight change at the 216 nm signal from 5^{o}C to 80^{o}C in Figure 2.3A might be associated with slight conformational changes of the turn region amino acids as the temperature increases. However, the data clearly show that these peptides remain unfolded under these conditions.

Due to the co-operativity between folding and self-assembly proceeding to hydrogel formation, β-sheet formation is concentration dependent [30]. Thus, the secondary structure of the peptides was also investigated at 1 wt% concentration at 37^{o}C using the same buffer conditions. Figure 2.3B shows that even at higher concentration, these peptides are incapable of folding and self-assembling at 37^{o}C. However, 0.5 wt% of MAX1 at 35^{o}C (pH 7.4, 50 mM BTP, 150 mM NaCl) is able to fold into beta sheet structure shown by the characteristic minima at 216 nm in

wavelength spectrum in Figure 2.3C. This lack of hairpin formation of the MAX1 turn variants could be likely due to the lower turn propensities of the VPDGT, VPIGT and YNGT sequences compared to the MAX1 V^{D}PPT turn. As will be shown later, other design changes can be made to accommodate these lower propensity turns to produce peptides capable of folding and self-assembling.

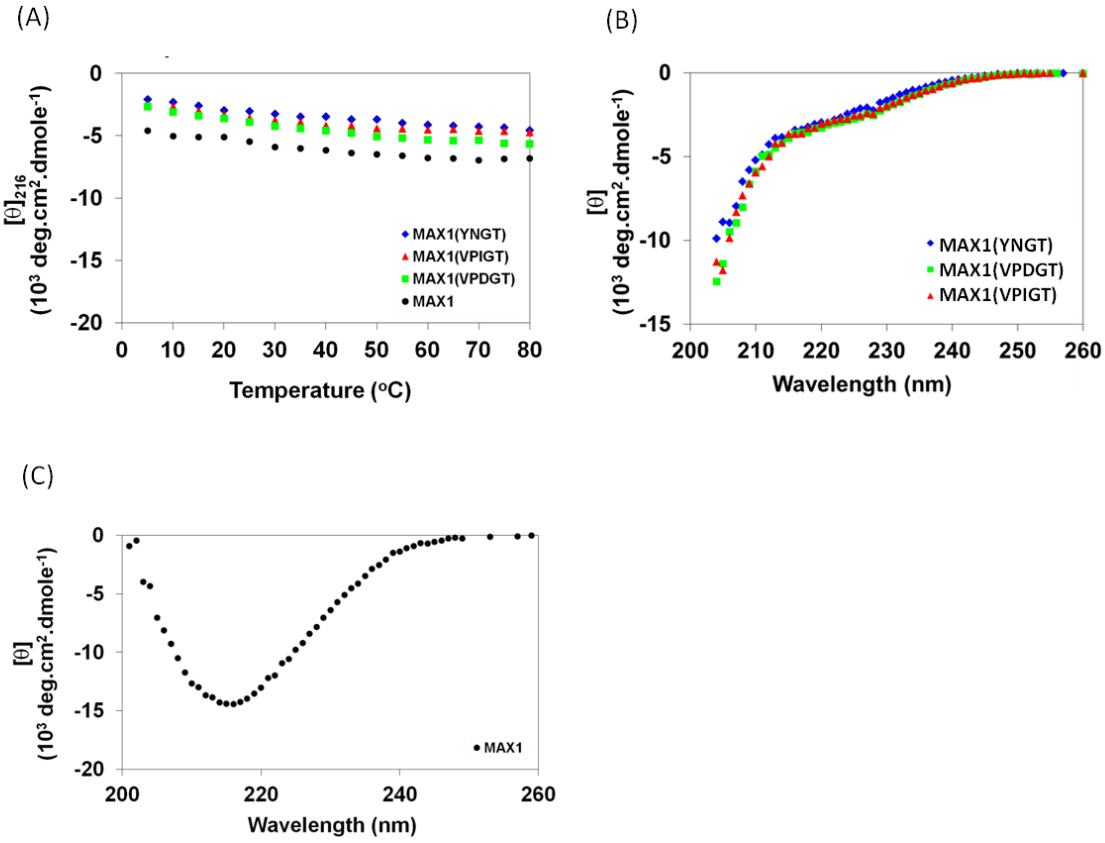


Figure 2.3 Temperature dependent CD spectroscopy. (A) Temperature dependent β-sheet formation of 150 μM MAX1(VPDGT), MAX1(VPIGT), MAX1(YNGT) and MAX1 by monitoring mean residue ellipticity at 216nm from 5°C to 80°C. (B) Wavelength spectra of 1 wt% MAX1(VPDGT), MAX1(VPIGT) and MAX1(YNGT) at 37°C after 1 hour. (C) Wavelength spectra of 0.5 wt% MAX1 at 35°C. (pH 7.4, 50 mM BTP, 150 mM NaCl).

MAX1(VPDGT), MAX1(VPIGT) and MAX1(YNGT) were expected not to form rigid hydrogels at 1 wt% concentration since the CD data at the same concentration (1 wt%) in the same buffer showed that the peptides do not fold (Figure 2.3B). The dynamic time sweep experiment shown in Figure 2.4, confirmed that the MAX1 turn variants do not form rigid hydrogels and remained viscous at 1 wt% concentration at physiologically relevant conditions. All three peptides form viscous solutions having very low storage moduli around 25 Pa which did not increase as a function of time.

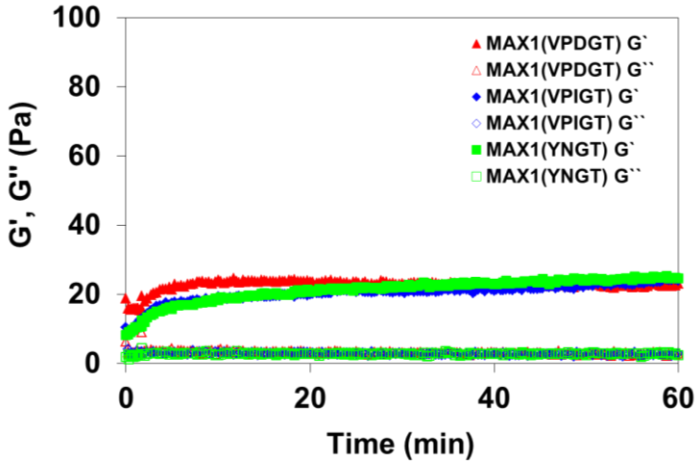


Figure 2.4 Mechanical properties of MAX1 turn variants by oscillatory rheology. Dynamic time sweep experiment showing the evolution of storage (G`) and loss (G``) moduli of 1 wt% peptides at pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C .

2.3 Change of Carboxyl C-terminus on MAX1(VPDGT)

Bacterial expression of these peptides will yield sequences having free C-terminal carboxy groups as opposed to the C-terminal amides derived from solid phase peptide synthesis employing PAL-based resins. Changing the C-terminal amide to a C-terminal carboxyl group on MAX1(VPDGT) not only mimics the recombinant peptide that would be produced by bacterial culturing but also decreases the net positive charge by 1 unit (Figure 2.5). The net charge of the peptide changes from +8 to +7 when the carboxy terminus is considered. The change in overall peptide charge might affect the characteristics of the peptide as discussed before. This change in C-terminal functionality was studied in the context of the MAX1(VPDGT) peptide, Figure 2.5.

MAX1(VPDGT) H_3N^+ - VKVKVKVK VPDGT KVKVKVKV - NH_2

MAX1(VPDGT)-C H_3N^+ - VKVKVKVK VPDGT KVKVKVKV - COO^-

Figure 2.5 Comparison of MAX1(VPDGT) and MAX1(VPDGT)-C sequences. The C-terminal amide of MAX1(VPDGT) was changed to a C-terminal carboxyl on MAX1(VPDGT)-C.

2.3.1 Characterization of MAX1(VPDGT)-C

The decrease in the net charge of the peptide did not alter its inability to fold at physiological conditions at 1wt % concentration (Figure 2.6). The wavelength spectra in Figure 2.6 clearly indicate that peptide remains unfolded at 37^{o}C.

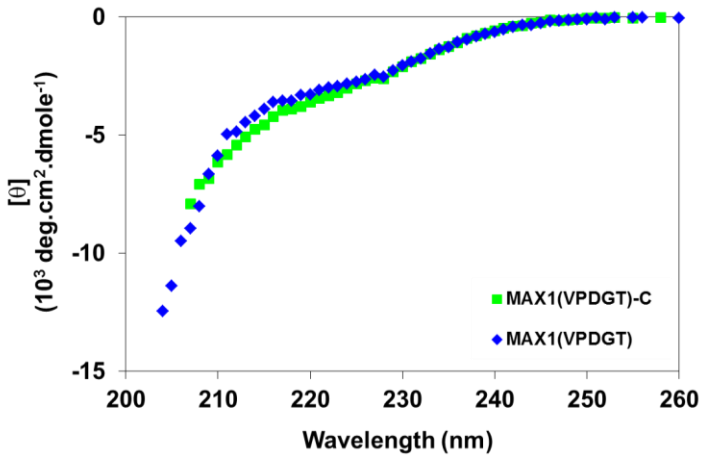


Figure 2.6 Wavelength CD spectra. 1 wt% MAX1(VPDGT) and MAX1(VPDGT)-C at 37^{o}C after 1 hour (pH 7.4, 50 mM BTP, 150 mM NaCl).

Although MAX1(VPDGT)-C remains unfolded at 1 wt% concentration at physiological conditions shown by Figure 2.6, rheology data given in Figure 2.7 showed that it was able to form a weak hydrogel. The storage modulus of the MAX1(VPDGT)-C gel was approximately 100 Pa while MAX1(VPDGT) remained a viscous solution at the end of 1 hour (Figure 2.7). This indicates that some form of cross-linked network was formed. The negative charges introduced at the turn region

aspartate and C-terminal carboxyl of the peptide could be forming intermolecular salt bridges with both the positively charged lysine residues as well as the amine group at the N-terminus resulting in the formation of an extended network of associated peptides capable of gelling the solution at high peptide concentrations. Because there was no β-sheet formation detected by CD, the mechanism of MAX1(VPDGT)-C hydrogel formation did not support the proposed hydrogelation mechanism explained in Chapter 1 and was not studied further.

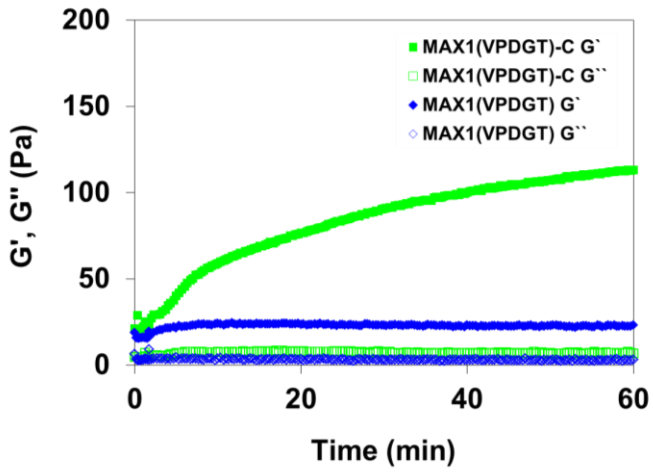


Figure 2.7 Mechanical properties by oscillatory rheology. Dynamic time sweep experiment showing the evolution of storage (G`) and loss (G``) moduli of 1wt% peptides at pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C .

2.4 Change of Lysine to Glutamic Acid at the Hydrophilic Face of MAX1(VPDGT)-C

In previous studies, the lysine at position 15 of MAX1 was replaced with a glutamate and the gels formed from the resulting peptide MAX1(K15E) were characterized and studied extensively for its use in cellular and protein encapsulation and injectable delivery [28, 29, 31]. Correspondingly, the lysine at position 16 of MAX1(VPDGT)-C was changed to glutamate and a peptide suitable for expression, EX1, was obtained (Figure 2.8).

MAX1(VPDGT)-C	**H_3N^+ - VKVKVKVK VPDGT KVKVKVKV - COO^-**
EX1	**H_3N^+ - VKVKVKVK VPDGT KVEVKVKV - COO^-**

Figure 2.8 Comparison of MAX1(VPDGT)-C and EX1 sequences. EX1 was derived from MAX1(VPDGT)-C by replacing the lysine at position 16 with glutamate.

2.4.1 Characterization of EX1

Further decrease in net charge by changing the lysine to glutamate yielded EX1 that can fold and self-assemble into β-sheet beyond 30^{o}C as characterized by temperature dependent CD at 150 μM concentration as shown in Figure 2.9A where the mean residue ellipticity was given as a function of temperature with comparison to MAX1(VPDGT) and MAX1(VPDGT)-C. At physiological conditions, 1 wt% EX1

was also characterized and the wavelength CD spectrum showed a minima at 216 nm confirming the formation of β-sheet structure (Figure 2.9B).

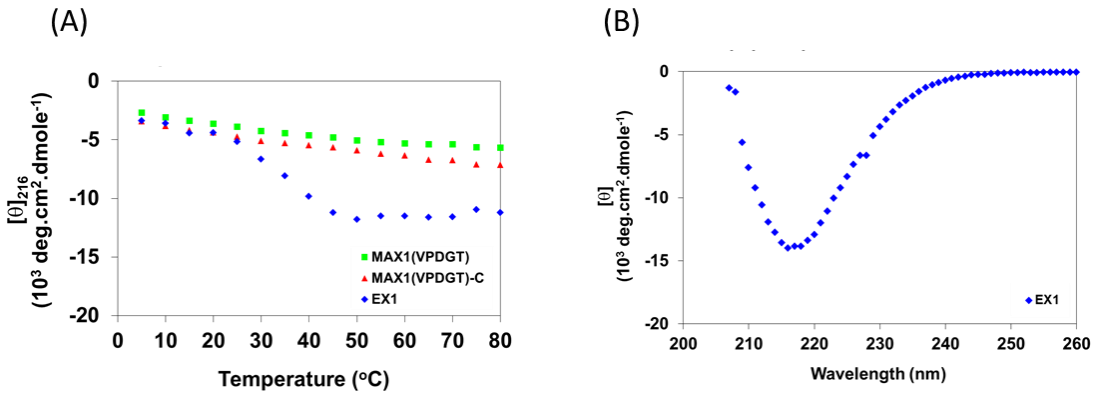


Figure 2.9 Temperature dependent CD spectroscopy (A) Temperature dependent β-sheet formation of 150 μM peptide solutions by monitoring mean residue ellipticity at 216 nm from 5°C to 80°C (pH 7.4, 50 mM BTP, 150 mM NaCl). (B) Wavelength spectra of 1 wt% EX1 at 37°C after 1 hour (pH 7.4, 50 mM BTP, 150 mM NaCl).

Substitution of glutamic acid in the peptide sequence significantly enhanced the self-supporting hydrogel forming properties of EX1. The gel formed from the peptide has a storage modulus around 1 kPA after 1 hour at physiologically relevant conditions (Figure 2.10A). This considerable difference in mechanical rigidity between EX1 and MAX1(VPDGT)-C may be due to the formation of cross-strand salt bridges between lysine and glutamate residues and a higher number of cross-links and fibril entanglements as a consequence of its faster assembling rate. Similar differences in gel mechanical properties were reported for gels prepared from MAX1 and MAX1(K15E) [29]. Figure 2.10B shows the shear-thinning and recovery properties of

EX1. During high strain (1000%), the hydrogel became completely viscous observed by the immediate decrease in G` and when the strain was reduced to 0.2%, the gel recovered it`s storage modulus by ~50% in a short time. When physical (non-covalent) cross-links are broken, the hydrogel shows liquid-like behavior and as the strain is decreased, physical cross-links are reformed. In the case of EX1 hydrogel recovery, the gel was not capable of restoring its rigidity to that of the original gel. The reason for this is not known. At any rate, the gel does not recover substantially to allow syringe based delivery. Dynamic frequency and strain sweep experiments (Figure 2.10 C and D) after 1 hour of gelation shows that the storage modulus of EX1 is not frequency dependent in the range of 0.1-100 rad/sec and a 1 wt% gel yields around 6% strain.

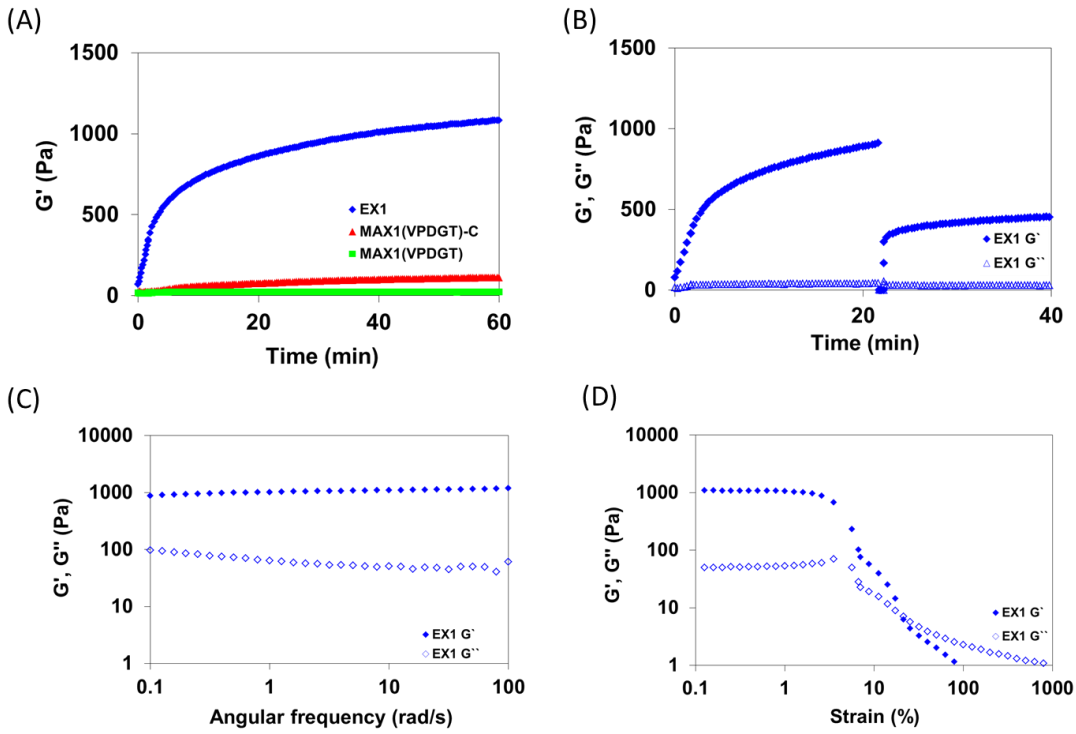


Figure 2.10 Mechanical properties by oscillatory rheology (A) Dynamic time sweep experiment showing the evolution of G` (storage modulus) of 1wt% peptides at pH 7.4, 50 mM BTP, 150 mM NaCl, 37°C . (B) Shear-thinning and recovery behavior of EX1. After 20 min of gelation, strain was increased to 1000% for 30 sec and back to 0.2% for another 20 min. Partial recovery of the elastic properties was observed. Linear viscoelastic regime of EX1 hydrogel shown by (C) dynamic frequency sweep and (D) dynamic strain sweep experiments.

2.5 C-terminal Carboxyl and K to E Change on MAX1(VPIGT) and MAX1(YNGT)

Peptide engineering on MAX1 yielded EX1, a peptide that is amenable to bacterial production that can fold and self-assemble into β-sheet structure and form a self-supporting hydrogel at physiological conditions. Although changing the C-

terminus of MAX1(VPDGT) to include a carboxyl group decreased the net charge of the peptide by 1 unit, the resulting peptide MAX1(VPDGT)-C was unable to fold and assemble and form a rigid hydrogel. Further reduction in the net change of the peptide by replacing a lysine residue with a glutamate, and the possibility of forming network stabilizing electrostatic interactions between glutamate and lysine residues, for the peptide EX1 resulted in a mechanically rigid, self-supporting hydrogel rich in β-sheet structure. Thus, both of these changes(e.g. c-terminal amide → carboxyl and Lys → Glu) were incorporated in the context of MAX1(VPIGT) and MAX1(YNGT) resulting in the peptides EX2 and EX3, respectively.

MAX1(VPIGT)	**H_3N^+ - VKVKVKVK VPIGT KVKVKVKV - NH_2**
EX2	**H_3N^+ - VKVKVKVK VPIGT KVEVKVKV - COO^-**
MAX1(YNGT)	**H_3N^+ - VKVKVKVK YNGT KVKVKVKV - NH_2**
EX3	**H_3N^+ - VKVKVKVK YNGT KVEVKVKV - COO^-**

Figure 2.11 Comparison of EX2 and EX3 peptide sequences. EX2 was derived from MAX1(VPIGT) and EX3 was derived from MAX1(YNGT) by changing the C-terminal amide to a carboxyl as well as changing the second lysine after the turn to glutamic acid.

2.5.1 Characterization of EX2 and EX3

Temperature dependent CD spectra of 150 μM EX2 and EX3 at 216 nm showed that at 5^{o}C, EX2 remained random coil whereas EX3 was partially folded (Figure 2.12A). As the temperature increases, EX3 folded into β-sheet structure at low

temperatures and retained its folded structure as the temperature increased. On the other hand, EX2 did not show any significant folding at dilute concentrations and stayed mostly unstructured even at higher temperatures (Figure 3.11A). Wavelength CD spectra of both peptides at 1 wt% at physiological conditions showed that the peptides folded in β-sheet structure by the observation of characteristic minima at 216 nm (Figure 3.11B). This result shows that higher concentrations of EX2 are needed to facilitate folding while EX3 adopts the folded state even at dilute concentrations.

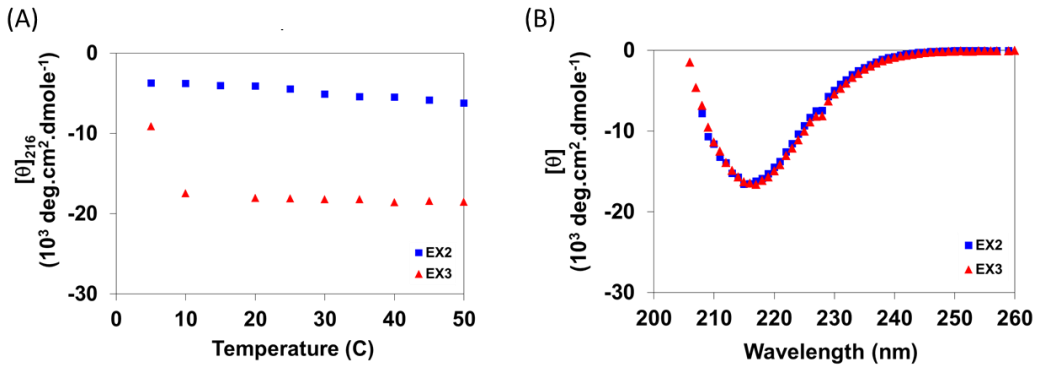


Figure 2.12 Temperature dependent CD spectroscopy (A) Temperature dependent β-sheet formation of 150 μM EX2 and EX3 by monitoring mean residue ellipticity at 216 nm from 5°C to 50°C (pH 7.4, 50 mM BTP, 150 mM NaCl). (B) Wavelength spectra of 1wt% EX2 and EX3 at 37°C after 1 hour (pH 7.4, 50 mM BTP, 150 mM NaCl).

The mechanical rheological properties of gels formed from EX2 and EX3 are shown in Figure 2.13 and are notably different when compared to gels formed from MAX1(VPIGT) and MAX1(YNGT). Storage moduli of 1 wt% EX2 and EX3 gels at 37°C at physiologically relevant conditions are more than an order of magnitude

greater (~1000 Pa vs. ~25 Pa) than gels formed by MAX1(VPIGT) and MAX1(YNGT). C.D. spectroscopy shows that EX3 has a greater propensity to fold and assemble as compared to EX2 (Figure 2.12A) and correspondingly, the time-sweep rheological experiments (Figure 2.13A) show that EX3 forms gels more quickly. The short lag time in the evolution of storage modulus at the beginning of rheology experiment shown in Figure 2.13A corresponds to the temperature ramp from 5^{o}C to 37^{o}C in the first 100 sec of the experiment. The greater lag time evident for EX2 is likely due to the requirement of a higher temperature needed to initiate gelation for this peptide compared to EX3. Although the turn region of EX2 (VPIGT) is more hydrophobic than the turn region of EX3 (YNGT), the turn propensity, rate of folding, and assembly for EX3 is faster than that of EX2. Figure 2.13B shows that 1 wt% EX2 recovers its elastic properties around 80% after it was shear-thinned whereas EX3 at the same conditions recovered more than 90% of its storage modulus. In figure 2.13C and D, dynamic frequency sweep experiments show that the storage moduli (G`) of the EX2 and EX3 gels are independent of the changes in the frequency and both hydrogels at 1 wt% concentration yield at more than 10% strain.

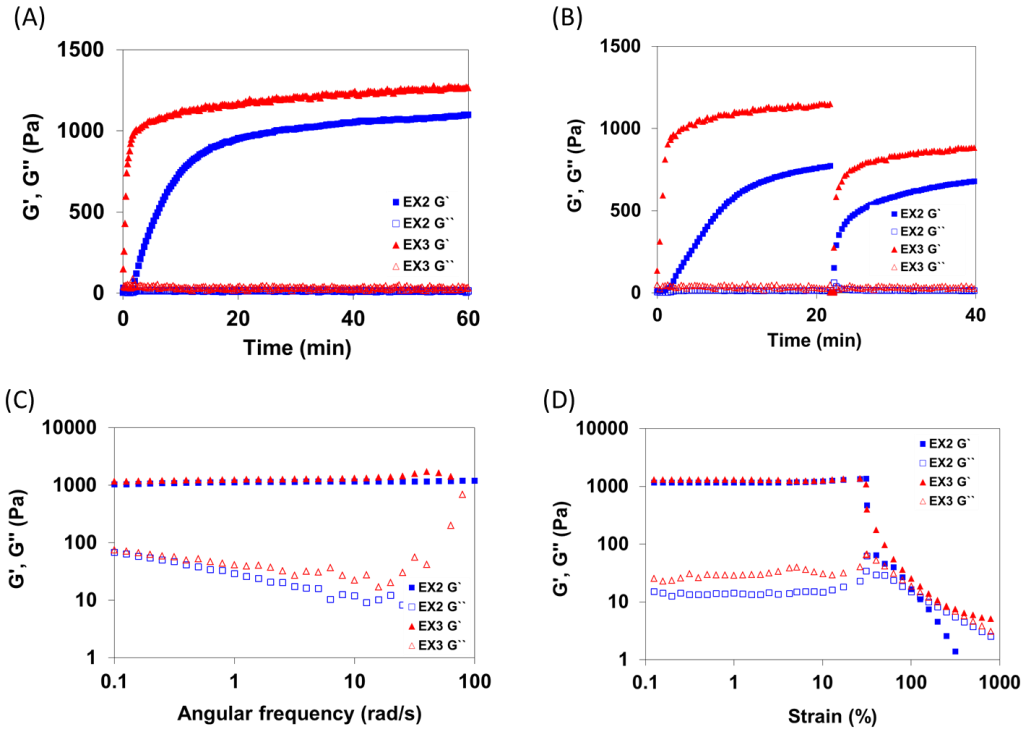


Figure 2.13 Mechanical properties of 1 wt% EX2 and EX3 by oscillatory rheology (A) Dynamic time sweep experiment showing the evolution of storage (G`) and loss (G``) moduli of 1 wt% peptides at pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C. (B) Shear-thin and recovery experiment with 1000% strain for 30 sec after 20 min of gelation followed by another 20min for recovery. (C) Dynamic frequency sweep and (D) Dynamic strain sweep experiments of 1 wt% peptide in pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C.

Local fibril morphologies of EX1, EX2 and EX3 were investigated by transmission electron microscopy (TEM) (Figure 2.14). Individual fibril widths were measured and monodispersity around 3.5 nm was observed. Fibril widths are in good agreement with the width of a folded hairpin in the self-assembled state as assessed by modeling. A slight degree of lamination was observed for EX1 fibrils and some fibril bundling was observed for EX2 and EX3. Interactions between the fibrils might be facilitated by residue-residue interactions in the self-assembled state. For example, the

negatively charged aspartate in the turn region of one EX1 molecule may be forming electrostatic interactions with the positively charged N-terminus or the hydrophilic lysine rich face of another folded hairpin molecule. Isoleucines displayed on the fibril edges due to their presence in the turn region of EX2 might also form hydrophobic contacts between two fibrils and such an interaction might result in fibril bundling. Another possibility is that asparagines originating from distinct turns could facilitate hydrogen bonding between their side chains facilitating interaction between two fibril edges which could result in bundling. The TEM data suggests that the variety of amino acids that can be incorporated into the turn region affects not only the turn propensity but also interfibrillar interactions.

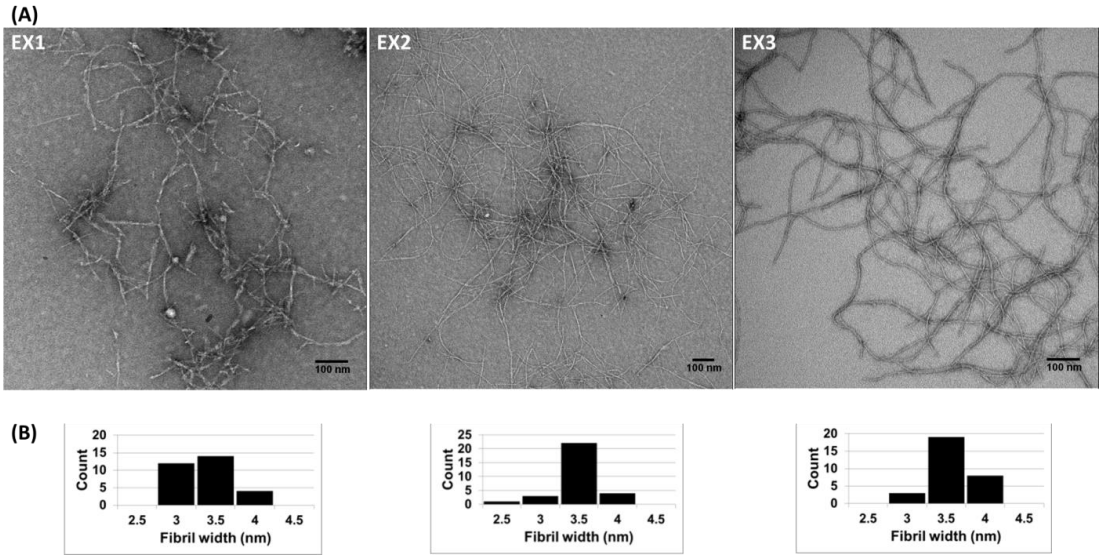


Figure 2.14 Local morphology of EX fibrils assessed by TEM. 1 wt% peptides were gelled at 37^{o}C for 1 hour. (A) Gels were diluted and deposited for TEM interrogation. (B) Individual fibril width distributions. TEM images were taken by Katelyn Nagy.

2.6 Conclusions

Peptide engineering performed on the parent peptide MAX1 afforded beta-hairpin peptides that can form rigid hydrogels rich in beta sheet structure at physiological conditions as well as amenable for bacterial production. Candidate peptides were characterized for their secondary structure and all EX peptides were found to fold into β-sheet structure evidenced by the wavelength CD spectra giving the characteristic minima at 216 nm at physiologically relevant conditions. The mechanical properties of the peptides` resultant gels were characterized by oscillatory rheology and the storage moduli of EX1, EX2 and EX3 were found to be greater than 1 kPa after 1 hour at 37^{o}C showing that each peptide formed self-supporting hydrogels. All three EX peptide gels can recover to some degree after being shear-thinned. The morphology of the fibrils formed by the self-assembling peptides was characterized by TEM. This data suggests that the EX peptides undergo similar folding and self-assembling mechanisms resulting in the formation of a fibril network that constitutes the formation of mechanically rigid gels. Individual fibril widths were measured and correspond to the measured width of a folded hairpin (3.5 nm) in the self-assembled state. Overall, three new sequences were generated and shown to be good candidates for bacterial expression studies.

2.7 Experimental Section

2.7.1 Solid phase peptide synthesis and purification

An automated ABI 433A peptide synthesizer was used to synthesize the peptides. Fmoc-based solid peptide chemistry with HCTU activation was performed. A trifluoroacetic acid: thioanisole: 1,2-ethanedithiol: anisole (90:5:3:2) cocktail under Argon atmosphere for 2 h was used to cleave the dried resin-bound peptides from the resin and for simultaneous side-chain deprotection. Precipitation by cold diethyl ether and lyophilization of crude peptides were followed by peptide purification that was carried out by reverse phase-HPLC at 40^{o}C equipped with a semi-preparative Vydac C18 column. HPLC solvents A and B were 0.1% TFA in water and 0.1% TFA in 9:1 acetonitrile: water, respectively. For MAX1(VPDGT) purification, linear gradients were employed as 0% solvent B over 1 min, 0-11% solvent B over 6 min, 11-18% solvent B over 14 min and then 18-33% solvent B over 60 min. For MAX1(VPIGT) purification, linear gradients were employed as 0% solvent B over 2 min, 0-13% solvent B over 13 min and then 13-33% solvent B over 40 min. For MAX1(YNGT) and MAX1(VPDGT)-C purifications, linear gradients were employed as 0% solvent B over 1 min, 0-11% solvent B over 11 min and then 11-41% solvent B over 60 min. For EX1 purification, linear gradients were employed as 0% solvent B over 2 min, 0-10% solvent B over 10 min, 10-27% solvent B over 34 min and then 27-37% solvent B over 40 min. For EX2 purification, linear gradients were employed as 0% solvent B

over 2 min, 0-15% solvent B over 13 min and then 15-35% solvent B over 40 min. For EX3 purification, linear gradients were employed as 0% solvent B over 1 min, 0-12% solvent B over 11 min, 12-25% solvent B over 26 min and then 25-35% solvent B over 40 min. Analytical HPLC and electrospray ionization (ESI-positive mode) were performed to verify the purity of the lyophilized peptides.

2.7.2 Circular dichroism

An Aviv model 410 circular dichroism spectrometer was employed to determine the secondary structure of the peptides at 150 μM or 1 wt% concentrations. Concentrations of peptide solutions were determined by using a small volume of stock solutions in UV spectroscopy by absorbance at 220 nm in a 1cm path length cell (ε = 15750 cm^{-1} M^{-1} for 20 amino acid peptides and ε = 16537 cm^{-1} M^{-1} for 21 amino acid peptides).

To prepare 150 μM peptide samples in 50 mM BTP, 150 mM NaCl pH 7.4 buffer, 300 μM stock solutions were prepared by UV spectroscopy. Lyophilized peptides were dissolved in chilled water and kept on ice until used. A chilled stock buffer of 100 mM BTP, 300 mM NaCl pH 7.4 was mixed with the peptide stock solution by inversion at equal volumes (~150 μL) in a 1 mm quartz cell to yield a final concentration of 150 μM peptide in 50 mM BTP, 150 mM NaCl pH 7.4 buffer. The cell was immediately placed into a 5^{o}C pre-equilibrated chamber. Temperature-

dependent wavelength scans from 260 to 200 nm were collected every 5°C from 5 to 80°C, with a 10 min equilibration time at each temperature interval.

For CD of 1 wt% samples, lyophilized peptides were dissolved in chilled water to make 2 wt% solutions and mixed with chilled stock buffer of 100 mM BTP, 300 mM NaCl pH 7.4 at equal volumes to yield a final concentration of 1 wt% peptide in 50 mM BTP, 150 mM NaCl pH 7.4 buffer. The resulting solution was immediately transferred to a 0.1 mm quartz cell. The cell was immediately placed into a 37^{o}C pre-equilibrated chamber. Wavelength scans from 260 to 200 nm were collected after 1 hour at 37^{o}C.

Mean residue ellipticity [θ] was calculated from the equation $[\theta] = \theta_{obs}/(10lcr)$, where θ_{obs} is the measured ellipticity in millidegrees, l is the path length of the cell (centimeters), c is the concentration (molar), and r is the number of residues in the sequence.

2.7.3 Oscillatory rheology

An AR-G2 Magnetic Bearing Rheometer (TA Instruments) with 25 mm or 8 mm diameter parallel plate geometry at a gap height of 0.5 mm was employed for dynamic time, frequency and strain sweep rheology experiments. A 2 wt% peptide stock solution was prepared by dissolving the lyophilized peptide in chilled water and

mixed with an equal volume of chilled stock buffer of 100 mM BTP, 300 mM NaCl pH 7.4 to yield a final concentration of 1 wt% peptide in 50 mM BTP, 150 mM NaCl pH 7.4 buffer. It was immediately transferred to the rheometer peltier plate at 5^{o}C. The temperature was ramped from 5°C to 37°C in 100 sec. The dynamic time sweep experiment, which measures the evolution of storage (G') and loss modulus (G'') with time, was performed for 1 hour at constant strain (0.2%) and frequency (6 rad/s). Independent dynamic frequency sweep (0.1−100 rad/s at 0.2% strain) and strain sweep (0.1−100% strain at 6 rad/s) experiments were performed to confirm that the time sweep measurements were executed within the linear viscoelastic regime. To test the recovery of elastic properties of 1 wt% hydrogels, shear-thinning and recovery experiments were performed. 1 wt% samples were prepared the same way. Following the temperature ramp from 5°C to 37 °C in 100 s and 20 min at 37°C with constant frequency and strain, the strain was increased to 1000% for 30 sec and reduced back to 0.2%. Over the next 20 min, the recovery of the hydrogel was monitored by measuring G` and G`` as a function of time.

2.7.4 Electron microscopy

1 wt% peptide samples in 50 mM BTP, 150 mM NaCl pH 7.4 buffer were prepared ~50 μL volume by mixing chilled 2 wt% peptide solution in water and chilled stock buffer of 100 mM BTP, 300 mM NaCl pH 7.4. Samples were placed in

37^{o}C incubator for 1 hour and left at room temperature until the TEM experiments were performed.

A small amount of hydrogel was diluted in water 100X and mixed intensely. Then, 5 μL was transferred onto 400 mesh carbon coated copper grid S6 and filter paper was used to dry remaining liquid on the grid. As a negative stain, 1% uranyl acetate was transferred onto the grid and filter paper was used to soak up extra stain. Imaging was performed immediately by using Hitachi H-7650 transmission electron microscopy at a voltage of 80 kV. ImageJ was used to measure the fibril widths.

REFERENCES

1. Radkov, A.D. and L.A. Moe, *Bacterial synthesis of D-amino acids.* Appl Microbiol Biotechnol, 2014.

2. Walsh, C.T., *The chemical versatility of natural-product assembly lines.* Accounts of Chemical Research, 2008. **41**(1): p. 4-10.

3. Finking, R. and M.A. Marahiel, *Biosynthesis of nonribosomal peptides.* Annual Review of Microbiology, 2004. **58**: p. 453-488.

4. Caboche, S., et al., *NORINE: a database of nonribosomal peptides.* Nucleic Acids Research, 2008. **36**: p. D326-D331.

5. Caboche, S., et al., *Diversity of Monomers in Nonribosomal Peptides: towards the Prediction of Origin and Biological Activity.* Journal of Bacteriology, 2010. **192**(19): p. 5143-5150.

6. Strieker, M., A. Tanovic, and M.A. Marahiel, *Nonribosomal peptide synthetases: structures and dynamics.* Curr Opin Struct Biol, 2010. **20**(2): p. 234-40.

7. Milnerwhite, E.J., *Beta-Bulges within Loops as Recurring Features of Protein-Structure.* Biochimica Et Biophysica Acta, 1987. **911**(2): p. 261-265.

8. Milnerwhite, E.J. and R. Poet, *Loops, Bulges, Turns and Hairpins in Proteins.* Trends in Biochemical Sciences, 1987. **12**(5): p. 189-192.

9. Wilmot, C.M. and J.M. Thornton, *Beta-Turns and Their Distortions - a Proposed New Nomenclature.* Protein Engineering, 1990. **3**(6): p. 479-493.

10. Sibanda, B.L. and J.M. Thornton, *Conformation of beta hairpins in protein structures: classification and diversity in homologous structures.* Methods Enzymol, 1991. **202**: p. 59-82.

11. Gunasekaran, K., C. Ramakrishnan, and P. Balaram, *beta-hairpins in proteins revisited: lessons for de novo design.* Protein Engineering, 1997. **10**(10): p. 1131-1141.

12. Nair, C.M., et al., *X-Ray Crystal-Structure of Pivaloyl-D-Pro-L-Pro-L-Ala-N-Methylamide - Observation of a Consecutive Beta-Turn Conformation.* Journal of the Chemical Society-Chemical Communications, 1979(24): p. 1183-1184.

13. Aravinda, S., et al., *Analysis of designed beta-hairpin peptides: molecular conformation and packing in crystals.* Organic & Biomolecular Chemistry, 2013. **11**(25): p. 4220-4231.

14. Chalmers, D.K. and G.R. Marshall, *Pro-D-Nme-Amino Acid and D-Pro-Nme-Amino Acid - Simple, Efficient Reverse-Turn Constraints.* Journal of the American Chemical Society, 1995. **117**(22): p. 5927-5937.

15. Haque, T.S., J.C. Little, and S.H. Gellman, *Stereochemical requirements for beta-hairpin formation: Model studies with four-residue peptides and depsipeptides.* Journal of the American Chemical Society, 1996. **118**(29): p. 6975-6985.

16. Rajagopal, K., Ph.D., *Rational peptide design for functional materials via molecular self-assembly*. 2007: University Of Delaware. p. 235.

17. Searle, M.S., D.H. Williams, and L.C. Packman, *A Short Linear Peptide Derived from the N-Terminal Sequence of Ubiquitin Folds into a Water-Stable Nonnative Beta-Hairpin.* Nature Structural Biology, 1995. **2**(11): p. 999-1006.

18. de Alba, E., et al., *Conformational investigation of designed short linear peptides able to fold into beta-hairpin structures in aqueous solution.* Fold Des, 1996. **1**(2): p. 133-44.

19. Blandl, T., A.G. Cochran, and N.J. Skelton, *Turn stability in beta-hairpin peptides: Investigation of peptides containing 3:5 type I G1 bulge turns.* Protein Sci, 2003. **12**(2): p. 237-47.

20. de Alba, E., M. Rico, and M.A. Jimenez, *The turn sequence directs beta-strand alignment in designed beta-hairpins.* Protein Sci, 1999. **8**(11): p. 2234-44.

21. Haque, T.S. and S.H. Gellman, *Insights on beta-hairpin stability in aqueous solution from peptides with enforced type I' and type II' beta-turns.* Journal of the American Chemical Society, 1997. **119**(9): p. 2303-2304.

22. Griffiths-Jones, S.R., et al., *NMR evidence for the nucleation of a beta-hairpin peptide conformation in water by an Asn-Gly type I ' beta-turn sequence.* Chemical Communications, 1998(7): p. 789-790.

23. Sharman, G.J. and M.S. Searle, *Dissecting the effects of cooperativity on the stabilisation of a de novo designed three stranded anti-parallel beta-sheet.* Chemical Communications, 1997(20): p. 1955-1956.

24. RamirezAlvarado, M., et al., *Role of beta-turn residues in beta-hairpin formation and stability in designed peptides.* Journal of Molecular Biology, 1997. **273**(4): p. 898-912.

25. Menach, E., K. Yasukawa, and K. Inouye, *Effects of site-directed mutagenesis of Asn116 in the beta-hairpin of the N-terminal domain of thermolysin on its activity and stability.* Journal of Biochemistry, 2012. **152**(3): p. 231-239.

26. Schneider, J.P., et al., *Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide.* Journal of the American Chemical Society, 2002. **124**(50): p. 15030-15037.

27. Ozbas, B., et al., *Salt-triggered peptide folding and consequent self-assembly into hydrogels with tunable modulus.* Macromolecules, 2004. **37**(19): p. 7331-7337.

28. Rajagopal, K., et al., *Tuning the pH Responsiveness of beta-Hairpin Peptide Folding, Self-Assembly, and Hydrogel Material Formation.* Biomacromolecules, 2009. **10**(9): p. 2619-2625.

29. Haines-Butterick, L., et al., *Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(19): p. 7791-7796.

30. Veerman, C., et al., *Gelation kinetics of beta-hairpin peptide hydrogel networks.* Macromolecules, 2006. **39**(19): p. 6608-6614.

31. Branco, M.C., et al., *Macromolecular diffusion and release from self-assembled beta-hairpin peptide hydrogels.* Biomaterials, 2009. **30**(7): p. 1339-1347.

Chapter 3

BACTERIAL EXPRESSION OF EX1 PEPTIDE

3.1 Introduction

Commonly, in a lab setting, peptides are obtained chemically by Fmoc-based or Boc-based solid phase peptide synthesis (SPPS) [1-3] which is efficient at small scale but scaling-up for industrial manufacturing presents challenges such as optimization, cost and waste management [4-7]. For ultimate translational studies, developing biological production methods as an alternative to chemical synthesis might allow relatively inexpensive mass production of peptides [8, 9]. Although chemical synthesis of hairpin peptides has been optimized [10], the synthesis of new and longer sequences or potentially bioactive protein-peptide fusions in large quantities may require additional optimization studies as well as higher raw material amounts and waste stream costs. On the other hand, recombinant peptide production is notoriously difficult especially for peptides prone to self-assembly. Yields are generally below 30 mg of pure peptide harvested from a liter of culture [11-13]. Recent studies by the McPherson group however, showed that high yield production of self-assembling peptides is achievable by optimizing culture conditions and downstream procedures [14].

Peptide production in bacteria has potential drawbacks such as susceptibility to proteolytic degradation due to their small size and potential toxicity if amphiphilic [15-17]. In *E. coli*, proteolytic degradation of proteins to peptide fragments are facilitated by ATP-dependent Clp and Lon proteases, and further digestion of peptide fragments into smaller peptides and amino acids is enabled by dipeptidases, amino and carbxypeptidases such as imidodipeptidase, oligopeptidase A and aminopeptidase N [18-21]. The general approach against degradation is to employ fusion partners to increase the size of the peptide [22, 23]. Another way of escaping from the bacterial degradation system is to promote aggregation so that enzymes cannot act on the aggregated target proteins [24, 25]. Such a strategy is also useful for expression of potentially toxic peptides. Fusion partners that are known to be insoluble when expressed in bacteria can be employed to prevent toxicity-related bacterial growth problems [26, 27]. For the expression of the 21-residue hydrogel forming β-hairpin peptide EX1, direct and fusion expression strategies were both considered. In the context of the fusion approach, in which the peptide is expressed on either termini of the fusion partner, fusion proteins are required to be cleaved to free the target peptide.

The possible alternative routes for recombinant EX1 (rEX1) production are outlined in Figure 3.1. DNA sequences (inserts) encoding EX1 peptide were designed for (i) native peptide expression that only yields EX1 peptide or (ii) fusion protein expression that enables the production of EX1 on the C-termini of various partner proteins. These constructs contain a cleavage site in-between the peptide and the

protein for further isolation of EX1. EX1 encoding oligonucleotide inserts for fusion protein expression were designed that enable either enzymatic cleavage or chemical cleavage of the peptide from its fusion partner. For enzymatic cleavage, a TEV protease cleavage site was incorporated and for chemical cleavage, a methionine residue was incorporated allowing cleavage to be enacted by cyanogen bromide. Both of these strategies will be discussed in more detail within this chapter.

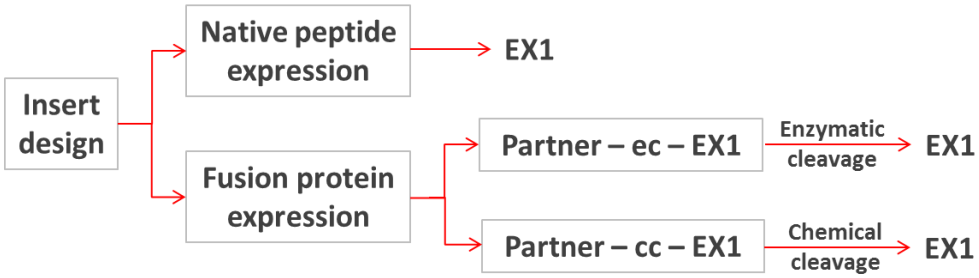


Figure 3.1 Insert design for EX1 expression. Inserts (DNA sequences) encoding EX1 were designed to investigate both native peptide expression and fusion protein expression with a partner protein. For soluble expression, an enzymatic cleavage site (ec) was incorporated, and for insoluble expression, a chemical cleavage site (cc) was incorporated into the insert design.

The overall method for rEX1 production is summarized in Figure 3.2. First, through cloning, EX1 inserts were transferred into an expression vector that had the genes to facilitate the inducible production of EX1 in bacteria. After transformation of the plasmid into the bacteria, the transformants were introduced into cell culture media for propagation and when the cells were at optimal metabolic activity, expression was induced over a period of time. Lastly the bacterial cells were harvested and lysed to analyze the extent of expression.

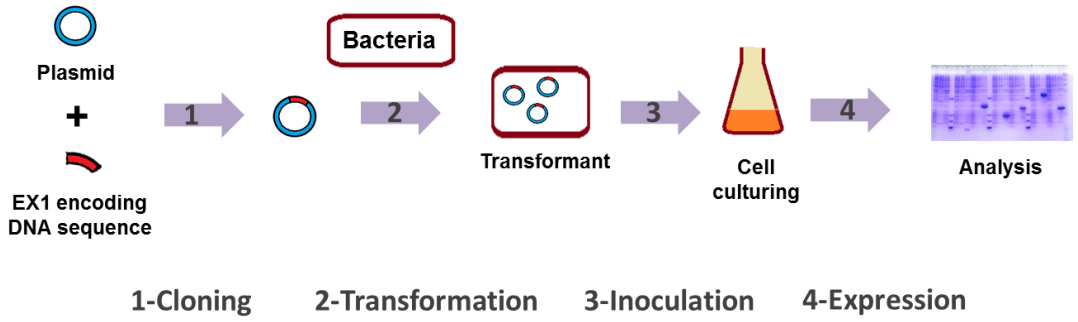


Figure 3.2 Overview of the EX1 expression method. The EX1 encoding oligonucleotide sequence was cloned and inserted into an appropriate plasmid. After sequence confirmation, the plasmid was transformed into a bacterial expression strain (*E. coli* Rosetta 2), and eventually inoculated into cell culture media for propagation and protein production. Bacterial cells were later harvested, lysed and analyzed for target protein expression.

3.2 Selection of the Cloning Method

For recombinant production, the bacteria *E. coli* is considered first among host organisms because expression in *E. coli* is fast, economical, scalable, and extremely well studied [28, 29]. There are many cloning methods available to insert a gene of interest into a vector to be transformed into bacteria. Cloning methods differ by cloning sites, size of cloning site, presence of compatible vectors, labor, time and cost [30]. Traditional cloning methods use restriction enzyme digestion and ligation which is inexpensive and compatible with numerous available vectors. On the other hand; these methods require extra steps such as gel purification to isolate the digested polymerase chain reaction (PCR) product and the digested vector. In addition, traditional methods involve screening a high number of colonies after transformation and performing more plasmid isolations until the correct sequence is encountered.

Alternative cloning methods include homologous recombination and site-specific recombination, which are highly reliable and easy to perform. Homologous recombination requires vector linearization prior to cloning. Site specific recombination is advantageous when subcloning an insert into multiple expression vectors is desired [30]. In this study, the Gateway cloning system which employs site-specific recombination was used. Gateway vectors with various fusion protein partners were kindly provided by Dominic Esposito (Protein Expression Lab, ATRF, NCI). Here, different proteins were investigated as fusion partners to assess which would afford the highest yields of peptide product.

Cloning by the Gateway system is summarized in Figure 3.3A. Entry clones enable subcloning of any insert (or gene) into multiple expression vectors that can be used for multiple downstream applications. For example, multiple expression vectors are useful if expression in different host organisms (e.g. *E. coli*, yeast, insect, or mammalian) is desired, or expression with different fusion partners (solubility enhancing or fluorescent) is desired, or if one wishes to incorporate epitopes (antibody detection) or tags (affinity purification). Site-specific recombination between certain attachment (att) sites where recombination is facilitated enables the transfer of the gene during cloning and subcloning [31]. Compatibility with the Gateway system is acquired by implementing attachment sites (attB1 and attB2) to the ends of the gene (insert) of interest. AttB flanking genes can be generated either by PCR or by adding the att sites in the sequence before DNA synthesis. Next, the donor vector containing

attP sites is used to make an entry clone in which the attL sites flank the gene, as shown in Figure 3.3B. This is accomplished through recombination between attB and attP sites by a BP clonase enzyme kit containing integrase (Int) and integration host factor (IHF) proteins. Once the entry clone is obtained, one additional step is needed to obtain any expression clone containing the desired gene. The recombination between attL sites (in the entry clone) and attR sites (in the expression clone) is effected by a LR clonase enzyme mix containing Int, excisionase (Xis) and IHF. This results in the gene of interest being inserted into the expression clone. Finally, transformants are selected through a combination of antibiotic resistance and growth inhibition protein encoded by the ccdB gene for the specific *E.coli* strain used [32].

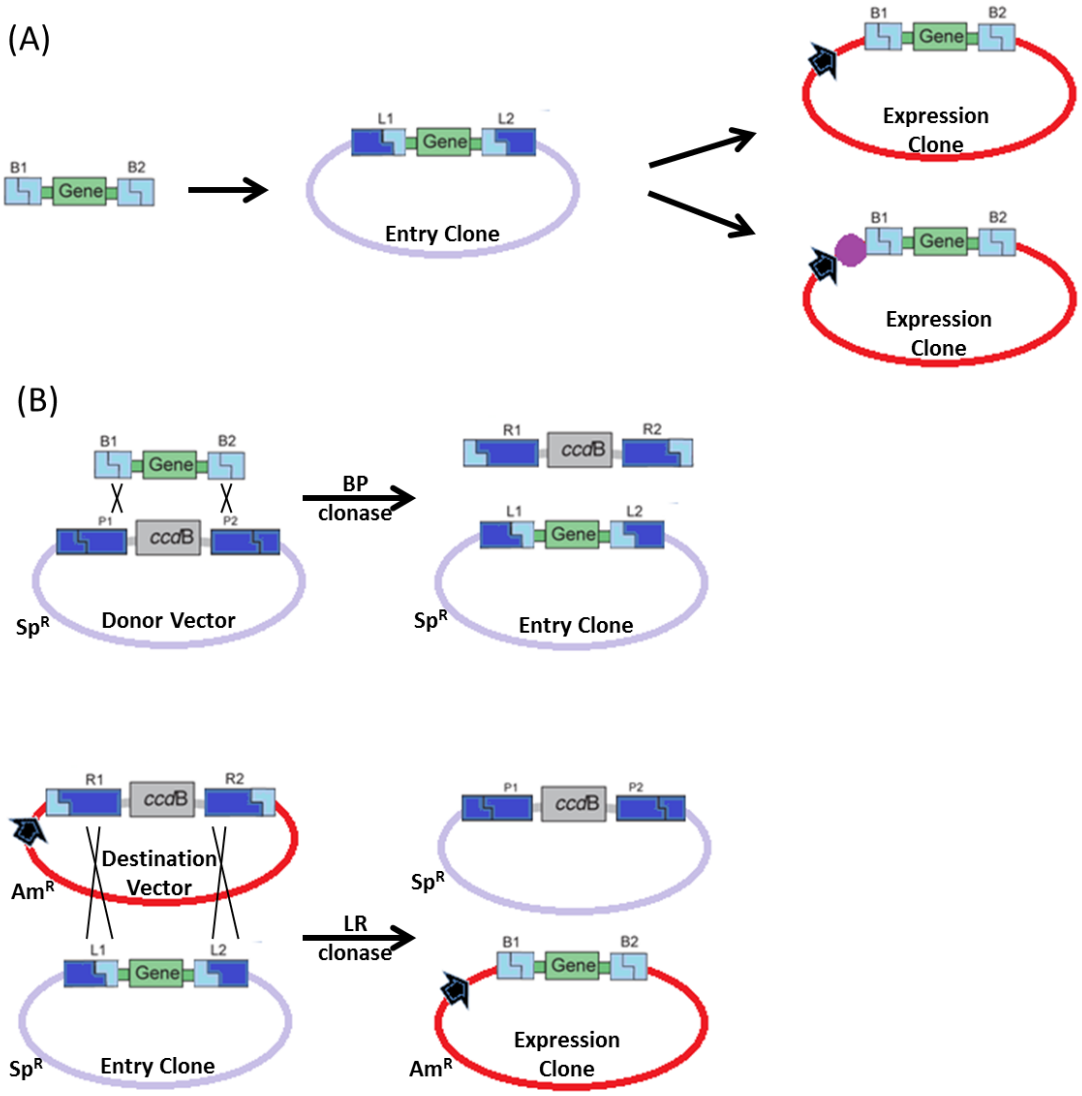


Figure 3.3 Gateway cloning system. (A) Overview of site-specific recombination cloning. An entry clone is first generated for parallel subcloning of a gene into various expression clones. Purple circle signifies different DNA constructs to enable a desired function. For example, different fusion protein partners could be incorporated. (B) The recombination sites (B, P, L, R) are shown in these cloning and subcloning steps. B and P sites generate L and R sites through a BP clonase reaction, and L and R sites generate B and P sites through a LR clonase reaction. (Sp: Spectinomycin, Am: Ampicillin, black arrows represent hypothetical promoters, X represents recombination between the specific sites).

3.3 Insert Design

Gateway compatible inserts for bacterial production of the EX1 peptide were designed to investigate both native expression, which would yield the exact amino acid sequence of EX1, and fusion expression, which would afford EX1 fused to a partner protein to be cleaved later. The cleavage site between the fusion partner and EX1 was also designed for either enzymatic or chemical cleavage of the fusion proteins. As will be shown, native expression of EX1 proved difficult, resulting in relatively poor yields. This could be due to the peptide being degraded or being toxic to the cell causing poor growth [33-35]. Next, peptide-protein fusions containing TEV protease cleavage sites were explored. Here, although the protein fusions were chosen to solubilize the peptide, production of the peptide-protein fusions were limited to inclusion bodies. Attempting to cleave the peptide from its fusion partner using TEV protease proved difficult due to the insoluble nature of the peptide-protein fusion. Finally, the TEV protease site was replaced with a methionine residue, which allowed the solubilization of the peptide-protein fusion with formic acid and its cleavage with CNBr [36].

3.3.1 Insert design for native peptide expression

In prokaryotic protein synthesis, the 3` end of 16S rRNA in the small ribosomal subunit 30S binds to the complementary sequence (Shine-Dalgarno /

ribosome binding site) located upstream (5` end) of the start codon during translation initiation [37]. In the design for expression of the native peptide (NatPep), shown in Table 3.1, the attB1 site-specific recombination site was followed by a ribosome binding site and a start codon (ATG). An oligonucleotide sequence encoding EX1 was placed between the start codon and the stop codon so that there would be no additional amino acids expressed with the peptide. Although protein synthesis in bacteria begins with methionine or *N*-formylmethionine, deformylases cleave *N*-formyl functionality and methionine aminopeptidases (MAP) cleave amino terminal methionine in the majority of proteins [38]. The stop codon was followed by an attB2 site-specific recombination site. The DNA sequence of the NatPep insert design and the expected amino acid sequence of the gene product are given in Table 3.1.

Table 3.1 NatPep insert design for native peptide expression. The DNA construct, DNA sequence and corresponding peptide sequence are shown.

DNA construct	attB1 – Ribosome Binding Site – ATG – EX1 Peptide – stop – attB2
DNA sequence	5'- GGGGACCACGCTGTATAAAAAAGTTGGC - GAAGGTGATCGTACC - ATG - GTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGA TGGCACGAAAGTTGAAGTGAAAGTTAAAGTG - TAA – CCAACTTTCTTGTACAAAGTTGTCCCC -3'
Amino acid sequence	VKVKVKVKVPDGTKVEVKVKV

3.3.2 Insert design for protein-peptide fusions contain a TEV protease cleavage site

Gateway destination vectors having different fusion partners upstream of the recombination region were initially selected. The ribosome binding site and start codon were already included upstream of the fusion protein genes. Destination vectors that have only a hexahistidine (6His) or a 6Histag followed by a fusion protein were used. The destination vector containing only the 6Histag and no other fusion protein was prepared to investigate whether or not a simple hexaHis tag could solubilize the peptide and enable efficient expression. For the constructs employing proteins, the fusion proteins that were employed included protein G B1 domain (GB1), KSI, thioredoxin (Trx) or glutathione-S-transferase (GST). Vectors were constructed such that the synthetic EX1 oligonucleotide sequence is downstream (3` end) of the fusion partner genes. Following expression, EX1 would be on the C-terminus of the fusion partners.

The specific fusion proteins were chosen based on their known ability to promote high-yield expression. The B1 domain of protein G (GB1) from *Streptococcus* is a 56-residue, highly soluble and stable protein that has been used for NMR spectroscopic analysis of fusion proteins [39-41] *E. coli* thioredoxin (TrxA) protein (~12 kDa, pI 4.7) as a fusion partner has been successfully employed for the recombinant production of many proteins due to its highly soluble nature and high thermal stability as well as having its N- and C-termini exposed at its surface [42]. Bacterial KSI protein (~13.4 kDa, pI 5.3) is a 125-residue long protein expressed as

inclusion bodies. It is used as a fusion partner for expression of peptides toxic for the bacteria [26, 36]. GST is a highly soluble 26 kDa protein (pI 5.3) that can facilitate affinity purification by using immobilized glutathione [23, 43].

With the exception of the KSI fusion, I envisioned these proteins to impart solubility to the peptide. Thus, a TevPep insert was designed as shown in Table 3.2 in such a way that an enzymatic Tobacco Etch Virus (TEV) cleavage site was included before the EX1 sequence for further enzymatic release of the peptide. The TEV site was also used for the hexaHis construct. TEV protease is capable of cleaving at the C-terminus of a sequence containing ENLYFQ. Enzymatic cleavage occurs between the glutamine and any following residue except Proline [44]. The TEV protease site is highly specific and the enzyme can be easily expressed and purified in a cost effective way [45].

Table 3.2 TevPep insert design for soluble fusion expression with an enzymatic (TEV) cleavage site. The DNA construct, DNA sequence and corresponding peptide sequence are shown.

DNA construct	attB1 – TEV cleavage site – EX1 Peptide – stop – attB2
DNA sequence	5'- GGGGACCACGCTGTATAAAAAAGTTGGC - GAAAACCTGTATTTTCAG - GTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGA TGGCACGAAAGTTGAAGTGAAAGTTAAAGTG - TAA - CCAACTTTCTTGTACAAAGTTGTCCCC -3'
Amino acid sequence	TSLYKKVGENLYFQVKVKVKVKVPDGTKVEVKVKV

The 6Histag-EX1 shown in Table 3.3 is the corresponding translation of the DNA sequences for the TevPep insert on the C-terminus of 6Histag with the attB1 site. The arrow after the glutamine (Q) shows the amide bond which is cleaved by TEV protease to release the free peptide.

Table 3.3 6Histag-EX1 fusion and its corresponding amino acid sequence. The DNA construct and the corresponding peptide sequence are shown. The green arrow shows where the TEV protease cleaves the fusion protein.

DNA construct	Histag – attB1 – TEV site – EX1 Peptide – stop
Amino acid sequence	MRSGSHHHHHHRSDITSLYKKVGENLYFQ↑VKVKVKVKVPDGTK VEVKVKV

For the constructs containing the fusion proteins, the proteins of interest are shown in Figure 3.4. The destination vector and entry clones used to prepare the final expression clones are also shown. Although the KSI fusion is expected to be insoluble, it was included in this set of fusion proteins to compare its expression level with the other soluble fusion proteins.

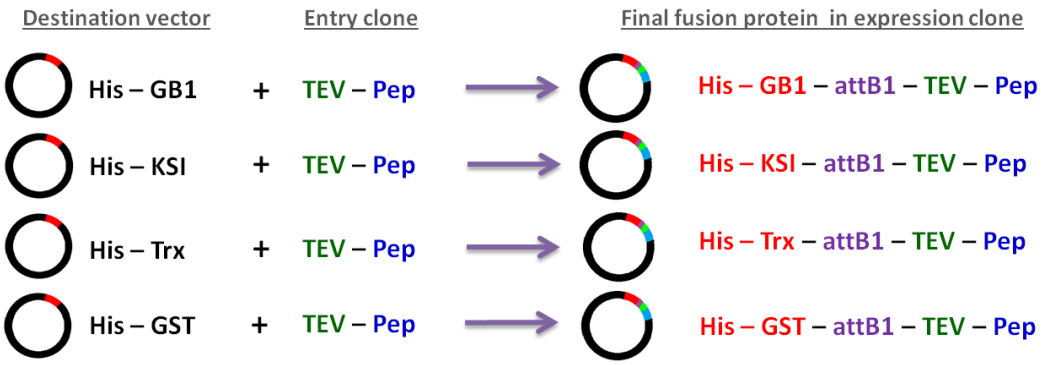


Figure 3.4 Generation of expression clones for soluble fusions. Subcloning of the TevPep insert into destination vectors with fusion partners to be cleaved by TEV protease are listed. TEV (green) represents the TEV cleavage site.

3.3.3 Insert design for peptide-protein fusions containing a Met chemical cleavage site

As will be shown, the originally designed TEV-containing protein fusion constructs proved to produce insoluble protein whose susceptibility to TEV protease was limited. This was surprising given the fact that the family of fusion protein partners were chosen for their solubilizing characteristics. This observation indicates that EX1 drives aggregation in the cell and expression is driven to the inclusion bodies. As such, new clones were designed where the TEV site was replaced with a Methionine residue. This allows the isolation of the EX1-protein fusion from the insoluble fraction of the cell lysate, solubilization of the fusion with formic acid and subsequent cleavage of the peptide from its fusion partner with CNBr.

Thus, a clone was generated that places a met-residue N-terminal to that first Valine residue of EX1 (Table 3.4). Chemical cleavage by CNBr in 70% formic acid, should release EX1.

Table 3.4 Generation of MetPep insert for insoluble fusion expression. The codon for glutamine was changed to methionine codon ATG and the fusion product is amenable to CNBr cleavage to free EX1.

DNA construct	attB1 – TEV cleavage site – EX1 Peptide – stop – attB2
Amino acid sequence	TSLYKKVGENLYFQVKVKVKVKVPDGTKVEVKVKV
MetPep insert product	TSLYKKVGENLYFMVKVKVKVKVPDGTKVEVKVKV

3.4 Pilot Expression of Native EX1 and EX1 in Fusion

In bacteria, the *lac* operon is important for the use of lactose as a carbon source and its regulatory region (promoter and operator sequences) serves as a control for the expression of target proteins (Figure 3.5). The *lacI* gene (upstream of the *lac* operon) is translated to give the lac repressor protein. In the absence of an inducer (lactose), the repressor binds to the *lac* operator region, preventing RNA polymerase from binding to the promoter region, and subsequently inhibits transcription of the lac genes [46].

Figure 3.5 The lacI, lac operon and control elements in bacterial chromosome. The *lacI* is located upstream of the *lac* operon and is transcribed independent of the *lac* operon. The promoter and operator are *lac* operon control elements upstream of the *lac* operon genes *lacZ, lacY* and *lacA*. Blue arrows represent the start and the direction of transcription for each gene.

For the control of target protein expression, the lac operator is coupled with the use of bacteriophage T7 RNA polymerase. This enzyme is highly selective for its promoter region and highly efficient for its transcription. T7 RNA polymerase inducible overexpression of target proteins can constitute more than half of all bacterial proteins in three hours [47]. Figure 3.6 shows how the lac operator, controlled by T7 RNA polymerase (RNAP) transcription, facilitates the transcription of the target gene upon induction. Bacterial chromosomes of *E. coli* DE3 strains contain T7 RNAP gene downstream of the lac promoter and operator region where lacI transcriptional repressor occupies and prevents the RNAP from binding to the promoter and starting transcription. When lacI repressor protein binds to isopropylthio-β-galactoside (IPTG, a non-hydrolyzable allolactose analog), it no longer stays bound to the operator region and T7 RNAP transcription starts. Production of T7 RNAP which is very specific for its promoter region upstream of the gene of interest, orchestrates the translational machinery for the target protein`s production. In the plasmid, the lac operator placed in-between the T7 promoter and gene of interest serves as an additional way to inhibit leaky expression of T7 RNAP. When IPTG is introduced to the cells, it binds not only to the lacI repressor protein

residing upstream of the DNA encoding T7 RNAP but also to binds the lacI repressor downstream of the T7 promoter region and allows T7 RNAP to initiate the transcription of the gene of interest [48].

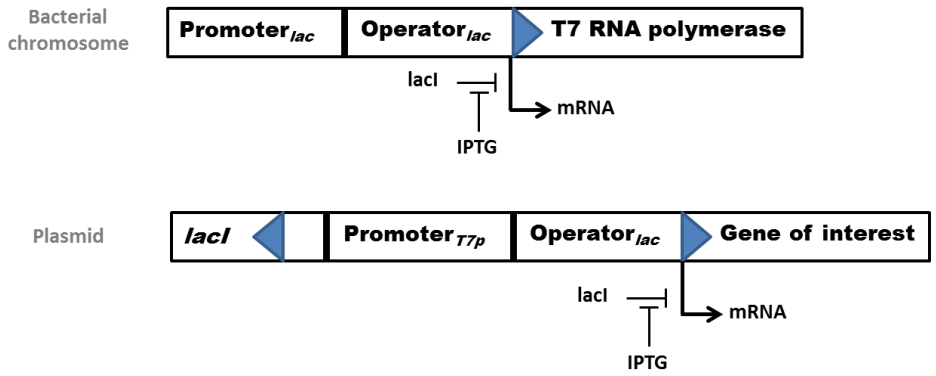


Figure 3.6 Schematic diagram of inducible over-expression of target proteins. Bacterial strain DE3 contains the T7 RNA polymerase (RNAP) under the control of lac regulatory elements. The transformed plasmid has a T7 promoter upstream of gene of interest. An additional lac operator downstream of the T7 promoter and a lacI upstream of T7 promoter serves as extra inhibition control of target gene transcription against basal T7 RNAP expression. Upon induction by IPTG, the lacI repressor protein is released from both lac operators on the bacterial chromosome and the plasmid, and T7 RNAP is produced which binds to the T7 RNAP promoter region (free of repressor protein) and starts the transcription of the target gene.

The NatPep insert and the inserts containing the TEV cleavage site were used for EX1 expression with the method shown in Figure 3.2. Bacteria containing the appropriate plasmid was grown in standard Luria Bertani (LB) medium in an orbital shaker (250 rpm) at 37^{o}C until an OD_{600} of ~0.6 (logarithmic growth phase) was reached. Expression was induced with 1 mM IPTG at 37^{o}C (250 rpm) for 4 hours. Cells were harvested and lysed, and Sodium Dodecyl Sulfate Poly-Acrylamide Gel

Electrophoresis (SDS-PAGE) analysis was performed on soluble and insoluble fractions of the lysate.

In Figure 3.7, SDS-PAGE analysis shows that there was no native peptide expression observed in neither the soluble nor insoluble fractions. However, all the fusion proteins starting from the 5.7 kDa 6Histag-EX1 to the GST-EX1 were successfully expressed (highlighted by the red rectangles in Figure 3.7). The expected molecular weights for each protein (noted below the protein bands) were extrapolated from the protein standards in the first lane, and corresponded nicely against each other. The 6Histag-EX1 fusion has a relatively small size and the relatively low band intensity suggests that its expression might have been affected by its possible poor stability due to its small size. Of note is the fact that, all of the fusions were found in the insoluble fraction. This indicates that, although the fusion proteins were designed to impart solubility, the EX1 peptide drives aggregation, shunting expressed protein-fusions to inclusion bodies. Even though the fusions were characterized by limited solubility, I next assessed whether or not enzymatic cleavage by TEV protease would be a viable option to release the peptide. Given the propensity of the EX1-protein fusions to aggregate, enzymatic cleavage seemed unlikely, but still worth a try.

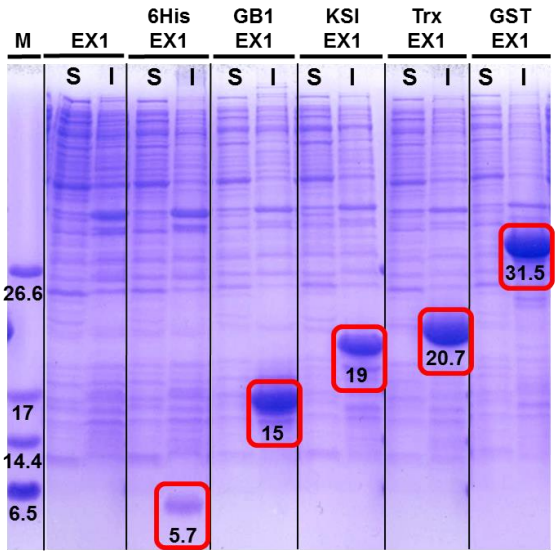


Figure 3.7 SDS-PAGE showing the pilot expression results. Following 4h of expression of fusion proteins containing enzymatic cleavage site at 37°C, cells were lysed and centrifuged. Soluble (S) and insoluble (I) fractions for each experiment, and a protein molecular weight standard (M) was loaded into the respective wells and separated by gel electrophoresis (18% gel, 200V, 60 min). Expressed fusion protein bands are shown in red squares with their expected molecular weights in Da.

3.5 Expression and Purification of Fusion Proteins with Enzymatic Cleavage Site

Once the SDS-PAGE analysis confirmed the conditions for the overexpression of the fusion proteins from the pilot expression, higher volume bacterial culturing at 0.8 L for each fusion protein was performed. For purification, the IMAC method was used: this method facilitates the purification of fusion proteins through the interaction between the 6Histag on the fusions and a metal cation fixed on a resin in the chromatography column [49]. Denaturing conditions were employed during purification to solubilize insoluble aggregates, including each respective EX1-protein

fusion, recovered after bacterial lysis. In figure 3.8, SDS-PAGE analysis shows that all the fusion proteins were successfully purified and exhibited single bands at the expected molecular weights. To recover the free EX1 peptide, fusion proteins must be cleavaed by TEV protease. In order to attempt enzymatic cleavage, the denaturing buffer in which the EX1-protein fusions were solubilized, had to be removed; TEV protease doesn't function under denaturing conditions. Upon dialysis to remove the denaturant, the fusion proteins precipitated and TEV protease was unable to efficiently cleave the fusion proteins. For this reason, no further studies were performed on fusion proteins with an enzymatic cleavage site. I then turned my attention to EX1-fusions containing a chemical cleavage site.

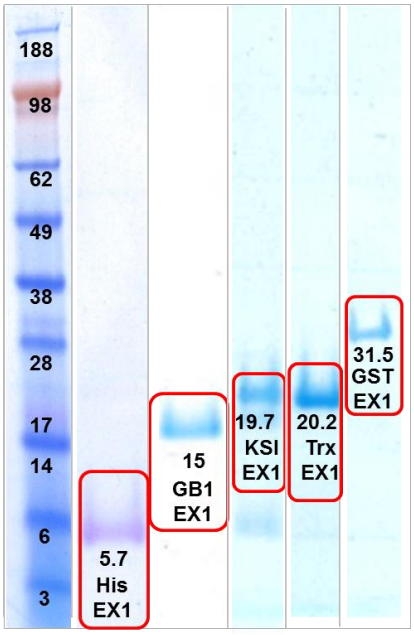


Figure 3.8 SDS-PAGE of the purified fusion proteins (enzymatic cleavage site). Fusion proteins with the enzymatic cleavage site were purified by metal affinity resin

under denaturing conditions, and analyzed by gel electrophoresis (4-12% gel, 200V, 35 min). samples from distinct gels were collected and collated for this figure.

3.6 Pilot Expression of Fusion Proteins with a Chemical Cleavage Site

For production of fusion proteins containing methionine on the N-terminus of EX1 (to facilitate CNBr cleavage under denaturing conditions), the TevPep insert was mutated by changing the glutamine codon to a methionine codon to yield the MetPep insert (Table 3.4). The expression method shown in Figure 3.2 was repeated with the MetPep insert for all the fusion proteins with the exception of GST since its molecular weight is relatively high compared to the other fusion proteins. Large molecular weight proteins result in a large protein:peptide ratio, diminishing yields of the final cleaved peptide. Following cloning, transformation, and pilot expression, SDS-PAGE analysis (Figure 3.9) showed that all the fusion proteins were successfully expressed and their corresponding molecular weights are in good aggrement with the protein weight standards except for the 6Histag fusion. It is unclear if the 6Histag fusion was expressed since precipitation was seen at the bottom of the gel-well suggesting that this sample could not enter the gel for separation, and neither its presence nor its molecular weight could be confirmed. The GB1-EX1 fusion alone was selected for further cleavage studies due to: (i) the relatively low molecular weight of GB1 resulting in a higher ratio of target peptide EX1 following cleavage and (ii) the relatively higher band intensity seen for GB1-EX1 suggesting higher expression for this fusion when compared to the other fusion proteins.

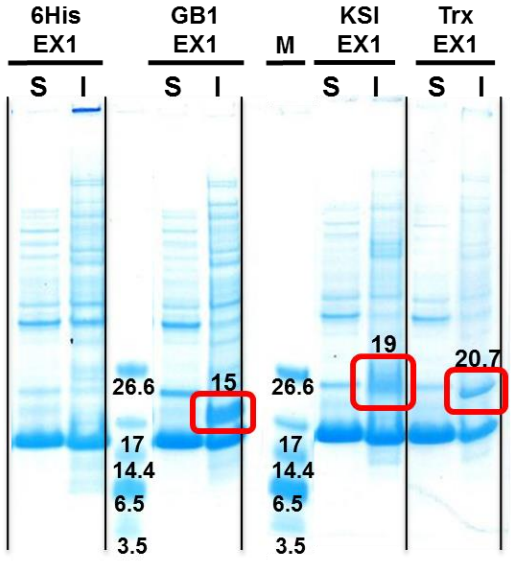


Figure 3.9 SDS-PAGE showing the pilot expression results of fusion proteins (chemical cleavage site). Following 4h expression (37^{o}C, 250 rpm) of fusion proteins containing the chemical cleavage site, cells were lysed and centrifuged. Soluble (S) and insoluble (I) fractions as well as a molecular weight standards (M) was loaded into the respective wells and separated (4-12% gel, 200V, 35 min). The expressed fusion proteins are highlighted within red boxes with expected molecular weights in kDa.

3.7 Expression of GB1-EX1 Fusion and rEX1 Purification

Bacterial culturing of the GB1-EX1 fusion was performed in a higher volume (0.8 L). After which, the cells were harvested and lysed. The resulting insoluble cell pellet was washed to remove remaining soluble proteins, dissolved in 8 M urea, 0.1% TFA, dialyzed in 0.1% TFA and finally lyophilized. After cell lysis, pellet washing and dialysis, small aliquots were saved for SDS-PAGE analysis (Figure 3.10). This

analysis showed that the majority of bacterial proteins were removed in the soluble fraction and the subsequent pellet washing steps helped the removal of remaining bacterial proteins. The insoluble fraction contained most of the GB1-EX1 fusion as confirmed by the presence of a large protein band at the expected size in Figure 3.10. During dialysis, some protein precipitation was observed and the fusion was found in both soluble and insoluble fractions of the dialysis. Low pH conditions during dialysis probably helped a portion of the fusion protein to stay in solution due to the fact that the EX1 peptide is unfolded and highly protonated at acidic conditions. Following dialysis, the soluble and insoluble fractions were combined, lyophilized, reconstituted in 0.1% TFA and analyzed by analytical reversed-phase high-performance liquid chromatographic (RP-HPLC).

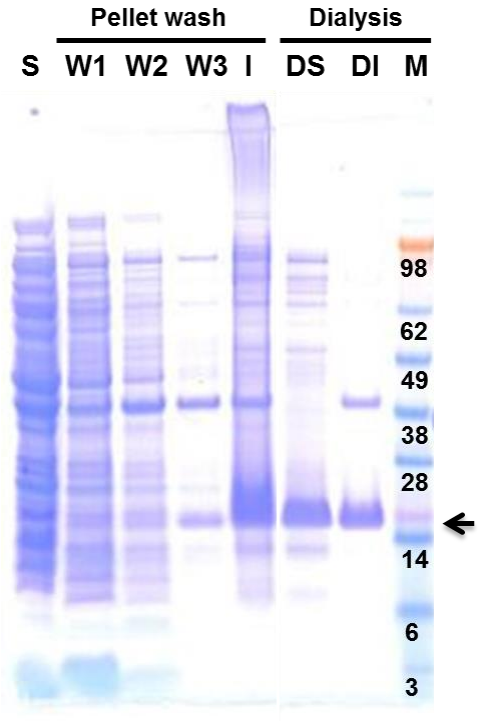


Figure 3.10 SDS-PAGE analysis of GB1-EX1 following cell lysis, pellet washing and dialysis. From left to right, lane 1: soluble fraction (S) after cell lysis, lane 2-4: fractions after pellet wash (W1-3), lane 5: insoluble fraction (I) after pellet wash, lane 6: soluble fraction after dialysis (DS), lane 7: insoluble fraction after dialysis (DI), and lane 8: protein molecular weight marker.

The analytical HPLC trace shows a large, predominant peak eluting around 45 min that corresponds to the GB1-EX1 fusion (Figure 3.11A). This shows that without using IMAC purification under denaturing conditions, the GB1-EX1 fusion protein can be recovered from the cell lysate with a relatively high purity. The fusion protein was later subjected to chemical cleavage by CNBr in 70% formic acid, which has the ability to solubilize the cell pellet and facilitate the cleavage at the C-terminus of methionine, freeing EX1 from GB1. In Figure 3.11B, an analytical RP-HPLC chromatogram shows that chemical cleavage affords rEX1, which has the same elution

time (~29 min) as chemically synthesized EX1. To purify the rEX1, RP-HPLC was employed. Figure 3.11C shows a chromatogram of purified rEX1. The molecular weight of rEX1 was confirmed by electrospray ionization mass spectrometry (ESI-MS), and the calculated and observed ion masses corresponded well, confirming that the purified rEX1 is of the correct mass.

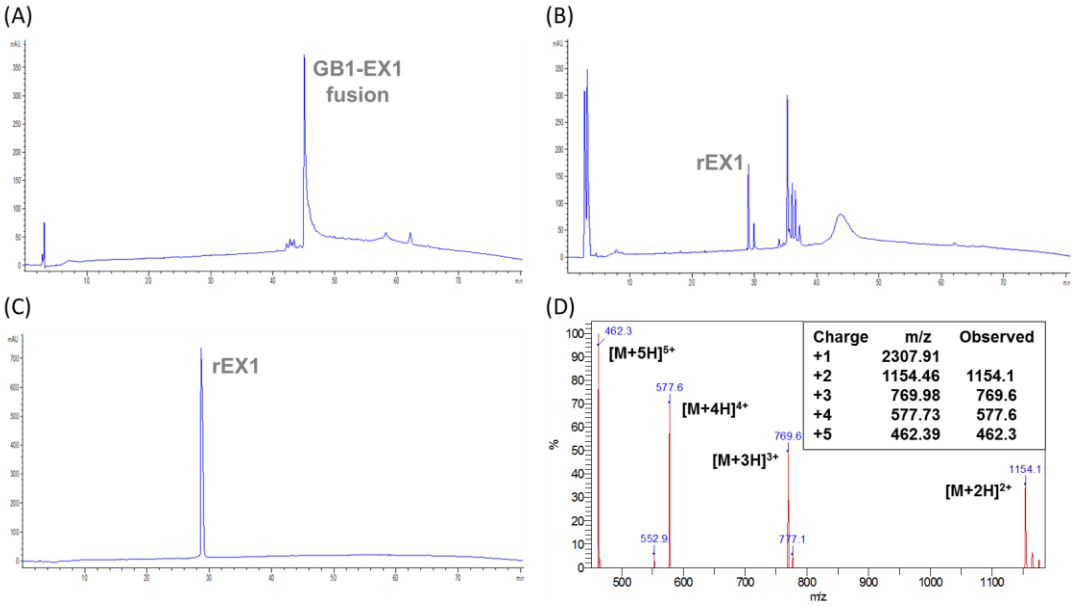


Figure 3.11 RP-HPLC and MS analysis of the fusion protein and rEX1. RP-HPLC chromatogram of (A) GB1-EX1 fusion, (B) crude material after CNBr cleavage of GB1-EX1 fusion and (C) purified rEX1. Absorbance at 220 nm was monitored as a function of retention time. 0-100% Std B in 100 min at 40^{o}C, Std A: 0.1 % TFA in water, Std B: 90% Acetonitrile, 9.9% water, 0.1 % TFA. (D) ESI-MS of pure rEX1. The inset is the comparison of calculated and observed ion masses.

3.8 Yield and Cost Analysis

The expression strategy employed in this chapter yielded 100 mg crude fusion protein and 4.6 mg pure rEX1 from a liter of bacterial culture at first attempt when the following steps were performed: culturing in LB media, cell lysis, insoluble cell pellet recovery and wash, dissolving in urea and dialysis, lyophilization, CNBr cleavage of GB1-EX1 fusion and HPLC purification of rEX1. The overall low yield of rEX1 could be related to inefficiencies and loss during chemical cleavage. Studies to improve the yield are discussed in the next chapter (Chapter 4).

A cost analysis was performed to compare the cost of chemical synthesis of EX1 and the cost of the bacterial production according to the yields obtained from this study. In the lab, the chemical synthesis of peptides is generally performed on an automated batch peptide synthesizer where a 0.25 mmol scale synthesis yields ~400 mg crude peptide material and ~150 mg pure peptide after semi-preparative HPLC purification. Total cost of obtaining 150 mg pure peptide is ~$530 with the consideration of the solvents used for synthesis and purification as well as synthesizer expenses: ~75% of the cost relates to synthesis and the remaining ~25% relates to peptide purification. On the other hand, the cost of production of pure recombinant peptide of the same amount (~150 mg) in a bacterial system (calculated with a yield of 4.6 mg peptide from a liter of culture within this study), is $953. When the theoretical yield, which is 15.6 mg peptide from a liter of culture, is considered, the

cost is reduced from $953 to $298 which is well below the cost of chemical synthesis. Thus, efforts towards improving the yield of bacterial production are worthwhile.

3.9 Conclusion

For the investigation of bacterial production of EX1, an expression method was developed starting from the cloning of the oligonucleotide sequence of EX1 into plasmids designed for various approaches towards expression such as native or fusion protein expression. This was followed by transformation of the clones into a bacterial strain in which expression of the target gene was induced by the concerted operation of genes on the bacterial chromosome and plasmid. Once the transformants containing the plasmids that carry the desired DNA constructs were obtained and cultured in media in small scale, the expression profiles of each of the transformants were analyzed.

Firstly, expression of the native EX1 sequence did not afford any peptide detectable by SDS-PAGE analysis whereas all the fusion proteins with an enzymatic cleavage site were expressed successfully in the insoluble fraction. These fusion proteins were purified by IMAC from the insoluble fraction of the cell lysate under denaturing conditions. Purified EX1 protein fusions were only soluble under denaturing conditions. These conditions prevented the use of TEV protease, which is necessary to cleave EX1 from its fusion partner. The focus of the study was therefore

redirected to the expression of fusion proteins having a chemical cleavage site that can be accommodated under denaturing conditions that permit protein solubility. The same expression method was repeated starting from the cloning of the oligonucleotide sequence containing a methionine codon which translates to methionine that can be cleaved at its C-terminus by CNBr in 70% formic acid. Formic acid acts to solubilize the insoluble fraction of the cell lysate which contains the EX1-protein fusions. Pilot expression resulted in fusion proteins occurring in the insoluble fraction as expected. Although all the fusion proteins enable expression of EX1, the GB1-EX1 fusion protein was selected for further evaluation because the relative expression level was highest and the ratio of the molecular weights of EX1 to fusion was highest for GB1 when compared to the KSI and Trx fusions. The fact that the GB1-EX1fusion is shunted to inclusion bodies is beneficial. The advantage of working with the insoluble cell pellet is that no purification is required since by washing the insoluble fraction a few times, highly pure fusion proteins can be obtained. Thus, instead of IMAC purification, the insoluble fraction was washed with Triton-X and dissolved in 8 M urea and 0.1% TFA. This low pH buffer is efficient in solubilizing the GB1-EX1 fusion. After centrifugation and dialysis to remove the cell debris and urea respectively, the protein sample was lyophilized in preparation for CNBr cleavage and subsequent HPLC purification to afford rEX1. Mass spectrometry confirmed the mass of rEX1 and a cost analysis was performed to investigate whether bacterial production provides a feasible alternative to chemical synthesis. The first attempt to prepare rEX1, resulted in a relatively low yield. Thus, the approximated cost of bacterial

expression was higher than that of chemical synthesis. It is therefore worthwhile to investigate whether or not the yield can be improved. If so, the cost could be substantially lower for bacterial expression as compared to synthesis. Overall, rEX1 was produced successfully in a bacterial expression system and yield improvement studies will be performed in the following chapter.

3.10 Experimental Section

3.10.1 Oligonucleotide sequences designed for cloning

Inserts were designed accordingly to be compatible with the Gateway cloning system by flanking attB sites at either ends. Oligonucleotides were ordered from Genscript and were provided cloned in a pUC57 vector.

DNA sequence of NatPep is 5'-
GGGGACCACGCTGTATAAAAAAGTTGGCGAAGGTGATCGTACCATGGTGA
AAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACGAAAGTTGAAGTGAAA
GTTAAAGTGTAACCAACTTTCTTGTACAAAGTTGTCCCC -3'

DNA sequence of TevPep is 5'-
GGGGACAACTTTGTACAAAAAAGTTGGCGAAAACCTGTACTTCCAGGTGA
AAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAAA
GTTAAAGTGTAACCAACTTTCTTGTACAAAGTTGTCCCC -3'

DNA sequence of MetPep gene is 5'-

GGGACAACTTTGTACAAAAAAGTTGGCGAAAACCTGTACTTCATGGTGAA

AGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAG

TTAAAGTGTAACCAACTTTCTTGTACAAAGTTGTCCCC -3'

3.10.2 DNA cloning and transformation

Inserts were supplied with 4 μg of lyophilized pUC57 plasmid DNA dissolved in 20 μL of water. For BP and LR reactions, the manufacturer's protocols were followed (Invitrogen). The following vectors were kindly provided by Dominic Esposito: donor vector pDONR253, destination vector pDEST521 for inserts with RBS (for NatPep insert), destination vectors with RBS and fusion partners (for TevPep and MetPep inserts), pDEST527 (6Histag), pDEST532 (6His-GB1), pDEST533 (6His-KSI), pDEST546 (6His-Trx), pDEST565 (6His-GST). Briefly, 0.5 μL of 200 ng/μL pUC57 containing NatPep and TevPep insert, 1 μL of 120 ng/μL of pDONR253, 4.5 μL of TE buffer and 2 μL of BP clonase were mixed and left at room temperature for at least 1 h to generate the entry clones. *Escherichia coli* DH5α strain was used for cloning, and 1 μL of BP clonase reaction was transferred into 20 μL of pre-aliquoted subcloning efficiency DH5α cells thawed on ice. After 30 min incubation on ice, 30 sec at 42^{o}C in a water bath and 2 min incubation on ice, 60 μL of SOC medium was added and cells were cultured in a shaker incubator at 250 rpm,

37°C for 1 h. From this culture, 50 μL was spread onto a LB agar plate with Spectinomycin (50 μg/mL) and incubated overnight at 37°C. Single colonies were picked and inoculated in 10 mL of LB media supplemented with Spectinomycin (50 μg/mL) and incubated overnight at 37°C (250 rpm), followed by plasmid DNA isolation using the Qiagen spin miniprep kit according to the manual. Plasmid DNA samples were submitted for sequencing and upon verification of the expected sequences, they were used for the LR cloning reaction.

For a 10 μL LR reaction, 50 ng of entry clone, 150 ng of destination vector that had the fusion partner of interest, 2 μL LR Clonase II and TE buffer pH 8 were mixed and left at room temperature for at least 1 h to generate the expression clones. From this LR clonase reaction, 1 μL was transferred into 20 μL of pre-aliquoted subcloning efficiency DH5α cells thawed on ice. Transformation, cell spreading on LB agar plates with ampicillin (100 μg/mL), single colony inoculation into 10 mL LB media with ampicillin (100 μg/mL), miniprep plasmid DNA isolation and sequence verification were performed as indicated before in the previous paragraph, followed by transformation into the expression strain Rossetta 2 cells and cell spreading onto LB agar plates with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). Glycerol stocks (8% glycerol) of Rossetta 2 cells with expression clones were prepared from overnight LB cultures inoculated with single colonies from LB agar plates.

3.10.3 Bacterial culturing

For Rossetta 2 cell cultures, 100 μg/mL carbenicilin and 34 μg/mL chloramphenicol were used in the media. For pilot expression experiments, 30 mL of media was used in a 250 mL baffled flask. For larger scale experiments, 800 mL of LB media was used in 2.8 L baffled flasks.

LB media was prepared by dissolving 20 g/L LB broth media (Sigma) in distilled water followed by autoclaving. For culturing in 30 mL LB media, 10 mL starter media with antibiotics was inoculated with the glycerol stock of Rossetta 2 cell (with plasmid of interest) and incubated on an orbital shaker at 37^{o}C (250 rpm) overnight. Fresh 30 mL LB media with antibiotics was then inoculated with 0.3 mL (1%) overnight culture and left on orbital shaker at 37^{o}C (250 rpm) until the OD_{600} reached 0.7-0.8. For culturing in 0.8 L LB media, 30 mL starter media with antibiotics was inoculated with the glycerol stock of Rossetta 2 cell (with plasmid of interest) and incubated at 37^{o}C (250 rpm) overnight. To 800 mL LB media with antibiotics, 8 mL (1%) of overnight culture was added and incubated at 37^{o}C (250 rpm) until the OD_{600} reached 0.7-0.8. A final concentration of 1 mM IPTG was used to induce the expression for 4h at 37^{o}C (250 rpm). After culturing, wet cell pellets obtained by centrifugation at 12000 x g for 10 min were stored at -80^{o}C prior to processing.

3.10.4 Cell lysis after pilot expression

For NatPep and TevPep fusions, an amount of culture normalized at OD_{600}:3 were centrifuged in a 1.5 mL tube at 13000 rpm for 5 min and the supernatant was discarded. Cell pellets were resuspended in 100 µL of lysis buffer A (50 mM Tris-HCl pH 7.5, 50 mM Dextrose, 1 mM EDTA, 1 mM $MgCl_2$, 30 U/µL Ready-Lyse (Epicentre) lysozyme solution) with 5 U Benzonase nuclease (Novagen) and incubated for 5 min at room temperature. To this, 100 µL of lysis buffer B (10 mM Tris-HCl pH 7.5, 50 mM KCl, 1 mM EDTA, 0.1% deoxycholate) was then added and incubated for 5 min at room temperature. Following centrifugation, soluble fractions were separated into new tubes and insoluble pellets were solubilized in 100 µL of 8 M urea. For pre-induction samples, 1 mL pre-induction cultures were centrifuged and cell pellets were resuspended in 20 µL Lysis buffer A for 5 min, then 20 µL Lysis buffer B was then added and vortexed.

For MetPep fusions, an amount of culture normalized at OD_{600}:3 was centrifuged in a 1.5 mL tube at 13000 rpm for 5 min and the supernatant was discarded. Cell pellets were resuspended in 250 µL of Lysis buffer C (50 mM Tris-HCl pH 7.5, 0.5x Bugbuster (Novagen), 1 mM EDTA, 1.5 mg Lysozyme (Sigma), 5 mM $MgCl_2$) with 2.5 U Bezonase nuclease (Novagen) and incubated for 5 min at room temperature. The cell lysate was then centrifuged for 5 min at 13000 rpm and the

soluble fraction was transferred into a new tube. To the insoluble fraction, 250 μL of 8 M Urea was added and vortexed as before.

3.10.5 SDS-PAGE analysis

For SDS-PAGE analysis of the pilot expression of NatPep and TevPep fusions, 10 μL pre-induction, soluble and insoluble samples were mixed with 10 μL Laemmli sample buffer with 0.35 M DTT before being heated at 95^{o}C for 3 min. In addition to the samples, 10 μL Polypeptide Standard diluted 20 times in Laemmli sample buffer with 0.35 M DTT, were loaded onto a Criterion 18% Tris-HCl precast gel and electrophoresed (200 V for 60 min).

For SDS-PAGE analysis of TevPep and MetPep fusion expressions, 20 μL of soluble and insoluble samples were mixed with 10 μL NuPAGE LDS sample buffer (3x) with 0.35 M DTT and heated as above. The samples as well as 10 μL of SeeBlue Plus2 pre-stained standard were loaded onto a NuPAGE 4-12% Bis-Tris gel and electrophoresed (200 V for 35 min). Gels were stained with a solution of 50 mg/L Coomassie Brilliant Blue G-250 and 3 mL/L glacial acetic acid in water.

3.10.6 Affinity (IMAC) purification of TevPep Fusions

Cell pellets (~2 g) stored at -80^{o}C were thawed and 6 mL lysis buffer A and 125 U benzonase (Novagen) was added. Cells were resuspended by vortexing and incubation on a rocking shaker for 15 min before 6 mL lysis buffer B was added and left on a rocking shaker for 20 min. The sample was centrifuged at 15000 g for 10 min and the soluble fraction was removed. To the insoluble fraction, 7 mL of wash buffer (lysis buffer A with 0.1% Triton-X100) was added, vortexed and centrifugation was performed as before. This pellet wash was repeated one more time before the cell pellet was resuspended in 15 mL denaturing equilibration buffer (eq. buf.) (50 mM, NaH_2PO_4, 6 M GuHCl, 300 mM NaCl, pH 7.5) and vortexed. After centrifugation at 12000g for 10 min, ~15 mL supernatant was mixed with 10 mL of Talon metal affinity resin (5 mL bed volume) in a gravity column on a rocking shaker for 20 min. The column was then drained to elute unbound proteins and50 mL eq. buf. was passed through the column. For elution, 40 mL of elution buffer (50 mM NaH_2PO_4, 6 M Gdn-HCl, 300 mM NaCl, 150 mM imidazole, pH 7.5) was used and the eluate was collected and later analyzed by SDS-PAGE. The resin was regenerated with 20 mM MES, 300 mM NaCl, pH 5 buffer according to the manufacturer's manual.

3.10.7 Purification of rEX1 from insoluble GB1-EX1 fusion by CNBr cleavage

Cell pellets (~2 g) stored at -80^{o}C were thawed and 9 mL lysis buffer A and 200U benzonase was added. Cells were resuspended as outlined before and 9 mL of

lysis buffer B was added and left on rocking shaker for 20 min. The sample was centrifuged at 15000 g for 15 min and the soluble fraction was removed. The insoluble fraction was treated with 10 mL wash buffer (50 mM Tris-HCl, 50 mM dextrose) and 1% TritonX-100 twice and then with just 10 mL wash buffer once, with vortexing and centrifugation steps in between each step. The washed cell pellet was dissolved in 8 M urea and 0.1% TFA, vortexed and centrifuged as before to separate the cell debris. The soluble fraction (in 8 M urea, 0.1% TFA) was transferred into a dialysis tube (3500 MWCO) and dialyzed against1 L 0.1% TFA in water at 4^{o}C overnight. The dialysis was repeated with fresh dialysis solution and the dialyzed protein sample was later lyophilized, reconstituted in 0.1% TFA and analyzed by reverse phase analytical HPLC.

Cleavage of GB1-EX1 fusion by CNBr was performed by adding 15 mg/mL CNBr in 70% formic acid to the protein sample to a final concentration of 3.5 mg/mL. The flask was purged with Argon and kept in the dark for 24 h before being diluted 7x in 0.1% TFA and lyophilized.

The lyophilized sample was dissolved in 0.1% TFA and centrifuged. A small amount of soluble fraction was analyzed by RP-HPLC and the rest of the peptide was purified by RP-HPLC (semi-preparative Vydac protein/peptide C18 column) at 40^{o}C employing an isocratic gradient for 1 min with 100% solvent A, then linear solvent B gradients from 0 - 13% B over 13 min, 13 - 27% over 28 min and finally from 27-

37% over 40 min where solvent A is 0.1% TFA in water and solvent B is 0.1% TFA in 9:1 acetonitrile:water. The corresponding peak for EX1 was collected and lyophilized. Analytical RP-HPLC and ESI-MS were performed to analyze the purity and the molecular weight of rEX1 peptide.

REFERENCES

1. Fields, G.B. and R.L. Noble, *Solid-Phase Peptide-Synthesis Utilizing 9-Fluorenylmethoxycarbonyl Amino-Acids.* International Journal of Peptide and Protein Research, 1990. **35**(3): p. 161-214.

2. Carpino, L.A. and G.Y. Han, *9-Fluorenylmethoxycarbonyl Amino-Protecting Group.* Journal of Organic Chemistry, 1972. **37**(22): p. 3404-&.

3. Carpino, L.A. and G.Y. Han, *9-Fluorenylmethoxycarbonyl Function, a New Base-Sensitive Amino-Protecting Group.* Journal of the American Chemical Society, 1970. **92**(19): p. 5748-&.

4. Bray, B.L., *Large-scale manufacture of peptide therapeutics by chemical synthesis.* Nature Reviews Drug Discovery, 2003. **2**(7): p. 587-593.

5. Kelley, W.S., *Therapeutic peptides: The devil is in the details.* Bio-Technology, 1996. **14**(1): p. 28-31.

6. Bruckdorfer, T., O. Marder, and F. Albericio, *From production of peptides in milligram amounts for research to multi-tons quantities for drugs of the future.* Current Pharmaceutical Biotechnology, 2004. **5**(1): p. 29-43.

7. Latham, P.W., *Therapeutic peptides revisited.* Nature Biotechnology, 1999. **17**(8): p. 755-757.

8. Mulder, K.C.L., et al., *Critical Aspects to be Considered Prior to Large-Scale Production of Peptides.* Current Protein & Peptide Science, 2013. **14**(7): p. 556-567.

9. Parachin, N.S., et al., *Expression systems for heterologous production of antimicrobial peptides.* Peptides, 2012. **38**(2): p. 446-456.

10. Nagarkar, R.P. and J.P. Schneider, *Synthesis and primary characterization of self-assembled peptide-based hydrogels.* Methods in molecular biology (Clifton, N.J.), 2008. **474**: p. 61-77.

11. Riley, J.M., et al., *Bioproduction and Characterization of a pH Responsive Self-Assembling Peptide.* Biotechnology and Bioengineering, 2009. **103**(2): p. 241-251.

12. Reed, D.C., et al., *Production and purification of self-assembling peptides in Ralstonia eutropha.* Protein Expression and Purification, 2006. **46**(2): p. 179-188.

13. van Hell, A.J., et al., *Self-assembly of recombinant amphiphilic oligopeptides into vesicles.* Biomacromolecules, 2007. **8**(9): p. 2753-2761.

14. Kyle, S., et al., *Recombinant self-assembling peptides as biomaterials for tissue engineering.* Biomaterials, 2010. **31**(36): p. 9395-9405.

15. Rozkov, A., et al., *Dynamics of proteolysis and its influence on the accumulation of intracellular recombinant proteins.* Enzyme Microb Technol, 2000. **27**(10): p. 743-748.

16. Shen, S.H., *Multiple joined genes prevent product degradation in Escherichia coli.* Proc Natl Acad Sci U S A, 1984. **81**(15): p. 4627-31.

17. Talmadge, K. and W. Gilbert, *Cellular Location Affects Protein Stability in Escherichia-Coli.* Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 1982. **79**(6): p. 1830-1833.

18. Park, M.S., et al., *Catalytic properties of the PepQ prolidase from Escherichia coli.* Arch Biochem Biophys, 2004. **429**(2): p. 224-30.

19. Chandu, D. and D. Nandi, *PepN is the major aminopeptidase in Escherichia coli: insights on substrate specificity and role during sodium-salicylate-induced stress.* Microbiology-Sgm, 2003. **149**: p. 3437-3447.

20. Miller, C.G., *Peptidases and proteases of Escherichia coli and Salmonella typhimurium.* Annu Rev Microbiol, 1975. **29**: p. 485-504.

21. Jain, R. and M.K. Chan, *Support for a potential role of E. coli oligopeptidase A in protein degradation.* Biochem Biophys Res Commun, 2007. **359**(3): p. 486-90.

22. Murby, M., M. Uhlen, and S. Stahl, *Upstream strategies to minimize proteolytic degradation upon recombinant production in Escherichia coli.* Protein Expr Purif, 1996. **7**(2): p. 129-36.

23. Terpe, K., *Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems.* Appl Microbiol Biotechnol, 2003. **60**(5): p. 523-33.

24. Cheng, Y.S., et al., *Stabilization of a degradable protein by its overexpression in Escherichia coli.* Gene, 1981. **14**(1-2): p. 121-30.

25. Hellebust, H., et al., *Different Approaches to Stabilize a Recombinant Fusion Protein.* Bio-Technology, 1989. **7**(2): p. 165-168.

26. Zorko, M. and R. Jerala, *Production of recombinant antimicrobial peptides in bacteria.* Methods Mol Biol, 2010. **618**: p. 61-76.

27. Pazgier, M. and J. Lubkowski, *Expression and purification of recombinant human alpha-defensins in Escherichia coli.* Protein Expr Purif, 2006. **49**(1): p. 1-8.

28. Chen, R., *Bacterial expression systems for recombinant protein production: E. coli and beyond.* Biotechnology Advances, 2012. **30**(5): p. 1102-1107.

29. Li, Y.F. and Z.X. Chen, *RAPD: a database of recombinantly-produced antimicrobial peptides.* Fems Microbiology Letters, 2008. **289**(2): p. 126-129.

30. Hartley, J.L., *Cloning technologies for protein expression and purification.* Current Opinion in Biotechnology, 2006. **17**(4): p. 359-366.

31. Hartley, J.L., G.F. Temple, and M.A. Brasch, *DNA cloning using in vitro site-specific recombination.* Genome Research, 2000. **10**(11): p. 1788-1795.

32. Bernard, P. and M. Couturier, *Cell Killing by the F-Plasmid Ccdb Protein Involves Poisoning of DNA-Topoisomerase-Ii Complexes.* Journal of Molecular Biology, 1992. **226**(3): p. 735-745.

33. Ingham, A.B. and R.J. Moore, *Recombinant production of antimicrobial peptides in heterologous microbial systems.* Biotechnology and Applied Biochemistry, 2007. **47**: p. 1-9.

34. Li, Y.F., *Carrier proteins for fusion expression of antimicrobial peptides in Escherichia coli.* Biotechnology and Applied Biochemistry, 2009. **54**: p. 1-9.

35. Lindhout, D.A., et al., *High-yield expression of isotopically labeled peptides for use in NMR studies.* Protein Science, 2003. **12**(8): p. 1786-1791.

36. Kuliopulos, A. and C.T. Walsh, *Production, Purification, and Cleavage of Tandem Repeats of Recombinant Peptides.* Journal of the American Chemical Society, 1994. **116**(11): p. 4599-4607.

37. Malys, N. and J.E.G. McCarthy, *Translation initiation: variations in the mechanism can be anticipated.* Cellular and Molecular Life Sciences, 2011. **68**(6): p. 991-1003.

38. Benbassat, A., et al., *Processing of the Initiation Methionine from Proteins - Properties of the Escherichia-Coli Methionine Aminopeptidase and Its Gene Structure.* Journal of Bacteriology, 1987. **169**(2): p. 751-757.

39. Huth, J.R., et al., *Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR.* Protein Science, 1997. **6**(11): p. 2359-2364.

40. Cheng, Y. and D.J. Patel, *An efficient system for small protein expression and refolding.* Biochemical and Biophysical Research Communications, 2004. **317**(2): p. 401-405.

41. Gronenborn, A.M., et al., *A Novel, Highly Stable Fold of the Immunoglobulin Binding Domain of Streptococcal Protein-G.* Science, 1991. **253**(5020): p. 657-661.

42. LaVallie, E.R., et al., *Thioredoxin and Related Proteins as Multifunctional Fusion Tags for Soluble Expression in E. coli*. 2002. p. 119-140.

43. Ray, M.V.L., et al., *Production of Recombinant Salmon-Calcitonin by in-Vitro Amidation of an Escherichia-Coli Produced Precursor Peptide.* Bio-Technology, 1993. **11**(1): p. 64-70.

44. Kapust, R.B., et al., *The P1 ' specificity of tobacco etch virus protease.* Biochemical and Biophysical Research Communications, 2002. **294**(5): p. 949-955.

45. Blommel, P.G. and B.G. Fox, *A combined approach to improving large-scale production of tobacco etch virus protease.* Protein Expression and Purification, 2007. **55**(1): p. 53-68.

46. Jacob, F. and J. Monod, *Genetic Regulatory Mechanisms in Synthesis of Proteins.* Journal of Molecular Biology, 1961. **3**(3): p. 318-&.

47. Studier, F.W. and B.A. Moffatt, *Use of Bacteriophage-T7 Rna-Polymerase to Direct Selective High-Level Expression of Cloned Genes.* Journal of Molecular Biology, 1986. **189**(1): p. 113-130.

48. Sweet, C.R., *Expression of recombinant proteins from lac promoters.* Methods Mol Biol, 2003. **235**: p. 277-88.

49. Cheung, R.C.F., J.H. Wong, and T.B. Ng, *Immobilized metal ion affinity chromatography: a review on its applications.* Applied Microbiology and Biotechnology, 2012. **96**(6): p. 1411-1420.

Chapter 4

BACTERIAL EXPRESSION OF EX PEPTIDES WITH HIGH YIELD

4.1 Introduction

The self-assembling β-hairpin peptide EX1, designed to form rigid hydrogels at physiological conditions, was successfully produced in a bacterial expression system as highlighted in Chapter 3. The 4.6 mg/L yield obtained however was not cost-competitive when compared to the chemical synthesis, and thus, alternative strategies to increase the yield of rEX1 were investigated. General strategies for increasing the yield encompass gene optimization for bacteria, promoter selection, bacterial strain screening, tandem repeat expression, fusion partner selection, optimization of culture media and optimization of downstream processing [1-7]. These strategies are evaluated below to identify those that may be improved the production of rEX1 in bacteria.

Gene optimization includes modulating parameters such as codon usage bias, guanine-cytosine (GC) content, DNA polymerase slippage sites and mRNA secondary structure to improve the expression of the target protein [8-11]. Codons are composed of three nucleotides and encode for a specific amino acid in addition to encoding for translation initiation (start-) and termination (stop-codons). Considering there are 20

amino acids and 64 possible codons, some amino acids are encoded by more than one codon and there are differences in the frequency of specific codon usage [12, 13]. Furthermore, even the frequency of codon usage for a specific amino acid changes among organisms from different species [14, 15]. Thus, a heterologous gene should be designed according to the codon usage of the expression host [16]. The expression system developed in Chapter 3 was already optimized with respect to codon usage for the production of EX1. Gene optimization also includes ensuring for optimum GC content which ultimately affects gene stability due to the number of hydrogen bonds formed between guanine and cytosine nucleotides [17]. During transcription, repeated sequences on the DNA can cause DNA slippage in which the addition or deletion of repeating units can occur, and following replication, longer or shorter sequences as well as frameshift mutations, which translate to undesired products, can be generated [18]. After transcription, mRNA can form undesired secondary structures and expression can be greatly reduced [19-22]. A common example is hairpin loop formation which renders the initiation codon inaccessible to the translation machinery and this eventually results in poor expression of the respective protein [23-25]. All the above parameters were already considered during optimization and design of the oligonucleotide sequences for the EX1 peptide, as well as for the respective destination vectors as described in the previous chapter. Thus, in view of increasing the yield of EX1 expression, gene optimization was already performed in Chapter 3.

The bacteriophage T7 promoter system, which recruits T7 RNA polymerase for mRNA production of downstream genes, stands out among other promoter systems since the T7 RNA polymerase is very specific and efficient [26-29]. The elongation rate of T7 RNA polymerase (200 nucleotide/sec) is around 8 times faster than that of *E. coli* RNA polymerase (25 nucleotide/sec) [30]. The LacI/P_{T7lac} system which contains the strong T7 promoter, generates high amounts of transcript that results in high amounts of protein production [31]. The destination vectors used in chapter 3 for the expression of the EX1 peptide also contains the LacI/P_{T7lac} promoter system, and thus this strategy for maximizing protein yield was implemented in Chapter 3.

Bacterial strains are generally chosen according to the protein that needs to be expressed. In cases where poor bacterial growth and viability are noted, a bacterial strain containing the T7 lysozyme gene can be used since it is a natural inhibitor of the T7 RNA polymerase and thus suppresses the basal expression of the polymerase before induction is initiated [32, 33]. In this way, cells can grow to reach a high cell density before they are induced for expression of the target protein [33].

When the genes to be expressed have rare codons that are produced in low amounts by bacteria, bacterial strains that express tRNAs of such rare codons are used [34, 35]. The strain used in Chapter 3 (Rosetta 2) for EX1 peptide expression is able to express rare tRNAs. Thus, with respect to fine-tuning the overall genetic strategy for

the production of EX1, all of the considerations discussed above were implemented in Chapter 3.

However, one approach to increase the yield of the EX1 peptide is to use tandem repeat expression in which a number of EX1 encoding sequences are placed on the same gene with a cleavage site in between them, and once the expression starts, more than one copy of EX1 peptide can be expressed from one gene transcript [36-41]. Such a tandem repeat expression approach may afford a higher production of EX1. This possibility was explored and is described in this chapter.

Yet another way to increase the yield of rEX1 is to decrease the fusion partner size, i.e. to increase the ratio of the EX1 peptide with respect to its fusion partner by making the fusion partner smaller in size. The fusion partner however should not be too small in size: results from chapter 3 showed that the 6His-EX1 was poorly expressed. Here the 6Histag can be considered a small fusion partner. The inability of 6His-EX1 to express well may be because its small size leads to poor cellular stability. All the larger fusion proteins (GB1-EX1, KSI-EX1, Trx-EX1 and GST-EX1) were expressed at relatively higher levels than that of 6His-EX1. An alternative EX1 fusion could be designed in such a way that the fusion size would be in between that of 6His-EX1 and GB1-EX1. This strategy is also explored in this chapter.

Culture media has a very important effect on the overall yield of recombinant proteins. Regular LB media is used in most experiments. Following the inoculation of the transformants into the media, the OD_{600} is generally monitored. At an OD_{600} of 0.6-0.8, the expression is induced by the addition of IPTG and after 4-6 hours of culturing, the cells are harvested [42]. Studier et. al. developed a different media: auto-induction media in which induction occurs according to a change in the carbon source use in the growth environment [43, 44]. Bacteria prefer to use glucose first as a carbon source and when glucose is depleted, lactose is utilized from the surrounding environment [45]. Auto induction media takes advantage of this switching in carbon source preference by including both glucose and lactose in the media. Cells grow until all the glucose is consumed. After which, the cell density is normally very high. At this time, lactose starts being metabolized and the "auto" induction of expression begins. The subsequent expression of T7 RNA polymerase is followed by target gene expression downstream of the T7 promoter region. The cell density monitored by OD_{600} at the end of culturing, can reach very high numbers compared to that obtained with the use of LB media. The use of auto-induction media will also be explored in this chapter.

Lastly, downstream processing starting from cell harvesting to the final recovery of pure peptide is an important set of steps that could have a significant effect on the overall peptide yield. Hence, the conditions under which the solubility of the EX1 peptide is controlled were modulated to optimize the downstream processing of

EX1-protein fusions especially for the pellet wash and solubilization steps which are integral parts of the isolation procedure.

4.2 Tandem Repeat Expression

In a tandem repeat expression approach, there is no ideal repeat number affording the highest yield, and different repeat numbers may result in different yields for different target peptides. So the general approach is to incorporate a different number of repeating units interspersed by a methionine codon, and screen for the best combination that gives the highest expression/yield. Riley *et al*. studied the expression of KSI fusion constructs that contained tandem repeats (1-6) of the P_{11}-4 peptide: the highest yield was achieved with a construct having 3 repeats, however the yield decreased when the number of repeats were increased to 4-6 [37]. Met residues are employed at the end of each peptide unit to enable CNBr cleavage which yields the individual peptides. A consequence of Met cleavage is that homoserine lactone (HSL) moiety is generated at the C-terminus of each peptide following the cleavage reaction. Thus, one must decide if this alteration in sequence can be tolerated for the intended use of the peptide.

For tandem repeat expression of the EX1 peptide in this study, 2, 3 and 6 repeats were chosen for investigation and the DNA constructs $(EX1)_2$, $(EX1)_3$ and $(EX1)_6$ were designed respectively as shown in Figure 4.1. These sequences will

eventually be cloned into a destination vector that starts with a hexahistidine tag (6Histag). Although this purification tag is not ultimately used to purify the peptide, I found that having the 6His gene upstream of the peptide genes helped in efficient transcription. In addition to these constructs, an alternative entry clone was constructed that has a ribosome binding site, 3 repeats of the EX1 oligonucleotide sequence followed by a 6Histag at the C-terminus as a fusion partner. This alternative sequence was prepared to investigate the effect of a positional difference when the gene starts with 3 repeats of EX1 instead of a 6Histag. Although not necessary, the C-terminus of the last EX1 of the repeat was followed by a Met-Gly so that CNBr cleavage could yield a rEX1-HSL unit. If this Met-Gly unit was not included, CNBr would yield both rEX1-HSL and rEX1 peptides and resolution of the two peptides could be problematic during the HPLC purification.

$(EX1)_2$	**attB1-(Met-EX1)$_2$-MetGly-stop-attB2**
$(EX1)_3$	**attB1-(Met-EX1)$_3$-MetGly-stop-attB2**
$(EX1)_6$	**attB1-(Met-EX1)$_6$-MetGly-stop-attB2**
$(EX1)_3$-6His	**attB1-RBS-(Met-EX1)$_3$-MetGly-6His-stop-attB2**

Figure 4.1 Design of oligonucleotides for tandem repeat expression. Different number of repeats were placed in between att sites. RBS: ribosome binding site.

The $(EX1)_2$, $(EX1)_3$ and $EX1)_6$ inserts were cloned into the destination vector pDEST527 containing an upstream 6Histag fusion partner. The $(EX1)_3$-6His construct was cloned into the destination vector pDEST521. The resulting expression clones:

6His-$(EX1)_2$, 6His-$(EX1)_3$, 6His-$(EX1)_6$ and $(EX1)_3$-6His are shown in Figure 4.2. After transformation and inoculation, pilot expression studies were performed.

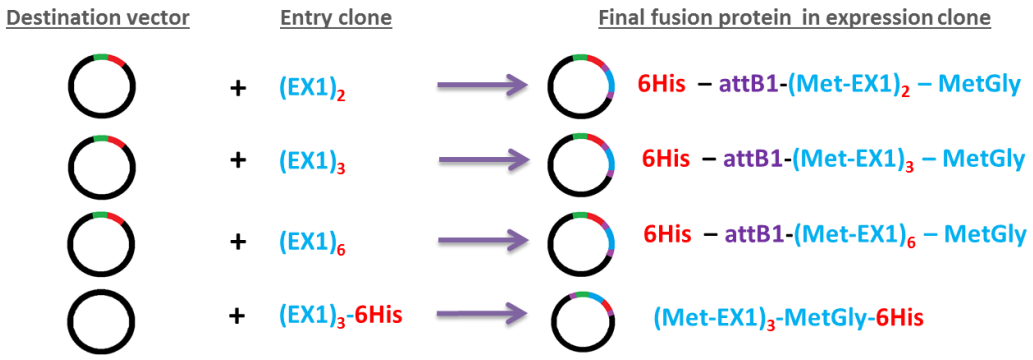


Figure 4.2 Generation of expression clones for EX1 tandem repeats. Subcloning of the $(EX1)_2$, $(EX1)_3$ and $(EX1)_6$ inserts into destination vector pDEST527 containing RBS-6Histag and subcloning of the $(EX1)_3$-6His insert into destination vector pDEST521 containing no RBS and tag. Green portion is RBS, red portion is 6Histag, blue portion is tandem repeats and purple portions are att sites.

For the first pilot expression study, 25 mL LB media was inoculated with 250 µL (1%) starter overnight culture and the cells were grown until the OD_{600} was ~0.6 for about 3 h before being induced with 1 mM IPTG. Culturing at 37°C for 4.5 h after induction resulted in poor growth as evaluated by measuring the OD at 600 nm. The final OD_{600} of 6His-$(EX1)_3$ and $(EX1)_3$-6His containing cultures were ~1.2 and that of the 6His-$(EX1)_2$ and 6His-$(EX1)_6$ containing cultures were ~2.2. Typically an OD_{600} of ~3-4 is seen with bacterial cultures grown at 37°C for 4.5 h post induction. The poor growth could be attributed to the EX1 peptide being toxic to the cells when expressed as tandem repeats. Previous studies (Chapter 3) with the 6His-EX1 fusion,

which only contains one EX1 unit also showed poor growth following induction. This may suggest that EX1, itself, is highly toxic to the bacteria.

A second pilot expression study was performed in a similar way and after IPTG induction, the culture was grown at a lower temperature of 25^{o}C for 11 h to reduce the rate of expression. The final cell density of the 6His-$(EX1)_3$ and $(EX1)_3$-6His containing cultures were once again low with OD_{600} values of ~1.3. Thus, the 6His-$(EX1)_3$ and $(EX1)_3$-6His containing cultures were unable to grow at either 25^{o}C or 37^{o}C following induction. On the other hand, the final cell density of cultures containing the 6His-$(EX1)_2$ and the 6His-$(EX1)_6$ had OD_{600} values of ~3.6.

Cells were harvested from cultures grown at 25^{o}C and lysed. The soluble and insoluble fractions were separated and analyzed by SDS-PAGE (Figure 4.3). Among the 4 different constructs, only the one containing two repeats was expressed and observed at the top of the gel due to precipitation. For the other three constructs; 6His-$(EX1)_3$, $(EX1)_3$-6His and 6His-$(EX1)_6$, no bands corresponding to the desired products were observed. Although normal growth was observed for the culture containing 6His-$(EX1)_6$, the gene was not transcribed/translated to protein. This might be due to its highly repetitive nature.

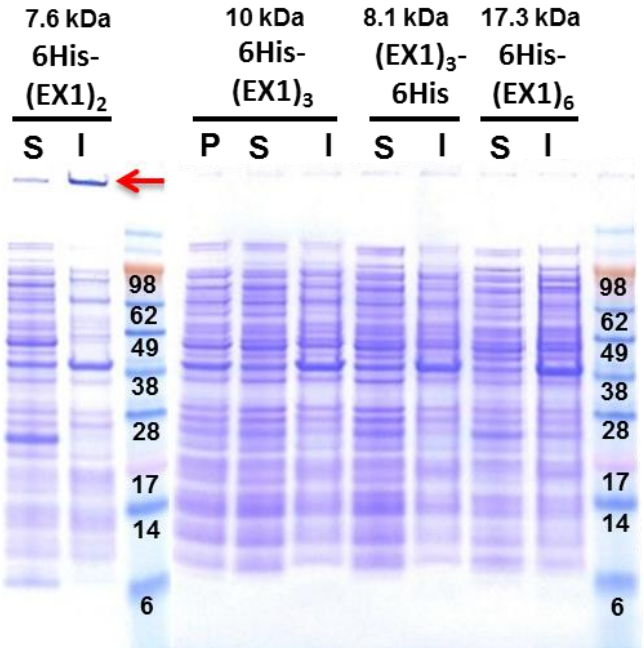


Figure 4.3 SDS-PAGE analysis after pilot expression of tandem repeats (25°C). After cell lysis, the separation of the soluble (S) and insoluble (I) fractions are analyzed. Expected MWs are shown above the tandem repeat names. P: Pre-induction sample, red arrow shows the precipitated protein containing 2 repeats.

A larger scale production, using 0.8 L of media was used to produce the 6His-$(EX1)_2$ containing transformants in order to purify and characterize the rEX1-HSL peptide. After the cell lysis and insoluble pellet wash, 8 M guanidine-HCl (0.5% TFA) was used to solubilize the pellet. Cell debris was then removed by centrifugation. The soluble fraction was filtered and injected onto an RP-HPLC column to purify the 6His-$(EX1)_2$ fusion and the eluted sample was then lyophilized in preparation for CNBr cleavage.

Previous studies in chapter 3 in which the washed cell pellet was directly cleaved in 70% formic acid containing CNBr resulted in a sticky precipitate after lyophilization. Due to very poor solubility of this precipitate purification and peptide recovery was very inefficient. Isolating the fusion protein by HPLC and lyophilization prior to chemical cleavage resulted in a powder that was easily solubilized in formic acid. Subsequent cleavage using CNBr and purification by HPLC afforded pure peptide. From one liter of LB media, the yield of pure rEX1-HSL was 14 mg. This low yield is most likely due to the fact that when the solubilized pellet (in 8 M guanidine-HCl) was injected onto the C18 RP-HPLC column, some of the peptide remained stuck on the column and did not elute. This highly adsorbed fraction could only be eluted with extensive column washings. Although not optimal, this isolation strategy yielded enough rEX1-HSL for further characterization. Figure 4.4A and B show the analytical HPLC chromatogram and corresponding mass spectral data for pure rEX1-HSL.

4.2.1 Characterization of rEX1-HSL

Temperature dependent CD spectra of 150 μM solutions of rEX1-HSL at 216 nm from 5^{o}C to 80^{o}C (pH 7.4, 50 mM BTP, 150 mM NaCl) is compared with that of chemically synthesized EX1 in Figure 4.4A. At temperatures greater than 20^{o}C, rEX1-HSL begins to fold into a β-sheet structure and at around 40^{o}C, it is mostly folded. Under these solution conditions, rEX1-HSL folds at slightly lower temperatures when

compared to EX1. Hairpin folding and assembly is partly driven by the hydrophobic effect. More hydrophobic peptides fold and assemble at lower temperatures when compared to more hydrophilic sequences. The folding of rEX1-HSL at lower temperatures can be attributed to its more hydrophobic nature when compared to EX1. Its more hydrophobic nature was also validated by its observed later elution time during analytical HPLC when compared to EX1.

Next, the gel forming properties of rEX1-HSL and EX1 were compared. Dynamic time sweep experiments in Figure 4.4B show the evolution of storage modulus as a function of time for rEX1-HSL at physiological conditions. At the end of one hour, rEX1-HSL storage modulus reached around 1.7 kPa which was greater than the storage modulus of EX1 (~1 kPa). In the time sweep experiment, the temperature is ramped from 5^{o}C to 37^{o}C in 100 sec to initiate folding and assembly of the peptides. A short lag time in the evolution of storage modulus was observed for rEX1-HSL. Because rEX1-HSL folds and assembles at a lower temperature compared to rEX1, it was expected that the evolution of G` should be faster for rEX1-HSL. This unexpected result can be due to the C-terminal HSL ring trying find a conformation that allows the formation of additional hydrogen bonds between N-terminal amine and the oxygen atoms of the HSL ring as depicted in Figure 4.5. Once such hydrogen bonds form, the hydrophobic face of the folded hairpin can be extended over the β and γ carbons of the HSL ring and this 'increased hydrophobic surface' can facilitate the formation of more branch points in the fibril network by acting as a surface from which additional fibrils

can grow. Frequency sweep experiments of rEX1-HSL and EX1 are shown in Figure 4.4C. These experiments show that both peptides provide gels whose storage and loss moduli are invariant with frequency. Panel F shows strain sweep experiments for those two gels. The rEX1-HSL gel yields at a higher % strain than the EX1 gel. The frequency and strain sweep experiments validate that the time-sweep experiments were performed in the linear regimes of strain and frequency. In conclusion, although the rEX1-HSL peptide and its corresponding gel have interesting folding, assembly and gel rheological properties, its relatively low expression level precluded further study.

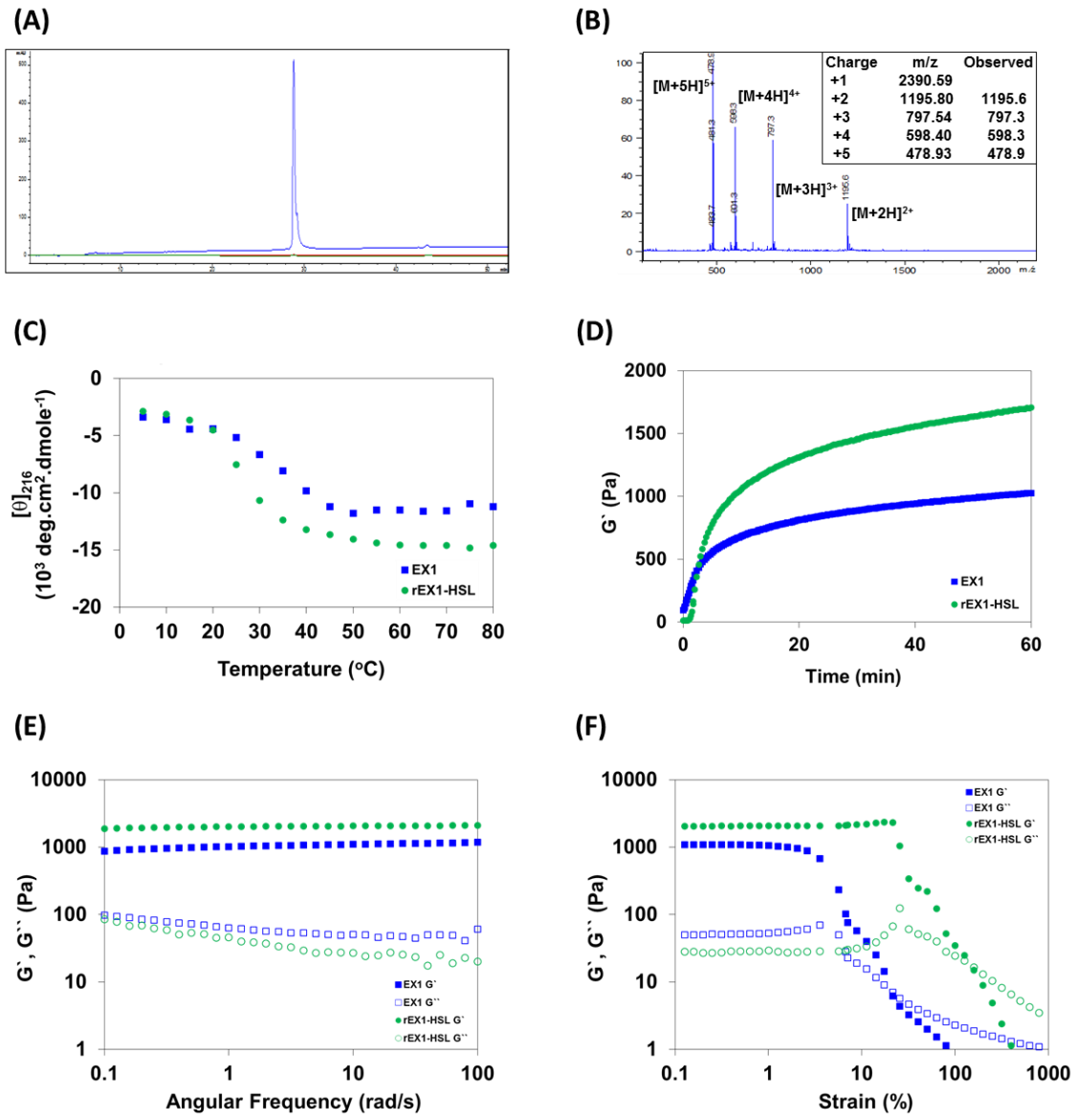


Figure 4.4 Temperature dependent β-sheet formation and mechanical properties of rEX1-HSL and EX1. (A) HPLC (analytical) of rEX1-HSL. (B) Mass spectrum of rEX1-HSL. (C) Mean residue ellipticity of 150 μM peptide solutions at 216 nm from 5°C to 80°C. (D) Dynamic time sweep experiment of 1 wt% peptides. (E) Dynamic frequency sweep and (F) dynamic strain sweep experiments after 1 h of gelation. Mechanical properties assessed at 37°C, pH 7.4, 50 mM BTP, 150 mM NaCl.

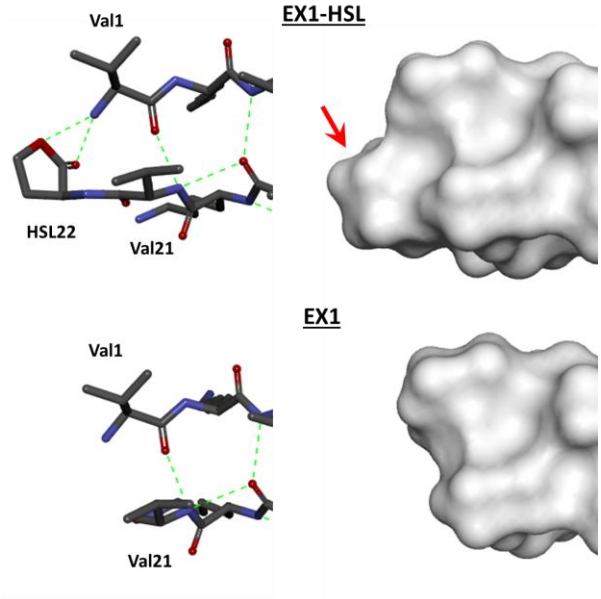


Figure 4.5 Comparison of EX1-HSL and EX1 termini. Stick representations and solvent accessibility surfaces of EX1-HSL and EX1. Red arrow shows the HSL ring that extends the surface of hydrophobic face of the peptide.

4.3 Design of a Low Molecular Weight Fusion Partner

The theoretical yield of rEX1 in a fusion construct is partly based on the relative molecular weight of its protein fusion partner. A smaller fusion partner yields more rEX1 but on the other hand, if the overall size of the fusion is too small, limited stability or degradation could be problematic. 6His-EX1 (5.7 kDa), which can be thought of as a construct having an extremely small fusion partner, namely the 6Histag, was shown to express but only in small amounts. GB1-EX1 (15 kDa), a construct having a large fusion partner was expressed at a higher level as indicated in Chapter 3. Hence, a fusion partner with a molecular weight intermediate of these two extremes was sought. The BH3 domain of the proapoptotic BAD (Bcl-2 Associated

Death) protein is a small α-helical peptide (BAD-BH3, 3.2 kDa, pI 8.6) that is small but structured [46]. After cloning into the 6Histag containing destination vector, the molecular weight of the fusion was 9.2 kDa which is expected to express relatively better than 6His-EX1. Figure 4.6 shows the oligonucleotide design of the BAD-EX1 insert to be cloned into the 6Histag containing destination vector. The fusion protein starting with 6Histag and followed by Bad-BH3, Met and EX1 was named 6His-BAD-EX1.

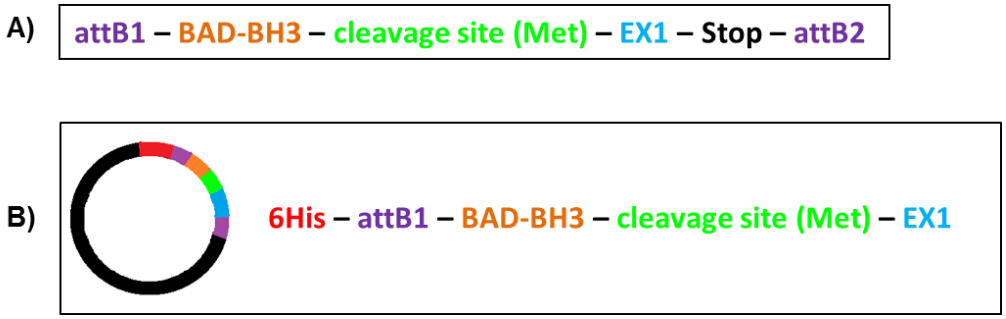


Figure 4.6 (A) Design of oligonucleotide insert with BAD-BH3. (B) The ultimate cloning vector

Cloning, transformation and inoculation into the culture media was performed for a pilot expression study. Figure 4.7 shows the soluble and insoluble fractions after cell lysis. The fusion protein was found in the insoluble fraction as expected and expressed at a relatively good level, taking into consideration that some precipitated protein could not enter the gel and remained at the top of the gel.

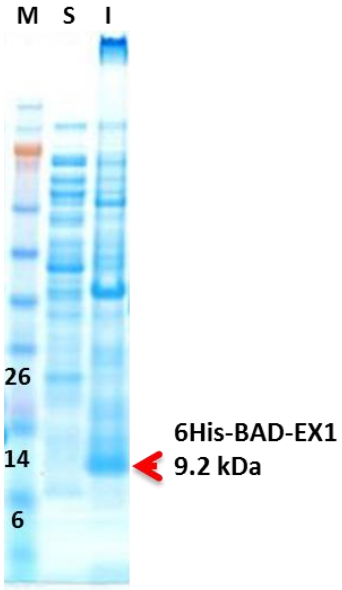


Figure 4.7 SDS-PAGE analysis of 6His-BAD-EX1 expression. Cell lysate was centrifuged and soluble (S) and insoluble (I) fractions were loaded onto the 4-12% gel. M: Marker

Given the favorable expression level in this pilot study, the 6His-BAD-EX1 system was further optimized.

4.4 Optimization of Bacterial Culture Media

Bacterial cultures typically employ standard LB medium, which has carbon and nitrogen sources originating from the yeast extract and tryptone. Following inoculations, the bacteria start to divide and propagate. When the metabolic activity is optimum during the logarithmic growth phase, expression of the desired protein is induced by the introduction of IPTG to the media. IPTG allows the production of T7 RNA polymerase, which binds to the T7 promoter region upstream of the gene of interest and facilitates its transcription. After induction, nearly all the bacterial

machinery works towards expressing the target protein and as a result, up to 50% of all bacterial protein will constitute the target protein [26]. This efficient system can be also employed with the auto induction media developed by Studier et. al. [43]. When auto-induction media is used, the bacteria preference for carbon source determines the induction time. When there is glucose in the environment, it is the bacteria's first choice as a carbon source until it is depleted. Once there is no more glucose available, lactose is the bacteria's next choice. Lactose is the natural inducer of the β-galactosidase gene in the *lac* operon in which the target gene is inserted. Hence, lactose intake from the surrounding environment is not only used as a carbon source, but also as the trigger for target gene expression. Exploring alterations to the culture media may be an avenue to realize better yields of peptide.

Figure 4.8A lists the final culture densities (monitored at OD_{600}) in LB, auto-induction (AI) and a modified version of the auto-induction (mAI) media. The LB media yielded relatively low density cultures after 4 hours post induction. With the auto-induction media, the cell density was 4 times higher at the end of 44 hours of culturing due to the spontaneous induction at the saturation density. The presence of glucose as an early energy source and glycerol as a late energy source provides additional energy to the cells when compared to LB media. Modified auto-induction media includes 8 times more of yeast extract and tryptone than that of regular auto-induction media. In addition, the modified media contains 4 times more glucose, glycerol and lactose. This increase in nutrients, in addition to the increase of carbon

source, promotes cell growth resulting in extremely high cell densities before the induction is triggered. Although it takes a longer time to reach these high cell densities, more cell mass and more protein per cell mass can be harvested. Figure 4.8B shows the SDS-PAGE analysis of the 6His-BAD-EX1 obtained from the same amount of bacteria cultured in LB, AI and mAI media. The band corresponding to the target protein (shown by the red arrow) obtained from the mAI media was significantly larger than the band obtained from the AI media, suggesting that extra glucose, glycerol and lactose in the mAI media increases target protein expression. Having optimized the media, I next set my attention to optimizing the purification protocol.

A)

Media	max OD_{600}
LB	2.5
AI	10.8
mAI	18.7

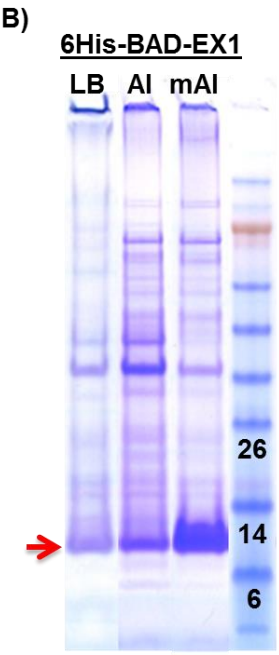


Figure 4.8 6His-BAD-EX1 expression in different media. (A) Cell density monitored at OD_{600} at the end of culturing in each media. (B) SDS-PAGE analysis of His-Bad-EX1 expression in LB, auto-induction (AI) and modified AI media (mAI).

4.5 Optimization of Purification

The isolation of the 6His-BAD-EX1 fusion, its chemical cleavage and the purification of the resulting rEX1 was performed by a process dependent on the folding and self-assembly characteristics of the parent EX1 peptide. By changing the environmental conditions of the EX1 containing fusion protein, the solubility of the fusion protein can be controlled. The charge repulsion between the lysine residues of the EX1 peptide in the fusion protein can keep the fusion soluble. However, these charge interactions can be screened with salt to induce precipitation of the fusion during the cell pelleting step directly after the cells are harvested. The pellet can then be washed to remove residual bacterial proteins. Once the cell pellet that contains insoluble fusion protein is mostly free of bacterial proteins, it can be suspended in low pH, low ionic strength buffer to promote the solubilization of the fusion protein. The solubilized fusion protein can be separated after sedimenting insoluble by-products. The resulting supernatant can be lyophilized for subsequent chemical cleavage and HPLC purification of the rEX1 peptide.

Figure 4.9 summarizes the overall downstream processing method. Inclusion bodies have relatively higher densities than that of bacterial proteins, and thus by centrifugation, they can be easily isolated. Sonicating and washing the pellet mass with detergent allows the separation of bacterial proteins from the insoluble fraction containing the desired peptide fusion. The washing step included 100 mM NaCl which

kept the peptide in an aggregated state. Thus, the washing step which includes sonication in the presence of detergent did not result in loss of fusion peptide. Additional washing steps were performed without detergent carried over into the insoluble fraction. The next step was to dissolve the fusion protein using a low pH and low salt buffer. Sonication aided the dissolution of the fusion. In this step, a 0.1% TFA solution was used to protonate the lysine residues of EX1, resulting in the solubilization of the fusion construct. Under these solution conditions, most of the fusion construct becomes soluble. The supernatant containing the soluble fusion was then separated by centrifugation from the insoluble fraction that was mostly composed of cell debris. The soluble fraction was then lyophilized to afford a crude powder that could be chemically cleaved by CNBr in small batches. After the rEX1 was cleaved from the fusion partner, it was purified by RP-HPLC. By using the method summarized in Figure 4.9, affinity purification of the fusion protein was avoided, and a straightforward path to pure rEX1 could be realized.

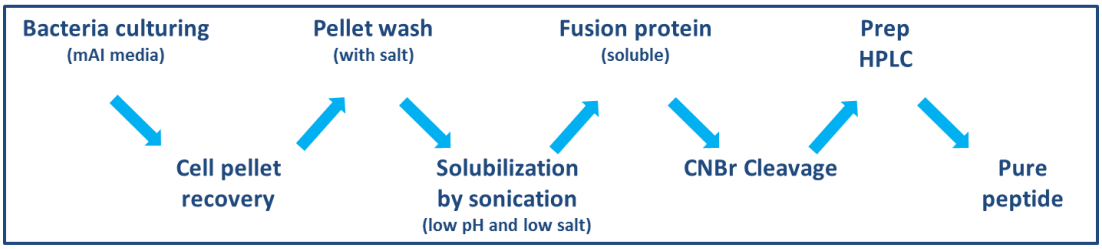


Figure 4.9 Schematic of downstream processing.

4.5.1 Final yield of rEX1 and cost analysis

After culturing the transformant bacteria containing the 6His-BAD-EX1 fusion (9.2 kDa) in modified auto-induction media, and following the downstream processing explained above, the overall yield was ~50 mg pure rEX1 obtained from 1 L of culture media. The first set of experiments yielded 4.6 mg rEX1 and the alternative studies investigated in this chapter succeeded in increasing the overall yield to more than 10 times the original yield.

A cost comparison was performed between chemical synthesis and bacterial production. The chemical synthesis yields ~150 mg pure peptide after semi-preparative HPLC purification from a 0.25 mmol scale synthesis with Fmoc chemistry on an automated peptide synthesizer and the cost is ~$530 including the cost of solvents used for synthesis and purification as well as synthesizer expenses. Synthesis of the crude peptide is ~75% of the cost and peptide purification is the remaining ~25% of the cost. On the other hand, the bacterial production cost of the same amount of peptide (150 mg) is ~$344. Just ~10% of this cost is due to the auto-induction media and the remaining 90% is for RP-HPLC purification. Thus, the majority of the cost associated with bacterial production is the cost of final purification. At any rate, bacterial production of EX1 proved cost effective.

The bacterial production of rEX1 was successful affording a high yield of peptide using the optimized conditions determined in this chapter. I next sought to determine if this optimized methodology could be extended to other β-hairpin peptide sequences. The sequence of rEX1 contains a five residue turn, namely –VPDGT-. In the next sections of this chapter, I applied optimized bacterial production methods to the production of two additional peptides, namely EX2 and EX3, peptides that contain different turn sequences.

4.6 Bacterial Production and Final Yields of EX2 and EX3

The EX2 peptide contains the five residue turn -VPIGT- in place of EX1`s turn. EX3 is the third peptide and has a four residue turn -YNGT-, replacing the five residue turns of EX1 and EX2. The expression of EX2 and EX3 begins by first designing the appropriate oligonucleotide sequences (Figure 4.10). In the original Bad-BH3 fusion design, there was an internal Met residue in the sequence of the α-helical Bad-BH3 fusion partner which complicated the chemical cleavage of peptide. CNBr cleaves at Met residues, thus the CNBr cleavage of 6His-BAD-EX1 was yielding 3 peptides. Hence the Met residue in the sequence of the α-helical Bad-BH3 was changed to Leu, which has a high propensity to adopt helical dihedral angles, resulting in a sequence not cleaved by CNBr. The changes to the turn region of EX1 were performed according to the turn region sequences of EX2 and EX3 shown underlined in Figure 4.10 by the second red arrow.

Figure 4.10 Final fusion sequences for rEX2 and rEX3 production. The changes performed on the sequence of Bad-EX1 to yield the fusion proteins 6His-BAD-EX2 and 6His-BAD-EX3 are shown by red arrows indicating an internal Met residue changed to Leu and the EX1 turn region changed to EX2 and EX3 turn regions. Met residue in red shows the CNBr cleavage site to afford the recombinant peptides.

BAD-EX2 and BAD-EX3 oligonucleotides were individually cloned into 6Histag containing destination vectors. Transformation into bacteria was followed by culturing in modified auto-induction media as previously outlined. At the end of culturing, cells were harvested by centrifugation and lysed to recover the insoluble pellet containing the fusion proteins. Pellet washes were followed by sonication at low pH (0.1% TFA) and low salt. However, this solution was unable to dissolve the majority of the inclusion bodies. Therefore, the TFA content was then increased to 1% and most of the inclusion bodies dissolved. The resulting solutions were lyophilized. CNBr cleavage was performed by dissolving the lyophilized 6His-BAD-EX2 and 6His-BAD-EX3 fusions in 70% formic acid. Isolation of the pure rEX2 and rEX3 peptides was enabled by RP-HPLC.

The final yields of pure rEX2 and rEX3 were 31 mg and 15 mg peptide from one liter of culture media, respectively. These yields are comparable, but lower when

compared to the rEX1 yield (50 mg/L). The relatively lower yields of rEX2 and rEX3 might be due to the higher propensity of EX2 and EX3 fusion proteins to aggregate resulting in lower solubility, which would detrimentally effect their resolubilization during their isolation. The change of methionine residue to a more hydrophobic leucine residue in the BAD fusion partner certainly increased the overall hydrophobicity of the EX2 and EX3 fusion proteins. In fact, after washing the pelleted inclusion bodies of the 6His-BAD-EX1 fusion, 0.1% TFA was able to fully solubilize the pellet. In contrast, EX2 and EX3 fusions required a higher concentration of TFA (1%). Even in 1% TFA, insufficient solubilization might explain the lower yields of rEX2 and rEX3. This suggests that more soluble fusion partners, which may have a positive impact on the recovery of the fusion proteins and lead to higher yields of the recombinant EX peptides, may be needed.

4.7 Comparison of Recombinant Peptides with Chemically Synthesized Peptides

Characterization of chemically synthesized (chem) and recombinant rEX peptides and their corresponding gels was performed by CD, rheology and transmission electron microscopy to assess their secondary structure, elastic properties and local fibril morphologies, respectively. The properties of chemically versus recombinantly-made peptides (and their gels) should be similar since their characteristics should be independent of the method of production.

4.7.1 Secondary structure characterization

Figure 4.11 shows the wavelength CD spectra of 1wt% EX peptide gels from chemical synthesis and bacterial production at physiological conditions after 1 h equilibration time. All the EX peptides contained in the gelled state have characteristic minima around 216 nm indicating that the peptides fold into β-sheet structure. Moreover, the spectra for each distinct peptide, whether produced chemically or by fermentation, are similar. For example, bacterially produced EX2 produces a gel having a similar CD spectrum to that of a gel prepared from chemically synthesized EX2, (red and orange data).

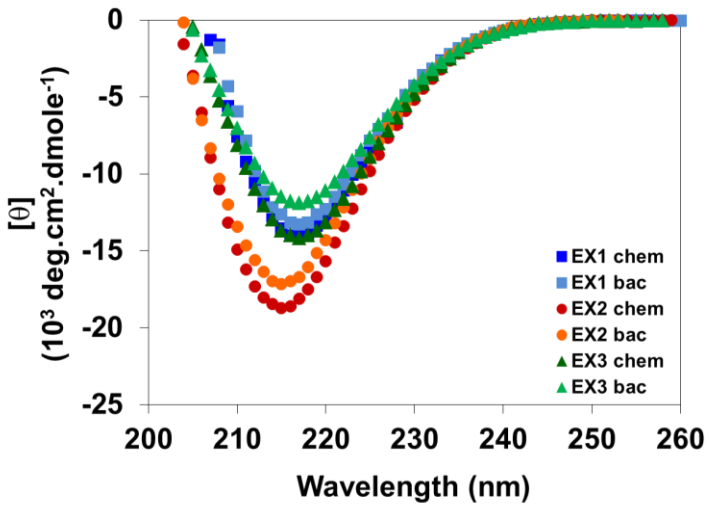


Figure 4.11 Wavelength CD spectra of EX peptides. Wavelength spectra of 1 wt% EX peptides obtained by chemical synthesis (chem) and bacterial production (bac) at 37°C after 1 hour (pH 7.4, 50 mM BTP, 150 mM NaCl).

4.7.2 Mechanical properties characterization

The elastic properties of the gels from the EX peptides were assessed by oscillatory rheology at physiological conditions by dynamic time, frequency and strain sweep experiments (Figure 4.12). EX peptide solutions (1 wt%) in buffer were prepared on ice and immediately transferred to a rheometer plate at 5^{o}C. The temperature was ramped to 37^{o}C in 100 sec to initiate folding and self-assembly and the evolution of storage moduli were monitored during the temperature ramp and the following 1 h (Figure 4.12A). The data show that after 1 hour of gelation, the storage moduli of the EX peptides reached around 1 kPa and each of the EX peptides showed a similar trend in the evolution of storage moduli as a function of time, independent of the method used to prepare the peptide (e.g. chemically vs. recombinantly). Then angular frequency sweep experiments shown in Figure 4.12C and D were performed by using chemically synthesized peptides and recombinant peptides, respectively. The data show that the storage moduli of the EX peptides were independent of the frequency from 0.1 to 100 rad/sec. The strain sweep experiments of chemically synthesized peptides and recombinant peptides are shown in Figure 4.12E and F, respectively. EX peptides yield between 5 and 30% strain. Independent shear-thinning and recovery experiments were performed to investigate the recovery of the elastic properties of the hydrogels after being subjected to shear-thinning (Figure 4.12B). The peptide solutions were prepared the same way and after 20 min of gelation, the strain was increased from 0.2% strain to 1000% for 30 sec to thin the gels. The immediate

decrease of G` is evident in Figure 4.11B. After thinning, the strain was reduced to 0.2% over 30 sec, and the storage moduli of the EX peptides was monitored as a function of time. The data show that each peptide gel was capable of recovering their solid-like properties after being thinned. EX2 and EX3 gels recovered 80% of their G` while EX1 recovered 50% of its original G`. Importantly, with respect to each individual peptide, its corresponding gel behaved similarly independent of the method of production of its parent peptide.

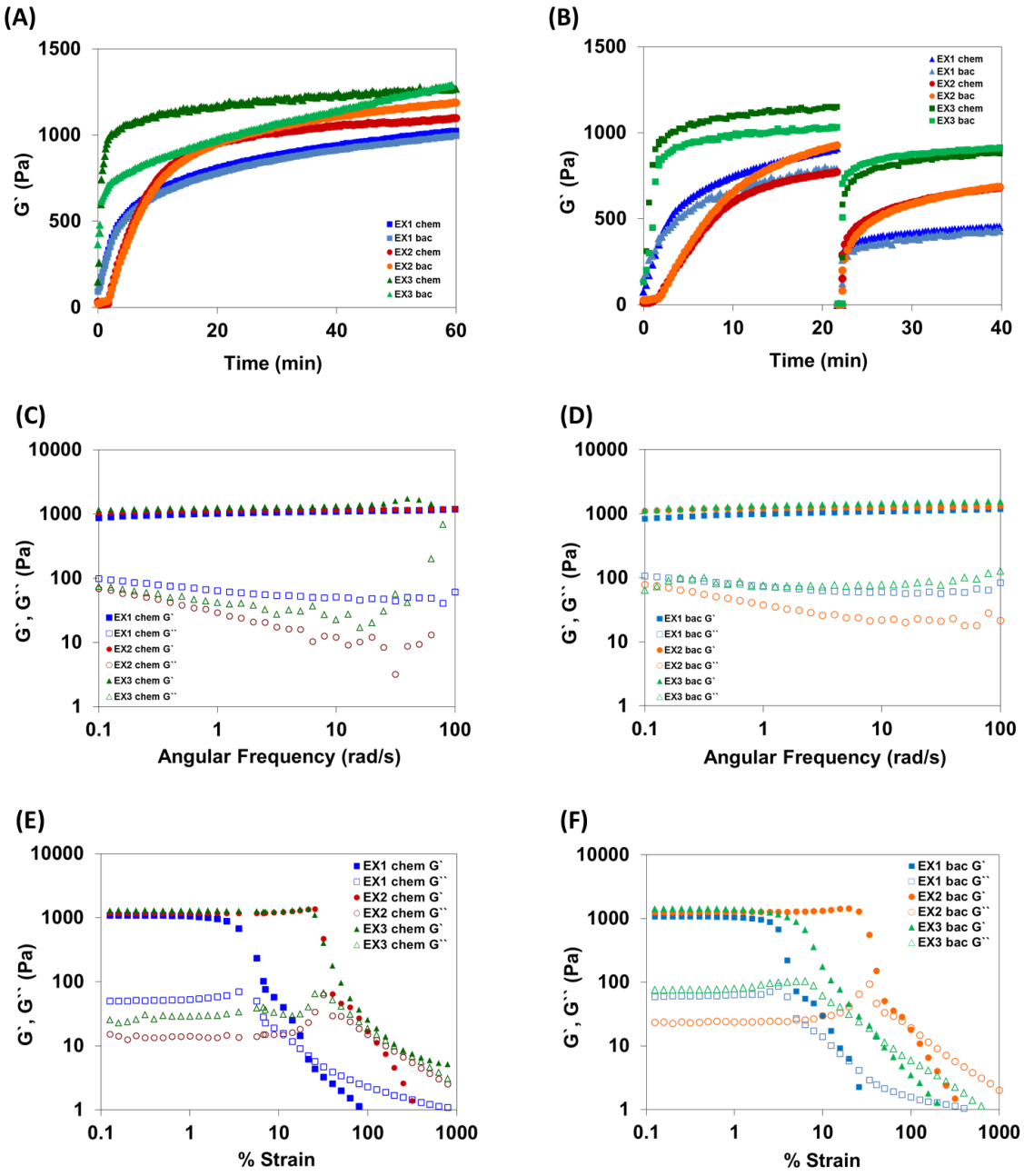


Figure 4.12 Comparison of mechanical properties of EX peptides by oscillatory rheology. Peptides characterized at 1 wt% concentration at 37°C, pH 7.4, 50 mM BTP, 150 mM NaCl. (A) Evolution of storage moduli of EX peptides as a function of time. (B) Shear-thin and recovery experiments with 1000% strain for 30 sec after 20 min of gelation followed by another 20 min for recovery. Dynamic frequency sweep experiments of (C) chemically synthesized EX peptides and (cont.)

(D) recombinant EX peptides after 1 hour of gelation. Dynamic strain sweep experiments of (E) chemically synthesized EX peptides and (F) recombinant EX peptides after 1 hour of gelation.

4.7.3 Characterization of the local morphology of fibrils produced by recombinantly expressed peptides

By transmission electron microscopy (TEM), local fibril morphologies of the EX peptides recovered from bacterial production were investigated and compared with chemically synthesized EX peptides in Figure 4.13. In Figure 4.13A, a slight degree of lamination for chemically synthesized EX1 and some fibril bundling for chemically synthesized EX2 and EX3 were observed, and in Figure 4.13C, similar morphologies were noted for the fibrils formed by the corresponding recombinant EX peptides. Thus, transmission electron microscopy images showed that the rEX peptides afforded similar fibril morphologies when compared to fibrils formed by the chemically synthesized EX peptides. Figure 4.13B and D shows the fibril width distributions for each EX peptide. The majority of fibrils had a width around 3.5 nm which corresponds well with the width of a folded hairpin peptide in the self-assembled state.

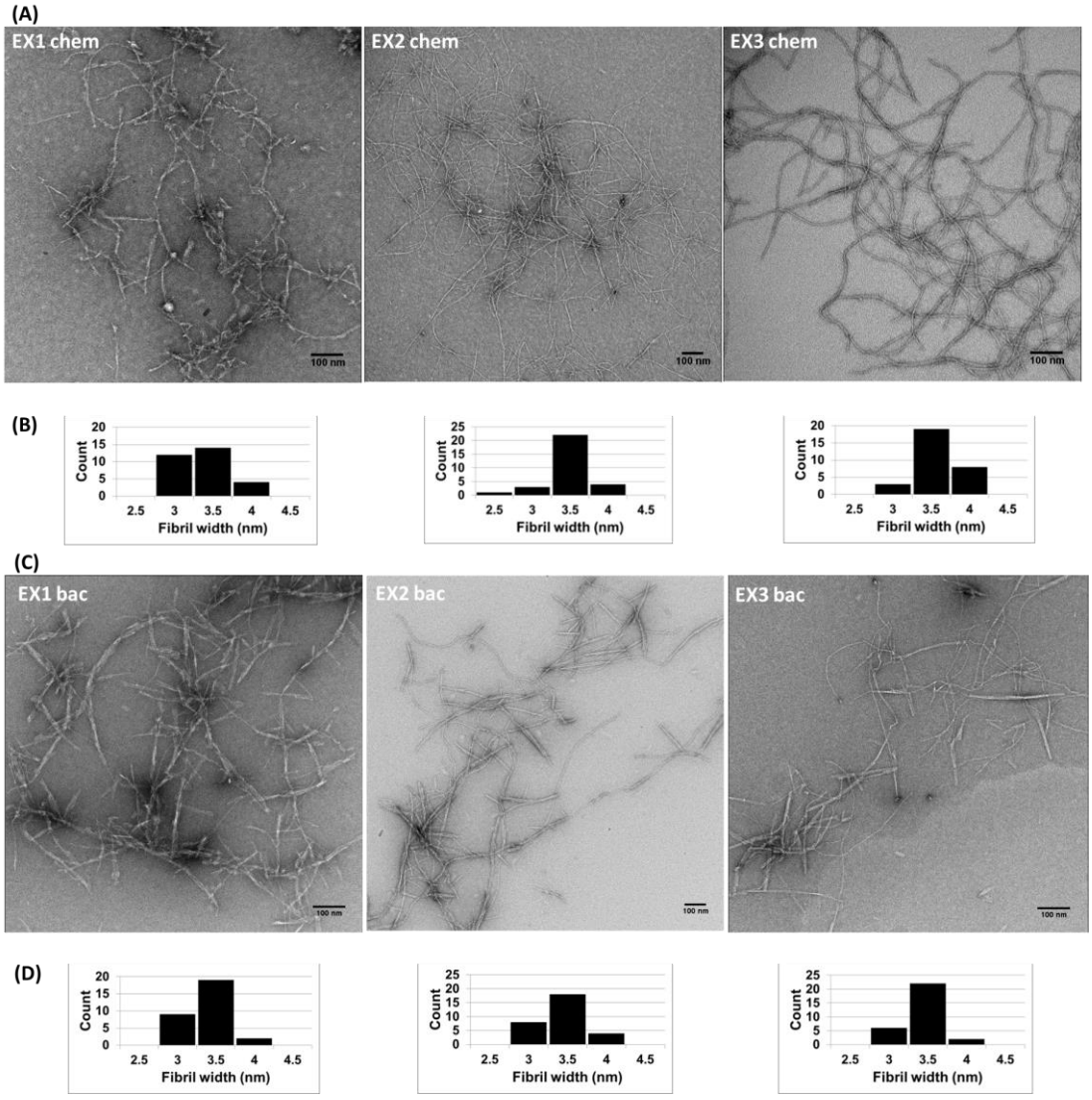


Figure 4.13 Local morphology of EX fibrils assessed by TEM. 1 wt% peptides were gelled at 37°C for 1 hour. Gels were diluted and deposited for TEM examination. TEM images show fibrils formed by EX peptides obtained by (A) chemical synthesis and (C) bacterial production together with their corresponding (B and D, respectively) individual fibril width distributions. Images were taken by Katelyn Nagy.

Therefore, the TEM investigation showed that the fibrils formed by each of EX peptides have similar morphologies independent of how they were produced.

4.8 Conclusions

Steps toward improving the yield of recombinantly expressed EX1 were taken. The optimization described in this chapter includes: 1) Assessing whether using tandem repeats is a viable option to increase yields of expressed peptide. 2) Optimizing the size of the fusion partner in the peptide-fusion construct, and 3) optimizing the culture media.

The tandem repeat expression approach was successful for a two-repeat containing construct (6His-$(EX1)_2$). However, the yield was only 14 mg from a liter of LB media. In addition, the tandem repeat method produces peptide having a homoserine lactone at its C-terminus. This additional functionality is not necessarily desirable.

Characterization studies of 1 wt% rEX1-HSL gels at physiological conditions afforded stiffer gels when compared to EX1 gels formed under the same conditions. This might be due to an increase in hydrophobicity of the peptide due to the HSL at the C-terminus. The HSL moiety might also form additional hydrogen bonds and extend the hydrophobic surface at the termini of the folded hairpin, which can lead to an increase in fibril branch during the evolution of the gel network. 1 wt% rEX1-HSL gels had a G` of 1700 Pa after 1 h of gelation. Dynamic frequency experiments shows that the storage modulus of rEX1-HSL is not frequency dependent in the range of 0.1-

100 rad/sec and strain sweep experiments show that a 1 wt% rEX1-HSL gel yields around 25% strain which is higher than that of EX1 (6% strain).

Although, the tandem repeat peptides proved interesting, their bacterial production proved inefficient. I next explored using smaller fusion partners as a viable means to increase the yield of bacterially produced peptide. A smaller fusion partner protein was designed and the new fusion 6His-BAD-EX1 was used for expression in modified auto-induction media. Together with the optimization in downstream processing, the overall yield of rEX1 was increased more than ten times from 4.6 mg to 50 mg pure peptide from a liter of culture. The ratio of the EX1 peptide in the overall fusion was increased by using a smaller size fusion partner, and the cell density at the end of culturing in modified auto-induction media was increased to around 7 times. Optimized downstream processing eliminated the HPLC purification step for the fusion proteins and facilitated the easier solubilization of the fusion protein by sonication at low pH. This step increased the recovery of the rEX1 peptide after CNBr cleavage and lyophilization. The overall method decreased the cost of producing EX1 in a bacterial system and is significantly competitive with the cost of chemical synthesis.

The same optimized culturing and isolation methods were applied for bacterial production of the EX2 and EX3 peptides. These two peptides have similar sequences as EX1 except that their turn sequences vary. The bacterial production of these two

peptides was investigated to determine if the optimized production methodology could be extended to other sequences. The yields of rEX2 and rEX3 from 1 liter of culture were 31 mg and 15 mg, which are a little lower, but still very good for the bacterial-based production of the peptides.

Lastly, the folding, self-assembly and material forming properties of rEX1, rEX2 and rEX3 were investigated and found to be similar to those of the corresponding peptides (and their gels) prepared by chemical synthesis.

Altogether, the data show that β-hairpin self-assembling peptides can be produced by fermentation in good yield at significantly reduced cost when compared to chemical synthesis.

4.9 Experimental Section

4.9.1 Oligonucleotide sequences designed for cloning

Inserts were designed accordingly to be compatible with the Gateway cloning system by flanking the attB sites at either ends. Oligonucleotides were ordered from Genscript and were provided cloned in a pUC57 vector.

DNA sequence of $(EX1)_2$ construct is: 5'-

GGGGACAACTTTGTACAAAAAAGTTGGCATGGCACCAAAGTTGAAGTGAA

AGTTAAAGTGATGGTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCA

CCAAAGTTGAAGTGAAAGTTAAAGTGATGGTGAAAGTTAAAGTGAAAGTT

AAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAGTTAAAGTGATGGGCTA

ATAGCCAACTTTCTTGTACAAAGTTGTCCCC -3'

DNA sequence of $(EX1)_3$ construct is: 5'-

GGGGACAACTTTGTACAAAAAAGTTGGCATGGTGAAAGTTAAAGTGAAAG

TTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAGTTAAAGTGATGGTG

AAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAA

AGTTAAAGTGATGGTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCA

CCAAAGTTGAAGTGAAAGTTAAAGTGATGGGCTAATAGCCAACTTTCTTGT

ACAAAGTTGTCCCC -3'

DNA sequence of $(EX1)_3$-6His gene is: 5'-

GGGGACAACTTTGTACAAAAAAGTTGGCTTAAGAAGGAGATATACATATG

GTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACCAAAGTTGAAGT

GAAAGTTAAAGTGATGGTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGATG

GCACCAAAGTTGAAGTGAAAGTTAAAGTGATGGTGAAAGTTAAAGTGAAA

GTTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAGTTAAAGTGATGGG

CCACCATCACCATCACCATTAATAGCCAACTTTCTTGTACAAAGTTGTCCC
C -3’

DNA sequence of $(EX1)_6$ construct is: 5’-
GGGGACAACTTTGTACAAAAAAGTTGGCATGGTGAAAGTTAAAGTGAAAG
TTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAGTTAAAGTGATGGTG
AAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAA
AGTTAAAGTGATGGTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCA
CCAAAGTTGAAGTGAAAGTTAAAGTGATGGTGAAAGTTAAAGTGAAAGTT
AAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAGTTAAAGTGATGGTGAA
AGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACCAAAGTTGAAGTGAAAG
TTAAAGTGATGGTGAAAGTTAAAGTGAAAGTTAAAGTGCCGGATGGCACC
AAAGTTGAAGTGAAAGTTAAAGTGATGGGCTAATAGCCAACTTTCTTGTA
CAAAGTTGTCCCC -3’

DNA sequence of BAD-EX1 construct is: 5’-
GGGGACAACTTTGTACAAAAAAGTTGGCGCTCCGCCGAACCTGTGGGCTG
CTCAACGCTATGGCCGCGAACTGCGCCGTATGTCAGATGAATTTGAAGGCT
CATTTAAACCGGGCGATGTGGAAAACCTGGGCGGCATGGTTAAAGTCAAA
GTCAAAGTCAAAGTCCCGGATGGCACCAAAGTGGAAGTCAAAGTCAAAGT
CTGACCAACTTTCTTGTACAAAGTTGTCCCC -3’

DNA sequence of BAD-EX2 construct is: 5'-

GGGGACAACTTTGTACAAAAAAGTTGGCGCTCCGCCGAACCTGTGGGCTG

CTCAACGCTATGGCCGCGAACTGCGCCGTCTGTCAGATGAATTTGAAGGCT

CATTTAAACCGGGCGATGTGGAAAACCTGGGCGGCATGGTTAAAGTCAAA

GTCAAAGTCAAAGTCCCGATTGGCACCAAAGTGGAAGTCAAAGTCAAAGT

CTGACCAACTTTCTTGTACAAAGTTGTCCCC -3'

DNA sequence of BAD-EX2 construct is: 5'-

GGGGACAACTTTGTACAAAAAAGTTGGCGCTCCGCCGAACCTGTGGGCTG

CTCAACGCTATGGCCGCGAACTGCGCCGTCTGTCAGATGAATTTGAAGGCT

CATTTAAACCGGGCGATGTGGAAAACCTGGGCGGCATGGTTAAAGTCAAA

GTCAAAGTCAAATATAACGGCACCAAAGTGGAAGTCAAAGTCAAAGTCTG

ACCAACTTTCTTGTACAAAGTTGTCCCC -3'

4.9.2 DNA cloning and transformation

Inserts supplied in 4 μg of lyophilized pUC57 plasmid DNA, were dissolved in 20 μL of water. For BP and LR reactions, the manufacturer's protocols were followed (Invitrogen). The following vectors were kindly provided by Dominic Esposito: donor vector pDONR253, destination vector pDEST521 for inserts with RBS (for $(EX1)_3$-His insert), destination vector pDEST527 with RBS and 6Histag (for $(EX1)_2$, $(EX1)_3$ and $(EX1)_6$ inserts as well as BAD-EX1, BAD-EX2, BAD-EX3 inserts). Briefly, 0.5

μL of 200 ng/μL pUC57 containing the corresponding inserts, 1 μL of 120 ng/μL of pDONR253, 4.5 μL of TE buffer and 2 μL of BP clonase were mixed and left at room temperature for at least 1 h to generate the entry clones. *Escherichia coli* DH5α strain was used for cloning, and 1 μL of BP clonase reaction was transferred into 20 μL of pre-aliquoted subcloning efficiency DH5α cells thawed on ice. After 30 min incubation on ice, 30 sec at 42^{o}C in a water bath and 2 min incubation on ice, 60 μL of SOC medium was added and cells were cultured at 37^{o}C (250 rpm) for 1 h. From this culture, 50 μL was spread onto a LB agar plate with Spectinomycin (50 μg/mL) and incubated 12-16 h (overnight) at 37^{o}C. Single colonies were picked and inoculated into 10 mL of LB media supplemented with Spectinomycin (50 μg/mL) and incubated overnight at 37^{o}C (250 rpm), followed by plasmid DNA isolation using the Qiagen spin miniprep kit according to the manual. Plasmid DNA samples were submitted for sequencing and upon verification of the expected sequences, they were used for the LR cloning reaction.

For the 10 μL LR reaction, 50 ng of entry clone, 150 ng of destination vector (with the fusion partner of interest), 2 μL LR Clonase II enzyme mix and TE buffer pH 8 were mixed and left at room temperature for at least 1 h to generate the expression clones. From this LR clonase reaction, 1 μL was transferred into 20 μL of pre-aliquoted subcloning efficiency DH5α cells thawed on ice. For the transformation of $(EX1)_6$ containing expression clone, chemically competent *E. coli* Stb3 cells (kindly

provided by Dominic Esposito) were used because this strain has a recombination defective phenotype and is less prone to recombination due to its *rec*A13 mutant genotype. RecA13 plays a role in conjugal recombination and post-replication recombinational repair [47], and this allowed the successful propagation of the expression clone since the bacteria is able to divide without any recombinational change in the number of repeats of $(EX1)_6$ construct. Transformation, cell spreading on LB agar plates with ampicillin (100 μg/mL), single colony inoculation into 10 mL LB media with ampicillin (100 μg/mL), miniprep plasmid DNA isolation and sequence verification were performed as indicated in the previous paragraph. This was followed by transformation into the expression strain Rossetta 2 cells, which were spread onto LB agar plates with ampicillin (100 μg/mL) and chloramphenicol (34 μg/mL). Glycerol stocks of Rossetta 2 cells with expression clones were prepared with 8% glycerol from overnight LB cultures inoculated with single colonies selected from LB agar plates.

4.9.3 Bacterial culturing

For Rossetta 2 cell cultures, 100 μg/mL carbenicillin and 34 μg/mL chloramphenicol were added to the media. For pilot expression experiments, 25 mL of media was used in a 250 mL baffled flask. For larger scale experiments, 800 mL of LB media was used in 2.8 L baffled flasks, and 400 mL of AI or mAI media was used in 2.8 L baffled flasks.

LB media was prepared by dissolving 20 g/L LB broth media (Sigma) in distilled water followed by autoclaving. For starter cultures, 10 mL LB media with antibiotics was inoculated with the glycerol stock of Rossetta 2 cell (with the plasmid of interest) and incubated at 37^{o}C (250 rpm) overnight. Fresh 25 mL LB media with antibiotics was then inoculated with 0.25 mL overnight culture and incubated at 37^{o}C (250 rpm) until the OD_{600} reached 0.7-0.8. For large scale experiments, 30 mL starter culture was prepared with the glycerol stock of Rossetta 2 cell (with the plasmid of interest) as outlined above. To 800 mL LB media with antibiotics, 8 mL of overnight culture was added and incubated at 37^{o}C (250 rpm) until the OD_{600} reached 0.7-0.8. A final concentration of 1 mM IPTG was used to induce the expression for 4 h at 37^{o}C (250 rpm). After culturing, wet cell pellets obtained by centrifugation at 12000 rpm for 10 min were stored at -80^{o}C overnight prior to processing.

Auto-induction (AI) media was prepared according to Studier's ZY-5052 media [43]: 5 g/L yeast extract and 10 g/L tryptone (ZY); 0.5% glycerol, 0.05% glucose and 0.2% α-D-Lactose (5052); 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl and 5 mM Na_2SO4 (M), 2 mM $MgSO_4$ and 0.2x trace elements. Modified AI (mAI) media was prepared similar to ZY-5052 media with the exception that it had 8 times of ZY and 4 times of 5052 in addition to same amounts of M, $MgSO_4$ and trace elements. For culturing in AI or mAI media, first 1 ml of LB media was inoculated with the glycerol stock and vortexed. Then 5 µl of the diluted glycerol stock was transferred and spread onto a LB agar plate with antibiotics for overnight incubation at

37°C. Three different colonies from the same plate were then individually inoculated into tubes with 2 mL LB media supplemented with antibiotics, M, 1% glucose, 2 mM MgSO4 and 0.2x trace elements, and incubated at 37°C (300 rpm) for 6 h. The most turbid culture was chosen and 200 µl was inoculated into 400 mL of AI or mAI media in 2.8 L baffled flasks, and incubated at37°C (300 rpm) for 64 h. After culturing, wet cell pellets obtained by centrifugation at 12000 rpm for 10 min, were stored at -80°C overnight prior to processing.

4.9.4 Cell Lysis after pilot tandem repeat expression and SDS-PAGE analysis

For the SDS-PAGE analysis of tandem repeat expression of the EX1 peptide, an amount of culture normalized at an OD_{600} of 3 was centrifuged in a 1.5 mL eppendorf at 13000 rpm for 5 min, and the supernatant was discarded. Cell pellets were resuspended in 140 µL of lysis buffer A (50mM Tris-HCl pH 7.5, 50 mM dextrose, 1 mM EDTA, 4 mM $MgCl_2$, 30U/µL Ready-Lyse (Epicentre) lysozyme solution) with 8U Benzonase nuclease (Novagen) and incubated for 10 min at room temperature. To this, 140 µL of lysis buffer B (10 mM Tris-HCl pH 7.5, 50 mM KCl, 1 mM EDTA, 0.1% deoxycholate) was then added and incubated for 10 min at room temperature. Following centrifugation, soluble fractions were transferred to new tubes and insoluble pellets were solubilized in 300 µL of 8 M urea. For pre-induction samples, 1 mL pre-induction cultures were centrifuged and cell pellets were

resuspended in 40 μL Lysis buffer A with 8 U Benzonase nuclease for 5 min, and 40 μL Lysis buffer B was then added and vortexed.

For SDS-PAGE analysis of pilot tandem repeat expressions, 12 μL of soluble or insoluble samples were mixed with 6 μL of 3x NuPAGE LDS sample buffer with 0.35 M DTT and heated at 95^{o}C for 5 min. The samples as well as 10 μL of SeeBlue Plus2 pre-stained standard were loaded onto a NuPAGE 12% Bis-Tris gel and electrophoresed (200 V for 35 min). Gels were stained with a solution of 50 mg/L Coomassie Brilliant Blue G-250 and 3 mL/L glacial acetic acid in water.

4.9.5 Purification of EX1-HSL from 6His-(EX1)$_2$ Fusion

Cell pellets (~3 g) stored at -80^{o}C were thawed and 30 mL lysis buffer (0.5 M NaCl, 50 mM Tris-HCl pH 7.5, 100 mM $MgSO_4$, 1 mM PMSF) was added. Cells were resuspended by sonication (QSonica XL-2000, 1/8" microprobe, 20 W, 4 min) and kept on ice to prevent overheating. The sample was centrifuged at 15000 rpm for 12 min and the soluble fraction was removed. The insoluble fraction was treated twice with 15 mL wash buffer (50 mM Tris-HCl pH 7.5) and 2% Triton-X-100, and then twice with 17 mL wash buffer only, with sonication and centrifugation steps in between each step. The washed cell pellet (~1.5 g) was dissolved in 12 mL of 8 M urea and 0.5% TFA, sonicated and centrifuged as before to separate the cell debris. The soluble fraction (in 8 M urea, 0.5% TFA) was filtered (0.2 μm nylon syringe

filter) and purified by RP-HPLC (semi-preparative Vydac protein/peptide C18 column) at 40^{o}C employing an isocratic gradient for 1 min with 100% solvent A, then for 14 min with 15% solvent B and finally for 20 min with 50% solvent B where solvent A is 0.1% TFA in water and solvent B is 0.1% TFA in 9:1 acetonitrile:water. The corresponding peak for 6His-$(EX1)_2$ fusion was collected and lyophilized.

Cleavage of 6His-$(EX1)_2$ fusion by CNBr was performed by adding 0.3 g/mL CNBr in 70% formic acid to the sample to a final concentration of 30 mg/mL. The flask was purged with Argon and kept in the dark for 6 h before being diluted 7x in 0.1% TFA and lyophilized.

The lyophilized sample was dissolved in 0.1% TFA, centrifuged and purified by RP-HPLC (semi-preparative Vydac protein/peptide C18 column) at 40^{o}C employing an isocratic gradient for 1 min with 100% solvent A, then linear solvent B gradients from 0 - 14% B over 14 min, 14 - 27% over 32 min and finally from 27-37% over 40 min. The corresponding peak for EX1-HSL was collected and lyophilized. Analytical RP-HPLC and ESI-MS were performed to analyze the purity and the molecular weight of rEX1 peptide.

4.9.6 Optimized purification of rEX peptides

Cell pellets (~1 g) stored at -80^{o}C were thawed and 40 mL wash buffer (0.1 M NaCl, 50 mM Tris-HCl pH 7.5, 1 mM EDTA) and 0.5% Triton-X100 was added.

Cells were resuspended by sonication (QSonica Q700, 1/2" tip, 60% Amp, 90 sec) and kept on ice to prevent overheating. The sample was centrifuged at 12000 rpm for 10 min and the soluble fraction was removed. The insoluble fraction was treated with 40 mL wash buffer and 0.5% TritonX-100 once more and then twice with 40 mL wash buffer only, with sonication and centrifugation steps in between each step. The washed cell pellet was solubilized by sonication at low pH. For the washed cell pellets of the BAD-EX1 fusion, 100 mL of 0.1% TFA was used and for that of BAD-EX2 and Bad-EX3 fusions, 100 mL of 1% TFA was used; cell pellets were sonicated (1/2" tip, 70% Amp, 3 min) and centrifuged at 15000 rpm for 15 min. Soluble fractions were lyophilized.

Cleavage of BAD-EX fusions by CNBr was performed by adding 0.1 g/mL CNBr in 70% formic acid to the sample to a final concentration of 60 mg/mL. The flask was purged with Argon and kept in the dark for 3 h before being diluted 10x in 0.1% TFA in 1:3 acetonitrile:water and lyophilized.

The lyophilized samples were dissolved in 0.1% TFA (for rEX1) or 1% TFA (for rEX2 and rEX3), sonicated, centrifuged and purified by RP-HPLC (semi-preparative Vydac protein/peptide C18 column) at 40^{o}C. For rEX1 purification, linear gradients were employed: 0% solvent B over 1 min, 0-12% solvent B over 12 min, 12-26% solvent B over 28 min and then 26-36% solvent B over 40 min. For rEX2 purification, linear gradients were employed: 0% solvent B over 1 min, 0-14% solvent

B over 10 min, 14-29% solvent B over 30 min and then 29-39% solvent B over 40 min. For rEX3 purification, linear gradients employed were: 0% solvent B over 1 min, 0-11% solvent B over 7 min, 11-26% solvent B over 30 min and then 26-36% solvent B over 40 min. Analytical HPLC and electrospray ionization (ESI-positive mode) were performed to verify the purity of the lyophilized peptides.

4.9.7 Characterization of rEX peptides

Secondary structures, mechanical properties and local fibril morphologies of rEX peptides were characterized by CD spectroscopy, oscillatory rheology and TEM, respectively as outlined in Chapter 2.

REFERENCES

1. Parachin, N.S., et al., *Expression systems for heterologous production of antimicrobial peptides.* Peptides, 2012. **38**(2): p. 446-456.

2. Jonasson, P., et al., *Genetic design for facilitated production and recovery of recombinant proteins in Escherichia coli.* Biotechnology and Applied Biochemistry, 2002. **35**: p. 91-105.

3. Makrides, S.C., *Strategies for achieving high-level expression of genes in Escherichia coli.* Microbiological Reviews, 1996. **60**(3): p. 512-&.

4. Gustafsson, C., S. Govindarajan, and J. Minshull, *Codon bias and heterologous protein expression.* Trends in Biotechnology, 2004. **22**(7): p. 346-353.

5. Varenne, S., et al., *Translation Is a Non-Uniform Process - Effect of Transfer-Rna Availability on the Rate of Elongation of Nascent Polypeptide-Chains.* Journal of Molecular Biology, 1984. **180**(3): p. 549-576.

6. Zhong, Z., et al., *Tandem repeat mhBD2 gene enhance the soluble fusion expression of hBD2 in Escherichia coli.* Appl Microbiol Biotechnol, 2006. **71**(5): p. 661-7.

7. Gutman, G.A. and G.W. Hatfield, *Nonrandom Utilization of Codon Pairs in Escherichia-Coli.* Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(10): p. 3699-3703.

8. Carlini, D.B., Y. Chen, and W. Stephan, *The relationship between third-codon position nucleotide content, codon bias, mRNA secondary structure and gene expression in the drosophilid alcohol dehydrogenase genes Adh and Adhr.* Genetics, 2001. **159**(2): p. 623-633.

9. Ivanovski, G., F. Gubensek, and J. Pungercar, *MRNA secondary structure can greatly affect production of recombinant phospholipase A(2) toxins in bacteria.* Toxicon, 2002. **40**(5): p. 543-549.

10. Boycheva, S., G. Chkodrov, and I. Ivanov, *Codon pairs in the genome of Escherichia coli.* Bioinformatics, 2003. **19**(8): p. 987-998.

11. Gursky, Y.G. and R.S. Beabealashvilli, *The Increase in Gene-Expression Induced by Introduction of Rare Codons into the C-Terminus of the Template.* Gene, 1994. **148**(1): p. 15-21.

12. Kane, J.F., *Effects of Rare Codon Clusters on High-Level Expression of Heterologous Proteins in Escherichia-Coli.* Current Opinion in Biotechnology, 1995. **6**(5): p. 494-500.

13. Ikemura, T., *Relationship between the Frequency of Codon Usage in E-Coli Genes and the Abundance of Cognate Transfer-Rna.* Japanese Journal of Genetics, 1979. **54**(6): p. 436-436.

14. Kurland, C.G., *Codon Bias and Gene-Expression.* Febs Letters, 1991. **285**(2): p. 165-169.

15. Menninger, J.R., *"Big Loser" codon pairs in E-coli have normal usage in other genomes.* Faseb Journal, 2006. **20**(4): p. A502-A502.

16. Bulmer, M., *Coevolution of Codon Usage and Transfer-Rna Abundance.* Nature, 1987. **325**(6106): p. 728-730.

17. Nishikubo, T., et al., *Improved heterologous gene expression in Escherichia coli by optimization of the AT-content of codons immediately downstream of the initiation codon.* Journal of Biotechnology, 2005. **120**(4): p. 341-346.

18. Baranov, P.V., et al., *Transcriptional slippage in bacteria: distribution in sequenced genomes and utilization in IS element gene expression.* Genome Biology, 2005. **6**(3).

19. Pandey, J.P., et al., *mRNA secondary structure modulates the translation of organophosphate hydrolase (OPH) in E-coli.* Molecular Biology Reports, 2009. **36**(3): p. 449-454.

20. Studer, S.M. and S. Joseph, *Unfolding of mRNA secondary structure by the bacterial translation initiation complex.* Molecular Cell, 2006. **22**(1): p. 105-115.

21. Zhang, W.C., et al., *mRNA secondary structure at start AUG codon is a key limiting factor for human protein expression in Escherichia coli.* Biochemical and Biophysical Research Communications, 2006. **349**(1): p. 69-78.

22. Kudla, G., et al., *Coding-Sequence Determinants of Gene Expression in Escherichia coli.* Science, 2009. **324**(5924): p. 255-258.

23. Wood, C.R., et al., *The Influence of Messenger-Rna Secondary Structure on Expression of an Immunoglobulin Heavy-Chain in Escherichia-Coli.* Nucleic Acids Research, 1984. **12**(9): p. 3937-3950.

24. Kimura, S. and T. Iyanagi, *High-level expression of porcine liver cytochrome P-450 reductase catalytic domain in Escherichia coli by modulating the predicted local secondary structure of mRNA.* Journal of Biochemistry, 2003. **134**(3): p. 403-413.

25. Puri, N., et al., *Effect of the codon following the ATG start site on the expression of ovine growth hormone in Escherichia coli.* Protein Expr Purif, 1999. **17**(2): p. 215-23.

26. Studier, F.W. and B.A. Moffatt, *Use of Bacteriophage-T7 Rna-Polymerase to Direct Selective High-Level Expression of Cloned Genes.* Journal of Molecular Biology, 1986. **189**(1): p. 113-130.

27. Takahashi, T., K.W. Ryan, and A. Portner, *A plasmid that improves the efficiency of foreign gene expression by intracellular T7 RNA polymerase.* Genet Anal Tech Appl, 1992. **9**(3): p. 91-5.

28. Minkley, E.G., Jr., *Transcription of the early region of bacteriophage T7: specificity and selectivity in the in vitro initiation of RNA synthesis.* J Mol Biol, 1974. **83**(3): p. 305-31.

29. Rong, M., et al., *Promoter specificity determinants of T7 RNA polymerase.* Proc Natl Acad Sci U S A, 1998. **95**(2): p. 515-9.

30. Iost, I., J. Guillerez, and M. Dreyfus, *Bacteriophage-T7 Rna-Polymerase Travels Far Ahead of Ribosomes Invivo.* Journal of Bacteriology, 1992. **174**(2): p. 619-622.

31. Tabor, S., *Expression using the T7 RNA polymerase/promoter system.* Curr Protoc Mol Biol, 2001. **Chapter 16**: p. Unit16 2.

32. Kumar, A. and S.S. Patel, *Inhibition of T7 RNA polymerase: Transcription initiation and transition from initiation to elongation are inhibited by T7 lysozyme via a ternary complex with RNA polymerase and promoter DNA.* Biochemistry, 1997. **36**(45): p. 13954-13962.

33. Studier, F.W., *Use of Bacteriophage-T7 Lysozyme to Improve an Inducible T7 Expression System.* Journal of Molecular Biology, 1991. **219**(1): p. 37-44.

34. Baca, A.M. and W.G.J. Hol, *Overcoming codon bias: A general method to overexpress Plasmodium genes in E-coli.* Faseb Journal, 1999. **13**(7): p. A1551-A1551.

35. Tourle, S., et al., *Increased levels of recombinant human proteins in E-Coli Rosetta that compensates for mammalian codon usage.* Molecular & Cellular Proteomics, 2006. **5**(10): p. S223-S223.

36. Kempe, T., et al., *Multiple-Copy Genes - Production and Modification of Monomeric Peptides from Large Multimeric Fusion Proteins.* Gene, 1985. **39**(2-3): p. 239-245.

37. Riley, J.M., et al., *Bioproduction and Characterization of a pH Responsive Self-Assembling Peptide.* Biotechnology and Bioengineering, 2009. **103**(2): p. 241-251.

38. Kuliopulos, A. and C.T. Walsh, *Production, Purification, and Cleavage of Tandem Repeats of Recombinant Peptides.* Journal of the American Chemical Society, 1994. **116**(11): p. 4599-4607.

39. Reed, D.C., et al., *Production and purification of self-assembling peptides in Ralstonia eutropha.* Protein Expr Purif, 2006. **46**(2): p. 179-88.

40. Sharpe, S., W.M. Yau, and R. Tycko, *Expression and purification of a recombinant peptide from the Alzheimer's beta-amyloid protein for solid-state NMR.* Protein Expr Purif, 2005. **42**(1): p. 200-10.

41. Lee, J.H., et al., *Acidic peptide-mediated expression of the antimicrobial peptide buforin II as tandem repeats in Escherichia coli.* Protein Expr Purif, 1998. **12**(1): p. 53-60.

42. Sivashanmugam, A., et al., *Practical protocols for production of very high yields of recombinant proteins using Escherichia coli.* Protein Sci, 2009. **18**(5): p. 936-48.

43. Studier, F.W., *Protein production by auto-induction in high-density shaking cultures.* Protein Expression and Purification, 2005. **41**(1): p. 207-234.

44. Studier, F.W., *Stable expression clones and auto-induction for protein production in E. coli.* Methods Mol Biol, 2014. **1091**: p. 17-32.

45. Shimizu, K., *Metabolic Regulation of a Bacterial Cell System with Emphasis on Escherichia coli Metabolism.* ISRN Biochemistry, 2013. **2013**: p. 47.

46. Chen, L., et al., *Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function.* Mol Cell, 2005. **17**(3): p. 393-403.

47. Lauder, S.D. and S.C. Kowalczykowski, *Negative Codominant Inhibition of Reca Protein Function - Biochemical-Properties of the Reca1, Reca13 and Reca56 Proteins and the Effect of Reca56 Protein on the Activities of the Wild-Type Reca Protein Function in-Vitro.* Journal of Molecular Biology, 1993. **234**(1): p. 72-86.

Chapter 5

PROBLEMS DURING PEPTIDE SYNTHESIS AND POTENTIAL OUTCOMES

5.1 Introduction

During solid phase peptide synthesis (SPPS), peptide-resin cleavage and side-chain deprotection, various side reactions and incomplete reactions may occur. The unintended variations on the target peptide sequence and structure due to such problems may result in production of an undesired peptide with significantly different characteristics [1-5]. In this chapter two different issues will be discussed that influenced the chemical production of peptides produced in this thesis. The first is succinimide (aspartimide) formation in the aspartate-containing EX1 peptide (Chapter 2) during synthesis and its effect on the peptide`s material forming characteristics. The second issue is the generation of a lysine deletion sequence during the synthesis of another peptide Ac-EX3, which was due to a malfunction of the automated peptide synthesizer. This unintentionally synthesized peptide was further investigated for its material forming properties that led to remarkable results. The sequences of peptides in question of concern are listed in Table 5.1.

Table 5.1 Intended and unintended peptides that have been investigated. Succinimide (suc) and isoaspartate (isoD) residues on EX1 sequence and K19 residue on Ac-EX3 sequence are shown in red and are underlined.

	Peptide	Sequence
Intended peptide 1	EX1	H_3N^+ - VKVKVKVKVPDGTVKVEVKVK - COO^-
Suc containing EX1	EX1(Suc)	H_3N^+ - VKVKVKVKVP*Suc*GTVKVEVKVK - COO^-
isoD containing EX1	EX1(isoD)	H_3N^+ - VKVKVKVKVP*isoD*GTVKVEVKVK - COO^-
Intended peptide 2	Ac-EX3	Ac - VKVKVKVKYNGTKVEVKVKV - COO^-
Lys19 deletion on Ac-EX3	Ac-EX3-K19del	Ac - VKVKVKVKYNGTKVEVKVV - COO^-

5.2 Succinimide and Isoaspartate Formation in EX1 Synthesis

In the case of aspartate-containing sequences, spontaneous succinimide formation can lead to aspartate isomerization at neutral and basic pH (Figure 5.1) [6]. An aspartic acid residue can form a succinimide by an intramolecular reaction between the protonated carboxyl group of its side chain and the amide nitrogen. Following this, the succinimide ring can be opened by water at either of its carbonyl carbons resulting in either the regeneration of the original residue or a newly formed β-amino acid, isoaspartate [7-9]. Aspartate isomerization is a covalent modification that is also observed ubiquitously in proteins and monoclonal antibodies [10-16]. Glycine and serine residues located C-terminal to aspartic acid facilitate isoaspartate formation in many model peptides and proteins [17-21]. In the case of isoaspartate formation, the backbone chain is extended by one methylene unit and neighboring

carbonyl carbon and α-carbons are separated by an additional ~1-1.5 Å. This structural change can affect the higher order structure of assemblies if the peptide is prone to self-assembly.

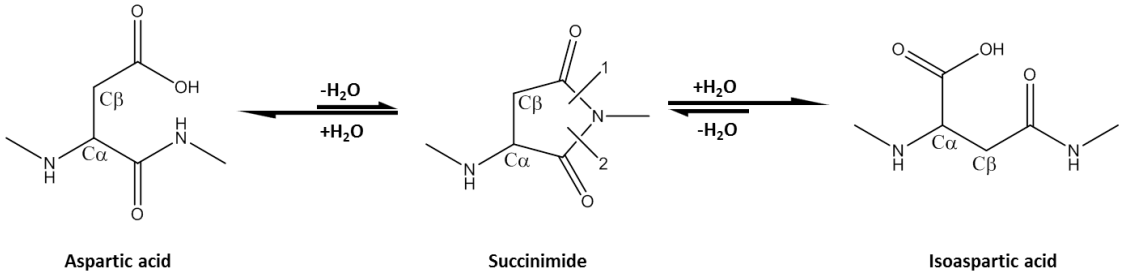


Figure 5.1 Isomerization of aspartate via succinimide formation. Succinimide hydrolysis yields (i) aspartate and (ii) isoaspartate.

One of the hydrogel forming β-hairpin peptides amenable for bacterial production, namely EX1 (H_3N^+-VKVKVKVK-VPDGT-KVEVKVKV-COO^-) has an aspartate residue in its turn region -VPDGT-. In the course of my studies, this peptide was produced by both chemical synthesis and by fermentation. Peptide produced by these two methods showed significant differences in their folding and assembling behavior. As discussed in this chapter, the chemically synthesized and recombinant peptides were subjected to further analysis to assess if isoaspartate formation was at play.

First, the analysis methods for the detection of succinimide and isoaspartate residues and their results will be discussed and then, the differences in the

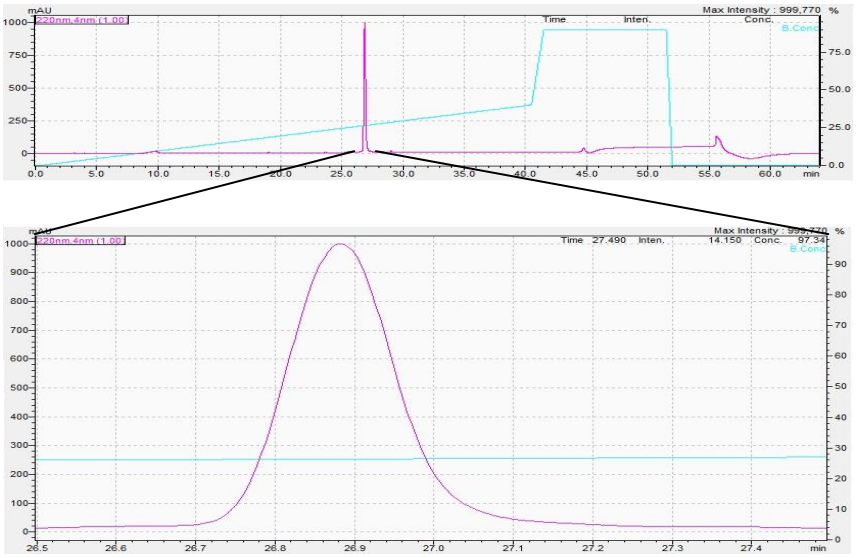
characteristics due to the succinimide and isoaspartate formation will be presented. Finally, a strategy to prevent succinimide formation during the chemical synthesis will be explained.

5.2.1 Detection of succinimide and isoaspartate in EX1 peptides

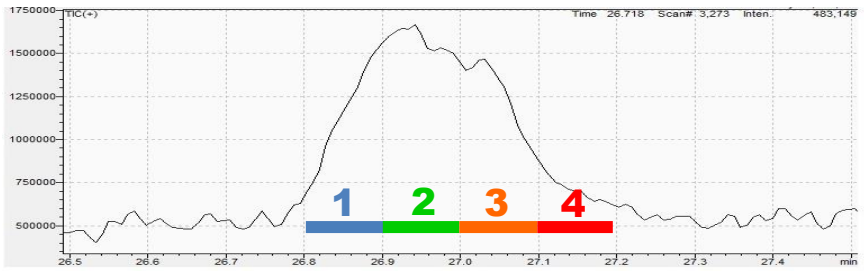
The purity of recombinant EX1 (rEX1) was analyzed by liquid chromatography-mass spectrometry (LC-MS) to detect any succinimide containing peptide. Succinimide formation results in an 18 Da decrease in the molecular weight due to the loss of water. The reverse phase HPLC (RP-HPLC) chromatogram at 220 nm in Figure 5.2A shows that the rEX1 elutes around 26.9 min and affords a symmetrical peak. Figure 5.2B shows the total ion chromatogram (TIC) in the time window that corresponds to rEX1 elution. The TIC time period between 26.8-27.2 min was divided into four 0.1 min segments, and each period of 0.1 min during the generation of ions was analyzed individually. In Figure 5.2C, it can be seen that the calculated and observed *m/z* correspond well. All the ions generated during electrospray ionization mass spectrometry (ESI-MS) corresponded to the expected molecular weight of the EX1 peptide for each 0.1 min time period. Only, in time segment 4 (Figure 5.2C), a very small amount of succinimide is observed at m/z=573, m/z=764 and m/z=1146 for z=4, z=3 and z=2, respectively. This suggests that the EX1 peptide produced recombinantly, in large part, does not undergo succinimide formation. This however, does not exclude the possibility that the peptide could have

undergone some succinimide formation, which was subsequently hydrolyzed to form the corresponding isoaspartate, which has the same mass as the aspartate-containing peptide.

(A)


mAU
220nm,4nm (1.00)
Max Intensity : 999,770
Time
Inten.
Conc.
B.Conc
Time 27.490 Inten. 14.150 Conc. 97.34
min

(B)


TIC(+)
Time 26.718 Scan# 3,273 Inten. 483,149
1
2
3
4
min

(C)

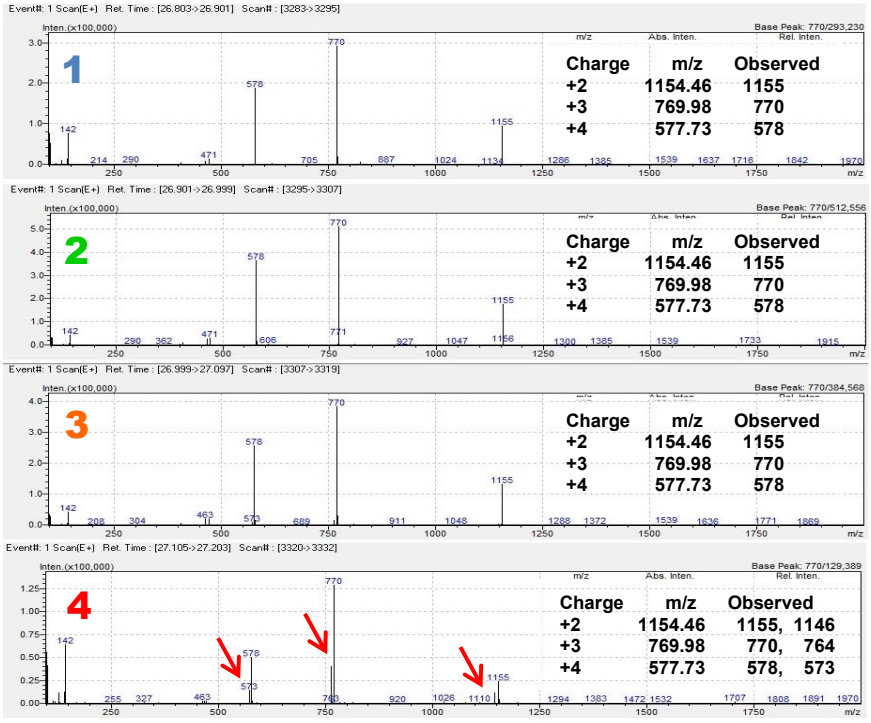
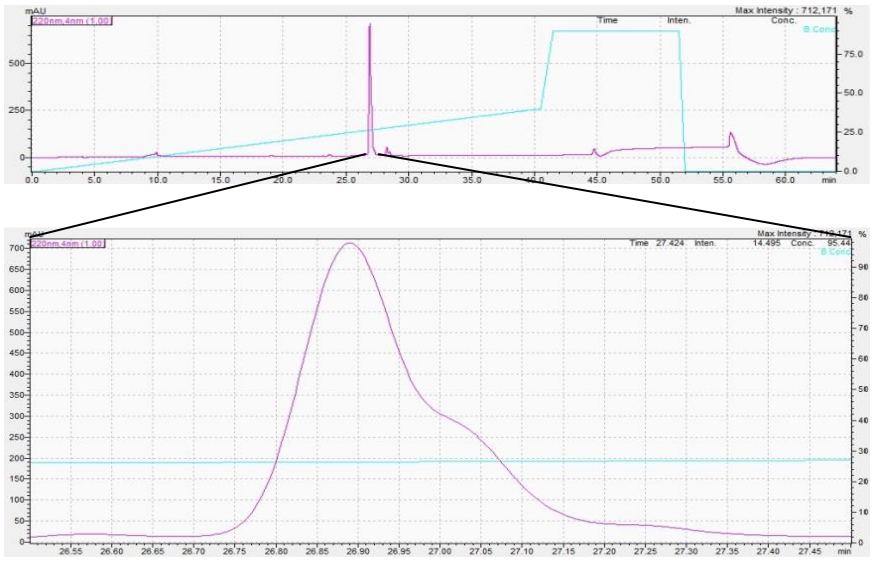

Event#: 1 Scan(E+) Ret. Time : [26.803->26.901] Scan# : [3283->3295]
Base Peak: 770/293,230
1
Charge m/z Observed
+2 1154.46 1155
+3 769.98 770
+4 577.73 578
Event#: 1 Scan(E+) Ret. Time : [26.901->26.999] Scan# : [3295->3307]
Base Peak: 770/512,556
2
Charge m/z Observed
+2 1154.46 1155
+3 769.98 770
+4 577.73 578
Event#: 1 Scan(E+) Ret. Time : [26.999->27.097] Scan# : [3307->3319]
Base Peak: 770/384,568
3
Charge m/z Observed
+2 1154.46 1155
+3 769.98 770
+4 577.73 578
Event#: 1 Scan(E+) Ret. Time : [27.105->27.203] Scan# : [3320->3332]
Base Peak: 770/129,389
4
Charge m/z Observed
+2 1154.46 1155, 1146
+3 769.98 770, 764
+4 577.73 578, 573

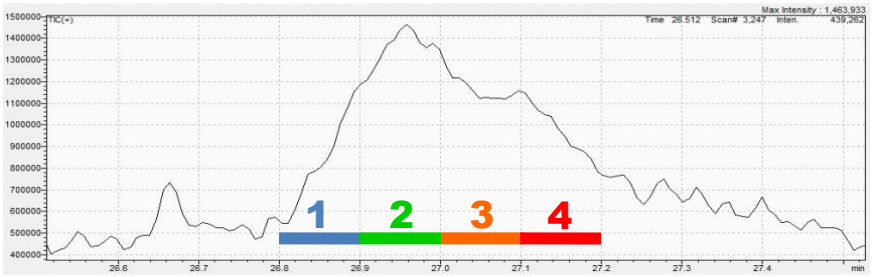
Figure 5.2 LC-MS analysis of recombinant EX1. (A) RP-HPLC chromatogram of recombinant EX1. Absorbance at 220 nm was monitored as a function of time. Gradient of 0-40% solvent B in 40 min at 40^{o}C is indicated by the blue trace. Solvent A: 0.1 % TFA in water, Solvent B: 90% acetonitrile, 9.9% water, 0.1 % TFA. (B) Corresponding TIC (m/z 150-2000) for the rEX1 peak aligned for the 0.1 min lag in detection. (C) Averaged ESI-MS of rEX1 for the indicated 0.1 min segment. The insert in each time range shows a comparison of calculated and observed ion masses. Masses indicative of succinimide containing product are indicated by red arrows.

Next, the same analysis was performed for the chemically synthesized EX1 peptide and the results are shown in Figure 5.3. The Analytical HPLC chromatogram in Figure 5.3A shows the peak corresponding to the EX1 peptide eluting around 26.9 min. When the peak is enlarged, a non-symmetrical peak with a shoulder around 27.05 min is observed. Figure 5.3B shows the total ion chromatogram around 27 min. The time range between 26.8-27.2 min was analyzed at every 0.1 min (region 1-4) for total ions that might have corresponded to a different peptide mass (Figure 5.3C). In region 1 and 2, there is a good correlation between the calculated and observed ions that correspond to the molecular weight of the EX1 peptide (2306.91 Da) whereas in region 3, there are additional ions that correspond to a MW that is 18 Da less than that of the EX1 peptide. In region 4, only ions corresponding to the MW of 18 Da less were observed. The 18 Da mass difference is due to the loss of water when the aspartate residue forms succinimide. Therefore, the chemically synthesized EX1 peptide contains mixture of aspartate and succinimide-containing peptides. Attempts to isolate each peptide were made by using different gradients on preparative HPLC but the two peptides did not resolve and always eluted together.

(A)



(B)


1
2
3
4

(C)


1
Charge m/z Observed
+2 1154.46 1155
+3 769.98 770
+4 577.73 578
2
Charge m/z Observed
+2 1154.46 1155
+3 769.98 770
+4 577.73 578
3
Charge m/z Observed
+2 1154.46 1155, 1146
+3 769.98 770, 764
+4 577.73 578, 573
4
Charge m/z Observed
+2 1154.46 1146
+3 769.98 764
+4 577.73 573

Figure 5.3 LC-MS analysis of chemically synthesized EX1. (A) RP-HPLC chromatogram of chemically synthesized EX1. Absorbance at 220 nm was monitored as a function of time. Gradient of 0-40% solvent B in 40 min at 40^{o}C is indicated by the blue trace. Solvent A: 0.1 % TFA in water, Solvent B: 90% acetonitrile, 9.9% water, 0.1 % TFA. (B) Corresponding TIC (m/z 150-2000) for the rEX1 peak aligned for the 0.1 min lag in detection. (C) Averaged ESI-MS of rEX1 for the indicated 0.1 min segment. The insert in each time range shows a comparison of calculated and observed ion masses. Masses indicative of succinimide containing product are indicated by red arrows.

Hydrolysis of the succinimide can lead to the formation of a mixture of aspartate and isoaspartate containing peptides. Therefore, the possible presence of isoaspartate content was investigated for the recombinant and chemically synthesized EX1 peptides. To identify isoaspartate formation, methods like NMR, Edman-based sequencing and protein isoaspartyl methyltransferase (PIMT)-mediated detection can be employed [22-24]. However, these methods can be insensitive and require high quantities of pure peptide [25]. On the other hand, peptides containing isoaspartate cannot be differentiated from aspartate by tandem mass spectrometry (LC-MS/MS) using conventional collision-induced disassociation (CID) methods [26]. With this technique, the amide bonds that are cleaved to produce ions for standard peptide sequencing have identical elemental composition and mass for both aspartate- and isoaspartate-containing peptides (Figure 5.4) [26]. Furthermore, EX1 peptides containing the different isomers of aspartate could not be resolved by analytical HPLC. Therefore, a complementary MS fragmentation technique, electron transfer dissociation (ETD) that primarily cleaves N—C_{α} bonds other than the amide linkage on the peptide backbone was employed to differentiate aspartic from isoaspartic acid.

As seen in Figure 5.4, ETD primarily cleaves N—C_α bonds in α-amino acid containing peptides. In the case of isoaspartate (β-amino acid), ETD fragmentation can also cleave the C_α—C_β bond. Thus, with isoAsp, an additional set of c and z ions, c + 57 (c`) and z – 57 (z`) serve as diagnostic ions [27-29].

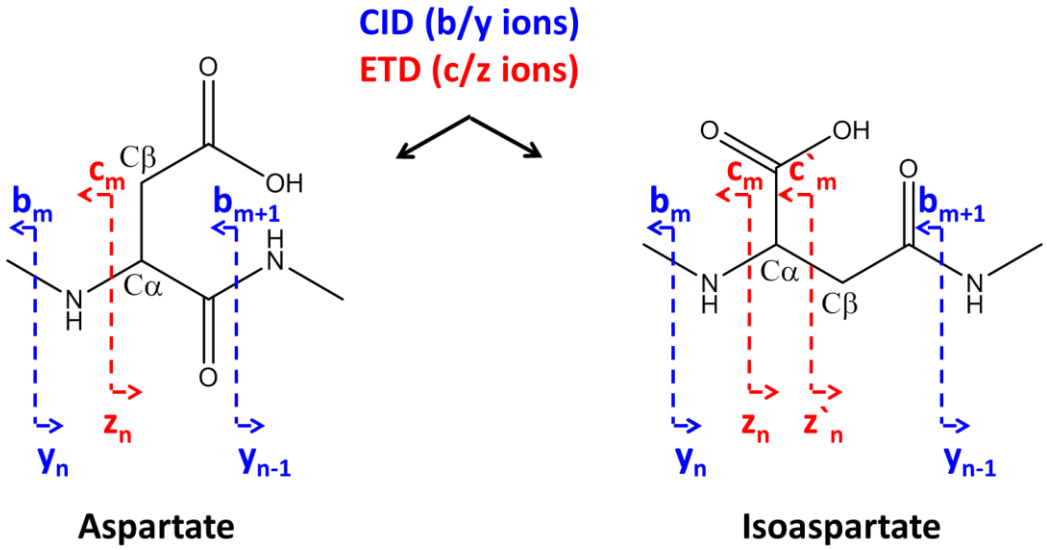


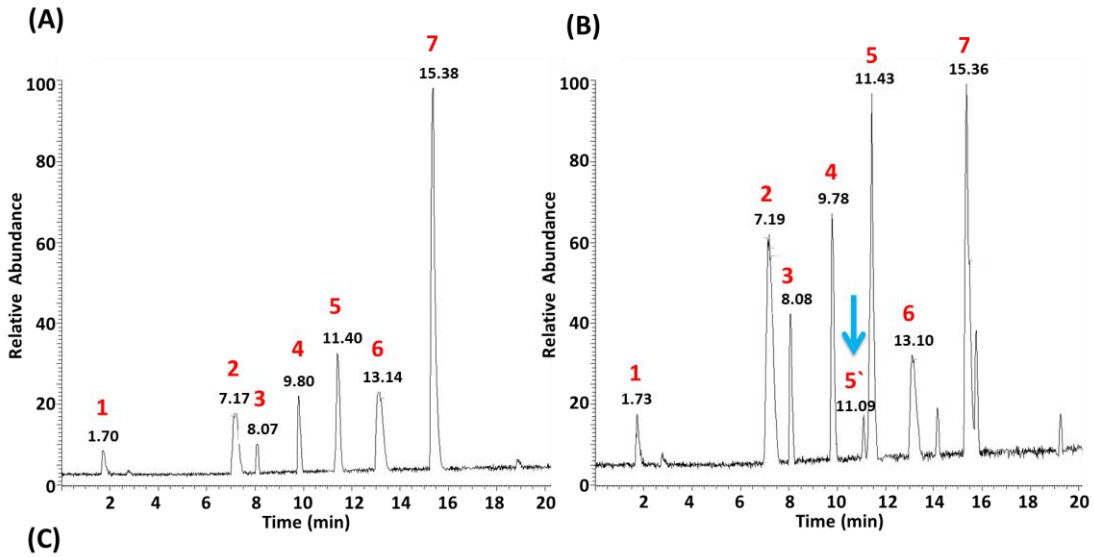
Figure 5.4 Comparison of CID and ETD fragmentation of aspartate and isoaspartate residues. During CID fragmentation, peptide bonds are cleaved and aspartate isomers cannot be identified by b and y type ions. However, during ETD fragmentation of peptides, primarily N—C_α bonds are cleaved, and c and z type ions are generated. In the presence of isoaspartate, C_α—C_β bonds may additionally be cleaved to generate c` (c+57) and z` (z-57) ions diagnostic for isoD. (m+n= peptide length)

In addition to the determination of isoaspartate presence, measurement of isoaspartate content is also desirable. For this, separation of two isomeric peptides is required. This can be more readily be achieved by using smaller isomeric peptides rather than larger ones so that the chemical and the chromatographic differences are

accentuated. Furthermore, the investigation of aspartate isomers in intact 21-residue EX1 peptide isomers by ETD would produce very complicated product ion spectra. Therefore, recombinant and chemically synthesized EX1 peptides were subjected to limited trypsin digestion to generate smaller tryptic peptides, among which would be some aspartate-containing peptides. ETD fragmentation generates fragment ions and it is most efficient with ions of higher charge. Thus, most abundant ions of highest charge are chosen to get the most of fragment ions. The peptides generated by limited trypsin digestion are shown in Figure 5.5C. Trypsin digestion, RP-HPLC of the trypsin digestion products and the ETD analysis were performed by Dr. James Kelley and Dr. Christopher Lai (Chemical Biology Lab, National Cancer Institute, Frederick, MD).

Seven primary tryptic peptides were detected and identified by ESI LC/MS analysis after limited digestion of recombinant (Figure 5.5A) and chemically synthesized (Figure 5.5B) EX1 peptides. Figure 5.5C shows the ions and their charge states obtained from ESI LC-MS analysis. Tryptic peptides **2**, **5**, **6** and **7** have the aspartyl residue. The additional peaks in Figure 5.5B with no time indicated are unidentified and do not correspond to any EX1 trypsin digestion products. The same tryptic peptides observed for recombinant EX1 are also observed for chemically synthesized EX1. One additional small peptide observed in Figure 5.5B is **5`** (VKVPDGTK) shown by blue arrow and it is detected at 11.09 min which is just before the detection of small peptide **5** (VKVPDGTK) detected at 11.43 min. Peptides

5 and **5`** are good candidates for detailed examination since they are chromatographically resolved, which will aid the determination of isoaspartate content. In addition, generation of triply charged (z=3) ions from octapeptides **5** and **5`** allows for a more efficient ETD fragmentation since the charge will be reduced during the ETD analysis. Therefore, the octapeptide is chosen for further ETD fragmentation to investigate the presence of isoaspartate by the investigation of c` and z` diagnostic ions. The presence of the diagnostic ions c` (c + 57) and z` (z – 57) will indicate which octapeptide has isoaspartate.



(C)

	VKVKVKVKVPDGTKVEVKVKV	m/z
1-	VK	246.181
2-	VPDGTK	616.330 308.668 (z=2)
3-	VEVK	474.292
4-	VKV	345.249
5 (5`)-	VKVPDGTK	422.250 (z=2) 281.836 (z=3)
6-	VPDGTKVEVK	536.306 (z=2) 357.873 (z=3)
7-	VKVPDGTKVEVK	649.888 (z=2) 433.594 (z=3) 325.448 (z=4)

Figure 5.5 Enhanced total ion chromatograms after limited tryptic digest of EX1 peptides. Enhanced TIC of (A) recombinant EX1 and (B) chemically synthesized EX1 after limited tryptic digest. (C) Tryptic peptides identified after limited digestion and their corresponding observed molecular species and charge states. The blue arrow in (B) indicates the elution position of tryptic peptide **5`**, which is isomeric to **5** and is not seen in (A).

The predicted ETD fragmentation pattern for octapeptides VKVPDGTK and VKVP*iso*DGTK are shown in Figure 5.6. For these octapeptide, the $c`_4$ ($c_4$ + 57) and $z`_4$ (z_4 – 57) ions are the diagnostic ions for isoaspartate identification. If these ions are

present, isoaspartate content will be confirmed. For this octapeptide, m/z for the c`$_4$ and z`$_4$ ions are 498.316 and 347.192, respectively.

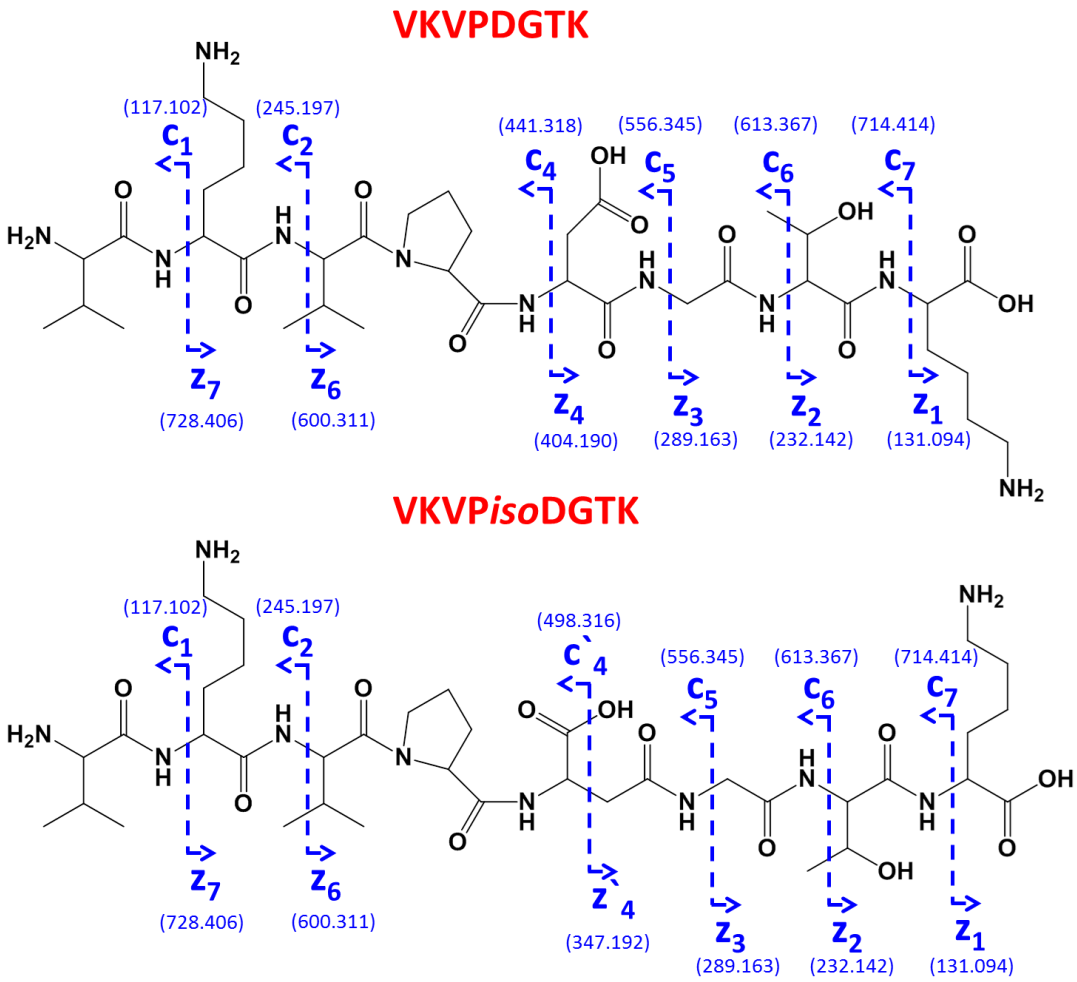


Figure 5.6 Predicted c and z fragment ions generated after ETD for VKVPDGTK and VKVPisoDGTK. The c`$_4$ (*m/z*=498.316) and z`$_4$ (*m/z*=347.192) ions are diagnostic for the presence of isoaspartate in the octapeptide.

The separation of the peptides from the limited trypsin digestion of chemically synthesized EX1 peptide is shown in the enhanced total ion chromatogram in Figure

5.7A (same as 5.5B). Figure 5.7B is MS/MS total ion chromatogram resulting from ETD fragmentation of $[MH_3]^{+3}$ (*m/z*=281.84) of the octapeptide VKVPDGTK. As expected, it is observed at 11.09 min and 11.43 min. Next, to locate isoaspartate-containing octapeptide VKVP*iso*DGTK, the diagnostic $z`_4$ and $c`_4$ ions are extracted from MS/MS TIC shown in Figure 5.7B. Figure 5.7C shows the high resolution (±0.005 Da) extracted ion chromatogram for $z`_4$ (*m/z*=347.192) ion and it is observed at 11.10 min which corresponds to tryptic peptide **5`** shown by the blue arrow in Figure 5.5B. Figure 5.7D shows the high resolution (±0.005 Da) extracted ion chromatogram for the other diagnostic ion $c`_4$ (*m/z*=498.316). It is also observed at the same time around 11.10 min. The presence of both diagnostic ions at 11.10 min confirms that the octapeptide **5`** in Figure 5.5B is the isoaspartate containing tryptic peptide and its content is likely to be ~15% when compared to the aspartate containing VKVPDGTK peptide **5**.

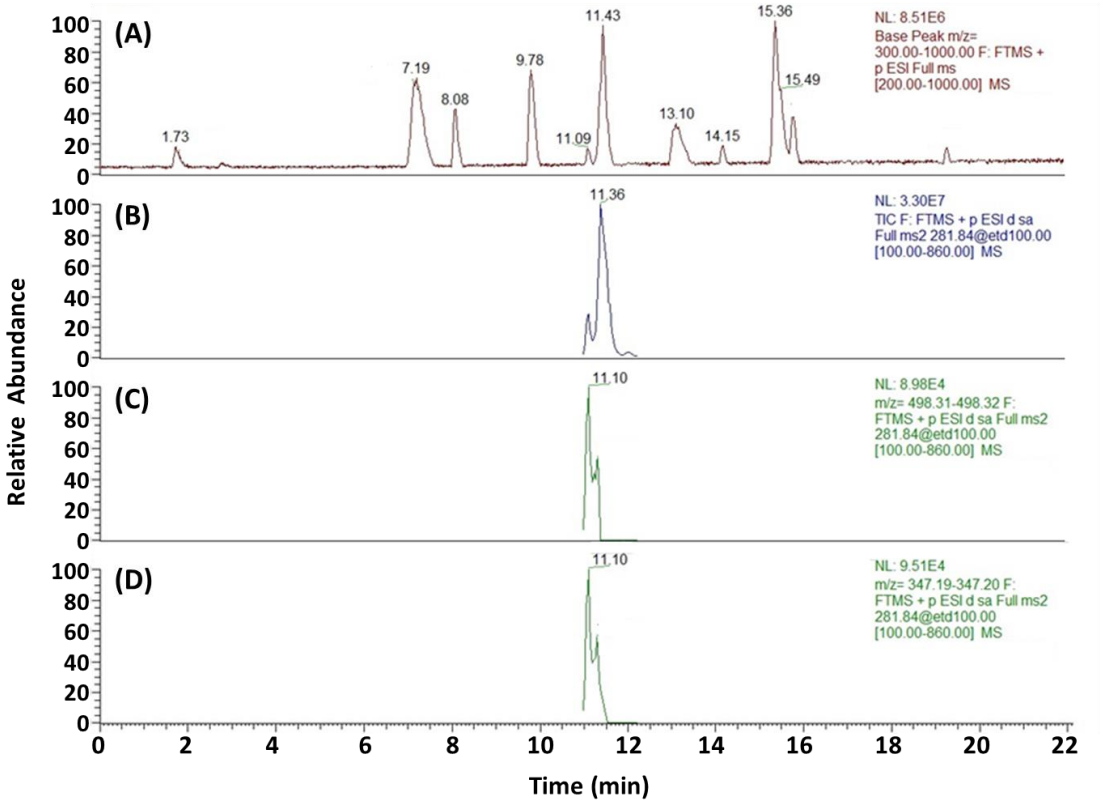


Figure 5.7 MS and MS2 ion chromatogram spectra generated during detection of diagnostic ions for VKVPisoDGTK fragment. (A) Enhanced TIC after mass spectrum of trypsin digested EX1 (chem) (B) MS^2 TIC for ETD of $[M+H_3]^{+3}$ (m/z=281.84). High resolution extracted ion chromatogram for (C) MS^2 m/z=347.19-347.20 and (D) MS^2 m/z=498.31-498.32.

Figure 5.8 shows the full scan-averaged MS^2 spectra for the two octapeptides **5** and **5`** identified in Figure 5.5. The time ranges selected are t_R=11.01-11.16 (Figure 5.8A) and t_R=11.30-11.57 (Figure 5.8B). When the time range 11.01-11.16 min was selected, the diagnostic ions were observed; however, none of the diagnostic ions were observed when the time range of 11.30-11.57 min was selected. These data also confirm that the isoaspartate containing VKVP*iso*DGTK peptide is the one eluting around 11.09 min and its content is likely to be ~15% when compared to the aspartate

containing VKVPDGTK peptide eluting around 11.43 min (Figure 5.7B). Recombinant EX1 was also confirmed for its aspartate content and the absence of succinimide or isoaspartate residues.

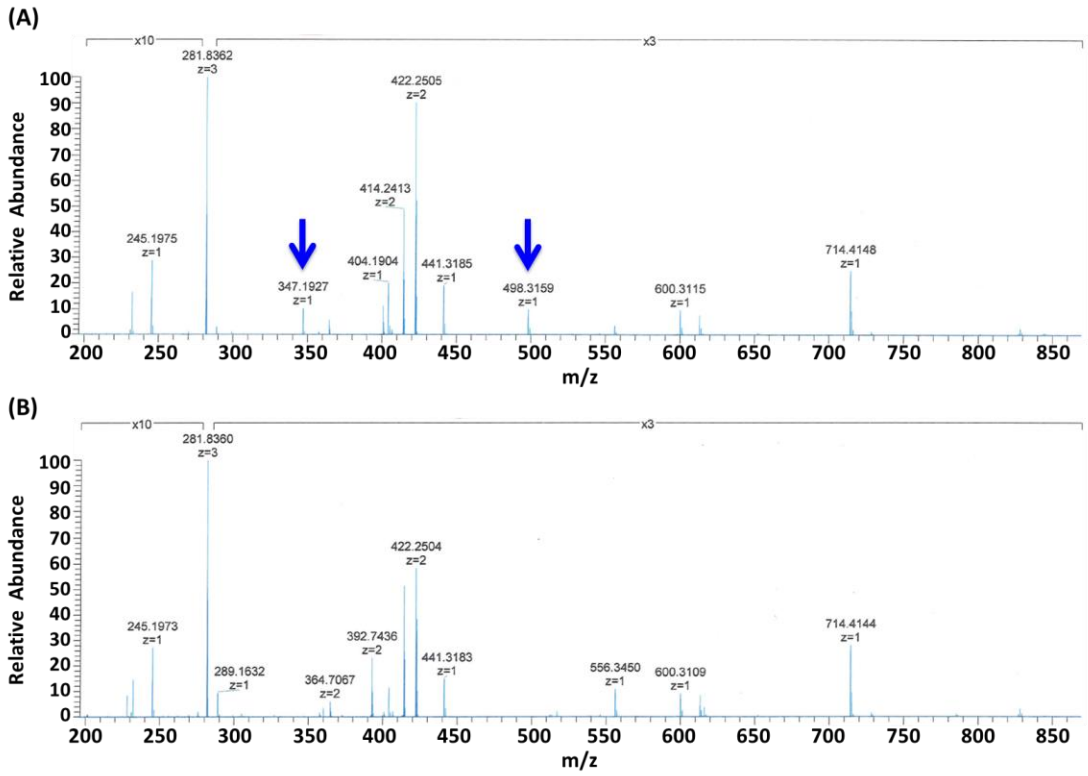


Figure 5.8 Full scan averaged MS2 spectra of octapeptides for peak (E) tR=11.01-11.16 and (F) tR=11.30-11.57 at trace B. Blue arrows show the diagnostic ions.

5.2.2 Characterization and comparison of pure recombinant and chemically synthesized (succinimide/isoaspartate containing) EX1 peptides

Liquid chromatography and mass spectrometry techniques showed that recombinant production of EX1 yielded the pure desired peptide nearly void of succinimide formation, whereas chemical synthesis resulted in EX1 as a mixture of peptides containing the expected aspartic acid, the succinimide, and the rearranged isoaspartate β-amino acid in its turn region. In this section, the effects of such impurities on the characteristics of the EX1 peptide and its gels will be presented.

Characterization of recombinant and chemically synthesized EX1 peptides and their corresponding gels was performed by circular dichroism (CD), rheology and transmission electron microscopy to assess peptide secondary structure, the elastic properties and their corresponding gels, and the local morphologies of the fibrils that constitute the gel networks. Wavelength CD spectra of 1wt% EX1 peptide gels formed by using recombinant and chemically synthesized EX1 peptides at physiological conditions after 1 h equilibration time is shown in Figure 5.9A. Pure recombinant EX1 peptide and chemically synthesized EX1 peptide, which has also succinimide and isoaspartate residues, in the gelled state have characteristic minima around 216 nm showing that the peptides fold and assemble into structures having high β-sheet content. Although C.D. spectroscopy shows little difference, the elastic properties of the EX1 peptide gels showed a significant difference. Dynamic time sweep

experiments assessed by oscillatory rheology at physiological conditions resulted in a storage modulus of 3.5 kPa at the end of 1 h for the chemically synthesized EX1 gel (Figure 5.9B). On the other hand, the pure recombinant EX1 peptide gel afforded a storage modulus of 1 kPa at the same conditions. Such an increase in the storage modulus might be due to the increased number of crosslinks when there are EX1 peptides containing succinimide or isoaspartate residues. The local morphologies of fibrils formed by the EX1 peptides were investigated by transmission electron microscopy (TEM) (Figure 5.9C). A slight degree of lamination between fibrils for pure recombinant EX1 was observed but chemically synthesized EX1 also having some succinimide and isoaspartate content, showed mostly long individual fibrils with a low degree of fibril bundling. Figure 5.9D shows the fibril width distributions for each EX1 peptide. Fibril widths of the corresponding EX1 peptides were observed around 3.5 nm which corresponds well with the width of a folded hairpin peptide in the self-assembled state.

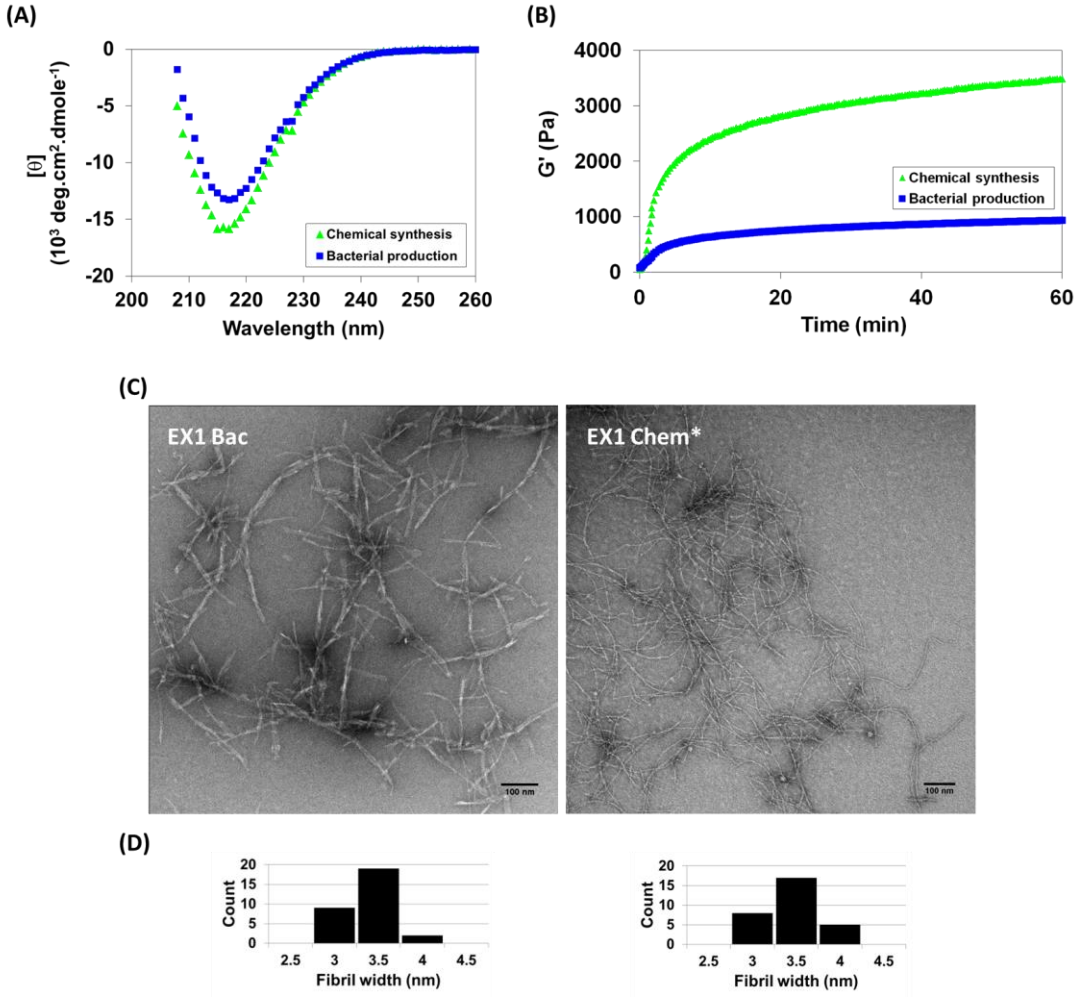


Figure 5.9 Comparison of characteristics of the pure recombinant and chemically synthesized (succinimide/isoaspartate containing) EX1 peptides. Peptides characterized at 1 wt% concentration at 37°C, pH 7.4, 50 mM BTP, 150 mM NaCl. (A) Wavelength CD spectra of EX1 peptides after 1 hour. (B) Evolution of storage moduli of EX1 peptide gels as a function of time. (C) Local morphology of EX1 fibrils assessed by TEM. Peptides were gelled for 1 hour, diluted and deposited for TEM examination. TEM images were taken by Katelyn Nagy (Schneider Lab). (D) Individual fibril width distributions underneath the corresponding TEM images.

The effect of succinimide and isoaspartate formation on the EX1 turn and hairpin structure in comparison with the aspartate containing EX1 hairpin was also studied. Energy minimized models by using Discovery Studio Suit v3.5 (Accelrys) and are shown in Figure 5.10. To construct the models, a model of EX1 containing aspartate at position 11 was first prepared. The turn region and the β-strands were pre-organized prior to minimization, having dihedral angles of their residues consistent with canonical β-sheet and type I + G1 turn structures. For the succinimide and isoaspartate-containing peptides, the aspartic acid of EX1 at position 11 was computationally replaced with a succinimide functional group or an isoaspartate β-amino acid. Small fibrils were formed by docking 4 hairpins in each layer of a bilayer (8 hairpins total). EX1 hairpins containing different residues at position 11 were not combined. Figure 5.10A shows the structures of individual aspartate, succinimide and isoaspartate residues. The succinimide forms a five-membered ring and isoaspartate formation results in one backbone having one additional methylene residue. Each of these assemblies was energy minimized by a protocol employing CHARMm force field and implicit water (dielectric set at 80). On examination of the minimized structures, it can be seen that the –VPDGT– turn of the EX1 peptide adopts a type I + G1 bulged turn where a stabilizing hydrogen bonding is formed between the amide hydrogen of valine at position *i* and the carbonyl oxygen of threonine at position *i*+4. A second bifurcated H bond is formed by the carbonyl oxygen of Valine at position *i* and the amide hydrogens of Glycine at position *i*+3 and the aspartate at position *i*+2

(Figure 5.10B). When aspartate is changed to succinimide and isoaspartate, the hydrogen bonding pattern in the turn region is disrupted due to the increased distance between the backbone atoms forming hydrogen bonds. When the aspartate residue is changed to succinimide and then to isoaspartate residues, the backbone around the turn region loses its regular tight reverse-turn shape and adopts a more rounded shape whose diameter is larger than the distance between the intramolecular hairpin strands (Figure 5.10C).

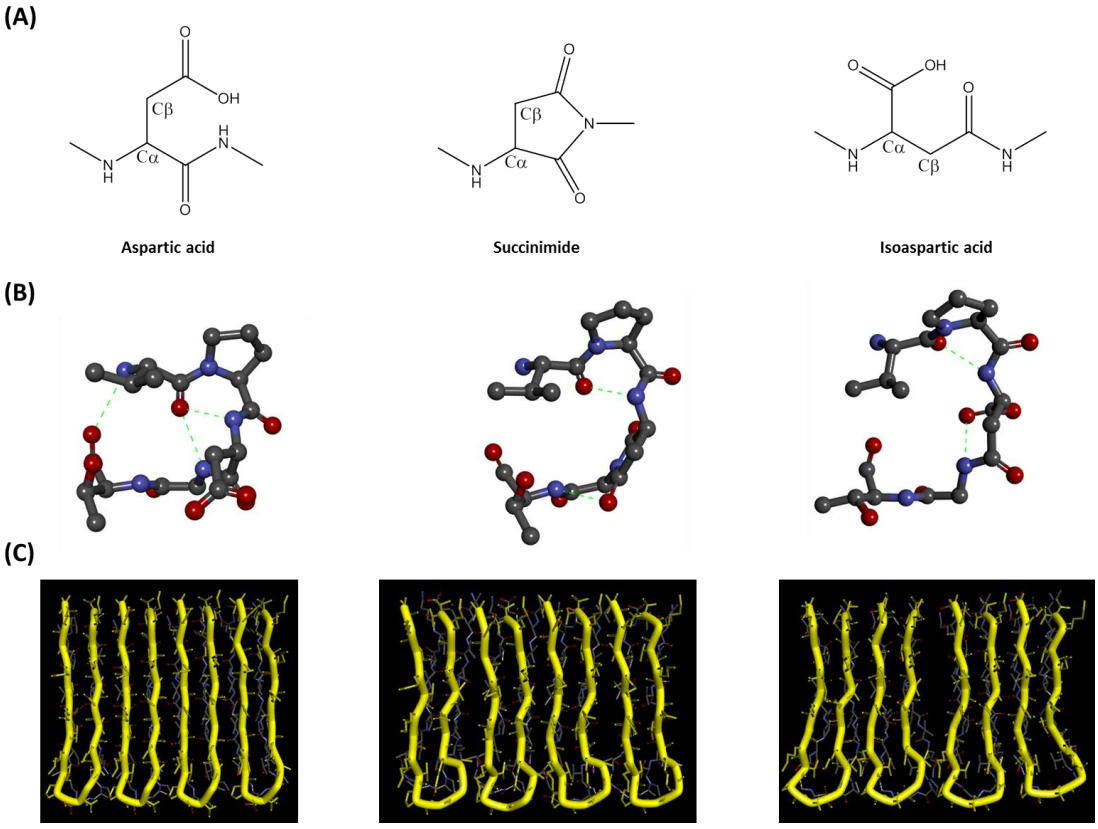


Figure 5.10 Effect of succinimide and isoaspartate formation on the turn and hairpin structure of EX1. (A) Structure of aspartate, succinimide and isoaspartate residues. Energy minimized models of (B) pentapeptide turn structures and (C) neighboring hairpin structures. Green dotted lines show hydrogen bonds and hairpin backbones are colored yellow.

The increased storage modulus of chemically synthesized EX1 gels having succinimide and isoaspartate residues, can be due to such irregular hairpin shapes promoting the nucleation of more branching points during the fibril growth along the fibril axis when compared to EX1 gels formed by pure recombinant EX1 peptide. More branching points result in higher number of cross-links that dictates the increase in the rigidity of the hydrogel.

TEM analysis showed that fibrils formed from pure EX1 were laminated whereas fibrils formed by succinimide- and isoaspartate-containing peptides were not. The differences observed in the local fibril morphologies might be explained by the structural changes at the turn region predicted by modelling. A slight degree of lamination in EX1 gels formed by pure recombinant EX1 peptide could be due to an ionic interaction between the negatively charged aspartate at the turn region and the positively charged N-terminus or the hydrophilic lysine rich face of another folded hairpin molecule in a neighboring fibril (Chapter 2). In the case of succinimide formation, the negative charge of aspartate is no longer available due to the formation of the five-membered succinimide ring. In the case of isoaspartate formation, there is still a negative charge but it is closer to the backbone and it might not be as accessible as in the case of aspartate residue for an ionic interaction with a positive charge on another hairpin molecule. Therefore, the absence of lamination for fibrils composed of chemically synthesized EX1 can be explained by the absence or inability of the negatively charged carboxylate to form interfibril electrostatic interactions.

5.2.3 Prevention of isoaspartate formation during chemical synthesis

During peptide synthesis, although the aspartate side chain is protected by *tert*Butyl group, succinimide formation is encountered with loss of the ester protecting group (Figure 5.11A) [30]. This is promoted by strong bases like piperidine used for the removal of the Fmoc group during automated peptide synthesis [31]. The Asp-Gly

sequence is especially susceptible to succinimide formation due to the absence of a side chain at the glycine residue which can hinder the nucleophilic attack of the nitrogen with the carbonyl carbon of the side-chain of the preceding residue. Succinimide formation is also dependent on the exposure time to piperidine so every deprotection steps that occurs after the aspartic acid residue is incorporated into the peptide is likely to increase the succinimide formation [32]. For complete protection against succinimide formation, the glycine nitrogen should be protected to prevent the ring closure reaction. The Fmoc-Asp(O*t*Bu)-(Dmb)Gly-OH dipeptide shown in Figure 5.11B has a 2,4-dimethoxy benzyl (Dmb) functionality protecting the glycine nitrogen. This protected dipeptide can be used in standard Fmoc-based solid phase peptide synthesis. Another approach can be to use pseudoproline dipeptides that are designed to disrupt secondary structure formation during peptide synthesis. Pseudoproline dipeptides have serine or threonine at their C-terminus and are protected as a TFA-labile oxazolidine which is isostructural with respect to proline. Figure 5.11C shows the structure of pseudoproline dipeptide Fmoc-Gly-Thr($\psi^{Me,Me}$pro)-OH, which can be used in the chemical synthesis of EX1. Although the nitrogen atom is not formally protected and still theoretically capable of attacking the carbonyl of the preceding residue side chain, this reaction can be disfavored by imposing a conformational constraint in the peptide backbone. The pseudoproline dipeptide imposes this constraint, making succinimide formation unlikely.

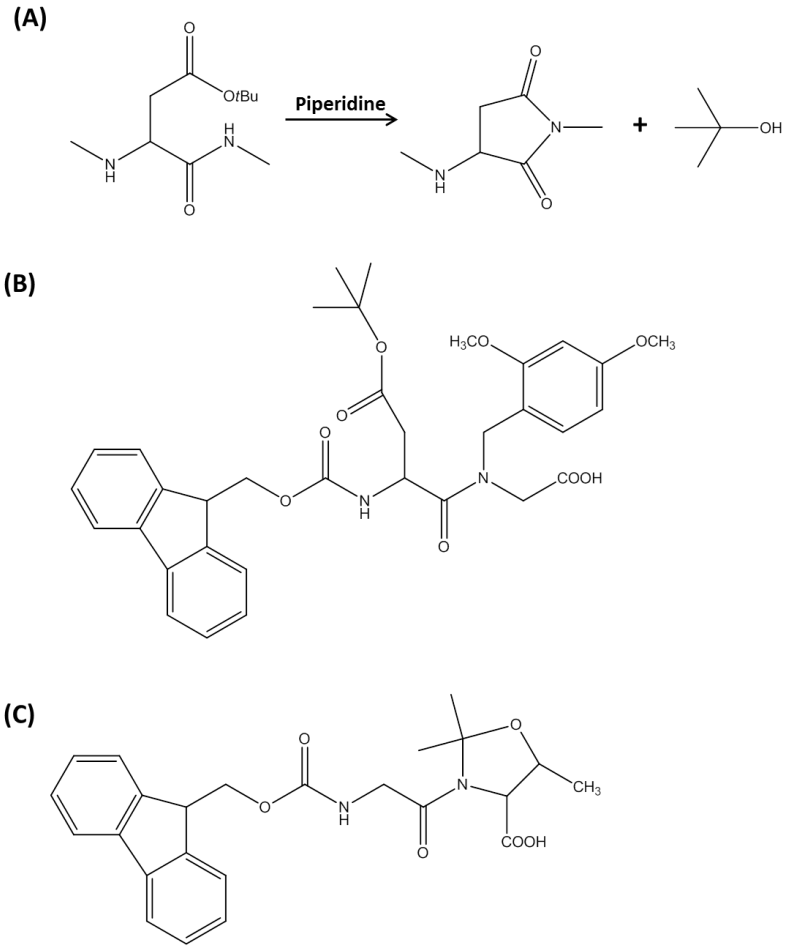


Figure 5.11 Succinimide formation and Fmoc-dipeptides to prevent succinimide formation during EX1 peptide synthesis. (A) Base-mediated succinimide formation (B) Fmoc-Asp(O*t*Bu)-(Dmb)Gly-OH (C) Fmoc-Gly-Thr($\psi^{Me,Me}$pro)-OH

The chemical synthesis of EX1 was repeated by using the pseudoproline dipeptide Fmoc-Gly-Thr($\psi^{Me,Me}$pro)-OH. The crude material was purified by RP-HPLC and analyzed by LC-MS. Although, EX1 was obtained as a single symmetrical peak and afforded just the expected molecular weight by mass spectrometry; trypsin digestion and EDT was performed for investigating isoaspartate presence. None of the

diagnostic ions for isoaspartate content were observed. Chemically synthesized pure EX1 peptide was characterized for secondary structure, elastic properties and local fibril morphologies (Chapter 2) and the corresponding comparisons with that of recombinant EX1 (Chapter 4) is explained in previous chapters. The results are in good agreement suggesting that the use of the pseudoproline dipeptide is successful in preventing succinimide and isoaspartate formation in the context of EX1. Thus either chemical synthesis using a pseudoproline dipeptide or bacterial expression affords EX1 with no succinimide and isoaspartate formation.

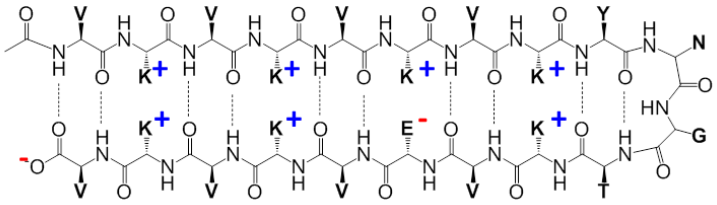
5.3 Effect of Lys Deletion at the Termini of Ac-EX3

Due to a malfunction of the automated peptide synthesizer, one of the preparations of the 20-residue Ac-EX3 peptide (Ac-VKVKVKVK-YNGT-KVEVKVKV-COO^-) resulted in a peptide whose molecular weight was consistent with the loss of one lysine residue. This deletion resulted in the loss of a positively charged residue in the peptide`s sequence. Decreasing the net charge of a β-hairpin amphiphile is known to result in faster peptide folding and self-assembly which consequently leads to more rigid hydrogels. For example, when a lysine residue on the solvent exposed hydrophilic face is changed to a glutamate residue, the resulting peptide, MAX8, folds and assembles quickly forming a more rigid hydrogel at physiological conditions [33]. When two lysine residues were changed to glutamate, the peptide rapidly precipitated and did not form a self-supporting hydrogel [34]. Each

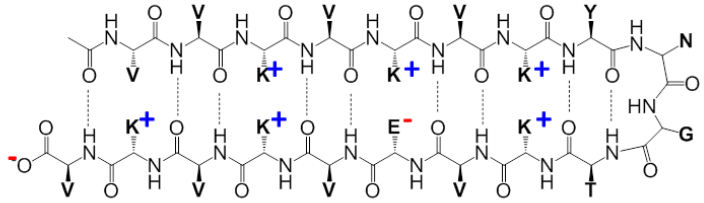
change from a lysine residue to a glutamate residue results in a decrease in net charge of the peptide by two units. However, deletion of a lysine residue decreases the net charge by one unit with respect to the parent Ac-EX3 peptide. In addition, for hairpin peptides that are amphiphilic, having distinct hydrophobic and hydrophilic faces, the position of the lysine deletion is important since it can disrupt the amphiphilic character of the folded peptide. For example, a lysine deletion in the middle of the hairpin would result in a frame shift of hydrophobicity and place a lysine side chain in the hydrophobic face of the hairpin. Such a change would be highly detrimental to the formation of the bilayers during self-assembly. However, lysine deletions at the peptide termini (C- or N-) might better accommodate this change where the amphiphilic faces would be unaffected. Hence lysine deletions at such positions may be better compensated for by the rest of the molecule.

The position of the deleted lysine residue was determined by tryptic peptide mapping and LC-MS. It was found that the C-terminal lysine at position 19 was deleted. Although this deletion sequence was unintentionally synthesized, it was worthy of further investigation for the effect the deletion had on the hydrogel forming characteristics of the peptide. A control peptide with a lysine deletion at the N-terminus (lysine at position 2) was also chosen to investigate if (i) the characteristics of the unintentionally synthesized 19-residue Ac-EX3 peptide are due to a change in the net charge from +5 to +4, or (ii) if there is any positional effects of the lysine deletion on the Ac-EX3 sequence. The 20-residue Ac-EX3, the Ac-EX3 with no lysine

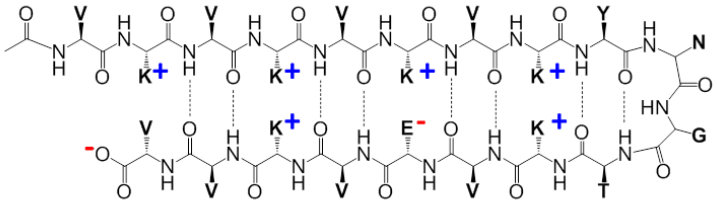
at position 19 (Ac-EX3-K19del) and the Ac-EX3 with no lysine at position 2 (Ac-EX3-K2del) were synthesized and purified for further investigation. The sequence, structure and net charge of each peptide are listed in Figure 5.12.



Ac-EX3 Ac - V<u>K</u>VKVKVK YNGT KVEVKV<u>K</u>V - COO⁻ +5



Ac-EX3-K2del Ac - VVKVKVK YNGT KVEVKVKV - COO⁻ +4



Ac-EX3-K19del Ac - VKVKVKVK YNGT KVEVKVV - COO⁻ +4

Figure 5.12 Structures of Ac-EX3, Ac-EX3-K2del and Ac-EX3-K19del. Peptide sequences and their net charges are provided below their respective structures. Deleted lysine residues on Ac-EX3 are underlined. Dotted lines show intramolecular hydrogen bonds.

5.3.1 Identification of the deleted lysine position

As stated earlier, the position of the deleted lysine residue was determined by trypsin digestion followed by LCMS. In this section is provided a brief description of that work. If the lysine deletion is closer to the N- or C-terminus of the peptide, the amphiphilicity would not be significantly disturbed. On the other hand, a lysine deletion in the middle of the strands would perturb the facial display of the hydrophilic and hydrophobic residues and consequently affect the folding and self-assembly characteristics of the peptide. To find the position of the deleted lysine residue, trypsin, which cleaves peptides at the C-termini of positively charged residues, was employed. Both the bonafide 20-residue Ac-EX3 and the 19-residue Ac-EX3 containing the deletion were purified and subjected to trypsin digestion and LC-MS analysis. When the total ion chromatograms (TIC) are compared in Figure 5.12, major differences are observed at the C-terminal tryptic fragments. Trypsin digestion of the 20-residue Ac-EX3 yielded a VKV fragment that corresponds to the C-terminal tripeptide (Figure 5.13A). When the TIC of the trypsin digestion of the 19-residue Ac-EX3 was analyzed, a VV dipeptide fragment was observed (Figure 5.13B). The results indicated that the lysine deletion occurred at position 19 of Ac-EX3 since the only possibility of the presence of the VV dipeptide following trypsin digestion was having no Lys19. This deletion sequence (Ac-EX3-K19del) was later synthesized intentionally to verify that the original peptide was indeed the K19 deletion.

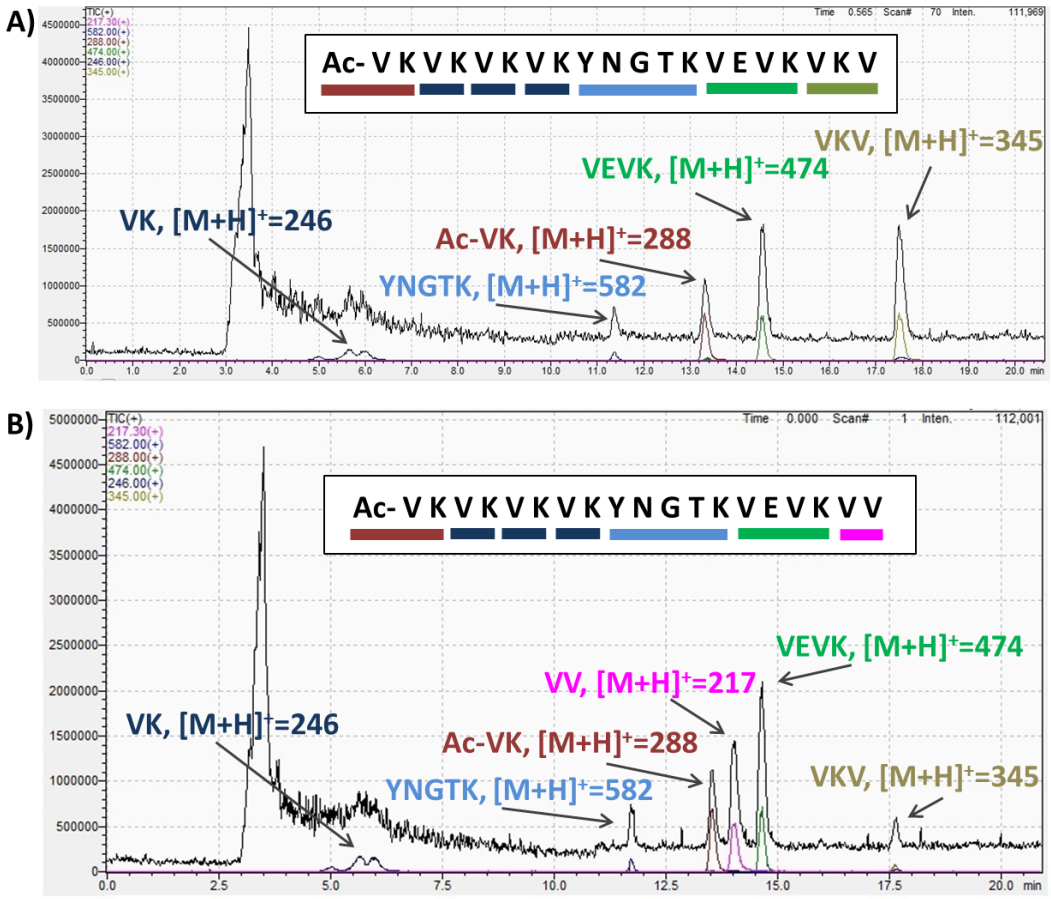


Figure 5.13 Total ion chromatogram of tryptic peptide fragments of (A) 20-residue and (B) 19-residue Ac-EX3 peptides. Each color designates a different tryptic peptide with the corresponding m/z. Inserts in each panel show peptide sequences according to the corresponding tryptic peptides identified.

5.3.2 Characterization of Ac-EX3, Ac-EX3-K2del and Ac-EX3-K19del

Temperature dependent CD spectra of 150 μM Ac-EX3 and its lysine deletion variants at 216 nm (pH 7.4, 50 mM BTP, 150 mM NaCl) are shown in Figure 5.14A. Ac-EX3 and Ac-EX3-K2del behaved very similarly throughout the experiment in the temperature range of 5 to 85°C. Around 25-30°C, Ac-EX3 and Ac-EX3-K2del

peptides start to fold into β-sheet structure and continue folding until around 50-55^{o}C, and they retained their folded structures as the temperature increased thereafter. On the other hand, Ac-EX3-K19del showed a different trend, staying in random coil conformation until around 70^{o}C. At higher temperatures, Ac-EX3-K19del started to fold into a β-sheet structure but it did not fully fold even at the end of the experiment at 85^{o}C. At a higher peptide concentration, 1wt% peptide gels of Ac-EX3 and its lysine deletion variants were formed and characterized for their secondary structures after 1 h at 37^{o}C (pH 7.4, 50 mM BTP, 150 mM NaCl). Each peptide gel showed a characteristic minimum at 216 nm (Figure 5.14B). This result shows that higher concentrations of Ac-EX3-K19del are needed to facilitate folding while Ac-EX3 and Ac-EX3-K2del adopt the folded state even at dilute concentrations at temperatures of 25-30^{o}C. The results at dilute concentrations suggest that the deletion of Lys2 does not have a significant effect on folding and assembly when compared to Ac-EX3. However, deletion of Lys19 notably changes the β-sheet forming characteristics of Ac-EX3-K19del at the same conditions. The reason that Ac-EX3-K19del folding and assembly requires higher temperatures might be due to the C-terminal VV portion of the peptide. This dipeptide is situated next to a negatively charged carboxylate functionality, which may be preventing hydrophobic collapse of the hairpins during bilayer formation. If this is the case, higher temperatures may be necessary to drive the hydrophobic effect, which is observed. On the other hand, at higher concentrations at physiological conditions, Ac-EX3-K19del folds into a β-sheet structure as Ac-EX3

and Ac-EX3-K19del do, suggesting that cooperativity in which preformed folded and self-assembled peptides act as nucleation sites and promote further folding and self-assembly, compensates for the extra thermal energy required for folding at dilute concentrations. In addition to the C-terminal carboxylate hindering the hydrophobic affect, it can also be hindering the association of β-strands during folding. On inspection of Figure 5.12, it can be seen that in the folded state, the carboxylate must be positioned directly across from the carbonyl oxygen of the Valine1 on the opposing β-strand. This results in charge-charge repulsion disfavoring the folded state. This might explain why Ac-EX3-K19del requires higher temperatures around 70^{o}C to start folding at dilute concentrations when compared to Ac-EX3 and Ac-EX3-K2del (Figure 5.14A). High temperatures drive the hydrophobic effect which ultimately overcomes the electrostatic repulsion. Interestingly, the change in the net charge of the peptide from +5 to +4 can be ruled out to explain the differences in the β-sheet forming characteristics because two peptides with lysine deletions at different positions should have behaved similar to each other and different when compared to Ac-EX3. Thus, the position of the lysine deletion on the peptide sequence is important.

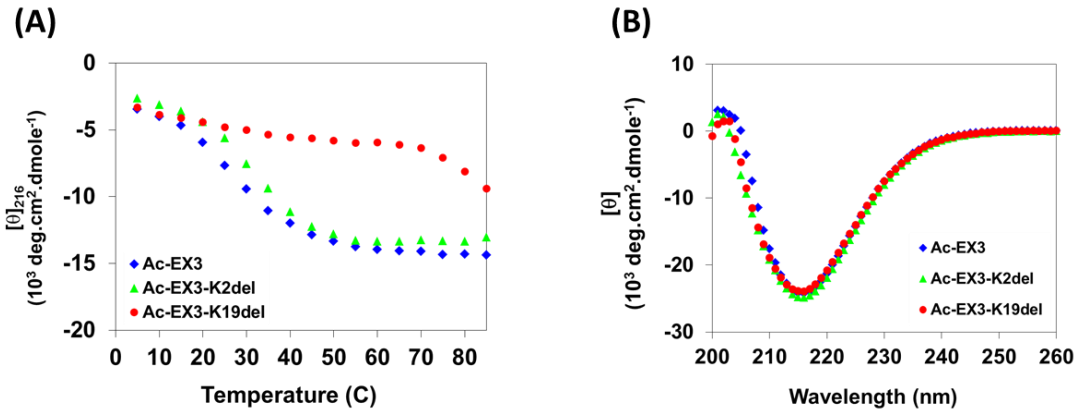


Figure 5.14 Wavelength CD spectroscopy of Ac-EX3 and its variants. (A) Temperature dependent β-sheet formation of 150 μM peptides by monitoring mean residue ellipticity at 216 nm from 5°C to 85°C (pH 7.4, 50 mM BTP, 150 mM NaCl). (B) Wavelength spectra of 1wt% gels at 37°C after 1 hour (pH 7.4, 50 mM BTP, 150 mM NaCl).

The mechanical rheological properties of gels formed from 1wt% Ac-EX3 and its lysine deletion variants at physiological conditions (37°C, pH 7.4, 50 mM BTP, 150 mM NaCl) are shown in Figure 5.15. The time-sweep rheological experiments where the gelation of a 1wt% peptide solution is monitored by the evolution of storage moduli as a function of time, for the three peptides are shown in Figure 5.15A. As observed in the secondary structure characterization by C.D., 1 wt% Ac-EX3 and Ac-EX3-K2del gels behave similarly and afford a storage modulus (G`) of 2 kPa at the end of 1 h. However, a 1wt% Ac-EX3-K19del gel affords a storage modulus of 4 kPa which is twice that of the Ac-EX3 and Ac-EX3-K2del gels. The significant increase in the storage modulus of the Ac-EX3-K19del gel suggests that an increased number of cross-links formed during the gelation when compared to the other two gels. When the

folded structures of the three peptides are compared (Figure 5.12), the lengths of the two strands of each peptide show some differences. The two strands of Ac-EX3 are almost the same length. The two strands of Ac-EX3-K2del have some differences but the most different strand lengths are observed for Ac-EX3-K19EC. During fibril growth, such a difference might have an effect in promoting fibril branching as lateral associations are formed during assembly. When the two strands of a hairpin molecule are at similar lengths, it is more likely to have regular fibril growth along the fibril axis. On the other hand, when the two strands are at different lengths, it is possible to mediate intermolecular H-bond formation that could cause irregular fibril growth and branching into different directions which results in more cross-link points and a more rigid gel. Independent shear-thinning and recovery experiments were performed to investigate the recovery of the elastic properties of the 1wt% hydrogels after being subjected to shear-thinning (Figure 5.15B). After 20 min of gelation in the same way as in the dynamic time sweep experiments, the strain was increased from 0.2% strain to 1000% for 30 sec to thin the gels. The immediate decrease of G` by 1000% strain during the course of thinning was followed by the immediate recovery of G` to almost the same extent in a few minutes. One minute after the shear ceased, the storage modulus recovered 90% for all three gels. When the total 40 min of experiment was analyzed, shear-thinning did not make any difference on the overall evolution of storage modulus for each peptide gel and excellent recovery properties are observed. In figure 5.15C and D, dynamic frequency sweep experiments show that the storage moduli (G`) of the Ac-EX3, Ac-EX3-K2del and Ac-EX3-K19del gels are independent

of the frequency from 0.1 to 100 rad/sec and all hydrogels at 1 wt% concentration yield around 10% strain.

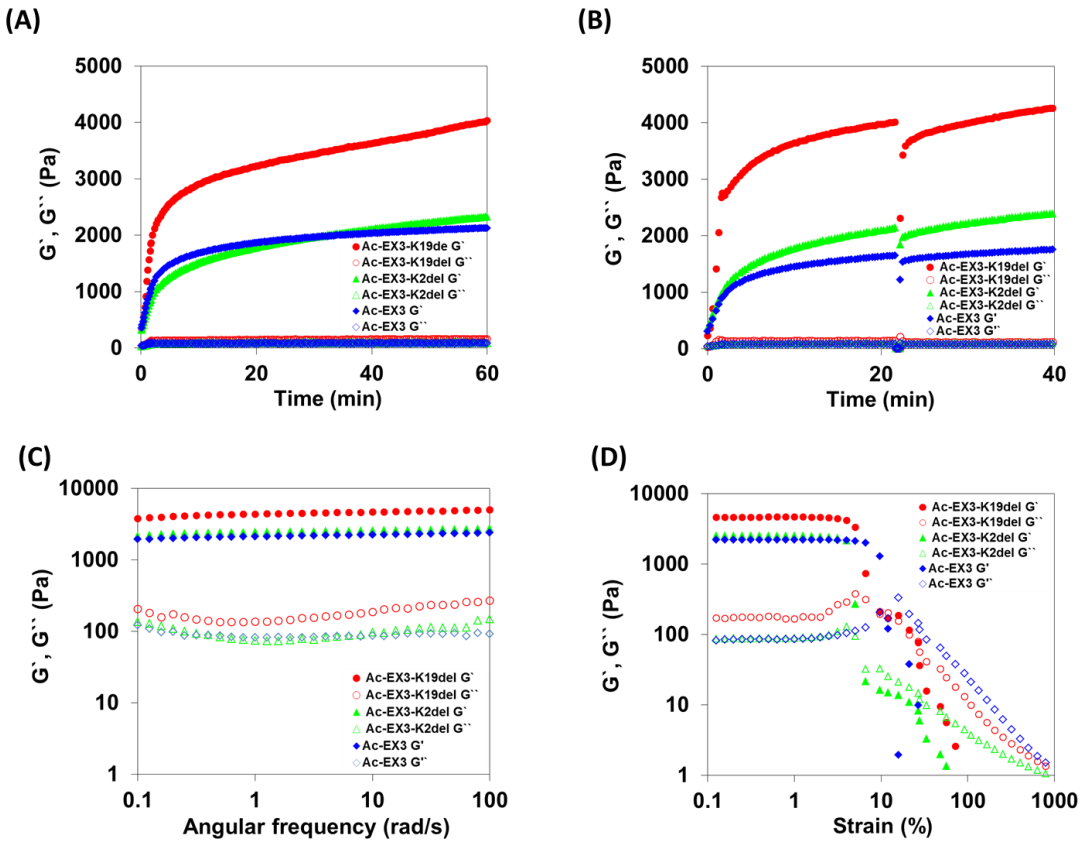


Figure 5.15 Mechanical properties of 1wt% Ac-EX3 and its variants by oscillatory rheology. (A) Dynamic time sweep experiment showing the evolution of storage (G`) and loss (G``) moduli of 1 wt% peptides at pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C. (B) Shear-thin and recovery experiment with 1000% strain for 30 sec after 20 min of gelation followed by another 20 min for recovery. (C) Dynamic frequency sweep. (D) Dynamic strain sweep experiments of 1 wt% peptides in pH 7.4, 50 mM BTP, 150 mM NaCl, 37^{o}C .

The local morphologies of the fibrils comprising the gels formed by the three peptides were investigated by transmission electron microscopy (TEM) in Figure 5.16. A slight degree of fibril bundling for Ac-EX3 and Ac-EX3-K2del was observed. Ac-EX3-K19del formed relatively larger bundles of fibrils. Again, Ac-EX3 and Ac-EX3-K2del had very similar fibril morphologies, whereas Ac-EX3-K19del resulted in a different morphology. The width of single fibrils formed in all three peptides however were around 3.5 nm which corresponds well with the width of a folded hairpin in the self-assembled state as assessed by modeling (Figure 5.16B). The bundling observed for the fibrils of all three peptides might be due to the hydrogen bonding between the side chains of asparagine residues originating from distinct turns that can facilitate interaction between two fibril edges. In the folded state, the portion between V3 and V18 of the Ac-EX3 peptide is common for all three peptides. The difference of the three hairpin structures are at the N- and C-termini region. Therefore, the interaction of the N- and C- termini region of Ac-EX3 and its variants might also contribute to differences in the local morphologies of the fibrils observed in the TEM images. The similar and low degree of bundling of the Ac-EX3 and Ac-EX3-K2del fibrils suggest that the interactions of distinct N- and C- terminal regions of two fibril edges are similar. However, in the case of the Ac-EX3-K19del fibrils, the interactions of distinct N- and C- terminal regions of fibril edges should be different than that of Ac-EX3 and Ac-EX3-K2del, and such differences result in a higher degree of bundling of the EX3-K19del fibrils. Such a difference can be related to the difference in the length of two strands of the hairpin molecule. As stated earlier, such differences were related to the

formation of branch points during fibril growth. These same differences can also be responsible for the interfibrillar interactions that form a higher level of bundling.

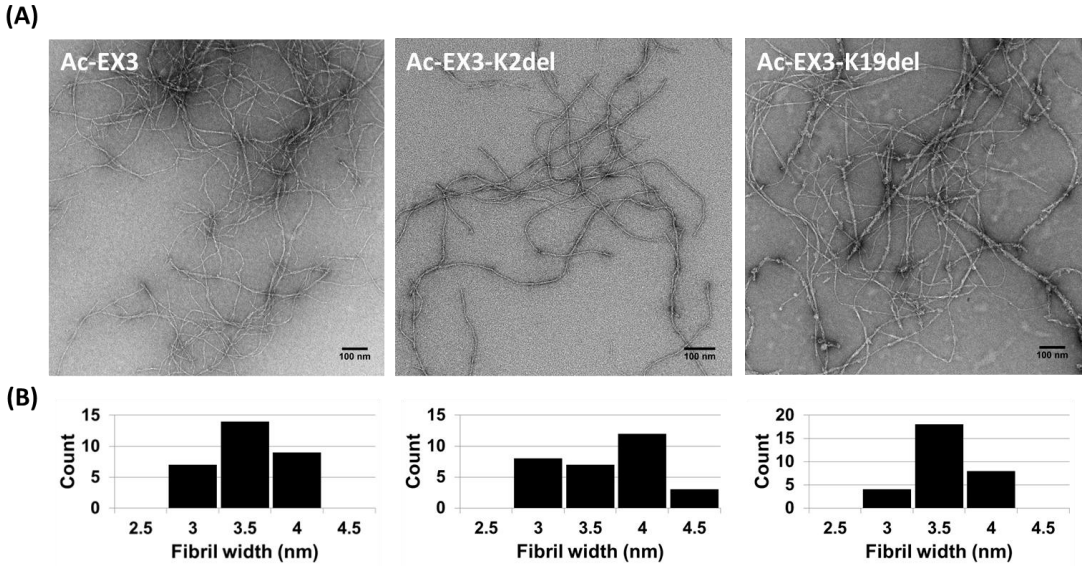


Figure 5.16 Local morphology of fibrils formed by Ac-EX3 and its variants assessed by TEM. The 1 wt% peptides were gelled at 37^{o}C for 1 hour. (A) Gels were diluted and deposited for TEM interrogation. (B) Individual fibril width distributions. TEM images were taken by Katelyn Nagy.

5.4 Conclusions

During peptide synthesis, undesired side reactions can result in changes in the target peptide sequence and structure and consequently the characteristics of the target peptide are altered. In this chapter, succinimide and isoaspartate formation, and lysine deletion in different target peptides are discussed together with their consequences on the peptide`s gel-forming behavior. Synthetic methods to prevent succinimide

formation and the consequence of lysine deletions at specific positions were also presented.

When the folding, assembly and gel characteristics of chemically synthesized EX1 were compared with that of recombinant EX1, significant differences were observed and each peptide was subjected to further investigation to identify the potential impurities causing such alterations in the peptide characteristics. The RP-HPLC chromatogram of the recombinant peptide afforded a single symmetrical peak, and detailed analysis of the peptide by mass spectroscopy showed that very little succinimide or isoaspartate formation accompanied the peptide`s bacterial production. However, chemically synthesized EX1 afforded a chromatogram having a shoulder indicating that an impurity having similar chromatographic behavior exists. Mass spectroscopy indicated that the shoulder contained a peptide having a molecular weight of 18 Da less than the target peptide EX1. The difference in the molecular weight was due to the loss of water indicated succinimide formation. This ring closing reaction can occur during synthesis and is promoted by piperidine at every deprotection step during the peptide synthesis. Through hydrolysis, the succinimide can either regenerate aspartate or form isoaspartate depending on how the ring is opened by water during hydrolysis. Unfortunately, the masses of the parent peptides having either aspartate or isoaspartate are the same, making the detection of the isoaspartate isomer difficult. Therefore, additional mass spectrometry analysis methods such as electron transfer disassociation (ETD) fragmentation were employed

to distinguish EX1 containing aspartate and isoaspartate. ETD fragmentation enables the detection of isoaspartate by cleaving the C_{α}—C_{β} bond of the β-amino acid and detecting the resulting fragments. The tryptic digestion of EX1 afforded a smaller portion of the peptide which was easier to analyze (octapeptide). The ETD results showed that when chemically synthesized EX1 was analyzed, diagnostic ions due to isoaspartate formation were observed. No diagnostic ion however was observed in the recombinant EX1 analysis. Furthermore, characterization of the recombinant and chemically synthesized EX1 peptides and their corresponding 1wt% gels were performed to assess their secondary structure, the elastic properties of their corresponding gels and the local fibril morphologies of their networks at physiological conditions. Both peptides adopt a β-sheet structure at 1 wt%, however there was a significant difference in the elastic properties of their gels. Pure recombinant EX1 gel afforded a storage modulus of 1 kPa whereas the chemically synthesized EX1 gel resulted in a more rigid hydrogel with a storage modulus of 3.5 kPa at the end of 1 hour of gelation at 37^{o}C. Local fibril morphologies imaged by TEM showed that recombinant EX1 fibrils had a slight degree of lamination and the chemically synthesized EX1 formed long individual fibrils. Individual fibril widths for both EX1 peptides were assessed around 3.5 nm which corresponds to the length of a folded hairpin in the self-assembled state. Modeling of the EX1 peptide with aspartate, succinimide and isoaspartate at position 11 suggested that the turn structure is perturbed when aspartate is changed to succinimide and isoaspartate. The aspartate containing turn adopts a regular type I` + G1 bulge turn, but succinimide and

isoaspartate changes the turn structure as well as the folded hairpin structures. Due to the presence of succinimide and isoaspartate, these differences might explain the increase in the storage modulus by forming irregular shapes that can promote nucleation of fibril branching during fibril growth, which consequently results in more cross-links and ultimately a more rigid gel. The difference of the turn region probably resulted in the difference in fibril morphology as well. The use of a pseudoproline dipeptide during the chemical synthesis of EX1 was successful in preventing succinimide or isoaspartate formation as confirmed by mass spectrometry. In addition, chemically synthesized pure EX1 behaved in a similar fashion compared to recombinant EX1 as shown in Chapter 4. Overall, it was shown that succinimide and isoaspartate formation during the chemical synthesis of EX1 altered the structure and the material forming characteristics of the peptide significantly. Bacterial production of EX1 was successful with very little or no alteration of the aspartate residue.

Following the synthesis of another peptide AcEX3, a lysine deletion sequence was observed due to a malfunctioning peptide synthesizer. The position of the deletion was determined by mass spectroscopy and found to occur at position 19. This peptide formed gels having very high moduli as compared to the parent Ac-EX3 peptide gels. As a result, the K19 deletion was studied further with the aid of a control peptide having a lysine deletion at position 2 to investigate if the change in the gel behavior was due to a decrease in the net charge of Ac-EX3 or due to the position of the deletion. Ac-EX3, Ac-EX3-K2del and Ac-EX3-K19del and their corresponding gels

were characterized for their secondary structure, elastic properties and local fibril morphologies at physiological conditions. The Ac-EX3 and Ac-EX3-K2del peptides behaved similarly suggesting that the deletion of lysine at position 2 had no significant effect. At dilute concentrations, Ac-EX3 and Ac-EX3-K2del folded into β-sheet structures at much lower temperatures than that of Ac-EX3-K19del. When the C-termini of the three peptides are compared, Ac-EX3 and Ac-EX3-K2del peptides have KV-COO^- and the Ac-EX3-K19del peptide has VV-COO^-. This VV-COO^- portion of the Ac-EX3-K19del peptide might have a role in preventing the peptide from folding and assembling at dilute concentrations and low temperatures. At 1wt% concentration, all the peptides showed similar β-sheet formation at 37^oC. Therefore, cooperativity during the folding and self-assembly process at higher concentrations can overcome the deleterious effects of the K19 deletion. However, the elastic properties of 1 wt% Ac-EX3-K19del gel at physiological conditions resulted in twice the storage modulus of that of the Ac-EX3 and Ac-EX3-K2del gels after 1 h of gelation (4 kPa vs. 2 kPa). This increased storage modulus may be due to an increase in the number of cross-links caused by the differences in the hairpin structure. The differences of the peptides are at the N- and C-termini region. In the case of Ac-EX3-K19del hairpin, the lengths of its strands are quite different. Such a difference during lateral assembly of hairpin peptides might promote nucleation of fibril branching that can result in more cross-links. Shear-thinning and recovery experiments showed that all three peptide gels have very good recovery properties. Local fibril morphologies of Ac-EX3 and Ac-EX3-K2del fibrils were similar and a low degree of bundling was observed. Ac-EX3-

K19del formed fibrils with more interfibrillar interactions and a relatively higher degree of bundling was observed. Fibril widths of all three peptide-gels were similar and similar to the length of a folded hairpin (3 nm). The higher degree of bundling observed for the Ac-EX3-K19del fibrils might arise from inter-fibril interactions that arise from interactions made between the termini of distinct hairpins across fibrils. Overall, the investigation of a lysine deletion position on Ac-EX3 due to an unintentionally synthesized peptide, and its characterization with comparison to another lysine deletion sequence, lead to new remarkable results: i.e. the changes at the N- and C-termini region of Ac-EX3 hairpin has an important effect on the elastic properties and macromolecular characteristics of its corresponding gel.

In conclusion, further investigation of the peptides with an altered structure either due to changes of a residue, or due to a deletion of a residue, during chemical synthesis lead to new results suggesting that the changes in hairpin turn region or hairpin N- and C-termini region has a significant impact on the hydrogel material properties.

5.5 Experimental Section

Production of EX1 peptides by chemical synthesis (Chapter 2) and by bacterial expression (Chapter 3-4) was outlined in previous chapters. Chemical synthesis of Ac-EX3 and its variants were synthesized the same way as outlined in Chapter 2.

Secondary structures, mechanical properties and local fibril morphologies of peptides were characterized by CD spectroscopy, oscillatory rheology and TEM, respectively as outlined in Chapter 2.

5.5.1 Purification of Ac-EX3, Ac-EX3-K2del and Ac-EX3-K19del

Peptide purification was carried out by reverse phase-HPLC at 40^{o}C equipped with a semi-preparative Vydac C18 column. HPLC solvents A and B were 0.1% TFA in water and 0.1% TFA in 9:1 acetonitrile: water, respectively. For Ac-EX3 and Ac-EX3-K19del purification, linear gradients were employed as 0% solvent B over 1 min, 0-15% solvent B over 8 min, 15-29% solvent B over 28 min and then 29-39% solvent B over 40 min. For Ac-EX3-K2del purification, linear gradients were employed as 0% solvent B over 1 min, 0-10% solvent B over 9 min, 10-25% solvent B over 30 min and then 25-35% solvent B over 40 min. Analytical HPLC and electrospray ionization (ESI-positive mode) were performed to verify the purity of the lyophilized peptides.

5.5.2 Limited tryptic digestion of EX1

A 100 μg aliquot of EX1 in aqueous solution was diluted with an equal volume of 10 mM NH_4HCO_3 (pH 7.8) before being added to 2 μg of specially purified trypsin (Thermo In-Solution Tryptic Digestion Kit) in a screw-cap Eppendorff vial. The resulting solution was then incubated at 37ºC for 4 h with periodic vortexing. A 50 μg aliquot of this digestion solution was then quenched by adding an equal volume of 2%

CH_3CN/H_2O containing 0.1% formic acid. The quenched digestion solution was further diluted to 1.00 ml with 2% CH_3CN/H_2O before analysis by LC/MS/MS (see below). The remaining tryptic digest solution was incubated overnight at 37º C before being quenched and treated in the same manner as above.

5.5.3 Mass spectrometry

Intact EX1 samples and their corresponding tryptic digests were analyzed by ESI LC/MS/MS on a Thermo-Fisher LTQ-Orbitrap-XL system in positive ion mode. A small-particle, C_{18} reversed phase, narrow-bore column was used with a shallow CH_3CN/H_2O gradient to separate the mixture of tryptic peptides. Primary ESI mass spectra were obtained at a resolution of 30,000 and were used to tentatively identify tryptic digest peptides. The sequence of these peptides was further confirmed by high resolution (15,000) MS/MS with ETD fragmentation utilizing data dependent selection of appropriate precursor ions. Fragmentation in the resulting ETD product ion mass spectrum was then related to structure on the basis of chemistry, chromatography, and accurate mass measurement.

5.5.4 Trypsin digestion of the 20-residue Ac-EX3 and 19-residue Ac-EX3

0.2 mg peptides were dissolved in 100 μL buffer (50 mM BTP, pH 7.4) including 2.5 μM trypsin. After 20 h at 37^{o}C, 50 μL of reaction solution was diluted

by the addition of 80 μL of 0.1% TFA in water and analyzed by LC-MS for tryptic peptides cleaved at the lysine C-terminus.

REFERENCES

1. Hsieh, K.H., M.M. Demaine, and S. Gurusidaiah, *Side reactions in solid-phase peptide synthesis and their applications.* International Journal of Peptide and Protein Research, 1996. **48**(3): p. 292-298.

2. King, D.S., C.G. Fields, and G.B. Fields, *A Cleavage Method Which Minimizes Side Reactions Following Fmoc Solid-Phase Peptide-Synthesis.* International Journal of Peptide and Protein Research, 1990. **36**(3): p. 255-266.

3. Stathopoulos, P., et al., *Side reactions in the SPPS of Cys-containing peptides.* Amino Acids, 2013. **44**(5): p. 1357-1363.

4. Dimarchi, R.D., et al., *Weak Acid-Catalyzed Pyrrolidone Carboxylic-Acid Formation from Glutamine during Solid-Phase Peptide-Synthesis - Minimization by Rapid Coupling.* International Journal of Peptide and Protein Research, 1982. **19**(1): p. 88-93.

5. Pedroso, E., et al., *Diketopiperazine Formation in Solid-Phase Peptide-Synthesis Using P-Alkoxybenzyl Ester Resins and Fmoc-Amino Acids.* Tetrahedron Letters, 1986. **27**(6): p. 743-746.

6. Stephenson, R.C. and S. Clarke, *Succinimide Formation from Aspartyl and Asparaginyl Peptides as a Model for the Spontaneous Degradation of Proteins.* Journal of Biological Chemistry, 1989. **264**(11): p. 6164-6170.

7. Sreedhara, A., et al., *Characterization of the Isomerization Products of Aspartate Residues at Two Different Sites in a Monoclonal Antibody.* Pharmaceutical Research, 2012. **29**(1): p. 187-197.

8. Mergler, M., et al., *The aspartimide problem in Fmoc-based SPPS. Part I.* Journal of Peptide Science, 2003. **9**(1): p. 36-46.

9. Nicolas, E., E. Pedroso, and E. Giralt, *Formation of Aspartimide Peptides in Asp-Gly Sequences.* Tetrahedron Letters, 1989. **30**(4): p. 497-500.

10. Perkins, M., et al., *Determination of the origin of charge heterogeneity in a murine monoclonal antibody.* Pharmaceutical Research, 2000. **17**(9): p. 1110-1117.

11. Toney, K., et al., *Aspartimide formation in the joining peptide sequence of porcine and mouse pro-opiomelanocortin.* J Biol Chem, 1993. **268**(2): p. 1024-31.

12. Orpiszewski, J., et al., *Protein aging hypothesis of Alzheimer disease.* FASEB J, 2000. **14**(9): p. 1255-63.

13. Reissner, K.J. and D.W. Aswad, *Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals?* Cell Mol Life Sci, 2003. **60**(7): p. 1281-95.

14. Clarke, S., *Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins.* Int J Pept Protein Res, 1987. **30**(6): p. 808-21.

15. Noguchi, S., K. Miyawaki, and Y. Satow, *Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose.* Journal of Molecular Biology, 1998. **278**(1): p. 231-238.

16. Yang, H., et al., *Alzheimer's disease and mild cognitive impairment are associated with elevated levels of isoaspartyl residues in blood plasma proteins.* J Alzheimers Dis, 2011. **27**(1): p. 113-8.

17. Oliyai, C. and R.T. Borchardt, *Chemical Pathways of Peptide Degradation .6. Effect of the Primary Sequence on the Pathways of Degradation of Aspartyl Residues in Model Hexapeptides.* Pharmaceutical Research, 1994. **11**(5): p. 751-758.

18. Harris, R.J., et al., *Identification of multiple sources of charge heterogeneity in a recombinant antibody.* J Chromatogr B Biomed Sci Appl, 2001. **752**(2): p. 233-45.

19. Chu, G.C., et al., *Accumulation of succinimide in a recombinant monoclonal antibody in mildly acidic buffers under elevated temperatures.* Pharm Res, 2007. **24**(6): p. 1145-56.

20. Mergler, M., et al., *The aspartimide problem in Fmoc-based SPPS. Part II.* Journal of Peptide Science, 2003. **9**(8): p. 518-526.

21. Lauer, J., C. Fields, and G. Fields, *Sequence dependence of aspartimide formation during 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis.* Letters in Peptide Science, 1995. **1**(4): p. 197-205.

22. Guttler, B.H.O., et al., *A quantitative analysis of spontaneous isoaspartate formation from N-terminal asparaginyl and aspartyl residues.* Amino Acids, 2013. **44**(4): p. 1205-1214.

23. Johnson, B.A., et al., *Formation of isoaspartate at two distinct sites during in vitro aging of human growth hormone.* Journal of Biological Chemistry, 1989. **264**(24): p. 14262-14271.

24. Zhang, W. and M.J. Czupryn, *Analysis of isoaspartate in a recombinant monoclonal antibody and its charge isoforms.* Journal of Pharmaceutical and Biomedical Analysis, 2003. **30**(5): p. 1479-1490.

25. O'Connor, P.B., et al., *Differentiation of aspartic and isoaspartic acids using electron transfer dissociation.* Journal of the American Society for Mass Spectrometry, 2006. **17**(1): p. 15-19.

26. Sobott, F., et al., *Comparison of CID Versus ETD Based MS/MS Fragmentation for the Analysis of Protein Ubiquitination.* Journal of the American Society for Mass Spectrometry, 2009. **20**(9): p. 1652-1659.

27. Hamidane, H.B., et al., *Electron Capture and Transfer Dissociation: Peptide Structure Analysis at Different Ion Internal Energy Levels.* Journal of the American Society for Mass Spectrometry, 2009. **20**(4): p. 567-575.

28. Chalkley, R.J., et al., *Statistical Analysis of Peptide Electron Transfer Dissociation Fragmentation Mass Spectrometry.* Analytical Chemistry, 2010. **82**(2): p. 579-584.

29. Kim, M.S. and A. Pandey, *Electron transfer dissociation mass spectrometry in proteomics.* Proteomics, 2012. **12**(4-5): p. 530-542.

30. Mergler, M. and F. Dick, *The aspartimide problem in Fmoc-based SPPS. Part III.* Journal of Peptide Science, 2005. **11**(10): p. 650-657.

31. Michels, T., et al., *Acid-mediated prevention of aspartimide formation in solid phase peptide synthesis.* Org Lett, 2012. **14**(20): p. 5218-21.

32. Isidro-Llobet, A., et al., *p-Nitrobenzyloxycarbonyl (pNZ) as a temporary N-alpha-protecting group in orthogonal solid-phase peptide synthesis - Avoiding diketopiperazine and aspartimide formation.* European Journal of Organic Chemistry, 2005(14): p. 3031-3039.

33. Haines-Butterick, L., et al., *Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells.* Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(19): p. 7791-7796.

34. Rajagopal, K., et al., *Tuning the pH Responsiveness of beta-Hairpin Peptide Folding, Self-Assembly, and Hydrogel Material Formation.* Biomacromolecules, 2009. **10**(9): p. 2619-2625.