2014

Design of Recombinant Tenebrio Molitor Antifreeze Protein for Purification Using Elastin-Like Polypeptide Tag

Joongseok Oh
Cleveland State University

DESIGN OF RECOMBINANT *TENEBRIO MOLITOR* ANTIFREEZE PROTEIN FOR PURIFICATION USING ELASTIN-LIKE POLYPEPTIDE TAG

JOONGSEOK OH

Bachelor of Science in Biological Sciences

Southern Illinois University at Carbondale

May 2005

submitted in partial fulfillment of requirements for the degree

MASTER OF SCIENCE IN BIOMEDICAL ENGINEERING

at the

CLEVELAND STATE UNIVERSITY

May 2014

We hereby approve this thesis for

JOONGSEOK OH

Candidate for the Master of Science in Biomedical Engineering degree for the

Department of Chemical and Biomedical Engineering

and the CLEVELAND STATE UNIVERSITY

College of Graduate Studies

Thesis Chairperson, Nolan B. Holland, Ph. D

Chemical & Biomedical Engineering_____
Department & Date

Thesis Committee Member, Joanne M. Belovich, Ph. D

Chemical & Biomedical Engineering_____
Department & Date

Thesis Committee Member, Xue-Long Sun, Ph. D

Chemistry_____
Department & Date

Student's Date of Defense: May 01, 2014

ACKNOWLEDGEMENT

First of all, I would like to sincerely thank my academic advisor, Dr. Nolan B. Holland, for providing me with the research opportunity in his laboratory, supporting me, sharing his expertise, and teaching me valuable lessons throughout my thesis research. His excellent guidance, patience, encouragement are deeply appreciated.

I would also like to thank my committee members Dr. Joanne M. Belovich and Dr. Xue-Long Sun for their invaluable support and advice.

I would to thank all my colleagues, Sen Bu, James Cole, Hakan Celik, and all the other students in Dr. Holland's research group for their friendship and invaluable insights.

I would like to give my special thanks to Ms. Becky Laird and Ms. Darlene Montgomery for their motivation and support at Cleveland State University.

Finally, I would like thank my wife, Jikyoung Jeong, for her trust and dedication. I would like to thank my family for their great support and endless love. Thanks to all my friends for their friendship and encouragement.

DESIGN OF RECOMBINANT *TENEBRIO MOLITOR* ANTIFREEZE PROTEIN FOR PURIFICATION USING ELASTIN-LIKE POLYPEPTIDE TAG

JOONGSEOK OH

ABSTRACT

Fusion protein technologies can aid to improve solubility of recombinant protein from microorganisms and help recombinant protein purification. Elastin-like polypeptides (ELP) as a fusion tag can be utilized to facilitate the purification of recombinant proteins because ELP can provide its thermally responsive behavior to ELP tagged proteins. An ELP tag can be used as the purification carrier through inverse transition cycling (ITC), which is a simple and non chromatographic separation process. The purification through ITC can reduce cost and can be quickly performed compared to other purification methods. However, we further considered ELP tag removal because it may possibly hinder the structure and function of fusion partners (or target proteins).

Tenebrio molitor antifreeze protein (TmAFP) is an insect AFP and it is classified as a hyperactive antifreeze protein because its activity is 10 to 100 times greater than AFPs from other sources. TmAFP as a target protein has been purified by cold finger purification or chromatographic separation but there are drawbacks.

The overall goal of this work is to design and implement a strategy for high yield of recombinant TmAFP in *E. coli* through purification using ELP tags. This includes ELP tagged TmAFP with a TEV cleavage site. DNA encoding this tag was designed and

successfully added to the N-terminus of the TmAFP gene. Then the ELP tags were added to the N-terminus of TmAFP and ELP tag DNA was successfully added to N-terminus of TEV protease. Of note, ELP[$(GVGVP)_{40}$] tagged DNA was successfully added to the N-terminus TEV protease but ELP[$(GVGVP)_{20}$] tagged DNA was not successfully added. Then the ELP tagged fusion proteins were successfully expressed. ELP tagged TmAFPs were successfully purified by using ITC while ELP tagged TEV protease was not successfully purified by ITC.

TABLE OF CONTENTS

Page No

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Recombinant proteins are finding many commercial applications as enzymes, bioactive compounds, and therapeutics and their relatively simple synthesis can possibly lead to reduced costs for specialty chemicals compared with multistep chemical synthesis methods [1] even though these proteins are produced in cells (often in yeast or bacteria) and must be purified from cell lysates. Important considerations for successful and efficient production of a recombinant protein are its solubility [2] and the existence of efficient purification methods that allow easy scale up of recombinant protein production [3].

It is often desired for biotechnogical purposes (or use) that recombinant proteins be soluble when proteins are expressed in bacterial cells [2]. However, recombinant proteins are not always soluble and often form inclusion bodies, which are insoluble accumulations, secondary to improper folding during protein expression [2].

Isolated inclusion bodies can be denatured and solubilized by using strong denaturants such as urea or guanidine hydrochloride. Then by using methods such as dilution, dialysis, diafiltration, gel filtration, or immobilization onto a solid support, the denaturant can be removed or decreased in concentration, allowing the protein to be refolded (renatured) [79]. According to Maxwell et al, refolding occurred with of 55% of

70 insoluble proteins by using denaturing/refolding protein purification [4]. However, denaturing inclusion bodies and refolding proteins may not be desired because it might result in improper folding, decreased yield, and decreased bioactivity [5]. Also, the denaturing and refolding processes can be expensive and time-consuming [5].

According to Schein, most proteins, such as (human interferon-γ and –α) are active and soluble with protein expression at ≤ 30 ºC in *E. coli* compared with 37 ºC because inclusion bodies were formed less from protein expression at low temperatures [6]. In fact, high soluble protein yield has been reported from protein expression at 4 ºC in *E. coli* [7]. However, lowering the expression temperature does not always improve solubility of proteins [8].

Fusion protein technologies have become more viable for recombinant protein expression and purification because fusion proteins can improve solubility of recombinant protein production from microorganisms and help recombinant protein purification [5 and 9]. For example, maltose binding protein (MBP) is 42 kDa, and is well-known to improve the solubility of recombinant proteins, and can be purified by using cross-linked amylase resins, which are not expensive [5]. HisHisHisHisHisHis (His_6 tag) is common for purifying recombinant proteins because His_6 tag can bind to Ni(II)-nitrilo triacetic acid (Ni-NTA), which can use immobilized metal affinity tag [9]. Multiple fusion techniques, such as the coexpression of MBP and His_6 systems, have been introduced to improve the solubility of protein expression and in assistance of protein purification [10 and 11]. ELP tags are another approach to improve solubility [12] and facilitate efficient protein purification [13] and will be discussed in detail later.

Fusion order can affect results of protein expression for some constructs [13] because it may affect the rate of degradation [9]. N-terminal fusion tags on target proteins are known for enhancing the yield and solubility of recombinant proteins [9]. Dyson et al found that MBP and Thierodoxin (Trx) showed a great improvement of soluble protein production from protein expression when both were placed at either the N-terminal or C-terminal via fusion [13]. However, MBP alone only enhanced soluble proteins for C-terminal fusion [13].

Even though fusion tags might improve solubility of recombinant proteins and assist in recombinant protein purification, fusion tag removal is an important consideration because fusion tags may possibly hinder the structure and function of fusion partners (or target proteins) [14]. To remove fusion tags, proteolytic enzymatic approaches are broadly used due to their specificity [13]. The cleavage site (or linker) is placed in between the fusion tag and fusion partner (or target proteins) [5]. TEV protease is often used as a removal tag because TEV provides high site specificity when compared to other proteases, such as factor Xa, enteropetidase, and thrombin [15].

1.1 ELP as a purification tag

Elastin has self aggregation behavior above a certain temperature, which has resulted in the development of ELP (elastin-like polypeptide) as an artificial recombinant polypeptides [16] designed and synthesized using protein engineering techniques [16 and 17]. The repeat unit of ELPs is (Gly-Xaa-Gly-Yaa-Pro) or (Yaa-Pro-Gly-Xaa-Gly); Xaa can be any amino acid except Proline [16] and Yaa can be any amino acid except Alanine [18]. Poly (GVGAP) has shown granular precipitate behavior at an increased temperature [18] that may ruin its elasticity and this granular precipitate behavior may be a sign of irreversible aggregation [19].

ELPs exhibit a phase change at a certain temperature, and this behavior of ELP is reversible: ELPs are soluble in aqueous solution below their transition temperature (T_t), and insoluble above the T_t as evident from solutions becoming turbid [16]. According to Urry, $(GVGVP)_{251}$ was observed to exhibit phase change behavior at a certain temperature: $(GVGVP)_{251}$ aggregated, as evident by its turbidity in aqueous solution above 25 ^oC (Figure 1.1 (A) and (B)) [18].

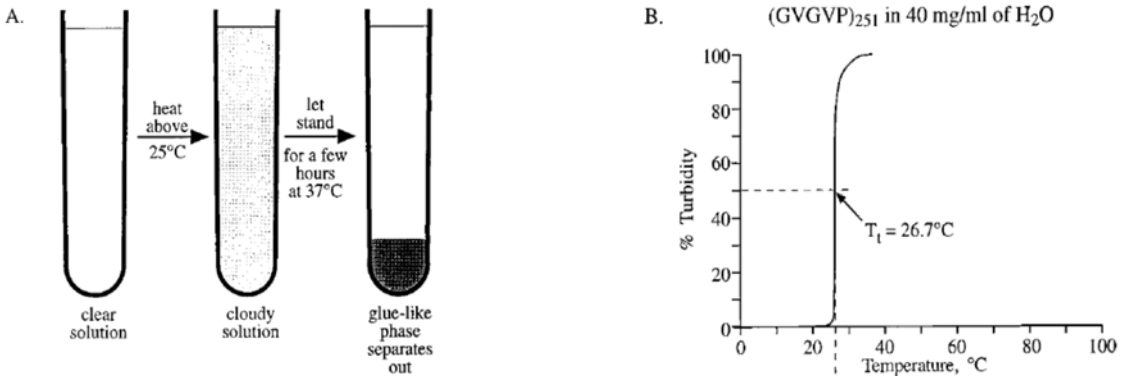


Figure 1.1. (A) Behavior of $(GVGVP)_{251}$, Phase separation by Heat. (B) Temperature profile, Turbidity vs. Temperature [18]. Reprinted with permission from (Urry, D. W. (1997). Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. *The Journal of Physical Chemistry B*, *101*, 11007-11028.). Copyright (2014) American Chemical Society.

This turbidity at temperatures above T_t is caused by aggregation of ELP molecules due to interacting of hydrophobic regions of ELPs [16]. Thermally induced conformation change of ELP may play a major role in this phase transition when above the transition temperature [20].

1.1.1 The aspects affecting T_t

The T_t is influenced by concentration, chain length [20 and 21], composition [20], salt concentration [22, 23, and 24], and pH [25 and 26]. Chilkoti and Meyer systemically showed the transition temperature changes due to concentration and chain length: Transition temperatures decreased as ELP concentration and chain length increase (Figure 1.2 (A)) [20].

Also, they mentioned that ELP with more hydrophobic constituents showed a lower transition temperature profile; ELP $[V_1A_8G_7]$ has a high T_t, ELP $[V_5A_2G_3]$ has a middle T_t, and ELP $[V_5]$ has a low T_t because ELP [V5] is the most hydrophobic amongst

them (Figure 1.2 (B)) [20]. ELP [$V_xA_yG_z$-n]: V is Valine, A is Alanine, and G is Glycine. x, y, z are ratio x:y:z of V:A:G. n is the number of pentapetides.

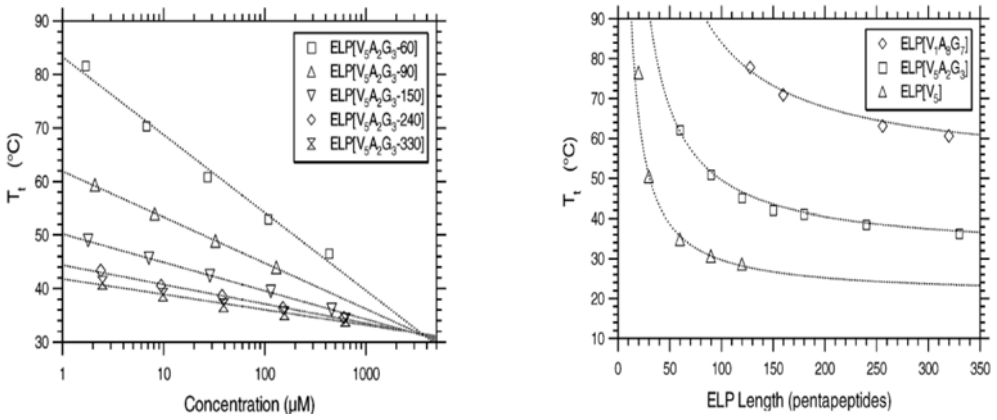


Figure 1.2. (A) T_t vs. Concentration [20]. **(B) T_t vs. ELP chain length** [20]. Reprinted with permission from (Meyer, D. E. & Chilkoti, A. (2004). Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules*, *5*, 846-851.). Copyright (2014) American Chemical Society.

Salt concentration can further alter transition temperatures [22, 23, and 24]. Tirrell et al described the transition temperature of $(VPGIG)_{40}$ with various salts and their concentrations [22]. The phase transition temperature of $(VPGIG)_{40}$ was decreased by the addition of salts and this change is caused by anions (Cl^-, SO_4^{2-}, or I^-) [22]. Reguera et al discovered the effect of NaCl concentration with a $(GVGVP)_{251}$ model polymer on transition temperature. Transition temperatures decreased as NaCl concentration increased (Figure 1.3 (A)) [23].

(A) (B)

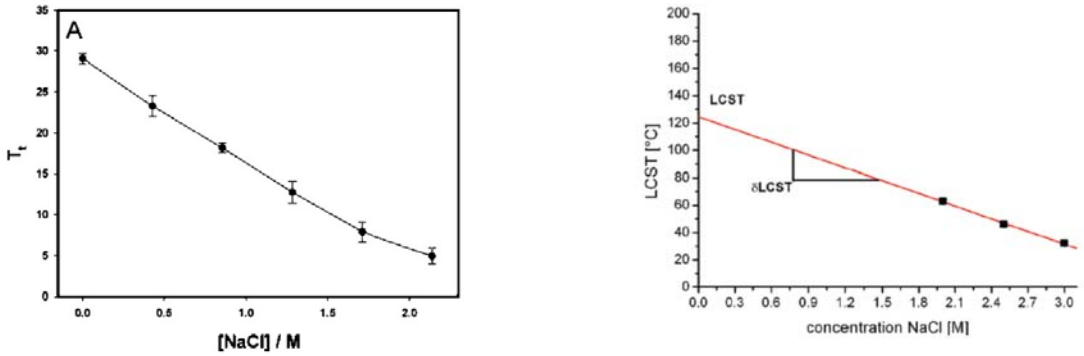


Figure 1.3. (A) T_t of $(GVGIP)_{40}$ vs. NaCl Concentration [23] **and (B) LCST of $(GVGVP)_{251}$ (^oC) vs. NaCl Concentration** [24]. Reprinted with permission from (Reguera, J., Urry, D. W., Parker, T. M., McPherson, D. T., & Rodriguez-Cabello, J. C. (2007). Effect of NaCI on the exothermic and endothermic components of the inverse temperature transition of a model elastin-like polymer. *Biomacromolecules*, *8*, 354-358.). Copyright (2014). American Chemical Society. [23]. Reprinted with permission from (Nuhn, H. & Klok, H. (2008). Secondary structure formation and LCST behavior of short elastin-like peptides. *Biomacromolecules*, *9*, 2755-2763.). Copyright (2014) American Chemical Society. [24].

Another research group examined LCST (lower critical solution temperature) behavior of short ELPs $[(GVGVP)_5]$ by NaCl concentration and LCST was decreased linearly by increased NaCl concentration (Figure 1.3 (B)) [24].

The solution pH can affect transition temperature of ELP [25 and 26]. Rodriguez-Cabello's lab group demonstrated the pH effect on five different pH-responsive ELPs $[(PGVGV)_2\text{-}(PGEGV)\text{-}(PGVGV)_2]n$, n=5, 9, 15, 30, 45 in the pH range of 2.0 to 5.5-6.0, and that T_t increased as pH increased [26]. Chilkoti's lab group examined the effect of pH with charged ELP showing that multiple glutamic acid residues can form acidic ELPs and histidine residues can form basic ELPs [25].

(A) (B)

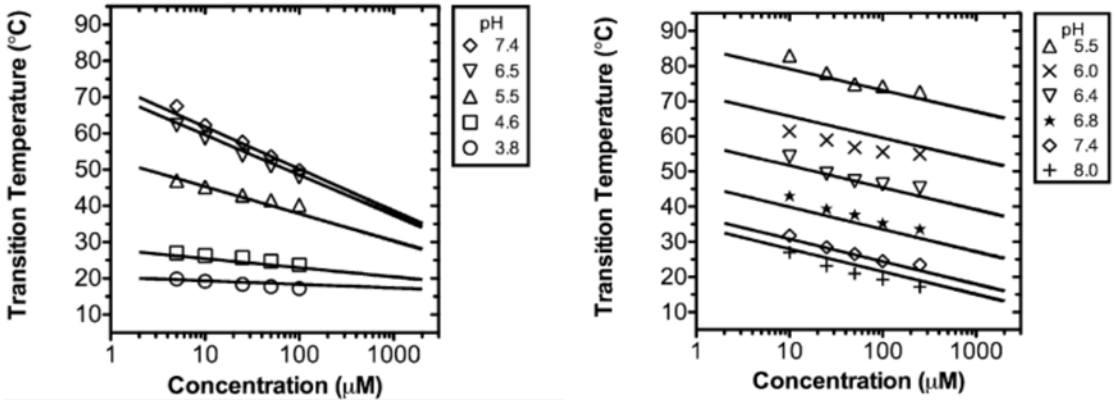


Figure 1.4. T_t vs. Concentration of (A) Acidic ELP pKa=5.29 and (B) Basic ELP pKa=6.22 [25]. (A) Ratio V:I:E=1:3:1, 160 pentapeptides (B) Ratio V:H:G:A=1:2:1:1, 120 pentapetides where V: Valine, I: Isoleucine, E: Glutamic acid, H: Histidine, A: Alanine, and G: Glycine. Reprinted with permission from (Mackay, J. A., Callahan, D. J., FitzGerald, K. N., & Chilkoti, A. (2010). Quantitative model of the phase behavior of recombinant pH-responsive elastin-like polypeptides. *Biomacromolecules*, *11*, 2873-2879.). Copyright (2014) American Chemical Society.

Transition temperature of acidic ELPs increased with increased pH (Figure 1.4 (A)) while transition temperature of basic ELPs increased with decreased pH (Figure 1.4 (B)). Both acid and basic ELPs behaved similarity and transition temperature was decreased by the increase of ELP concentration [25].

1.2 Recombinant protein purification by using ELP tag

1.2.1 Inverse Thermal Cycling

ELPs are thermally sensitive proteins, and ELP gene can be fused genetically to target protein DNA. When ELP and the target protein are expressed simultaneously, ELPs as fusion tags can facilitate recombinant protein purification. ELP retains its thermally response behavior as a fusion protein. Inverse thermal cycling (ITC) is a

method where ELP fusion proteins can be purified by the phase change by cycling above or below their T_t [52].

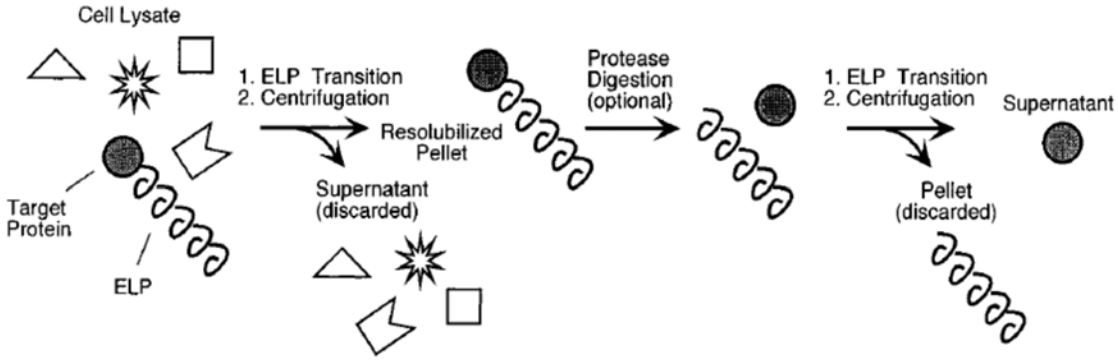


Figure 1.5. Scheme of ITC purification [52]. (Meyer, D. E., Trabbic-Carlson, K., & Chilkoti, A. (2001), *Biotechnology Progress*, Protein purification by fusion with an environmentally responsive elastin-like polypeptide: Effect of polypeptide length on the purification of thioredoxin, *17*, 720-728.).

ITC is a non-chromatographic separation that is relatively simple [52] because the substances, which are in different phases of ELP fusion protein depending on the temperature being above or below its T_t, can be separated by changing the temperature above and below its T_t and selecting out the appropriate phase [53]. ELP fusion protein phase separates when increasing temperature above its T_t, but ELP fusion protein is soluble at temperatures below its T_t [52]. The aggregated ELP fusion protein can be resolubilized below its T_t [52]. Ideally, the ELP tag does not alter the stability and function of the target protein during ITC [3, 52, and 54]. The ELP tag can be removed from the target protein by a proteolytic cleavage sequence between the ELP tag and target protein by using a site-specific protease [52].

1.2.2 Advantages of ITC

ITC can provide several advantages: ITC does not need expensive chromatographic resins and equipment to purify recombinant protein and therefor can reduce the cost to scale-up from a laboratory to an industry scale. ITC needs small temperature or ionic strength change, is fast and easily performed with several short centrifugation [3]. Moreover, ITC is simple and can obtain more than 95% purity within a few rounds of ITC. This in comparison to chromatography which might be time consuming and may exhibit a limited yield due to the resin's loading capacity [12]. According to Trabbic-Carlson et al, target protein yields from both ITC and IMAC (Immobilized Metal Affinity binding Chromatography) by using His tag are similar, but IMAC is much less efficient in protein purification. Also, His fusion proteins were less stable then ELP fusion protein [55].

1.2.3 Considerations for using ELP tag

To use ELP fusion protein for recombinant protein purification, the primary considerations are the T_t [3, 20-26, 52, 56, and 57] and fusion order [12]. The T_t of ELP fusion protein is desired within a few degrees (^oC) higher than protein expression temperature: ELP fusion protein remains in a soluble state during protein expression, and protein purification is possible with only a small increase of temperature [3]. As mentioned earlier, T_t of ELP can be altered by its length and concentration [20 and 21], composition [20], salt concentration [22, 23, and 24], and pH [25 and 26], and those aspects can also affect T_t of an ELP fusion protein [3].

The large molecular weight ELP tag fusion protein is shown to lower T_t when compared to that of small molecular weight ELP tag, but small molecular ELP tag is known to increase fusion protein yield compared to large molecular ELP tag [3 and 52] even though the length (or molecular weight) of ELP might be not related to the efficiency of recombinant protein purification [3]. If the T_t of a small molecular weight ELP tag fusion protein is too high, the T_t can be lowered by adding NaCl [3 and 52]. However, the addition of NaCl does not modify ELP fusion protein behavior [52].

The surface hydrophobicity of a target protein might shift the T_t of free ELP because ELP might interact with the hydrophobic surface of a target protein resulting in a lower T_t [56]. The higher hydrophocity of proteins tend to resulting in lower T_t, thus imparting a ΔT_t effect [56].

Chilkoti and Meyer reported that the ELP fusion protein exhibits a T_t shift when compared with free ELP but the T_t shift of thioredoxin-ELP fusion protein is much smaller versus that of tendamistat and ELP. It might imply that the interaction between ELP tag and thioredoxin was not significant compared to tendamistat and ELP. The hydrophobic region of tendamistat and ELP tag might interact more thus resulting in a greater T_t depression than thioredoxin-ELP fusion protein (Figure 1.6) [3].

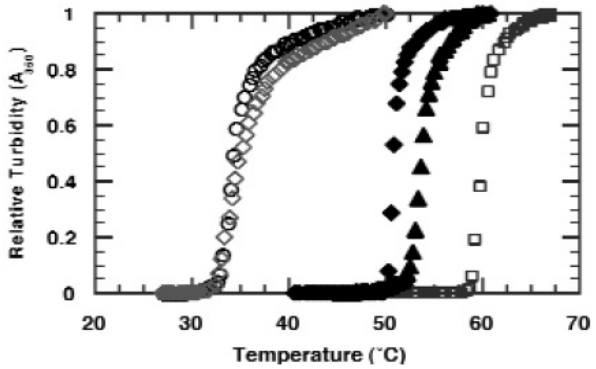


Figure 1.6. Turbidity profile, where (♦) free ELP; (▲) thioredoxin-ELP; (○) thioredoxin-ELP-tendamistat; (◊) ELP-tendamistat; and (□) thioredoxin-ELP [3]. Reprinted by permission from Macmillan Publishers Ltd: *Nature Biotechnology*, Meyer, D. E. & Chilkoti, A. (1999). Purification of recombinant proteins by fusion with thermally-responsive polypeptide. *17*, 1112-1115, Copyright (2014).

Lim et al stated that cationic ELP could interact with negatively charged target proteins and described this interaction as 'greater salt sensitivity' of cationic ELP. Adding a small amount of NaCl to small cationic ELP resulted in decreased phase transition temperature [57]. Also, they suggested that anionic ELP with cationic target protein might show similar behavior [57].

Fusion order is where a fusion tag is placed on either the N-terminal or C-terminal of a target protein. Fusion order has been shown to affect the solubility, expression level, or activity of a fusion protein in some cases [12]. According to Christensen et al, most protein-ELP showed higher expression level (with increased yield), and better specific activity than ELP-protein: Protein-ELP is a C-terminal fusion of ELP, and ELP-protein is an N-terminal fusion of ELP to protein (Figure 1.7) [12]. Even though ELP-protein fusion (or N-terminal fusion to protein) showed lower level of soluble protein expression, this N-terminal fusion has been reported to have little inclusion body formation. This is

due to an ELP tag in N-terminal fusion might help mis-folded fusion protein achieve a soluble state [12]. Mis-folded proteins often are aggregated and become insoluble [12].

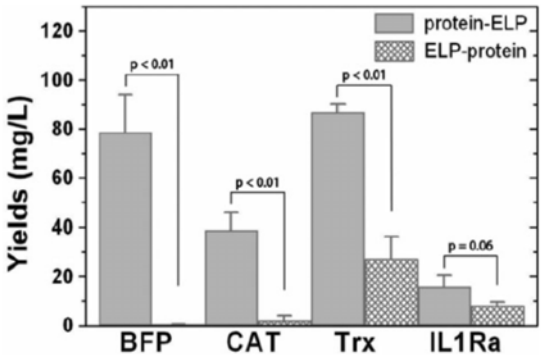


Figure 1.7. Fusion protein yield depending on fusion order; BFP is blue fluorescent protein, CAT is chloramphenicol acetyltransferase, Trx is thioredoxin, and IL1Ra is interleukin-1 receptor antagonist [12]. (Christensen, T., Amiram, M., Dagher, S., Trabbc-Carlson, K., Shamji, M. F., Setton, L. A., and Chilkoti, A. (2009), *Protein Science*, Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins, *18*, 1377-1387.).

1.3 TmAFP

1.3.1 Antifreeze protein

Many living creatures survive in extreme environments by adaptative behaviors. Subzero temperatures is a harsh environment but many organisms have adapted to survive in a subzero habitat [27]. Ice formation within cells can cause the interruption of structural organization resulting in cell death [28]. Some organisms, which are freeze tolerant, have adapted to a subzero environment by synthesizing special substances that prevent ice formation within their bodies, such as antifreeze protein (AFP) otherwise known as thermal hysteresis proteins [27].

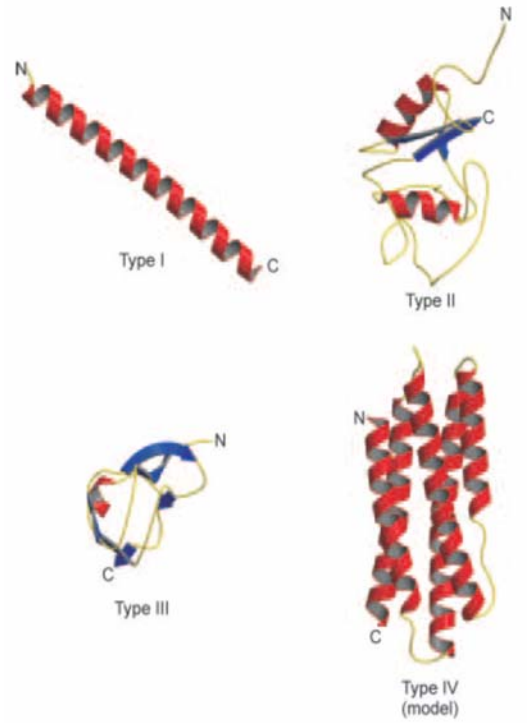
AFPs are capable of binding to ice crystals which lead to a lower melting point of ice [29]. AFPs are capable of inhibiting growth and recrystallization of ice crystal within an organisms body in the subzero temperature because AFPs have ice-binding domains [29]. In this manner, the interaction between AFPs and ice crystal can cause a non-equilibrium freezing-point depression, which is the difference between the melting point and freezing point of ice. This difference is called thermal hysteresis and can be used for measuring activity of AFP [30]. AFPs have different behavior compared with other chemical antifreeze materials because AFPs exhibit non-colligative freezing point depression due to the interaction between AFPs and ice crystals [30].

Antifreeze proteins are divided into two major classes, antifreeze proteins and antifreeze glycoproteins (AFGP) [27]. Antifreeze proteins have been found in several different species: fish [31-38], insects [37, 39-42, 50, and 51], plants [42 and 44], and other organisms [47-49].

Fish AFPs are classified as Type I [31 and 32], II [33 and 34], III [35], IV [36], and AFGP [37 and 38]. Type I APFs were found in plasma and skin tissues of winter flounder and the sculpins which live in polar oceans [32].Type I AFPs are alanine rich and are mostly single α-helical elongated proteins [32], and have a molar mass of 3.3 – 4.5 kDa [31]. Type II AFPs are found in sea raven (*Hemitripterus americanus*), herring (*Clupea hargengus harengus*), smelt (*Osmerus merdox*) [33], and long snout poacher (*Brachyopsis rostratus*) [34]. Type II AFPs are globular shaped, have a molar mass of 14-24 kDa, and are cysteine rich proteins [34]. Type II AFPs are categorized by the Ca^{2+} dependence of their antifreeze activity; herring and smelt AFPs are Ca^{2+} dependent, and sea raven and long snout poacher AFPs are Ca^{2+} independent [34]. Type III AFPs were first found in the ocean pout (*Macrozoarces americanus*), and are globular shaped proteins with a 7 kDa molecular weight [35]. Type IV AFPs are found in longhorn sculpin (LHS) and shorthorn sculpin (SHS); type IV AFPs from LHS are 12.3 kDa, Glutamine rich, with an α-helix structure, and type IV AFPs from SHS are 12.2 kDa, and shown to have 87% similarity to that of LHS [36]. Also, LSH and SHS can both produce type I and type IV AFPs which might enable them respond quickly to temperature changes as an evolutionary adaption [36].

AFGPs are found in the blood serum of Antartic notothenoids and Artic cod [37]. AFGP have a molar mass of 2.6 kDa to 33.7 kDa, and have an $(Ala\text{-}Ala\text{-}Thr)_n$ repeat unit [37]. Arctic psychrophilic yeast also contain AFGP that are 26 kDa in size [38].

(A) (B)



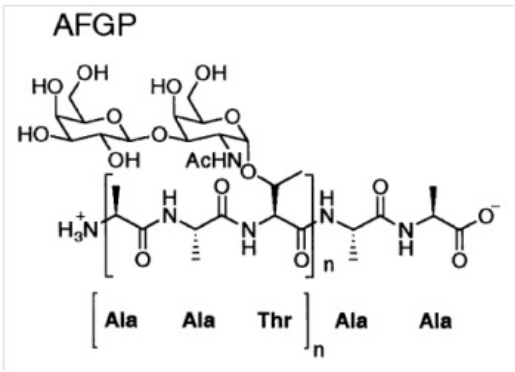


Figure 1.8. (A) Fish AFP structures and model [40] **and (B) Antifreeze glycoprotein structure** [37]. (Graether, S. P. & Sykes, B. D. (2004), *European Journal of Biochemistry*, Cold survival in freeze-intolerant insects: The structure and function of β-helical antifreeze proteins, *271*, 3285-3296.) [40]. (Harding, M. M., Anderberg, P. I., & Haymet, A. D. J. (2003), *European Journal of Biochemistry*, 'Antifreeze' glycoproteins from polar fish, *270*, 1381-1392.) [37].

Two major different insect antifreeze proteins have been vigorously studied, the spruce budworm (*Choristoneura fumiferama*)-sbwAFP or CfAFP and yellow mealworm (*Tenebrio moliter*) – TmAFP [39 and 40]. SbwAFP is a left handed, β-helix protein with 15 residues per coil, 4 disulfide bonds, TXT (Thr-X-Thr, X can be any amino acid) repeat unit, and are a triangular prism shape [40]. TmAFP is a right handed, β-helix protein with 12 residues per coil, 8 disulfide bonds, less than 9 kDa, contain a TCTxSxxCxxAx (x can be any amino acid) repeat unit [39], with a flattened cylinder shape, and is a more regular structure than sbwAFP [40].

(A) (B)

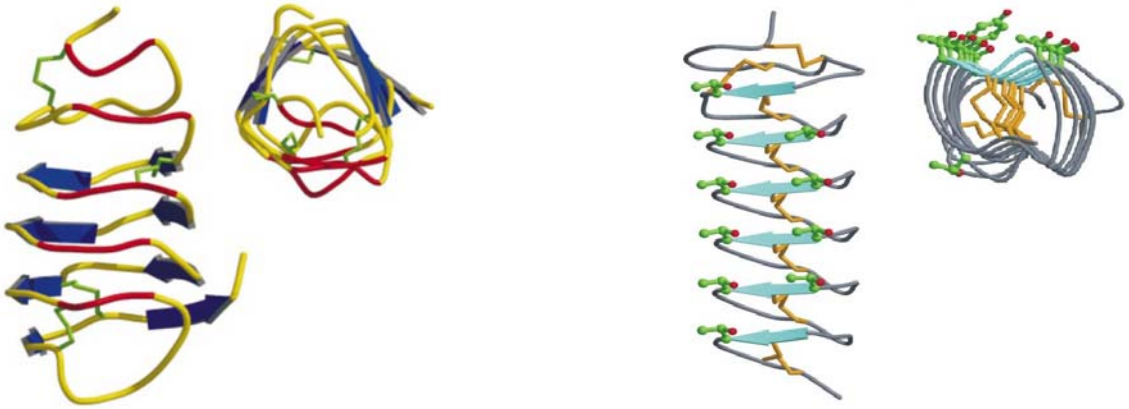


Figure 1.9. (A) Structure of sbwAFP [41] **and (B) Structure of TmAFP** [42]. Reprinted by permission from Macmillan Publishers Ltd: *Nature*, Graether, S. P., Kulper, M., Gagné, S. M., Walker, V. K., Jia, Z., Sykes, B. D., and Davies, P. L. (2000). β-Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect, *406*, 325-328., copyright (2014) [41]. (Marshall, C. B., Daley, M. E., Graham, L. A., Sykes, B. D., and Davies, P. L. (2002), *FEBS Letters*, Identification of the ice-binding face of antifreeze protein from *Tenebrio molitor, 529*, 261-267.) [42].

The TmAFP and sbwAFP are classified as hyperactive antifreeze proteins because the activities of TmAFP and sbwAFP are 10 to 100 times greater than of fish antifreeze protein [50], which is attributed to the fact that they can bind to both the basal and prism plane of ice crystals [40 and 51].

1.4 TEV protease as a cutter

1.4.1 TEV protease

TEV protease is commonly used for tag removal [58]. Dougherty and Parks mentioned that the 27 kDa of TEV protease can be expressed as a recombinant TEV protease protein in *E. coli* [59].

The optimal temperature of TEV protease activity is 30 ºC [15 and 60], and TEV protease activity is decreased above 34 ºC because of possible unfolding of the protease [15]. TEV protease activity is decreased with increasing NaCl concentration [15].

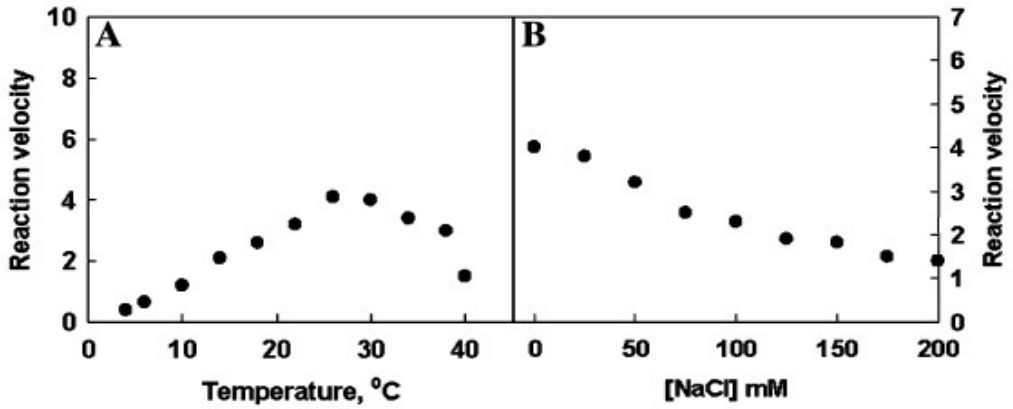


Figure 1.10. (A) Optimal Temperature, 30 ºC is the optimal temperature of TEV protease activity **(B) Optimal Salt Concentration**, TEV protease activity decreased by addition of NaCl. The figures were modified from the original pictures [15]. Reprinted from *Protein Expression and Purification*, *38*, Nallamsetty, S., Kapust, R. B., Tözsér, J., Cherry, S., Tropea, J. E., Copeland, T. D., and Waugh, D. S. (2004), Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. 108-115, Copyright (2014), with permission from Elsevier.

TEV protease is becoming more common for tag removal by cleavage [58] because TEV protease has several advantages [61, 62, and 63]. TEV protease can be expressed in *E. coli* which can increase its yield and improve its solubility. The increase in high yield of TEV protease by modifications of its expression condition may result in

cheaper production costs [61]. Phan et al reported that the structure of TEV protease leads to highly restricted specificity for the TEV recognition sequence [62]. The crystal structure of TEV protease is comprised of two-domain antiparellel β-barrel fold (which is found on trypsin-like serine protease) with His 46, Asp 81, and Cys 151 of catalytic triad residue placed in the contact face between two domains [62]. Also, the extended β-sheet interactions between TEV protease and its substrate backbone were observed [62]. TEV protease is active in various buffers [63] even though its activity was affected by NaCl [15] and some detergents [64]. Mohanty et al found that some detergents could reduce the activity of TEV protease [64]. However, Sun et al discovered that the cleavage activity of TEV protease showed high activity in four commonly used elution buffers and tolerance to various detergents and denaturants [63].

1.4.2 TEV protease recognition site

According to Carrington and Dougherty, the recognition site of TEV protease was known as EXXYXQS or EXXYXQG (Glu-Xaa-Xaa-Tyr-Xaa-Gln-Ser or Gly, where X or Xaa can be any amino acid), which is cleaved between Gln and Ser or Gly [65]. The position of residues were named Glu(P6)-Xaa(P5)-Xaa(P4)-Tyr(P3)-Xaa(P2)-Gln(P1)-Ser or Gly (P1') [65]. Subsequently, research by Tözsèr et al discovered the followings: First, Gln in P1 position is required. Second, Gly, Ser, Ala are small aliphatic resides which showed strong preference in P1' position. Third, Phe in P2 and Glu in P6 showed strong preference. Fourth, P5 residue is not a significant determinant. Fifth, Leu, Ile, and Val are preferred in the P4 position because they are all branched-chain aliphatic amino acids [66]. Also, Leu is very favorable in P4 position because of the hydrophobic

interaction between TEV protease and its substrate [66]. The most prevalent and used recognition sequence of TEV protease is known as ENLYFQS(Glu-Asn-Leu-Tyr-Phe-Gln-Ser) or ENLYFQG(Glu-Asn-Leu-Tyr-Phe-Gln-Gly) [67]. The study of P1' position showed that Ser and Gly residue in recognition sequence of TEV protease could be more catalytically efficient due to their paucity of side chain residues [67].

1.4.3 Issues of TEV protease expression

TEV protease has been reported to have two major issues in recombinant protein; autolysis [68, 69, and 70] and low solubility [61, 71, and 72].

Parks et al discovered that the recombinant fusion protein (~28 kDa) with a seven-His tag and TEV protease showed self-cleavage at a specific internal site, which was shown by an excised C-terminal 24 amino acid region of TEV protease amino acid sequences. This truncated protease was much less active than the full, untruncated TEV protease [68]. Since the C-terminal residues 235-242 of TEV protease are involved in the binding of TEV protease recognition site to its substrate, the activity of TEV protease would be reduced if the self-cleavage (or auto cleavage) takes places between its residue 218 and 219 [69]. Kapust et al discovered that some mutations of amino acid in certain residues of TEV protease resulted in reduced autolysis. The residue 219 S was replaced with D, E, V, and P (which were S219D, S219E, S219V, and S219P respectively), and the residue 217 F was replaced with K (which was F217K) [70]. S219V showed ~100 fold slower rate of truncated protease formation than that of wild -type protease, and S219V was more catalytically efficient than wild-type protease [70].

Another issue is that TEV protease expression results in low solubility of recombinant protein [71]. In silico design by using the PoPMuSic program, the mutations in both L56V and S135G (eg., V instead of L at 56, and G instead of S at 135) showed improvement of the solubility of TEV protease with mutations leading to enhanced surface polarity of TEV protease (Figure 1.11) [71] .

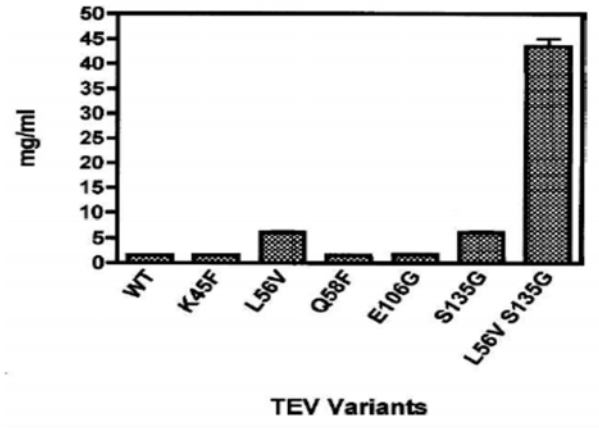


Figure 1.11. Solubility vs. TEV Variants [71]. (Cabrita, L. D., Gilis, D., Robertson, A. L., Dehouck, Y., Rooman, M., and Bottomley, S. P. (2007), *Protein Science*, Enhancing the stability and solubility of TEV protease using in silico design, *16*, 2360-2367.).

According to Fang et al, the TEV protease expression showed little improvement of soluble protein production at 28 ºC versus 37 ºC, but both co-expression of fusion protein (His-TEV protease) and the lowering expression temperature showed a drastic improvement of soluble fusion protein production at 28 ºC versus 37 ºC. It seemed that this co-expression and lower expression temperature work synergistically together [61]. In addition, MBP (maltose-binding protein), which is a fusion tag and known as solubility improvement of its fusion partner, is fused to TEV protease, and the fusion protein's solubility is enhanced even more compared to other solubility enhancer fusion tags because MBP may contribute to the appropriate folding of TEV protease [72].

1.5 The goal of this work

TmAFP as a target protein was modified previously by adding a trimer forming domain and was purified by cold finger purification [39 and 73]. Other researchers modified it using His_6 tag for purification by chromatographic separation [39]. However, the cold finger purification takes time to purify AFP, more than 8-16 hours [74] and the chromatographic separation is not easy to scale up [3].

The overall goal of this work was to design and implement a strategy for high yield of recombinant TmAFP in *E. coli* through purification using ELP tags. This strategy included TEV cleavage sites and ELP tags that were successfully added to the N-terminus of TmAFP. Then ELP tags were added to the N-terminus of TmAFP in pET20b and ELP tagged TmAFP fusion proteins which were successfully expressed in Origami host cells. The ELP tag DNA were successfully added to the N-terminus of TEV protease in pET20b and ELP tagged TEV protease fusion proteins were successfully expressed in BL21* host cells. ELP tagged TmAFPs were successfully purified by using ITC while ELP tagged TEV protease was not successfully purified by using ITC. ELP tags were functional through ITC purification and appeared to aid in target protein purification. This purification method may be applicable for scale-up mass production of recombinant TmAFP.

CHAPTER II

MATERIALS AND METHODS

2.1 Synthesis of plasmid DNAs

In order to accomplish the goal of this work, the genes of ELP tags and TEV protease cleavage sites were sequentially added using molecular biology techniques to the N-terminus of TmAFP gene resulting in ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP. In addition, ELP tags were added to N-terminus of TEV Protease resulting in ELP-TEV protease.

2.1.1 Design and synthesis of TEV protease cleavage site: TEV$_G$ (ENLYFQG) and TEV$_S$ (ENLYFQS)

In order to remove the ELP tag in the final protein, TEV protease cleavage sites were engineered in between ELP and TmAFP. The common cleavage sites of TEV protease were reported as amino acid sequences of ENLYFQG and ENLYFQS [67]. To synthesis DNA for TEV protease cleavage sites, a set of primers were designed to be compatible with restriction enzyme cut sites of NdeI, PflMI, and BglI, which are commonly used in our lab for DNA synthesis.

TEV$_G$ forward and TEV$_G$ reverse primers were designed and were ordered from Invitrogen.

TEV$_G$ forward primer (5’ to 3’)

TATGGGCCACGGCGTGGAAAACCTGTACTTCCAGGGGCCGGGC

TEV$_G$ reverse primer (5’ to 3’)

CGGCCCCTGGAAGTACAGGTTTTCCACGCCGTGGCCCA

TEV$_S$ forward and TEV S reverse primers were designed and were ordered from Invitrogen.

TEV$_S$ forward primer (5’ to 3’)

TATGGGCCACGGCGTGGAAAACCTGTACTTCCAGTCGCCGGGC

TEV$_S$ reverse primer (5’ to 3’)

CGGCGACTGGAAGTACAGGTTTTCCACGCCGTGGCCCA

The annealing process of each set of forward and reverse primer were performed to synthesize TEV$_G$ and TEV$_S$ double stranded DNA by heating to 95 ^oC and cooling gradually in a thermo cycler (Thermo Electron Corp.). The annealing process protocol is described in Appendix (A.1). The following double stranded DNA sequences can be obtained by the annealing process;

ATG is the start DNA code.

TEV$_G$ (TEV protease cleavage site, ENLYFQ/G)

```
 NdeI        PflMI      TEV protease cleavage site BglI(SfiI)
5′ ^TATGGGCCACGGC^GTG GAA AAC CTG TAC TTC CAG GGGCCGGGC^ 3′
3′    ACCCGGTG^CCGCAC CTT TTG GAC ATG AAG GTC CCCGGC^    5′
                      Glu Asn Leu Tyr Phe Gln Gly
                       E   N   L   Y   F   Q   G
```

TEVS (TEV protease cleavage site, ENLYFQ/S)

```
 NdeI         PflMI     TEV protease cleavage site     BglI
5′ ^TATGGGCCACGGC^GTG GAA AAC CTG TAC TTC CAG TCGCCGGGC^ 3′
3′     ACCCGGTG^CCGCAC CTT TTG GAC ATG AAG GTC AGCGGC^    5′
                       Glu Asn Leu Tyr Phe Gln Ser
                        E   N   L   Y   F   Q   S
```

2.1.2 Construction of ELP-TEVG-TmAFP and ELP-TEVS-TmAFP DNA

Genes encoding ELP[(GVGVP)$_{20}$]-TEVG-TmAFP, ELP[(GVGVP)$_{20}$]-TEVS-TmAFP, ELP[(GVGVP)$_{40}$]-TEVG-TmAFP, and ELP[(GVGVP)$_{40}$]-TEVS-TmAFP were prepared. The TmAFP-pET25b plasmid was obtained in our lab.

The DNA sequences for TmAFP (5’ to 3’), 252 bps:

CAGTGCACCGGTGGTGCTGACTGCACCAGCTGCACCGCTGCTTGCACCGGTTGC
GGTAACTGCCCGAACGCTGTTACTTGCACTAACAGCCAGCACTGCGTTAAAGCT
ACCACCTGCACCGGTAGCACCGACTGCAACACCGCTGTTACCTGCACTAACTCT
AAAGACTGCTTCGAAGCTCAGACCTGCACCGACTCTACCAACTGCTACAAAGCT
ACCGCTTGCACCAACTCTACCGGCTGCCCGGGTCAC

The amino acid sequences for TmAFP:

QCTGGADCTSCTAACTGCGNCPNAVTCTNSQHCVKATTCTGSTDCNTAVTCTNS
KDCFEAQTCTDSTNCYKATACTNSTGCPGH

PflMI, a restriction enzyme, was commonly used in our lab for DNA synthesis. Since PflMI cut site was discovered in the middle of pET25b expression vector, it can lead to an additional PflMI cut site which could lead to inability to use PflMI for further DNA synthesis. Therefore, pET25b was eventually replaced by pET20b for this research.

TmAFP-pET25b plasmid was double digested by the restriction enzymes NdeI and EcoRI that was used as an insert for ligation. The 40 µl of reaction mixture was added with 2 µl NdeI, 2 µl EcoRI (New England Biolabs), 15 µl DNA, 4 µl 10x buffer EcoRI (New England Biolabs), and 16 µl Nuclease free water. The reaction mixture was incubated in a 37 ^{o}C water bath for 3 hours. pET20b plasmid was obtained in our lab and also double digested with NdeI and EcoRI that was used as a vector. The 40 µl of reaction mixture was added with 2 µl NdeI, 2 µl EcoRI, 15 µl DNA, 4 µl 10x buffer EcoRI, and 16 µl Nuclease free water. The reaction mixture was incubated in a 37 ^{o}C water bath for 3 hours. The double digested DNA was mixed with 6x load dye (Promega), and then loaded on a 1.5 % agarose gel with TAE buffer at 80V for 40 minutes. The protocol of preparation of 1.5 % agarose gel is described in Appendix (A.2). Under UV light, the gel (the size on position) was cut out, and then the gel was purified by using a gel extraction kit (Qiagen). Then, TmAFP insert and pET20b vector were ligated together. The 21 µl of ligation mixture contained 1 µl vector DNA, 3 µl insert DNA, 10 µl quick ligase buffer (New England Biolabs), 1 µl quick ligase (New England Biolabs), and 6 µl nuclease free water. The ligation mixture was kept in a 5 min incubation at room temperature. The ligated DNA was inserted into BL21* competent cells. BL21* competent cells from -80 ^{o}C freezer were thawed out on ice. The 10 µl of ligation mixture and 50 µl of BL21* competent cells were mixed and heat shocked for 2

minutes at 37 ºC. It was then incubated at 37 ºC for 1 hour with addition of 950 µl of LB media. The 350 µl of transformation mixture was then placed on a LB agar plate with ampicillin (100 µg/ml) from Fisher Scientific. The LB agar plates were placed in an incubator at 37 ºC for overnight. When the colonies were observed, the colonies were carefully selected and transferred into 6 ml of LB medium with ampicillin. The cultures were incubated at 37 ºC with shaking 300 rpm overnight. The frozen stocks were made with the mixture of 400 µl of 50 % glycerol and 500 µl of bacteria cell culture, and the frozen stocks were stored in a -80 ºC freezer. The plasmid DNA was isolated from an overnight culture by using a miniprep kit (Qiagen). The isolated plasmid DNA was screened with PCR screening via a thermo cycler (Thermo Electron Corp.) to determine whether the ligation was successful or not. The PCR was performed with a T7 forward primer and T7 reverse primer. The PCR protocol for screening is described in Appendix (A.3). After PCR was performed, the PCR samples were loaded onto the agarose gel as previously described. If ligation was successful, the plasmid DNA (vector and insert) was located on a higher position as much as an insert size, versus a vector only plasmid. Isolated DNA from positive screens was sent to Genomic Lab at the Cleveland Clinic Foundation (CCF) for DNA sequencing. These isolated DNAs were confirmed to be TmAFP-pET20b plasmids.

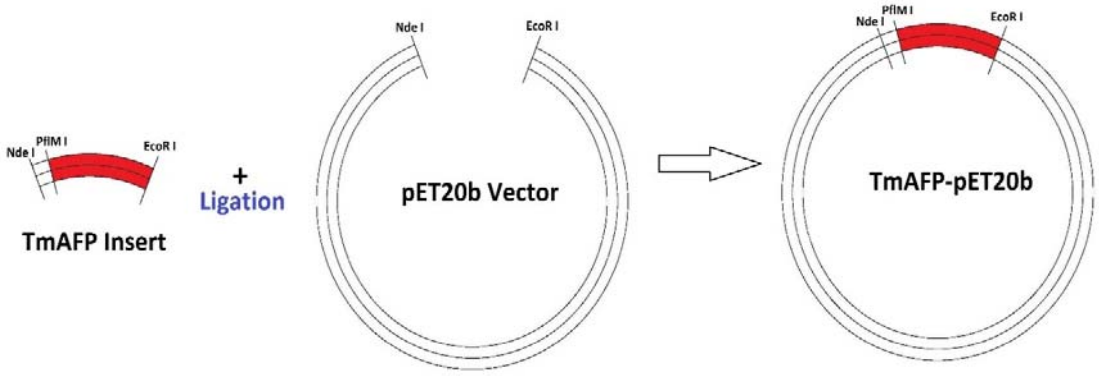


Figure 2.1. Schematic diagram of the synthesis of TmAFP-pET20b.

TmAFP-pET20b plasmid was double digested by the restriction enzymes NdeI (New England Biolabs) and SfiI (New England Biolabs) sequentially and used as a vector for ligation. The 50 µl of reaction mixture contained 2 µl NdeI, 10 µl DNA, 5 µl 10 buffer 2 (New England Biolabs), and 31 µl Nuclease-free water. The reaction mixture was incubated in a 37 ^{o}C water bath for 2 hours. Then, 2 µl SfiI was added into the reaction mixture, and incubated in a 55 ^{o}C dry bath for 2 additional hours. The double digested DNA was then loaded onto the agarose gel and purified by using a gel extraction kit (Qiagen). Ligation of TEV_G or TEV_S inserts and TmAFP-pET20b vector was performed. Transformation into BL21* competent cells was carried out. The colonies were carefully selected and grown in LB medium with ampicillin. The frozen stocks were made and stored. The plasmid DNA was isolated and screened with PCR screening. The isolated DNA was sent to the Genomic Lab at the Cleveland Clinic Foundation (CCF) for DNA sequencing.

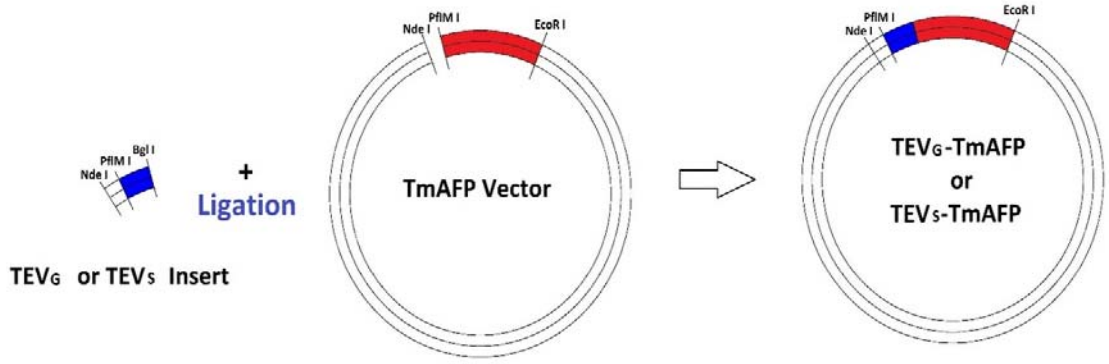


Figure 2.2. Schematic diagram of the synthesis of TEV_G-TmAFP and TEV_S-TmAFP.

The neutral ELPs, which were $(GVGVP)_{20}$ and $(GVGVP)_{40}$, were chosen and ELP plasmids were obtained in our lab. ELP-pET20b plasmids were isolated from overnight cultures by using a miniprep kit (Qiagen). ELP plasmids were double digested with NdeI and BglI as inserts for ligation. The 40 µl of double digestion mixture contained 2 µl NdeI, 2 µl BglI, 4 µl 10x buffer 3 (New England Biolabs), 17 µl nuclease free water, and 15 µl DNA. The reaction mixture was incubated in a 37 ^{o}C water bath for 3 hours. Also TEV_G-TmAFP and TEV_S-TmAFP in pET20b plasmid were double digested with NdeI and PflMI for vectors. The 40 µl of reaction mixture contained 2µl NdeI, 2 µl PflMI, 0.4 µl 100x BSA (New England Biolabs), 4 µl 10x buffer 2 (New England Biolabs), 16.6 µl nuclease free water, and 15 µl DNA. The double digestion mixture was incubated in a 37 ^{o}C water bath for 3 hours. The double digested DNA for vectors and inserts were loaded onto the agarose gel, and purified by using a gel extraction kit (Qiagen). The ligation and transformation were performed by the previously mentioned protocols. The colonies were picked and grown in LB media with

ampicillin. The frozen stocks were made and stored in a -80 ^{o}C freezer. The plasmid DNA was isolated and PCR screening was carried out. The isolated DNA was sent to the Genomic Lab at the Cleveland Clinic Foundation for DNA sequencing.

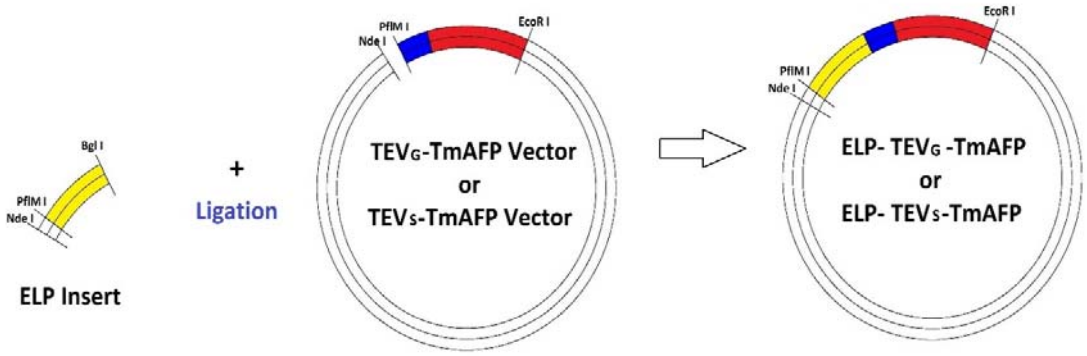


Figure 2.3. Schematic diagram of the synthesis of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP.

Finally, ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP DNA plasmids were transformed into origami competent cells since origami host cells can facilitate forming the proper pairing of disulfide bonds in TmAFP during protein expression [39].

2.1.3 Construction of ELP-TEV Protease DNA

ELP[(GVGVP)$_{20}$]-TEV Protease and ELP[(GVGVP)$_{40}$]-TEV Protease were synthesized. TEV Protease plasmid DNA, Plasmid 8827: pRK793 was ordered from Addgene.

DNA sequences of TEV Protease, 708bp

GGAGAAAGCTTGTTTAAGGGGCCGCGTGATTACAACCCGATATCGAGCACCATT
TGTCATTTGACGAATGAATCTGATGGGCACACAACATCGTTGTATGGTATTGGA
TTTGGTCCCTTCATCATTACAAACAAGCACTTGTTTAGAAGAAATAATGGAACA
CTGTTGGTCCAATCACTACATGGTGTATTCAAGGTCAAGAACACCACGACTTTG
CAACAACACCTCATTGATGGGAGGGACATGATAATTATTCGCATGCCTAAGGAT
TTCCCACCATTTCCTCAAAAGCTGAAATTTAGAGAGCCACAAAGGGAAGAGCG**C**
ATATGTCTTGTGACAACCAACTTCCAAACTAAGAGCATGTCTAGCATGGTGTCA
GACACTAGTTGCACATTCCCTTCATCTGATGGCATATTCTGGAAGCATTGGATT
CAAACCAAGGATGGGCAGTGTGGCAGTCCATTAGTATCAACTAGAGATGGGTTC
ATTGTTGGTATACACTCAGCATCGAATTTCACCAACACAAACAATTATTTCACA
AGCGTGCCGAAAAACTTCATGGAATTGTTGACAAATCAGGAGGCGCAGCAGTGG
GTTAGTGGTTGGCGATTAAATGCTGACTCAGTATTGTGGGGGGGCCATAAAGTT
TTCATGGTGAAACCTGAAGAGCCTTTTCAGCCAGTTAAGGAAGCGACTCAACTC
ATGAAT

Amino Acid Sequences of TEV Protease (236 aa):

GESLFKGPRDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNGT
LLVQSLHGVFKVKNTTTLQQHLIDGRDMIIIRMPKDFPPFPQKLKFREPQREER
ICLVTTNFQTKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGF
IVGIHSASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKV
FMVKPEEPFQPVKEATQLMN

The restriction enzyme cut sites were checked.

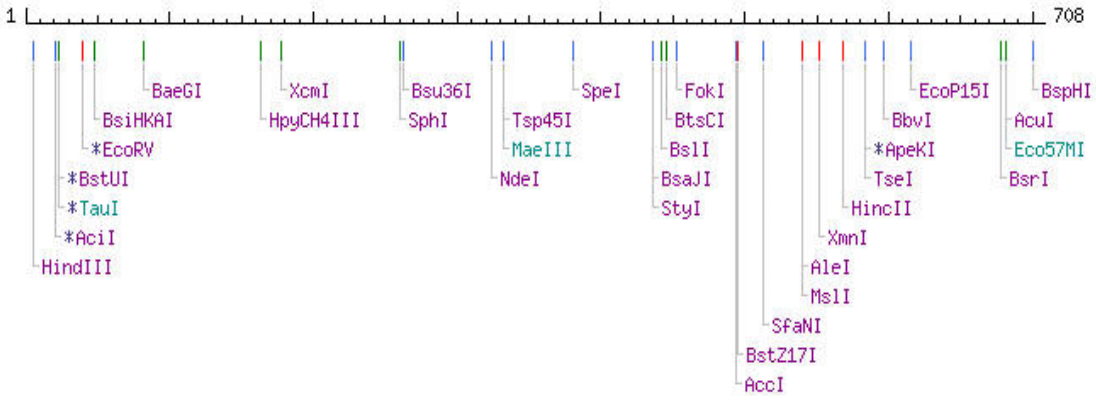


Figure 2.4. Restriction enzyme recognition site map of TEV protease by using NEB cutter 2.0. Retrieved from New England Biolabs web site http://tools.neb.com/NEBcutter2/index.php [75].

As shown in the above restriction enzyme recognition site map, a NdeI cut site is located in the middle of the TEV protease DNA sequence (bolded DNA sequences). NdeI is commonly used in our lab when adding fusion tags. Therefore, the NdeI cut site was removed from the original DNA of TEV protease by mutation in order to not excise an area of interest within the TEV protease DNA sequences when adding the fusion tags. Compared with the original TEV protease, the new DNA sequences were different, but the amino acid sequence coded for was the same as that of the original TEV protease.

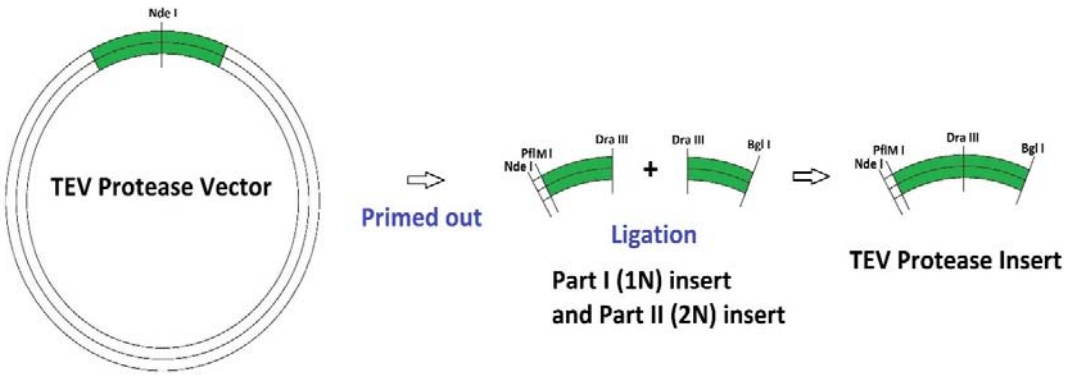


Figure 2.5. Schematic diagram of the synthesis of the mutated TEV Protease insert.

The NdeI cut site was replaced by a DraIII cut site because DraIII was not commonly used in our lab. Also, PflMI cut site needed to be modified because DraIII recognizes the PflMI cut site that was used in our lab. For this mutational process, two pairs of DNA primers were designed, and these primers were ordered from Invitrogen. The original DNA of TEV protease was extracted by PCR amplification using a thermo cycler (Thermo Electron Corp.) using the two pairs of primers. Two parts were prepared from the original DNA of TEV protease sequentially (Appendix A.3); labeled as 1N (Part I) and 2N (Part II). For 2N (Part II), TEV Dra forward and TEV Bgl reverse primers were used:

TEV Dra forward primer (5' to 3')

TTTGGGCATATGGGGCCACAACGTGAAGAGCGTATATGTCTTGTGAC

TEV Bgl reverse primer (5' to 3')

GACTCAACTCATGAATTAATGACTGCCAGGCTGGCCCTCTT

For 1N, TEV NoDra forward and TEV Dra reverse primers were used:

TEV NoDra forward primer (5' to 3')

TTTTTTCATATGGGCCAGGGCGTGGGAGAAAGCTTG

TEV Dra reverse primer (5' to 3')

TTCCTTCACGTTGTGGCTCTCTAAATTTCAGC

The PCR products were run on an agarose gel and purified using a gel extraction kit (Qiagen). Both 1N and 2N were single digested with Dra III. Each 40 µl of reaction mixture contained 4 µl 10x buffer 3, 0.4 µl 100x BSA, 18.6 µl nuclease free water, and 2 µl DraIII (New England Biolabs). The reaction mixtures were incubated in a 37 ^{o}C water bath for 3 hours followed by gel purification. Purified 1N and 2N were ligated with *E. coli* DNA ligase. The 20 µl of ligation mixture contained 5 µl of 1N, 5 µl of 2N, 2 µl *E. coli* DNA ligase buffer (New England Biolabs), 1 µl *E. coli* DNA ligase (New England Biolabs), and 7 µl nuclease free water. The ligation mixture was incubated in a 4 ^{o}C cold room for 1 hour, and then heat inactivation was carried out by using a thermo cycler (Thermo Electron Corp.) for 20 minutes at 65 ^{o}C.

The expected DNA sequences of the mutated TEV protease (5’ -> 3’), 708bp

GGAGAAAGCTTGTTTAAGGGGCCGCGTGATTACAACCCGATATCGAGCACCATT

TGTCATTTGACGAATGAATCTGATGGGCACACAACATCGTTGTATGGTATTGGA

TTTGGTCCCTTCATCATTACAAACAAGCACTTGTTTAGAAGAAATAATGGAACA

CTGTTGGTCCAATCACTACATGGTGTATTCAAGGTCAAGAACACCACGACTTTG

CAACAACACCTCATTGATGGGAGGGACATGATAATTATTCGCATGCCTAAGGAT

TTCCCACCATTTCCTCAAAAGCTGAAATTTAGAGAGCCACAACGTGAAGAGCGT

ATCTGCCTTGTGACAACCAACTTCCAAACTAAGAGCATGTCTAGCATGGTGTCA
GACACTAGTTGCACATTCCCTTCATCTGATGGCATATTCTGGAAGCATTGGATT
CAAACCAAGGATGGGCAGTGTGGCAGTCCATTAGTATCAACTAGAGATGGGTTC
ATTGTTGGTATACACTCAGCATCGAATTTCACCAACACAAACAATTATTTCACA
AGCGTGCCGAAAAACTTCATGGAATTGTTGACAAATCAGGAGGCGCAGCAGTGG
GTTAGTGGTTGGCGATTAAATGCTGACTCAGTATTGTGGGGGGGCCATAAAGTT
TTCATGGTGAAACCTGAAGAGCCTTTTCAGCCAGTTAAGGAAGCGACTCAACTC
ATGAAT

The amino acid sequences of the mutated TEV Protease (236 aa):

GESLFKGPRDYNPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNGT
LLVQSLHGVFKVKNTTTLQQHLIDGRDMIIIRMPKDFPPFPQKLKFREPQREER
ICLVTTNFQTKSMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGF
IVGIHSASNFTNTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKV
FMVKPEEPFQPVKEATQLMN

The restriction enzyme cut sites of the mutated TEV protease were checked.

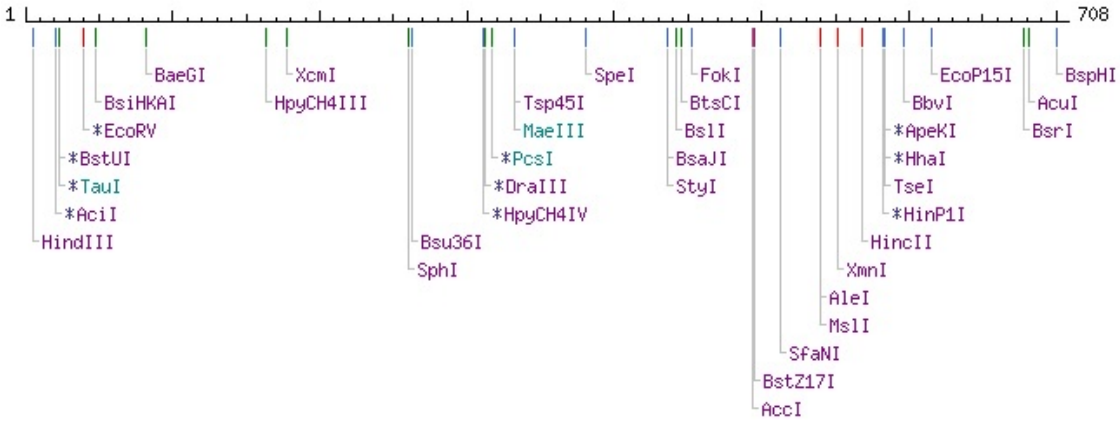


Figure 2.6. Restriction enzyme recognition site map of the mutated TEV protease by using NEB cutter 2.0. Retrieved from New England Biolabs web site http://tools.neb.com/NEBcutter2/index.php [75].

The ligation of 1N and 2N was amplified by PCR (Appendix A.3) using TEV NoDra forward and TEV Bgl reverse primers. The PCR product was gel purified. The mutated TEV protease was double digested with NdeI and BglI. The double digested TEV protease DNA was gel purified for insert. pET20b SfiI plasmid was double digested with NdeI and SfiI, and gel purified for vector. The ligation between mutated TEV protease insert and pET20b SfiI vector was carried out followed by transformation into BL21* competent cells. The plasmid DNA was isolated as previously described, and PCR screening and DNA sequencing were performed.

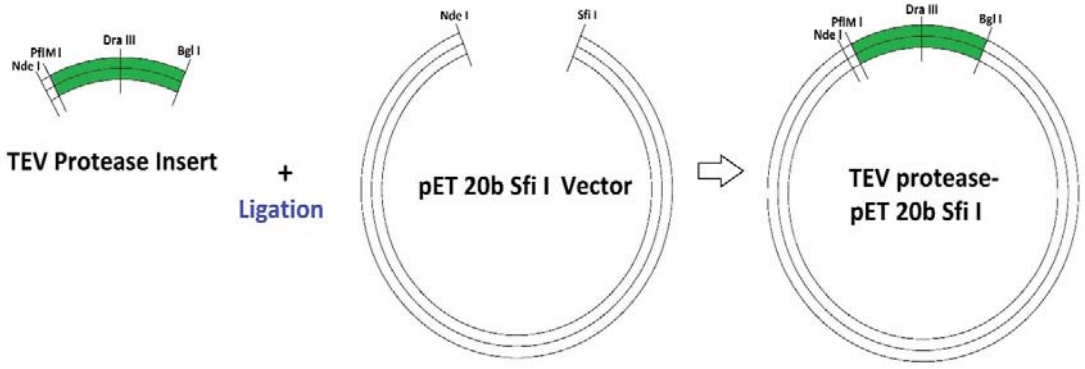


Figure 2.7. Schematic diagram of the synthesis of TEV Protease-pET 20b Sfi I.

The ELP inserts $(GVGVP)_{20}$ and $(GVGVP)_{40}$ were prepared. TEV Protease-pET 20b Sfi I vector was also prepared. TEV Protease-pET 20b Sfi I plasmid was double digested with NdeI and PflMI. The double digested DNA was loaded onto the gel and purified. ELP inserts and TEV Protease-pET 20b Sfi I vector were ligated and transformed into BL 21* competent cells. The frozen stocks were made and stored in a -80 ^{o}C freezer. The isolated plasmid was PCR screened and DNA sequences were obtained from the Genomic Lab at the Cleveland Clinic Foundation.

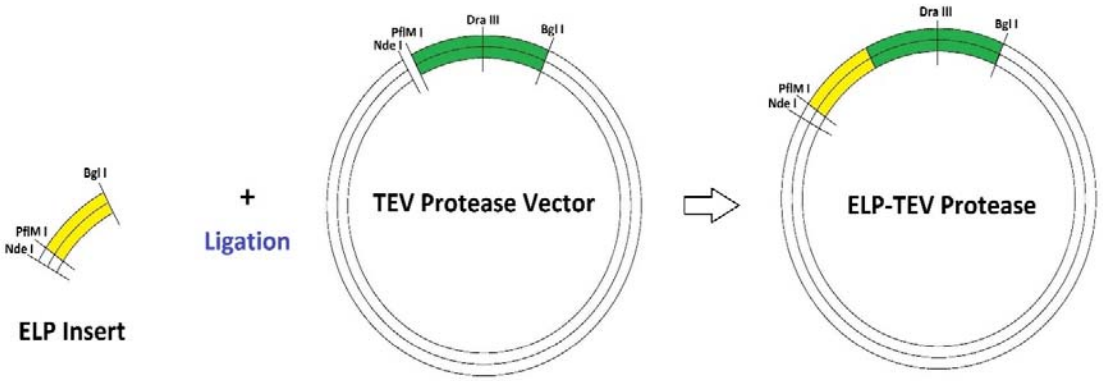


Figure 2.8. Schematic diagram of the synthesis of ELP-TEV Protease.

2.2 Protein Expression

2.2.1 ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP in Origami host cell

ELP-TEV$_G$-TmAFP-pET20b and ELP-TEV$_S$-TmAFP-pET20b in Origami host cell were used for protein expression. The cells from a frozen stock at -80 ^oC were grown in 10 ml LB media with ampicillin (100 µg/ml) at 37 ^oC with shaking 300 rpm. The 10 ml overnight cultures were inoculated and grown in 1 L LB medium with ampicillin (100 µg /ml). Cell growth was monitored by optical density measured at 600 nm wave length (OD_{600}) using a UV-Vis spectrophotometer (Thermo Scientific Biomate 3). When the OD_{600} reached 0.6, 0.5 mM of Isopropyl-β-D-thiogalatopyranoside (IPTG) from Promega was added to each 1 L culture for induction. The incubation was continued for 48 hours more at 15 ^oC with concurrent shaking at 300 rpm.

2.2.2 ELP-TEV protease in BL21 host cell*

The ELP-TEV protease in pET20b expression vector was used for protein expression. The cells from a frozen stock at -80 ^{o}C was grown in 10 ml LB media at 4 ^{o}C cold room with ampicillin (100 µg/ml) at 37 ^{o}C with shaking at 300 rpm. The 10 ml overnight cultures were inoculated and grown in 1 L LB medium with ampicillin (100 µg /ml) to OD_{600nm} 0.6. IPTG (1 mM) was added to each 1 L culture for induction, and the incubation was continued at 37 ^{o}C for 6 hours more with concurrent shaking at 300 rpm.

2.3 Protein Purification

After protein expression, the 1 L culture was transferred into 3 large centrifuge tubes and then centrifuged at 3000 xg for 20 min at 4 ^{o}C by using a floor centrifuge (J2-21 centrifuge from Beckman). The media was discarded, and the cell pellets were stored in -20 ^{o}C freezer for subsequent usage. The frozen pellets were thawed at room temperature, and 10 ml of PBS was added into each tube. The cell pellets were resuspended, transferred into a 50 ml falcon tube, and incubated on ice for 10 minutes. Sonication was performed (Appendix A.4) using a 550 Sonic Dismembrator from Fisher Scientific. After sonication, the sample was cooled down on ice for 15 minutes.

2.3.1 Inverse Transition Cycle (ITC)

ELP-TEV$_G$-TmAFP, ELP-TEV$_S$-TmAFP, and ELP-TEV protease proteins were purified by ITC [52]. The cold spin and warm spin were defined by the transition temperature (T_t) of the ELP tagged protein: The cold spin was performed below the T_t, and the warm spin was performed above the T_t. Insoluble substances are separated with

the cold spins because ELP tagged protein is soluble in aqueous solution below the T_t. Soluble substances are separated with the use of warm spins because ELP tagged protein are insoluble in aqueous solution above the T_t. The protocol of ITC purification is described in Appendix (A.5).

2.3.2 Cold finger purification

For antifreeze proteins (ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP), active proteins can be purified by cold finger purification since antifreeze proteins bind to ice and become incorporated in the growing ice, while others are excluded. This general method for cold finger purification has been previously described (Figure 2.9) [74], and the protocol is reported in Appendix A.6. The entire cold finger purification was performed in a 4 ^oC room. The cold finger was connected with a temperature controlled water bath (VWR Scientific model 1156) filled with antifreezer. The temperature of the water bath was pre-set at -0.5 ^oC, and the cold finger was covered with a thin layer of ice in a double distilled water beaker filled with a piece of ice. The sonicated sample was diluted up to 80 ml with double distilled water added, and a magnetic stir bar was placed in the sample solution. The diluted protein solution was switched to a double distilled water beaker, and the ice covered cold finger was placed in the protein solution. The temperature was gradually lowered from -0.5 to -2.0 ^oC over 8 hours (i.e. -0.1°C / 30min) with stirring at 350 rpm. The sample solution (liquid fraction) was removed. Then the ice on the cold finger was set in a 1 ^oC water bath until it melted down and the collected solution is the ice fraction.

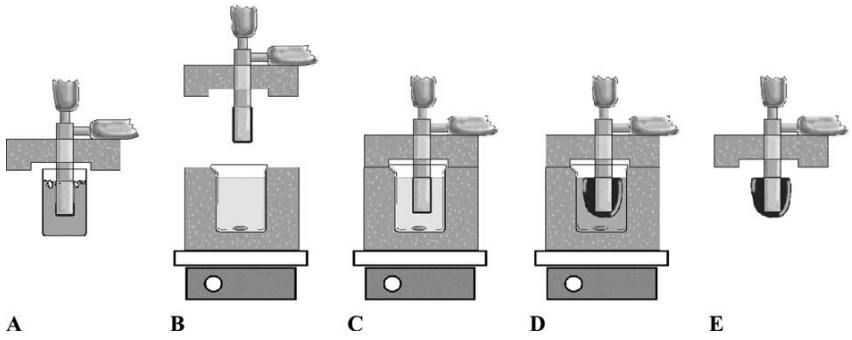


Figure 2.9. The schematic diagram of cold finger purification. (A) seeding the cold finger in cooled distilled water with a piece of ice; (B) A water filled beaker is replaced with sample beaker after thin ice layer is formed on the cold finger; (C) the cold finger is placed in sample solution; (D) the ice is growing until temperature reaches -2.0 ^{o}C; (E) the ice fraction containing the bound AFPs is collected [74]. Reprinted from *Biochemical and Biophysical Research Communication*, *300*, Kuiper, M. J., Lankin, C., Gauthier, S. Y., Walker, V. K., and Davies, P. L. (2003), Purification of antifreeze proteins by adsorption to ice, 646, Copyright (2014), with permission from Elsevier.

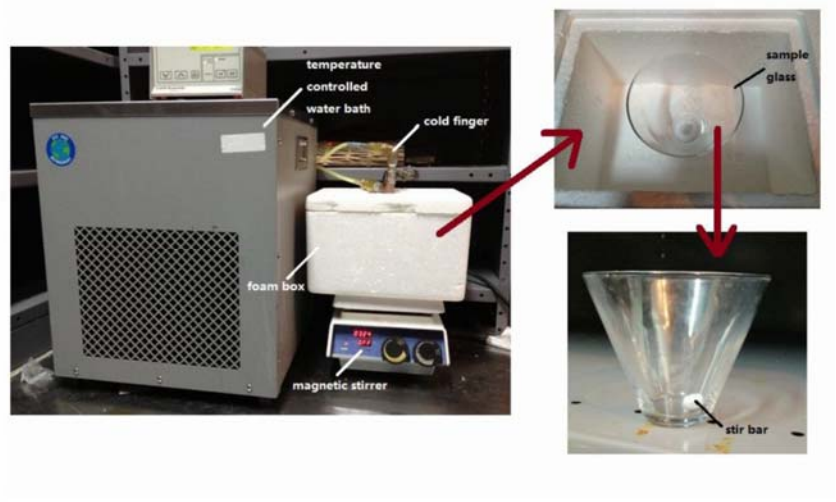


Figure 2.10. Cold finger purification apparatus. Figure was modified from the original figure with permission of *Sen Bu.* [73].

2.4 Characterization of Proteins

To elucidate the protein, protein samples were loaded onto a 4-20 % Precise protein gel (Thermo Scientific) by using a SDS electrophoresis kit (Mini Protein 3 cell from BIO-RAD) at 100 V for 1 hour with BupH^{TM} Tris-HEPES-SDS running buffer from Thermo Scientific. The protein gel was stained by using GelCodeBlue (Thermo Scientific) at room temperature for 2 hours, and then washed with double distilled water several times and double distilled water was added overnight as a destaining process. A scanner (Epson Perfection 1650) was used to scan and image the destained gel.

2.4.1 Protein concentration measurements

Protein concentration was calculated from absorbance at 280 nm wavelength using a calculating molecular extinction coefficient [76]. Absorbance was measured in a 1cm diameter quartz cuvette using UV spectrophotometer (Biomate 3 from Thermo Scientific).

2.4.2 Activity of TEV protease

ELP tagged TEV protease that is active can cleave at TEV protease cleavage sites, which were placed between the ELP tag and TmAFP. ELP tags of ELP-TEV protease, ELP-TEV$_G$-TmAFP, and ELP-TEV$_S$-TmAFP were identical to purified proteins by using the same temperature setting during ITC. The cleavage reaction by TEV protease was performed in PBS with an overnight incubation at 4 ^oC or 3 hours incubation at room temperature in the 1:5 concentration ratio of enzyme to substrate. After incubation, the reaction mixtures were characterized by SDS-PAGE.

2.4.3 Availability of TEV protease cleavage site

ProTEV from Promega was used to measure the availability of the TEV protease cleavage site placed in between ELP and TmAFP. The 100 µl of reaction volume contained 5 µl ProTEV buffer, 1 µl 100 mM DTT (Promega), 2 µl Pro TEV plus protease, 20 µg fusion protein, and filled with nuclease free water up to a 100 µl total volume. The reaction mixture was incubated at 4 ^{o}C overnight or at 30 ^{o}C for 1 hour. After incubation, the reaction mixtures were loaded on protein gels.

2.4.4 Activity of TmAFP

Activity of TmAFP as the target protein was observed by using cold finger purification as previously mentioned. The samples were loaded on protein gels.

CHAPTER III

RESULTS AND DISCUSSION

3.1 Synthesis of plasmid DNAs

3.1.1 Synthesis of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP

Plasmid DNA sequencing confirmed that the synthesis of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP DNAs were accomplished in pET20b plasmids. The yellow highlighted sequences are ELP, the blue highlighted sequences are TEV$_G$ and TEV$_S$, and the red highlighted sequences are TmAFP.

ELP[(GVGVP)$_{20}$]-TEV$_G$-TmAFP (5' → 3')

ATGGGCCACGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCG
GGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGT
GTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGC
GTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTT
CCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTG
GGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGAAAACCTG
TACTTCCAGGGGCCGGGCTGGCCACAGTGCACCGGTGGTGCTGACTGCACCAGC
TGCACCGCTGCTTGCACCGGTTGCGGTAACTGCCCGAACGCTGTTACTTGCACT
AACAGCCAGCACTGCGTTAAAGCTACCACCTGCACCGGTAGCACCGACTGCAAC
ACCGCTGTTACCTGCACTAACTCTAAAGACTGCTTCGAAGCTCAGACCTGCACC

GACTCTACCAACTGCTACAAAGCTACCGCTTGCACCAACTCTACCGGCTGCCCG

GGTCAC

ELP[(GVGVP)$_{20}$]-TEV$_S$-TmAFP (5’ → 3’)

ATGGGCCACGGCGTGGGTGTTCCGGGAGTAGGTGTCCCAGATGTGGGCGTACCG

GGCGTTGATGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGT

GTCCCAGATGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGC

GTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTT

CCTGGTGTCGGCGTGCCGAGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTG

GGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGAAAACCTG

TACTTCCAGTCGCCGGGCTGGCCACAGTGCACCGGTGGTGCTGACTGCACCAGC

TGCACCGCTGCTTGCACCGGTTGCGGTAACTGCCCGAACGCTGTTACTTGCACT

AACAGCCAGCACTGCGTTAAAGCTACCACCTGCACCGGTAGCACCGACTGCAAC

ACCGCTGTTACCTGCACTAACTCTAAAGACTGCTTCGAAGCTCAGACCTGCACC

GACTCTACCAACTGCTACAAAGCTACCGCTTGCACCAACTCTACCGGCTGCCCG

GGTCAC

ELP[(GVGVP)$_{40}$]-TEV$_G$-TmAFP (5’ → 3’)

ATGGGCCACGGCGTGGGTGTTCCGGGAGTAGGTGTCCCAGATGTGGGCGTACCG

GGCGTTGATGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGT

GTCCCAGATGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGC

GTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTT

CCTGGTGTCGGCGTGCCGAGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTG

GGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCG

GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGC

GTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGC

GTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTC

CCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTG

GGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT

GGTGTCGGCGTGCCGGGCGTGGAAAACCTGTACTTCCAGGGGCCGGGCTGGCCA

CAGTGCACCGGTGGTGCTGACTGCACCAGCTGCACCGCTGCTTGCACCGGTTGC

GGTAACTGCCCGAACGCTGTTACTTGCACTAACAGCCAGCACTGCGTTAAAGCT

ACCACCTGCACCGGTAGCACCGACTGCAACACCGCTGTTACCTGCACTAACTCT

AAAGACTGCTTCGAAGCTCAGACCTGCACCGACTCTACCAACTGCTACAAAGCT

ACCGCTTGCACCAACTCTACCGGCTGCCCGGGTCAC

ELP[(GVGVP)$_{40}$]-TEV$_S$-TmAFP (5’ → 3’)

ATGGGCCACGGCGTGGGTGTTCCGGGAGTAGGTGTCCCAGATGTGGGCGTACCG

GGCGTTGATGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGT

GTCCCAGATGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGC

GTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTT

CCTGGTGTCGGCGTGCCGAGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTG

GGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCG

GGAGTAGGTGTCCCAGATGTGGGCGTACCGGGCGTTGATGTTCCTGGTGTCGGC

GTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGATGTGGGCGTACCGGGC

GTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTC

CCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGAGCGTG

GGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT

GGTGTCGGCGTGCCGGGCGTGGAAAACCTGTACTTCCAGTCGCCGGGCTGGCCA
CAGTGCACCGGTGGTGCTGACTGCACCAGCTGCACCGCTGCTTGCACCGGTTGC
GGTAACTGCCCGAACGCTGTTACTTGCACTAACAGCCAGCACTGCGTTAAAGCT
ACCACCTGCACCGGTAGCACCGACTGCAACACCGCTGTTACCTGCACTAACTCT
AAAGACTGCTTCGAAGCTCAGACCTGCACCGACTCTACCAACTGCTACAAAGCT
ACCGCTTGCACCAACTCTACCGGCTGCCCGGGTCAC

Also, the corresponding amino acid sequences were obtained from the confirmed DNA sequences using DNA to protein translation [77]. The yellow highlighted sequences are ELP, the blue highlighted sequences are TEV_G and TEV_S, and the red highlighted sequences are TmAFP.

ELP[$(GVGVP)_{20}$]-TEV_G-TmAFP)

MGHGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPG
VGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVENL
YFQGPGWPQCTGGADCTSCTAACTGCGNCPNAVTCTNSQHCVKATTCTGSTDCN
TAVTCTNSKDCFEAQTCTDSTNCYKATACTNSTGCPGH

ELP[$(GVGVP)_{20}$]-TEV_S-TmAFP

MGHGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPG
VGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVENL
YFQSPGWPQCTGGADCTSCTAACTGCGNCPNAVTCTNSQHCVKATTCTGSTDCN
TAVTCTNSKDCFEAQTCTDSTNCYKATACTNSTGCPGH

ELP[(GVGVP)$_{40}$]-TEV$_{G}$-TmAFP

MGHGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPG
VGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVP
GVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV
PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVENLYFQGPGWP
QCTGGADCTSCTAACTGCGNCPNAVTCTNSQHCVKATTCTGSTDCNTAVTCTNS
KDCFEAQTCTDSTNCYKATACTNSTGCPGH

ELP[(GVGVP)$_{40}$]-TEV$_{S}$-TmAFP

MGHGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPG
VGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVP
GVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV
PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVENLYFQSPGWP
QCTGGADCTSCTAACTGCGNCPNAVTCTNSQHCVKATTCTGSTDCNTAVTCTNS
KDCFEAQTCTDSTNCYKATACTNSTGCPGH

3.1.2 Synthesis of ELP-TEV Protease

The synthesis of the gene for ELP[(GVGVP)$_{40}$]-TEV protease was successful. The DNA sequence of ELP[(GVGVP)$_{40}$]-TEV protease was accomplished in pET20b plasmids: The yellow highlighted sequences are ELP and the green highlighted sequences are TEV protease.

ELP[(GVGVP)$_{40}$]-TEV protease (5’ → 3’)

ATGGGCCACGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCG

GGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGT

GTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGC

GTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTT

CCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTG

GGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCG

GGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGC

GTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGC

GTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTGGGTGTTCCGGGCGTAGGTGTC

CCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCTGGTGTCGGCGTGCCGGGCGTG

GGTGTTCCGGGCGTAGGTGTCCCAGGTGTGGGCGTACCGGGCGTTGGTGTTCCT

GGTGTCGGCGTGCCGGGCGTGGGAGAAAGCTTGTTTAAGGGGCCGCGTGATTAC

AACCCGATATCGAGCACCATTTGTCATTTGACGAATGAATCTGATGGGCACACA

ACATCGTTGTATGGTATTGGATTTGGTCCCTTCATCATTACAAACAAGCACTTG

TTTAGAAGAAATAATGGAACACTGTTGGTCCAATCACTACATGGTGTATTCAAG

GTCAAGAACACCACGACTTTGCAACAACACCTCATTGATGGGAGGGACATGATA

ATTATTCGCATGCCTAAGGATTTCCCACCATTTCCTCAAAAGCTGAAATTTAGA

GAGCCACAACGTGAAGAGCGTATCTGCCTTGTGACAACCAACTTCCAAACTAAG

AGCATGTCTAGCATGGTGTCAGACACTAGTTGCACATTCCCTTCATCTGATGGC

ATATTCTGGAAGCATTGGATTCAAACCAAGGATGGGCAGTGTGGCAGTCCATTA

GTATCAACTAGAGATGGGTTCATTGTTGGTATACACTCAGCATCGAATTTCACC

AACACAAACAATTATTTCACAAGCGTGCCGAAAAACTTCATGGAATTGTTGACA

AATCAGGAGGCGCAGCAGTGGGTTAGTGGTTGGCGATTAAATGCTGACTCAGTA

TTGTGGGGGGGCCATAAAGTTTTCATGGTGAAACCTGAAGAGCCTTTTCAGCCA

GTTAAGGAAGCGACTCAACTCATGAAT

Also, the corresponding amino acid sequences were determined from DNA sequences using DNA to protein translation [77]. The yellow highlighted sequences are ELP and the green highlighted sequences are TEV protease.

ELP[(GVGVP)$_{40}$]-TEV protease

MGHGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPG

VGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVP

GVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGV

PGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGVPGVGESLFKGPRDY

NPISSTICHLTNESDGHTTSLYGIGFGPFIITNKHLFRRNNGTLLVQSLHGVFK

VKNTTTLQQHLIDGRDMIIIRMPKDFPPFPQKLKFREPQREERICLVTTNFQTK

SMSSMVSDTSCTFPSSDGIFWKHWIQTKDGQCGSPLVSTRDGFIVGIHSASNFT

NTNNYFTSVPKNFMELLTNQEAQQWVSGWRLNADSVLWGGHKVFMVKPEEPFQP

VKEATQLMN

3.2 Protein Expression

3.2.1 Protein expression of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP

Since ELP and TmAFP were previously expressed independently in our lab, the protein expressions of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP were not investigated. The molecular weight of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP were

calculated based on the amino acid sequence using the Protein Information Resource [78].

Table 3.1. The expected molecular weight of ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP.

Sample	Molecular Weight
ELP[(GVGVP)$_{20}$]-TEV$_G$-TmAFP	18.3 kDa
ELP[(GVGVP)$_{20}$]-TEV$_S$-TmAFP	18.4 kDa
ELP[(GVGVP)$_{40}$]-TEV$_G$-TmAFP	26.5 kDa
ELP[(GVGVP)$_{40}$]-TEV$_S$-TmAFP	26.6 kDa

3.2.2 Protein expression of ELP[(GVGVP)$_{40}$]-TEV protease

The expected molecular weight of ELP[(GVGVP)$_{40}$]-TEV protease was calculated from Protein Information Resource [78] and it is 43.7 kDa. The protein expression of ELP[(GVGVP)$_{40}$]-TEV protease was investigated.

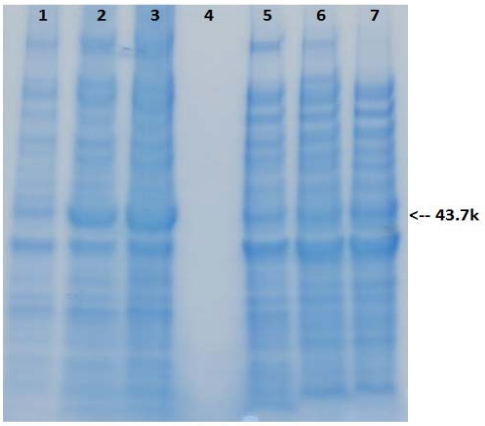


Figure 3.1. B-PER of ELP[(GVGVP)$_{40}$]-TEV protease. Lane 1 to 3 were an insoluble fraction, lane 4 was blank, and lane 5 to 7 were a soluble fraction. Lane 1 and 5 were collected before induction, lane 2 and 6 were collected at 3 h after induction, and lane 3 and 7 were collected at 6 h after induction.

As expression time lapsed after induction, the protein bands at 43.7 kDa became more intense (darker) in both the insoluble and soluble fractions. This implies that ELP[(GVGVP)$_{40}$]-TEV protease was expressed as both insoluble and soluble protein. The recombinant TEV protease has been reported to be insoluble in aqueous solution [71], while ELP is known to facilitate the increase in solubility of fusion proteins in aqueous solutions [12]. However, it seemed that ELP might not completely make ELP[(GVGVP)$_{40}$]-TEV protease soluble in aqueous solution.

3.3 Protein Purification

3.3.1 ITC Purification

ELP-TEV$_G$-TmAFP, ELP-TEV$_S$-TmAFP, and ELP-TEV protease were purified by ITC purification.

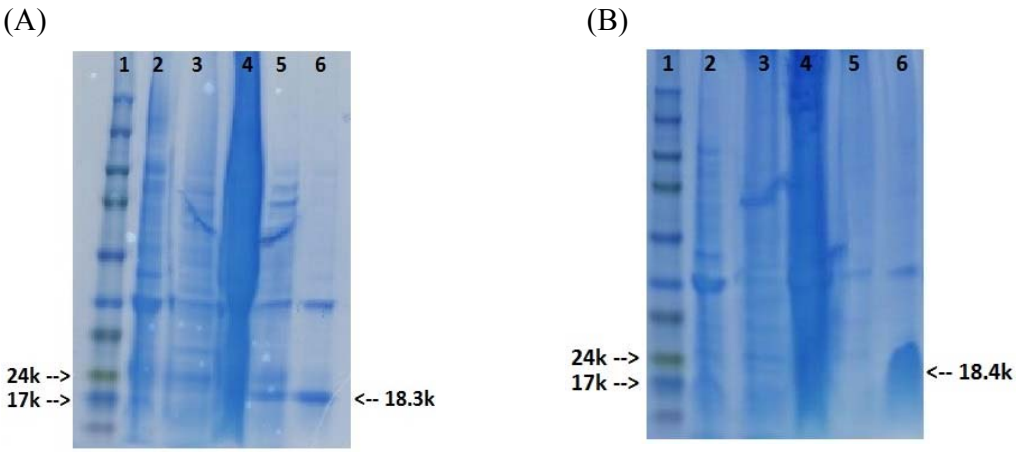


Figure 3.2. (A) ITC purification of ELP[(GVGVP)20]-TEV$_G$-TmAFP (18.3 kDa) and (B) ITC purification of ELP[(GVGVP)20]-TEV$_S$-TmAFP (18.4 kDa). Lane 1 is the molecular weight marker, lane 2 is the first cold centrifuge pellet, lane 3 is the first warm centrifuge supernatant, lane 4 is the second cold centrifuge pellet, lane 5 is the second warm centrifuge supernatant and lane 6 is the third cold centrifuge supernatant, which is the final purified product.

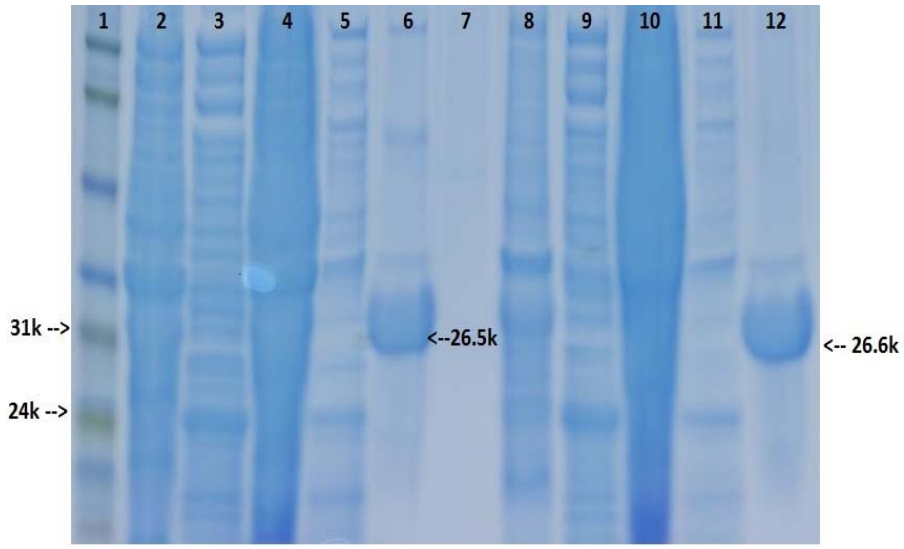


Figure 3.3. ITC purification of ELP[(GVGVP)$_{40}$]-TEV$_G$-TmAFP : 26.5 kDa and ITC purification of ELP[(GVGVP)$_{40}$]-TEV$_S$-TmAFP : 26.6 kDa. Lane 1 is the molecular weight marker, lane 7 is blank, lane 2 and 8 are the first cold centrifuge pellet, lane 3 and 9 are the first warm centrifuge supernatant, lane 4 and 10 are the second cold centrifuge pellet, lane 5 and 11 are the second warm centrifuge supernatant and lane 6 and 12 are the third cold centrifuge supernatant, which is the final purified product.

The protein concentrations by ITC purification were measured.

Table 3.2. The protein concentration by ITC purification.

Sample	C (M)	Yield (mg) in 1 L
ELP[(GVGVP)$_{20}$]-TEV$_G$-TmAFP	6.79×10^{-5}	3.74
ELP[(GVGVP)$_{20}$]-TEV$_S$-TmAFP	1.08×10^{-4}	5.59
ELP[(GVGVP)$_{40}$]-TEV$_G$-TmAFP	1.81×10^{-4}	14.41
ELP[(GVGVP)$_{40}$]-TEVs-TmAFP	2.52×10^{-4}	20.08

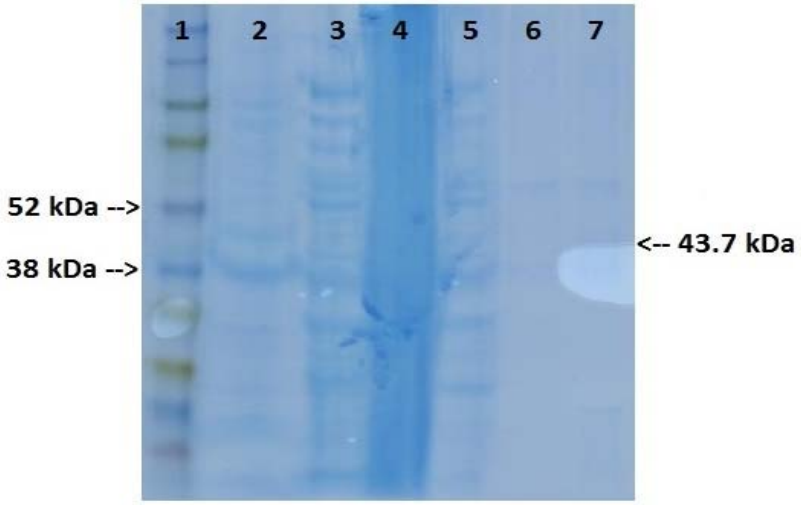


Figure 3.4. ITC purification of ELP[(GVGVP)$_{40}$]-TEV protease : 43.7 kDa. Lane 1 is the molecular weight marker, lane 2 is the first cold centrifuge pellet, lane 3 is the first warm centrifuge supernatant, lane 4 is the second cold centrifuge pellet, lane 5 is the second warm centrifuge supernatant, lane 6 is the third cold centrifuge pellet, and lane 7 is the third cold centrifuge supernatant, which is the final purified product.

ELP-TEV$_G$-TmAFP and ELP-TEV$_S$-TmAFP were successfully purified by using ITC purification, while ELP[(GVGVP)$_{40}$]-TEV protease was not successfully purified by ITC purification. Because the optimal temperature of TEV protease is 30 ^oC [15], it seemed that ELP[(GVGVP)$_{40}$]-TEV protease might be denatured during ITC purification above its optimal temperature (T_t of ELP[(GVGVP)$_{40}$]-TEV protease is around 40 ^oC). ELP[(GVGVP)$_{40}$]-TEV protease might need an alternative purification method(s) or additional purification steps.

3.4 Protein Characterization

3.4.1 The functionality of TEV protease cleavage site; TEV G and TEV S

With ProTEV protease from Promega, which is 50 kDa in size, the ability to cleave at the TEV protease cleavage sites between ELP and TmAFP were examined. ELP-TEV_G-TmAFP and ELP-TEV_S-TmAFP were examined for this reaction.

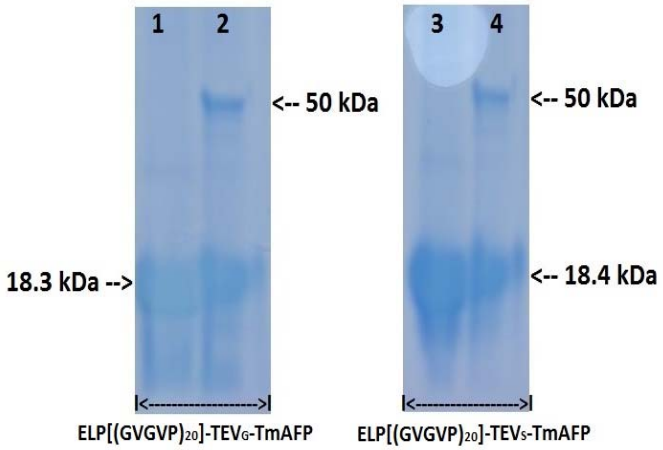


Figure 3.5. The functionality test of TEV protease cleavage sites: TEV_G and TEV_S : ELP[$(GVGVP)_{20}$]-TEV_G-TmAFP : 18.3 kDa and ELP[$(GVGVP)_{20}$]-TEV_S-TmAFP : 18.4 kDa. Lane 1 and 3 are ELP fusion protein only and lane 2 and 4 are samples of a mixture solution of ELP fusion protein and ProTEV that were incubated at room temperature for 3 hours. Lane 1 and 3 are samples that were intentionally kept at the same concentration as the original stock.

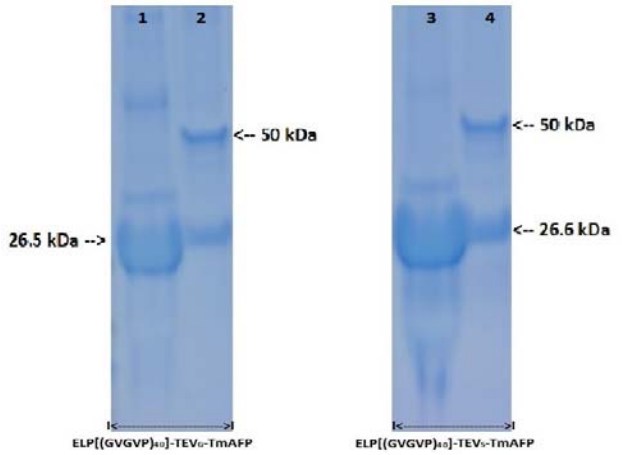


Figure 3.6. The functionality test of TEV protease cleavage sites; TEV_G and TEV_S : ELP[$(GVGVP)_{40}$]-TEV_G-TmAFP : 26.5 kDa and ELP[$(GVGVP)_{40}$]-TEV_S-TmAFP : 26.6 kDa. Lane 1 and 3 are ELP fusion protein only, lane 2 and 4 are samples of a mixture solution of ELP fusion protein and ProTEV that were incubated at room temperature for 3 hours. Lane 1 and 3 are samples that were intentionally kept at the same concentration as the original stock.

The TEV cleavage reactions were not observed between ELP[$(GVGVP)_{20}$] and TmAFP or between ELP[$(GVGVP)_{40}$] and TmAFP because the size of ProTEV and ELP tagged TmAFP fusion proteins were only observed on protein gel.

3.4.2 The activity of TmAFP

Cold finger purification was performed to observe the activity of TmAFP.

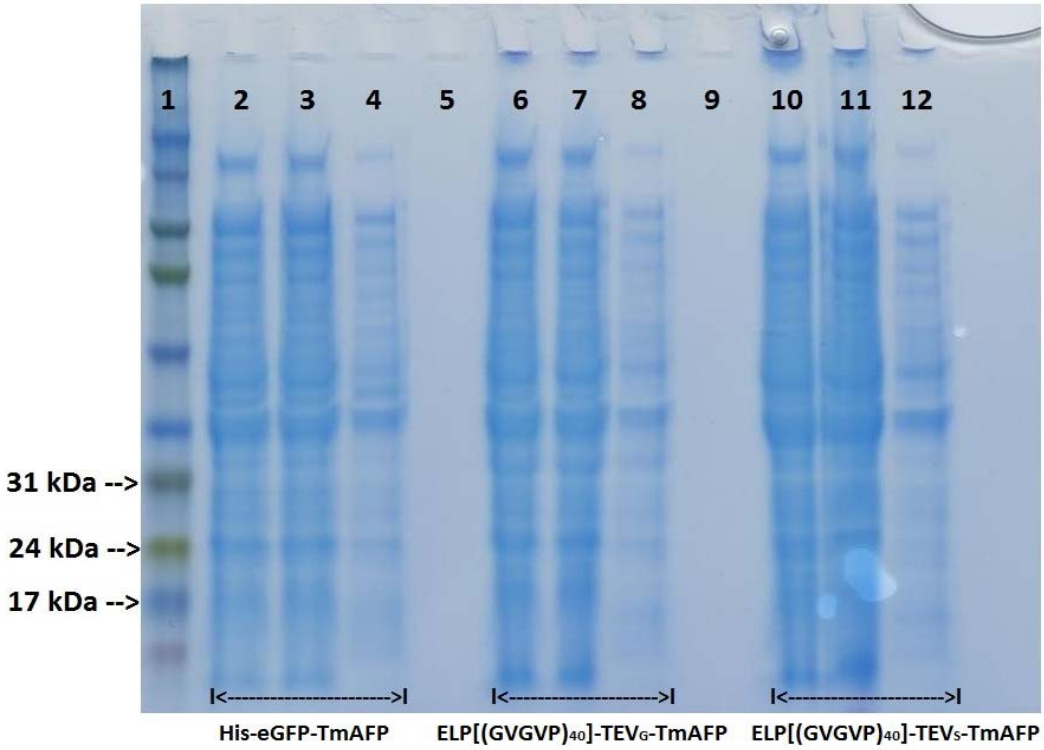


Figure 3.7. Cold finger purification. Lane 5 and 9 are blank. Lane 2, 6, and 10 are diluted unpurified proteins. Lane 3, 7, and 11 are liquid fraction after cold finger purification. Lane 4, 8, and 12 are ice fraction after cold finger purification.

There were no fusion proteins observed on the protein gel at His-eGFP-TmAFP, 19.6 kDa, ELP[(GVGVP)$_{40}$]-TEV$_G$-TmAFP, 26.5 kDa, and ELP[(GVGVP)$_{40}$]-TEVs-TmAFP, 26.6 kDa. The paucity of protein expressions was observed several times from the bacterial cultures obtained from frozen stocks which had previously been stored in a -80 ^{o}C freezer that became malfunctional. After the DNA sequences of fusion proteins were verified, the multiple expression of fusion proteins were performed. However, there were no fusion proteins expressed. It may need re-transformation into newer host competent cells.

CHAPTER IV

CONCLUSION

In this research, elastin-like polypeptide (ELP) tags were used to implement a strategy to purify recombinant *Tenebrio molitor* antifreeze protein (TmAFP) from *E. coli*. This purification method would allow an efficient purification method compared to other purification methods for recombinant TmAFP.

In order to implement this approach, we successfully prepared genes encoding TmAFP and TEV protease with N-terminal ELP fusion tags. The TmAFP constructs included a TEV cleavage site for removal of the ELP fusion tag. TEV cleavage sites and ELP tag DNAs were designed and successfully added to the N-terminus of TmAFP sequentially. ELP[$(GVGVP)_{40}$] tag DNA was successfully added to the N-terminus of TEV protease while ELP[$(GVGVP)_{20}$] tag DNA was not successfully added to the N-terminus of TEV protease.

There was mixed success in the expression and purification of the constructs. ELP tagged fusion proteins were successfully expressed at least once. ELP tagged fusion TmAFPs were successfully purified by using ITC. However, ELP tagged fusion TmAFPs were not successfully purified by cold finger purification because of failed expression. ELP tagged fusion TEV protease was not successfully purified by ITC purification because ELP[$(GVGVP)_{40}$]-TEV protease showed poor solubility in an aqueous solution.

The purification process may need to be improved to increase its solubility or an alternative purification method(s) may need to be used. The TEV protease cleavage reaction was not observed between ELP and TmAFP by ProTEV from Promega.

In conclusion, the genetic engineering of ELP tagged fusion proteins was overall successful. However, the expression and purification process need to be improved. ELP tags were functional through ITC purification and appear to aid in target protein purification. The functionality of the TEV protease cut site was not observed, but because of a lack of a positive control this conclusion is not certain. The activity of TmAFP was not measurable because of negative results from poor protein expression.

In the future, this purification method may be applicable for scale-up mass production of recombinant TmAFP. However, the activity of ELP tagged TEV protease needs to be measured. The activity of TmAFP without an ELP tag after TEV protease cleavage reaction needs to be measured as well. The comparison of several factors such as activity, yield, efficient purification, and solubility of ELP tagged and untagged TEV protease and TmAFP need to be considered. These factors should be further investigated in subsequent experiments.

REFERENCES

[1] McPherson, D. T, Xu, J., & Urry, D. W. (1996). Product purification by reversible phase transition following *Escherichia coli* expression of genes encoding up to 251 repeats of the elastomeric pentapeptide GVGVP. *Protein expression and purification, 7*, 51-57.

[2] Sørensen, H. P. & Mortensen, K. K. (2005). Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories*, *4*(1), 1-8.

[3] Meyer, D. E. & Chilkoti, A. (1999). Purification of recombinant proteins by fusion with thermally-responsive polypeptide. *Nature Biotechnology*, *17*, 1112-1115.

[4] Maxwell, K. L., Bona, D., Liu, C., Arrowsmith, C. H., & Edwards, A. M. (2003). Refolding out of guanidine hydrochloride is an effective approach for high-throughput structural studies of small proteins. *Protein Science*, *12*, 2073-2080.

[5] Walls, D. & Loughran, S. T. (2011). Tagging recombinant proteins to enhance solubility and aid purification. *Protein Chromatography: Methods and Protocols, Methods in Molecular Biology*, *681*, 151-175.

[6] Schein, C. H. (1989). Production of soluble recombinant proteins in bacteria. *Bio/technology*, *7*, 1141-1149.

[7] San-Miguel, T., Pérez-Bermúdez, P., & Gavidia, I. (2013). Production of soluble eukaryotic recombinant proteins in *E. coli* is favoured in early log-phase cultures induced at low temperature. *SpringerPlus*, *2*(89), 1-4.

[8] Hammarström, M., Hellgren, N., Van Den Berg, S, Berglund, H., & Härd T. (2002). Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*. *Protein Science*, *11*, 313-321.

[9] Waugh, D. S. (2005). Reprint of: Making the most of affinity tags. *TRENDS in Biotechnology*, *23*, 316-320.

[10] Nallamsetty, S. & Waugh, D. S. (2007). A generic protocol for the expression and purification of recombinant proteins in *Escherichia coli* using a combinatorial His_6-maltose binding protein fusion tag. *Nature Protocols*, *2*, 383-391.

[11] Nallamsetty, S., Austin, B. P., Penrose, K. J., & Waugh, D. S. (2005). Gateway vectors for the production of combinatorially-tagged His_6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli. Protein Science*, *14*, 2964-2971.

[12] Christensen, T., Amiram, M., Dagher, S., Trabbic-Carlson, K., Shamji, M. F., Setton, L. A., & Chilkoti, A. (2009). Fusion order controls expression level and activity of elastin-like polypeptide fusion proteins. *Protein Science*, *18*, 1377-1387.

[13] Dyson, M. R., Shadbolt, S. P., Vincent, K. J., Perera, R. L., & McCafferty, J. (2004). Production of soluble mammalian proteins in *Escherichia coli*: identification of protein features that correlate with successful expression. *BMC Biotechnology*, *4*(32), 32-49.

[14] Waugh, D. S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Protein Expression and Purification, xxx*, 1-11.

[15] Nallamsetty, S., Kapust, R. B., Tözsér, J., Cherry, S., Tropea, J. E., Copeland, T. D., & Waugh, D. S. (2004). Efficient site-specific processing of fusion proteins by tobacco vein mottling virus protease in vivo and in vitro. *Protein Expression and Purification*, *38*, 108-115.

[16] van Eldijk, M. B., McGann, C. L., Kiick, J. L., & van Hest, J. C.M. (2012). Elastomeric polypeptides. *Topics in Current Chemistry*, *310*, 71-116.

[17] Rodríguez-Cabello, J. C., Martín, L., Alonso, M., Arias, F. J., & Testera, A. M. (2009). "Recombinamers" as advanced materials for the post-oil age. *Polymer*, *50*, 5159-5169.

[18] Urry, D. W. (1997). Physical chemistry of biological free energy transduction as demonstrated by elastic protein-based polymers. *The Journal of Physical Chemistry B*, *101*, 11007-11028.

[19] Urry, D. W. (1992). Free energy transduction in polypeptides and proteins based on inverse temperature transitions. *Progress in Biophysics and Molecular Biology*, *57*, 23-57.

[20] Meyer, D. E. & Chilkoti, A. (2004). Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules*, *5*, 846-851.

[21] Urry, D. W., Trapane, T. L., & Prasad, K. U. (1985). Phase-structure transitions of the elastin polypentapeptide-water system within the framework of composition-temperature studies. *Biopolymers*, *24*, 2345-2356.

[22] Yamaoka, T., Tamura, T., Seto, Y., Tada, T., Kunugi, S., & Tirrell, D. A. (2003). Mechanism for the phase transition of a genetically engineered elastin model peptide $(VPGIG)_{40}$ in aqueous solution. *Biomacromolecules*, *4*, 1680-1685.

[23] Reguera, J., Urry, D. W., Parker, T. M., McPherson, D. T., & Rodríguez-Cabello, J. C. (2007). Effect of NaCI on the exothermic and endothermic components of the inverse temperature transition of a model elastin-like polymer. *Biomacromolecules*, *8*, 354-358.

[24] Nuhn, H. & Klok, H. (2008). Secondary structure formation and LCST behavior of short elastin-like peptides. *Biomacromolecules*, *9*, 2755-2763.

[25] Mackay, J. A., Callahan, D. J., FitzGerald, K. N., & Chilkoti, A. (2010). Quantitative model of the phase behavior of recombinant pH-responsive elastin-like polypeptides. *Biomacromolecules*, *11*, 2873-2879.

[26] Girotti, A., Reguera, J., Arias, F. J., Alonso, M., Testera, A. M., & Rodríguez-Cabello, J. C. (2004). Influence of the molecular weight on the inverse temperature transition of a model genetically engineered elastin-like pH-responsive polymer. *Macromolecules*, *37*, 3396-3400.

[27] Venketesh, S. & Dayananda, C. (2008). Properties, potentials, and prospects of antifreeze proteins. *Critical Reviews in Biotechnology*, *19*, 57-82.

[28] Costanzo, J. P. & Lee, R. E. (1994). Biophysical and physiological responses promoting freeze tolerance in vertebrates. *Neural Information Processing Systems Foundation*, *9*, 252-256.

[29] Davies, P. L. & Sykes, B. D. (1997). Antifreeze proteins. *Current Opinion in Structural Biology*, *7*, 828-834.

[30] Jia, Z. & Davies, P. L. (2002). Anifreeze proteins: an unusual receptor-ligand interaction. *TRENDS in Biochemical Sciences*, *27*, 101-106.

[31] Harding, M. M., Ward, L. G., & Haymet, A. D. J. (1999). Type I 'antifreeze' proteins: structure-activity studies and mechanisms of ice growth inhibition. *European Journal of Biochemistry*, *264*, 653-665.

[32] Patel, S. N. & Graether, S. P. (2010). Structures and ice-binding faces of the alanine-rich type I antifreeze proteins. *Biochemistry and Cell Biology*, *88*, 223-229.

[33] Gronwald, W., Loewen, M. C., Lix, B., Daugulis, A. J., Sönnichsen, F. D., Davies, P. L., & Sykes, B. D. (1998). The solution structure of type II antifreeze protein reveals a new member of the lectin family. *Biochemistry*, *37*(14), 4712-4721.

[34] Liu, Y., Li, Z., Lin, Q., Kosinski, J., Seetharaman, J., Bujnicki, J. M., Sivaraman, J., & Hew, C. (2007). Structure and evolutionary origin of Ca^{2+}-dependent herring type II antifreeze protein. *PlosOne*, *2*(6), e548.

[35] Takamichi, M., Nishimiya, Y., Miura, A., & Tsuda, S. (2009). Fully active QAE isoform confers thermal hysteresis activity on a defective SP isoform of type III antifreeze protein. *The Federation of European Biochemical Societies Journal*, *276*, 1471-1479.

[36] Gauthier, S. Y., Scotter, A. J., Lin, F., Baardsnes, J., Fletcher, G. L., & Davies, P. L. (2008). A re-evaluation of the role of type IV antifreeze protein. *Cryobiology*, *57*, 292-296.

[37] Harding, M. M., Anderberg, P. I., & Haymet, A. D. J. (2003). 'Antifreeze' glycoproteins from polar fish. *European Journal of Biochemistry*, *270*, 1381-1392.

[38] Lee J. K., Park, K. S., Park, S., Park, H., Song, Y. H., Kang, S., & Kim, H. J. (2010). An extracellular ice-binding glycoprotein from an arctic psychrophilic yeast. *Cryobiology*, *60*, 222-228.

[39] Bar, M., Bar-Ziv, R., Scherf, T., & Fass, D. (2006). Efficient production of a folded and functional, highly disulfide-bonded β-helix antifreeze protein in bacteria. *Protein Expression and Purification*, *48*, 243-252.

[40] Graether, S. P. & Sykes, B. D. (2004). Cold survival in freeze-intolerant insects: The structure and function of β-helical antifreeze proteins. *European Journal of Biochemistry*, *271*, 3285-3296.

[41] Graether, S. P., Kuiper, M. J., Gagné, S. M., Walker, V. K., Jia, Z., Sykes, B. D., & Davies, P. L. (2000). β-Helix structure and ice-binding properties of a hyperactive antifreeze protein from an insect. *Nature*, *406*, 325-328.

[42] Marshall, C. B., Daley, M. E., Graham, L. A., Sykes, B. D., & Davies, P. L. (2002). Identification of the ice-binding face of antifreeze protein from *Tenebrio molitor*. *FEBS Letters*, *529*, 261-267.

[43] Griffith, M. & Yaish, M. W.F. (2004). Antifreeze proteins in overwintering plants: a tale of two activities. *TRENDS in Plant Science*, *9*(8), 399-405.

[44] Atici, Ö. & Nalbantoğlu, B. (2003). Antifreeze proteins in higher plants. *Phytochemistry*, *64*, 1187-1196.

[45] Yamashita, Y., Nakamura, N., Omiya, K., Nishikawa, J., Kawahara, H., & Obata, H. (2002). Identification of an antifreeze lipoprotein from *Moraxella* sp. of Antarctic origin. *Bioscience, Biotechnology, and Biochemistry*, *66*(2), 239-247.

[46] Xiao, N., Suzuki, K., Nishimiya, Y., Kondo, H., Miura, A., Tsuda, S., & Hoshino, T. (2010). Comparison of functional properties of two fungal antifreeze proteins from *Antarctomyces psychrotrophicus* and *Typhula ishikariensis. The Federation of European Biochemical Societies Journal*, *277*, 394-403.

[47] Raymond, J. A. & Fritsen, C. H. (2001). Semipurification and ice recrystallization inhibition activity of ice-active substances associated with Antarctic photosynthetic organisms. *Cryobiology*, *43*, 63-70.

[48] Muryoi, N., Sato, M., Kaneko, S., Kawahara, H., Obata, H., Yaish, M. W. F., Griffith, M., & Glick, B. R. (2004). Cloning and expression of *afpA*, a gene encoding an antifreeze protein from the arctic plant growth-promoting rhizobacterium *Pseudomonas putida* GR12-2. *Journal of Bacteriology*, *186*, 5661-5671.

[49] Garnham, C. P., Gilbert, J. A., Hartman, C. P., Campbell, R. L., Laybourn-Parry, J., & Davies, P. L. (2008) A Ca^{2+}-dependent bacterial antifreeze protein domain has a novel β-helical ice-binding fold. *Biochemical Journal*, *411*, 171-180.

[50] Scotter, A. J., Marshall, C. B., Graham, L. A., Gilbert, J. A., Garnham, C. P., & Davies, P. L. (2006). The basis for hyperactivity of antifreeze proteins. *Cryobiology*, *53*, 229-239.

[51] Lin, F., Davies, P. L., & Graham, L. A. (2011). The Thr- and Ala- rich hyperactive antifreeze protein from inchworm folds as a flat silk-like β-helix. *Biochemistry*, *50*, 4467-4478.

[52] Meyer, D. E., Trabbic-Carlson, K., & Chilkoti, A. (2001). Protein purification by fusion with an environmentally responsive elastin-like polypeptide: Effect of polypeptide length on the purification of thioredoxin. *Biotechnology Progress*, *17*, 720-728.

[53] Chow, D., Nunalee, M. L., Lim, D. W., Simnick, A. J., & Chilkoti, A. (2008). Peptide-based biopolymers in biomedicine and biotechnology. *Materials Science and Engineering R Reports*, *62*(4), 125-155.

[54] Kang, H., Kim, J., Chang, W., Kim, E., & Koo, Y. (2007). Heterologous expression and optimized one-step separation of levansucrase *via* elastin-like polypeptides tagging system. *Journal of Microbiology and Biotechnology*, *17*(11), 1751-1757.

[55] Trabbic-Carlson, K., Liu, L., Kim, B. & Chilkoti, A. (2004). Expression and purification of recombinant proteins from *Escherichia coli*: Comparison of an elastin-like polypeptide fusion with an oligohistidine fusion. *Protein Science*, *13*, 3274-3284.

[56] Trabbic-Carlson, K., Meyer, D. E., Liu, L., Piervincenzi, R., Nath, N., LaBean, T., & Chilkoti, A. (2004). Effect of protein fusion on the transition temperature of an environmentally responsive elastin-like polypeptide: a role for surface hydrophobicity? *Protein Engineering, Design & Selection*, *17*, 57-66.

[57] Lim, D. W., Trabbic-Carlson, K., Mackay, J. A., & Chilkoti, A. (2007). Improved non-chromatographic purification of a recombinant protein by cationic elastin-like polypeptides. *Biomacromolecules*, *8*, 1417-1424.

[58] Polayes, D. A., Parks, T. D., Johnston, S. A., & Dougherty, W. G. (1998). Application of TEV protease in protein production. *Methods in Molecular Medicine*, *13*, 169-183.

[59] Dougherty, W. G. & Parks, T. D. (1991). Post-translational processing of the tobacco etch virus 49-kDa small nuclear inclusion polyprotein: Identification of an internal cleavage site and delimitation of VPg and proteinase domains. *Virology*, *183*, 449-456.

[60] Raran-Kurussi, S., Tözsér, J., Cherry, S., Tropea, J. E., & Waugh, D. S. (2013). Differential temperature dependence of tobacco etch virus and rhinovirus 3C proteases. *Analytical Biochemistry*, *436*, 142-144.

[61] Fang, L., Jia, K., Tang, Y., Ma, D., Yu, M., & Hua, Z. (2007). An improved strategy for high-level production of TEV protease in Escherichia coli and its purification and characterization. *Protein Expression and Purification*, *51*, 102-109.

[62] Phan, J., Zdanov, A., Evdokimov, A. G., Tropea, J. E., Peters III, H. K., Kapust, R. B., Li, M., Wlodawer, A., & Waugh, D. S. (2002). Structural basis for the substrate specificity of tobacco etch virus protease. *The Journal of Biological Chemistry*, *277*, 50564-50572.

[63] Sun, C., Liang, J., Shi, R., Gao, X., Zhang, R., Hong, F., Yuan, Q., & Wang, S. (2012). Tobacco etch virus protease retains its activity in various buffers and in the presence of diverse additives. *Protein Expression and Purification*, *82*, 226-231.

[64] Mohanty, A. K., Simmons, C. R., & Wiener, M. C. (2003). Inhibition of tobacco etch virus protease activity by detergents. *Protein Expression and Purification*, *27*, 109-114.

[65] Carrington, J. C. & Dougherty, W. G. (1988). A viral cleavage site cassette: Identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proceedings of the National Academy of Sciences of the USA*, *85*, 3391-3395.

[66] Tözsér, J., Tropea, J. E., Cherry, S., Bagossi, P., Copeland, T. D., Wlodawer, A., & Waugh, D. S. (2004). Comparison of the substrate specificity of two potyvirus proteases. *The Federation of European Biochemical Societies Journal*, *272*, 514-523.

[67] Kapust, R. B., Tözsér, J., Copeland, T. D., & Waugh, D. S. (2002). The P1' specificity of tobacco etch virus protease. *Biochemical and Biophysical Research Communications*, *294*, 949-955.

[68] Parks, T. D., Howard, E. D., Wolpert, T. J., Arp, D. J., & Dougherty, W. G. (1995). Expression and purification of a recombinant tobacco etch virus NIa proteinase: Biochemical analyses of the full-length and a naturally occurring truncated proteinase form. *Virology*, *210*, 194-201.

[69] Nunn, C. M., Jeeves, M., Cliff, M. J., Urquhart, G. T., George, R. R., Chao, L. H., Tscuchia, Y., & Djordjevic, S. (2005). Crystal structure of tobacco etch virus protease shows the protein C terminus bound within the active site. *Journal of Molecular Biology*, *350*, 145-155.

[70] Kapust, R. B., Tözsér, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D., & Waugh, D. S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Engineering*, *14*, 993-1000.

[71] Cabrita, L. D., Gilis, D., Robertson, A. L., Dehouck, Y., Rooman, M., & Bottomley, S. P. (2007). Enhancing the stability and solubility of TEV protease using in silico design. *Protein Science*, *16*, 2360-2367.

[72] Kapust, R. B. & Waugh, D. S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science*, *8*, 1668-1674.

[73] Bu, S. (2011), Modification, expression, and purification of hyperactive antifreeze proteins from insect *Tenebrio molitor*, Cleveland State University.

[74] Kuiper, M. J., Lankin, C., Gauthier, S. Y., Walker, V. K., & Davies, P. L. (2003). Purification of antifreeze proteins by adsorption to ice. *Biochemical and Biophysical Research Communication*, *300*, 645-648.

[75] NBEcutter v2.0 from New England Biolabs website. Retrieved from http://tools.neb.com/NEBcutter2/index.php

[76] Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry, 182*, 319-326.

[77] DNA to protein translation from http://insilico.ehu.es/translate/.

[78] Protein molecular weight marker from PIR from

http://pir.georgetown.edu/pirwww/search/comp_mw.shtml

[79] Clark, E. D. (1998). Refolding of recombinant proteins. *Current Opinion in Biotechnology*, *9*, 157-163.

Appendices

A.1 The protocol of DNA annealing

1. The 50 µl of reaction mixture, which contains 4 µl of forward primer, 4 µl of reverse primer, 21 µl of Quick ligation buffer (New England Biolabs), and 21 µl of nuclease free water (New England Biolabs), are added into PCR tubes.
2. Place PCR tubes in a thermo cycler (Thermo Electron Corp.)
    - The sample cycle C27 is selected.

The description of sample cycle C27

- Heat up to 95 ^{o}C for 5 min and cool down to 25 ^{o}C over 70 cycles.
- Keep at 4 ^{o}C for 55 minutes

A.2 The protocol of 1.5% agarose gel preparation for DNA sample gel electrophoresis

1. Add 75 ml 1X TAE buffer and 1.125 g agarose (molecular biology grade) in a beaker.
2. Stir (or heat up) mixture until a clean solution is formed.
3. If heat is applied, allow it to cool down.
4. Add 7.5 µl ethidium bromide and mix well by stirring gently at vent hood.
5. Pour the solution into gel casting cassette.
6. Place the combs for sample loads.
7. Allow it dry at room temperature.
8. Remove the combs.
9. Place the gel in a gel tank.
10. Pour 1X TAE buffer in a gel tank until the gel is submerged.

A.3 The protocol of PCR

The 25 μl of PCR for screening:

1. The 25 μl of PCR mixture, which contains 12.5 μl PCR master mix (Promega), 2.5 μl T7 forward primer, 2.5 μl T7 reverse primer, 6.5 μl nuclease free water, and 1 μl plasmid DNA, is added into PCR tubes.
2. Place PCR tubes in a thermo cycler (Thermo Electron Corp.).
   - The sample cycle C16 is selected.

The 100 μl of PCR for priming out and amplification

1. The 100 μl of PCR mixture, which contains 50 μl PCR master mix, 26 μl nuclease free water, 10 μl forward primer, 10 μl reverse primer, and 4 μl plasmid DNA, is added into PCR tubes.
2. Place PCR tubes in a thermo cycler (Thermo Electron Corp.).
   - The sample cycle C16 is selected.

   The description of sample cycle C16

   - 2 minutes at 95 °C for melting the DNA.
   - 30 cycles of:
     - 1 min at 72 °C for annealing the primers to the DNA.
     - 30 seconds at approximately 5 °C below the melting point of the primers for primer extension.
     - 1 minute at 72 °C for primer extension.

- Refrigeration at 4 °C for several hours until the sample is removed from the thermocycler.

A.4 The protocol of sonication

1. Resuspend the pellets with added 10 ml of PBS.

2. Transfer the resuspended samples into a 50 ml falcon tube.

3. Cool down on ice for 15 minutes.

4. Place the falcon tube into a beaker containing ice and water.

5. Before sonication, clean the probe of the sonicator with ethanol and distilled water.

6. Place the probe of the sonicator in the faclon tube.

7. Turn on the sonicator.

8. Press the Tune key, and turn the amplitude control knob to the settings of 3, 6, and 10. At each setting, adjust the amplitude control knob to keep the bar graph less than 50 % in order to maintain a minimum meter reading less than 10 %.

9. Press the stop key and turn the amplitude control knob back to zero.

10. After sonication, clean the probe of the sonicator with ethanol and distilled water again.

11. Cool down the falcon tube on ice for 20 min.

A.5 The protocol of ITC purification

1. The sonicated sample is transferred into a centrifuge tube for the cold centrifugation.
2. Cold centrifugation
    - Place a centrifuge tube in rotor and balance a centrifuge tube.
    - Tighten the lid and set the centrifuge for 20 minutes at 20,000 xg.
3. The supernatant is descanted into a centrifuge tube for the warm centrifugation.
4. Incubate the sample at above T_t of the ELP tagged protein for 2 hours.
5. Warm centrifugation
    - Place a centrifuge tube in rotor and balance a centrifuge tube.
    - Tighten the lid and set the centrifuge for 20 minutes at 20,000 xg.
6. The warm supernatant is discarded, a warm pellet is resuspended with added 5 ml of PBS on ice.
7. Step 2-5 is repeated.
8. The warm supernatant is discarded, a warm pellet is resuspended with added 3 ml of PBS on ice.
9. Final cold spin is performed, and the cold supernatant is collected in a 50 ml falcon tube as it is the final protein solution.

A.6 The protocol of Cold finger purification

1. Set the water bath to -0.5 ^{o}C.

2. Place the cold finger in a cooled water beaker until a thin layer of ice is formed on the cold finger.

3. Replace a water beaker to sample solution glass, and place a magnetic stir bar in sample solution.

4. Lower temperature gradually from -0.5 to -2.0 ^{o}C for 8 hours (i.e. -0.1 °C /30min) with stirring 350 rpm.

5. Wash ice on the cold finger with distilled water and replace sample solution beaker into an empty beaker.

6. Harvest sample beaker, which is the liquid fraction

7. Set the water bath to 1 °C and harvest the ice on the cold finger until it detaches from the cold finger.

8. Let the ice melt down and collect the solution, which is the ice fraction.

A.7 The bacterial protein extraction reagent, B-PER, protocol

1. Collect 1 ml of bacterial cells in 1.5 ml microcentrifuge tubes before and after induction as expression time lapses.

2. Centrifuge bacterial cells at 14,000 rpm in a microcentrifuge for 5 minutes and discard the supernatant.

3. Re-suspend the bacterial cell pellet in 100 µl B-PER reagent by vigorously vortexing.

4. Centrifuge at 14,000 rpm in a microcentrifuge for 5 minutes and collect the supernatant (soluble fraction). The pellet is considered to be the insoluble fraction.

5. Re-suspend the pellet in 100 µl B-PER reagent by vigorously vortexing again.

6. Use 15 µl of both the soluble and insoluble fraction for SDS-PAGE.