Expression of Temporin G, an antimicrobial peptide from *Rana temporaria* using an Intein self-cleavage system in *Escherichia coli*

Sabah Mushrafin Kagri

Thesis submitted to The University of Nottingham for the degree of Master of Research

July 2020

Abstract

Temporins are a class of antimicrobial peptides (AMPs) naturally secreted from the skin of a common frog *Rana temporaria* especially under stressful condition. Temporin-G (TMPG) is a subclass of temporins with potent broad-spectrum antimicrobial activities. Although solid phase peptide synthesis (SPPS) is a common method used for the chemical synthesis of peptides, this method is technically demanding, involve the use of hazardous chemicals, and requires a well setup facility and the associated equipment for the overall operation which could be challenging for many laboratories. One of the promising alternatives for AMPs biosynthesis is by using the bacterial expression system. Expression of AMPs in *Escherichia coli* had been previously documented; however, the expressed AMPs would carry the vector-borne DNA sequences flanking the AMP-encoding gene. This gives rise to the undesired amino acid residues co-translated with the AMP which would alter the characteristics of the peptide expressed. Hence, we aim to construct an AMP recombinant expression system using the IMPACT-TWIN plasmid expression vector (intein mediated purification via affinity chitin-binding tag-two intein) for the biosynthesis of TMPG in its native form. The TMPG-encoding gene was cloned into pTWIN1 vector via N-terminal fusion to produce the pTWIN1-CBD-Int-TMPG vector. Transformants carrying pTWIN1-TMPG recombinant plasmid produced a band size of 277bp and DNA sequencing of the pTWIN1-TMPG amplicon further confirmed the orientation and sequence identity of the cloned TMPG encoding gene. Subsequently, the TMPG encoding gene was transformed into *E. coli* BL21 (DE3) cells and optimized for peptide expression by one-factor-at-a-time (OFAT) approach. That is, optimization of the various parameters influencing peptide expression in the host cell such as Isopropyl β- d-1-thiogalactopyranoside (IPTG) concentration, induction temperature, duration of induction, and the effect of agitation. Following the OFAT approach, the optimum condition for the induction of expression was determined to be using 0.9 mM IPTG at 25^{o}C for 48h, coupled with shaking at 180rpm. The optimized induction produced a significantly higher protein expression which was 8-fold more than the pre-optimized conditions in the target pellet fraction and 7-fold more in the clarified cell lysate. Intein mediated purification system represents a promising bacterial recombinant expression system for the biosynthesis of native AMPs and a high peptide expression level was achieved following optimization for potential bioproduction of TMPG in the future.

Acknowledgment

It is a great honor and accomplishment for me to have the opportunity to carry out my Master's in Research degree (MRes) in University of Nottingham Malaysia under faculty of Bioscience. I would like to thank all the people who helped me make it through this master's degree in University of Nottingham, Malaysia campus. Special thanks for my main supervisor, Dr. Le Cheng Foh for his constant guidance, support and valuable time to help me through the difficult challenges, I wouldn't have accomplished this thesis without his help, and Thanks to my co-supervisor Prof. Dr. Sandy Loh for her constant support in making sure everything went well.

The lab technicians – kak Siti, kak Asma and kak Khadija made my lab life much easier and more tolerable by helping me out all the time with any lab related issue I faced. I would also like to thank all my friends and family members for their mental support, unconditional love, and positivity towards my work and whole master's journey.

List of Abbreviation

%	Percent
× g	Relative centrifugal force (rcf)
°C	Degree Celsius
µ	Micro
BSA	Bovine serum albumin
PBS	Phosphate Buffered-Saline
CBD	Chitin binding domain
ddH_2O	Double distilled water
LB	Luria-Bertani
LB/Amp	LB agar supplemented with 100ug/ml Ampicillin
AMP	Antimicrobial Peptide
RE	Restriction enzyme
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
OD_{600}	Optical density at the wavelength of 600 nm
MCS	Multiple cloning site
E. coli	*Escherichia coli*
R.temporaria	*Rana temporaria*
R.esculenta	*Rana esculenta*
R.erythraea	*Rana erythraea*
B.cereus	*Bacillus cereus*
S.aureus	*Staphylococcus aureus*
kDa	Kilo Dalton
DMSO	Dimethyl sulfoxide
PMSF	Phenylmethylsulfonyl fluoride
EDTA	Ethylenediaminetetraacetic acid
IPTG	Isopropyl B-D-1-thiogalactopyranoside
TMPG	Temporin
MHB	Mueller Hinton Broth
SOC	Super Optimal broth with Catabolite repression

APS	Ammonium persulfate
TEMED	Tetramethyl ethylenediamine
TBE buffer	Tris-borate-EDTA buffer
NB	Nutrient broth
SPPS	Solid phase peptide synthesis (SPPS)

Reagent list

Reagent	Brand
Coomassie Blue	Sigma-Aldrich
Glycerol	Bio-Rad
Bovine serum Albumin	Sigma-Aldrich
SOB medium	Conda
Agarose powder	1st base
Ampicillin	Nacalai tesque
Ethanol	Merck
Methanol	RCI labscan
EDTA, disodium salt dihydrate	R & M, UK
Sodium Dodecyl Sulfate	Sigma-Aldrich
Tris base	Bio-Rad
Ammonium Persulphate	Sigma-Aldrich
Sodium Hydroxide	Merck
Nutrient broth	Merck
LB agar broth	Lennox
TEMED	Fisher Scientific
Sybr safe dye	Invitrogen, USA
PMSF	Thermo Fisher
Bradford reagent	Sigma-Aldrich
Triton X-100	Fisher- scientific
B-marcaptoethanol	Bio Basic Inc
Anti CBD Monoclonal Antibody	NEB
Secondary antibody (Anti-mouse lgG)	R and D systems
Sap1 and Pst1-HF	NEB
IPTG	Promega
Polypeptide SDS-PAGE standard ladder	Bio-Rad

3-Color high Range Protein Marker	SMOBIO
T4 DNA Ligase	NEB
MHB	Oxoid
Nutrient broth	Oxoid
LB agar	Lennox
LB broth (miller)	Merck

List of figures

Figure 1.1: Mechanism of T7 RNA polymerase expression……..17

Figure 1.2: IMPACT pTWIN1 vector map, showing the multiple cloning site, the origin of replication and the restriction enzyme sites……20

Figure 1.3: Fusion of the intein tag to the N-terminus of the target peptide in pTWIN1 vector. The release of the peptide was induced by the change in pH and temperature……..21

Figure 3.1: Colony PCR screening of pTWIN1 vector and pTWIN1-TMPG transformants……..31

Figure 3.2: Confirmation of sequence identity of the inserted TMPG-encoding gene using NCBI BLASTn tool……...32

Figure 3.3: Optimization of cell lysis using probe sonication at varying pulse time, sonication cycle, and incorporation of free thaw step……34

Figure 3.4: Effect of varying IPTG concentrations on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction……..36

Figure 3.5: Effect of varying temperature on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction……37,38

Figure 3.6: Effect of different induction duration on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction……..39

Figure 3.7: Effect of shaking on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction……40

Figure 3.8: The flow chart of the overall optimization process tested in this study……41

Figure 3.9: Comparison of recombinant protein expression profiles before and after optimization……..42

List of tables

Table 1.1: Primary structure of the Temporin peptides ……..15

Table 2.1: Lists the equipment, software, detection and purification kit used to carry out this study ……..23,24

Table 2.2: List of all oligonucleotides used in this study ……25

Table 2.3: Composition of stacking and resolving gels ……29

Table 3.1: Absorbance and concentration values of the protein before and after optimization ……42

Table of Contents

Chapter 1: Introduction and literature review 11

1.1 Research problems and significance of study 11

1.2 Characteristics and properties of AMPs 12

1.3 Synthesis and production of AMPs 13

1.4 Different peptide expression systems 16

1.5 Expression of AMP in *E. coli* cells 17

1.6 IMPACT-TWIN system 19

1.7 General aim and specific objectives 21

Chapter 2: Materials and Methods 23

2.1 Equipment, Software and kits 23

2.2 Bacterial strains and plasmid expression vector 24

2.3 Preparation of chemically competent *E. coli* cells 24

2.4 Gene design and oligonucleotide synthesis 24

2.5 Oligonucleotide assembly and construction of pTWIN1-TMPG vector 25

2.6 Expression of CBP-Int-TMPG protein construct 27

2.7 Cell Lysis 27

2.8 Optimization of CBP-Int-TMPG protein construct expression 27

2.9 SDS-PAGE gel electrophoresis for protein separation 28

2.10 Western immunoblot assay 29

2.11 Bradford assay 30

Chapter 3: Results 31

3.1 Construction of recombinant pTWIN1-TMPG expression plasmid 31

3.2 Efficiency of cell lysis by sonication 33

3.3 Optimization of recombinant TMPG-Int-CBD protein construct expression 35

3.3.1 Effect of IPTG concentration on protein expression 35

3.3.2 Effect of induction temperature on protein expression 37

3.3.3 Effect of induction duration on protein expression 38

3.3.4 Effect of agitation on protein expression 40

3.4 Comparison of recombinant protein expression before and after optimization 41

Chapter 4 : Discussion 43

Chapter 5: Conclusions and Future Recommendations 48

References 49

Appendices 54

Chapter 1: Introduction and literature review

1.1 Research problems and significance of study

Antibiotics are antimicrobial agents used in the treatment of infections by inhibiting the growth or killing the causative microorganisms. With the rise of antibiotic-resistant bacteria (ARB), simple infections could now turn lethal with ineffective antibiotic treatment. By 2050, 30 million people are expected to be at risks of mortality due to ARB (Kim, et al., 2019). Antibiotic resistance is the ability of microorganisms to develop tolerance against the inhibitory effect of antibiotics. There is an urgent need to develop new therapeutic agents to fight infections caused by microorganisms and to ensure the continuous development of new antibiotics in the drug discovery pipeline. Antimicrobial peptides (AMPs) have gained great interest in the past decade for their role as a potential antimicrobial agent. Almost all living organisms, including bacteria and fungi, produce AMPs as part of their native first line defense mechanisms and thus have coined the term, host-defense peptides. These small molecules represent a promising alternative to conventional antibiotics, in particular infections involving ARB.

AMPs are not the same as peptide-based antibiotics. The latter are small peptide-derived molecules exhibiting antibacterial activity which are produced via a process called non-ribosomal peptide synthesis (NRPS). Some examples of peptide-based antibiotics include the beta-lactams penicillin and the glycopeptides vancomycin. On the other hand, AMPs are produced by the classical ribosomal translation mechanism from DNA, mRNA, and eventually into protein (peptide). These compounds are initially synthesized in the form of pre-protein which is then processed to the mature active AMP. For example, LL-37 is an AMP produced as a precursor protein which undergoes modification to become the mature human cathelicidin AMP (Andersson, et al., 2016). Antibiotics and AMPs differ in their chemical composition. Antibiotics are more diverse and complex in structure, while not limited by biosynthesis based on ribosomal translation alone and commonly contain a mixture of amino acids with other non-canonical amino acids. In contrast, AMPs usually contain only the natural amino acids with minimal residual/functional group modifications.

A common method used for small- and large-scale bioproduction of AMPs is through the chemical solid phase peptide synthesis (SPPS) method. In modern laboratories, SPPS is performed using the automated peptide synthesizer. However, one major drawback is the high cost associated with the availability of facilities, equipment, technical expertise, and economy of scale. Therefore,

availability of SPPS is often limited to well-funded laboratories or commercial service providers. Although individual researchers can outsource such demand to the commercial service providers, however, the high cost of peptide synthesis service represents a significant bottleneck in drug discovery using AMPs. One of the main reasons is that drug design and generation of novel AMP analogues, as what we have previously experienced, would involve the testing of as many candidates as possible in order to identify the few leads of interest from screening. Therefore, the "start-up" cost required could be staggering especially for many smaller research groups. Even with such a diverse application of AMPs, there are many practical limitations associated with the direct extraction of AMPs from the native hosts. For instances, the naturally occurring AMPs are not usually abundantly expressed, slow rate of synthesis, laborious and low yield from purification, and prone to proteolytic degradation. Therefore, the diverse applications would undoubtedly continue to drive for the scalable production of AMPs. For these reasons, we recognize the need to develop a suitable approach for the biosynthesis of AMPs which lay the foundation for the current study. The AMP selected as the candidate in this study was temporin G (TMPG) for its potent, rapid killing kinetic, and broad spectrum antibacterial activity previously reported. Here we aim to develop a suitable biosynthesis approach for the recombinant expression of TMPG in *Escherichia coli* as the host bacterial expression system.

1.2 Characteristics and properties of AMPs

AMPs are mainly small-sized cationic peptides containing lysine (Lys) and arginine (Arg) amino acid residues in the primary structure backbone, giving the AMPs its characteristic cationicity. However, some AMPs with anionic properties have also been reported in some papers. Based on the 3D structure, AMPs are divided into four groups: alpha-helical (e.g. human cathelicidin LL-37), beta-sheet (e.g. human a-defensins), extended length structure, and hairpin/loop antimicrobial peptides (e.g. indolicidin). According to the antimicrobial peptide database, more than 1500 AMPs had been discovered to date (http://aps.unmc.edu/AP/main.php). Based on the antimicrobial activities, they can be classified as antiviral, antibacterial, antifungal, anticancer, and antiparasitic peptides. Other promising properties such as mechanisms of action, broad spectrum antibacterial activity, target specificity, tendency to develop resistance, small size, biodegradability, and others have made AMPs a highly promising class of alternative antimicrobials for therapeutic applications. Unlike conventional antibiotics, many AMPs are able to exert non-specific disruption to the bacterial cell membrane and targeting multiple cellular targets simultaneously. Such unique actions have made AMPs an efficient bactericidal molecule and less prone to resistance by the target pathogens.

Moreover, the antimicrobial potency of AMPs could be further synergized with combination treatment using antibiotics. AMPs have also been used in the food and agriculture sectors.

AMPs have gained great interest in recent years for their potential activity used for developing various antibiotics. They have a wide spectrum activity against pathogenic microorganisms as well as against cancerous cells. According to the surveys available, it was shown that 90% of the AMPs are cationic in nature and only 10% are anionic peptides. AMPs are mainly cationic due to the presence of lysine (Lys) or arginine (Arg) residue, which allows them to interact with the anionic phospholipid layer and lipopolysaccharides (LPS) of bacterial cell membranes. Hence in return, these AMPs kill bacteria by targeting some important pathways such as protein synthesis and cell wall formation (Sang, et al., 2018). AMPs are also shown to have activity against intracellular targets without causing extensive membrane damage. The human beta-defensin AMP were translocated across the outer and inner membranes of *E. coli* without lysing the cell suggesting the antimicrobial activity depends on cytoplasmic interactions (Andersson, et al., 2016). Another broad-spectrum AMP, indolicidin (a bovine cathelicidin), was shown to inhibit DNA synthesis without causing cell lysis (Subbalakshmi & Sitaram, 1998).

AMPs are found in all kingdoms of life including both living prokaryotic (e.g., bacteria) and eukaryotic organisms (e.g., protozoan, fungi, plants, insects and animals). The AMPs in animals are present as the first line of innate immunity defense, found in the tissues and organs those are exposed to pathogens such as virus, bacteria and fungi (Bahar & Ren, 2013). The first animal originated AMP was isolated from leukocyte cells of the rabbit in 1956 and was called defensin. Epithelial cells, phagocytes, lymphocytes and lymphs are other eukaryotic cells involved in the production of AMPs (Narayana & Chen, 2015). Such peptides have also been found in a wide range of amphibian species and it has been shown to have a clear connection between frog skin peptides and their related ability against pathogens.

1.3 Synthesis and production of AMPs

Synthesis of peptide has been a significant importance in protein chemistry to study protein function and identify and characterize the proteins. Solid phase peptide synthesis (SPPS) is a synthetic strategy first introduced by Bruce Merrifield for synthesizing peptide (Made, et al., 2014). In this method, a chemically unreactive polystyrol (resins or beads) is used as a solid support and the first step involves coupling the C-terminus of the first amino acid to this solid support. The resin used in this method acts as a protecting group for the C-terminal. SPPS is basically a method involving repeated cycles of deprotection-wash-coupling-wash to generate the peptide by adding one amino

acid at a time via coupling reaction. Using a filtration procedure, the liquid reagents and by products are washed away helping to retain the immobilized protein. To prevent side reaction and undesired chemical bond formation between the amine and the carboxylic acid functional group, it is mandatory to use the protecting group strategy like Boc (*tert*-butyloxycarbonyl) and Fmoc (fluorenyl methoxycarbonyl) for SPPS. Boc is used as protective group and can be removed using trifluoroacetic acid (TFA). The Fmoc group on the other hand serves as a N-terminal α-amino protecting group and to allow the growth of the peptide chain it is removed with a strong base such as piperidine (Stawikowski & Fields, 2002).

Nevertheless, SPPS is observed with several disadvantages such as swelling of the matrix, protecting amino acid residues and the type of linker used also needs to be taken into consideration. The repetitive stepwise synthesis of the peptide is a tedious procedure as well. SPPS doesn't favor the synthesis of peptides with more than 50 amino acids, because larger peptides may cause impurities and inter/intramolecular aggregation reducing the overall yield. Even though short AMP's can be prepared by chemical synthesis, there is a constant need to produce high quality peptides with low cost, therefore a potential alternative to SPPS like recombinant protein technology is needed (Geldart, et al., 2016).

Under stressful condition or upon injury, the frogs of the *Ranidae* family synthesize AMPs which are eventually released in the skin in high concentrations. These secreted peptides play a vital role in maintaining the immunity of the frog against pathogenic microorganisms. Among these defensive group of peptides, the temporin family which consists of 8 to 27 amino acids in length is the smallest AMPs found in nature. Temporins are highly hydrophobic, relatively short and weakly charged peptides with low net positive charge. The property of having low cationic property makes temporins unique compared to AMPs from other sources (Mangoni & Shai, 2009). Their secondary structure is C-terminally α-helical and in linear arrangement, which contributes to their cytotoxicity in destabilizing the membranes of microbial or cancerous cells in hydrophobic environment. Most of the temporins have antimicrobial activity specially against gram positive bacteria such as methicillin-resistant *Staphylococcus aureus* (MRSA) and *Enterococcus faecalis* but some are also active against gram negative bacteria (Wu, et al., 2017).

Temporin-like peptides were first identified from the skin of the Asian frog *Rana erythraea* and the European hybrid frog *Rana esculenta*. These peptides were referred as Vespa-like peptides due to their structural and sequence similarity to the peptides isolated from venom of different wasp species. Some of these peptides, termed temporins, were isolated from the skin of a common frog, *Rana temporaria, and* shown to have antimicrobial activity (Mangoni & Shai, 2009). Compared

to other AMPs from frog skin, temporins are post translationally amidated at the C-terminus (carboxyl end) with low net positive charge due to the presence of one or two basic residues. Temporins are synthesized as precursors, where the N-terminal domain and acidic propeptide region containing Lys-Arg site are highly conserved. The matured temporin is generated by cleavage at the Lys-Arg prohormone processing site, and the Glycine (Gly) residue present in the C-terminus serves as a substrate for peptidyl-glycine α-amidating monooxygenase to produce C-terminally amidated temporins. Generally, temporins are shown to be active against gram positive bacteria (e.g., *Staphylococcus* and *Streptococcus* spp.), with MIC ranging from 2.5-20 mM (Patocka, et al., 2018).

TMPG is one such modified small AMP that is derived from the skin of *R. temporaria* which species is widespread throughout European region. This is a semi-aquatic amphibian of the Ranidae family which has the highest distribution among all other frog families. However, not much research has been done specifically on TMPG compared to other temporin peptides (Rinaldi & Conlon, 2013). Table 1.1 shows the list of the temporin peptides. One attempt to characterize the biological property of TMPG was made long ago by Simmaco, et al., (1996), where temporin A, B, F, G, and L showed antimicrobial activity against *B. megaterium* and *E. coli* D21. In this study, the recovery of the TMPG was too little to further analyze the properties.

Table 1.1: Primary structure of the Temporin peptides (Simmaco, et al., 1996).

Peptide	Sequence
Temporin A	FLPLIGRVLSGIL-NH2
Temporin B	LLPIVGNLLKSLL-NH_2
Temporin C	LLPILGNLLNGLL-NH_2
Temporin D	VLPIIGNLLNSLL-NH_2
Temporin E	VLPIIGNLLNSLL-NH_2
Temporin F	FLPLIGKVLSGIL-NH_2
Temporin G	**FFPVIGRILNGIL-NH_2**
Temporin H	LSP---NLLKSLL-NH_2
Temporin K	LLP---NLLKSLL-NH_2
Temporin L	FVQWESKFLGRIL-NH_2

1.4 Different peptide expression systems

Recombinant DNA technology is among the most effective approach to produce proteins, peptides and enzymes. Recombinant production of peptides involves cloning the desired gene encoding the target peptide in a specific vector. The gene then translates into a peptide/protein using the host cellular mechanism. The use of recombinant protein technology to produce peptides or proteins has been considered as one of the most effective methods mainly due to its advantages such as reduced production cost and scale-up feasibility. There are many host systems available to express recombinant peptides, some of which includes bacteria, fungi, unicellular algae and yeast. Among them, the most commonly used host system for *in vivo* biosynthesis of AMP is the bacterial host system which includes *Lactic Acid Bacteria* (LAB), *Escherichia coli* (*E. coli*) and *Pseudomonas* expression systems. Besides, *Pichia pastoris* has been used as the potential host for yeast system. Among bacteria, *E. coli* is the most common bacterium used as a host due to its long history of safe usage in industries.

The bacterial expression system is broadly utilized due to its advantages compared to other expression system in case of producing small antimicrobial peptides (Gaiser, et al., 2011). *E. coli* strain K-12 and its derivatives DH1, DH5α, MG1655, RV308 and W3110 are most widely used strains but were known to produce high amount of acetate which effects the growth of the cell. Therefore, *E. coli* B strains were used instead because of their low acetate accumulation, from which the *E. coli* BL21 strain was selected. It compromises of T7 RNA polymerase promoter which allows expression from a T7 promoter expression system. Engineered *E. coli* strain (BL21 DE3) carries gene that encodes for the RNA polymerase of bacteriophage T7. This polymerase is different from *E. coli* RNA polymerase and the two are used differently. The activity of T7 RNA polymerase is controlled by isopropyl B-D-1-thiogalactopyranoside (IPTG) inducible promoter known as *lacUV5*. Normally the *lac* promoter is bounded by *lac* repressor, lacl, which represses transcription and prevents the T7 polymerase to be synthesized (Overton, 2014). However, in the presence of IPTG, T7 polymerase gene is transcribed and translated as the lacl is released from the *lac* promoter region. In return the T7 polymerase activates the T7 promoter, which is situated in the expression plasmid, downstream of which the gene of interest (peptide gene) is cloned as shown in Figure 1.1.

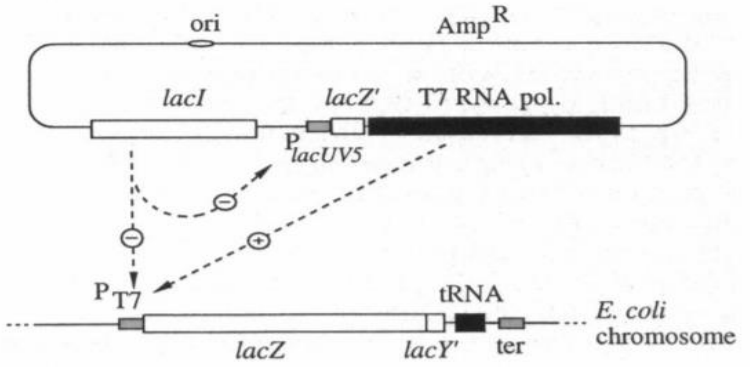


Figure 1.1: Mechanism of T7 RNA polymerase expression in the bacterial plasmid. The plasmid contains genes encoding for T7 RNA polymerase, *lacI*, and *lacUV5* promoter. The + and – symbols stand for positive and negative controls by the T7 RNA polymerase and *lac* repressor, respectively (Makarova, et al., 1995).

1.5 Expression of AMP in *E. coli* cells

Previously *E. coli* has been used to produce several short peptides which includes lactoferricin, dermicin, moricidin, defensins, and buforin. It is also one of the first organisms whose genome was sequenced and well-characterized with established protocol (Overton, 2014). The growth rate of *E. coli* cells is rapid and in ideal environmental conditions such as glucose salt media, the time it takes to multiply is about 20 mins only (Rosano & Ceccarelli, 2014). According to Overton, (2014), *E. coli* is recognized by the drug authorities such as the drug production companies, as a cost-effective host for production of recombinant proteins. In a few cases, the AMP producing organism can be itself be a target organism for the produced AMP. Some of the ways to overcome this problem are by fusing it with another protein, or, using an inducible promoter to balance amount of peptide produced. For the success of peptide expression using recombinant technology, several other factors such as vector system, codon usage, the signal peptide and the fusion partner should also be taken into consideration.

In the past work, cationic antimicrobial peptide had been effectively expressed in *E. coli* cells and purified utilizing various purification methods. Some of which include the 6X histidine tag, enzymatic and chemical cleavage, use of fusion tags and sometimes directly expressed. Using chemical and enzymatic cleavages, the peptide also undergoes side chain modifications and nonspecific cleavages when incubated for a longer time using high temperature. Although excessive yields of products may also result from the approaches mentioned previously, these products often form insoluble inclusion body or aggregates. These inclusion body are made up of misfolded and denatured proteins which are not in its active form. The changes in the disulfide bridge of the aggregates cause the proteins to remain inactive despite going through the denaturation and renaturation process (Wong, et al., 2018). To overcome the problem of inclusion bodies formation, researchers have also made use of fusion tags which fall into two categories: solubility tags and

affinity tags. Examples of solubility tags include glutathione S-transferase (GST), maltose binding protein (MBP), small ubiquitin modifying protein (SUMO) and thioredoxin (TRX), that are fused with the peptide thereby help solubilizing the peptide. The use of tags and fusion partners are used because they improve the production yield of protein, solubility, folding and aids the purification process. SUMO was shown to be a good carrier for the soluble expression of peptides and with the presence of specific and robust SUMO protease, this adds on the advantage of the system (Li, 2011). SUMO and MBP translocate the protease targeted protein to different cellular location to prevent the protein from being degraded. GST and MBP tags play a dual role as both solubility and affinity tags and are affinity-purified by glutathione and amylose resin, respectively (Li, 2011).

Other affinity tags such as the poly-His, calmodulin-binding peptide (CBP), chitin-binding domain (CBD), cysteine protease domain tag, and strep tag have also been used to purify and extract the peptide. The tagged recombinant proteins can easily be detected by western immunoblot assay (WIA) due to the availability of commercial antibodies for the tags. The use of tags also allows one step purification due to the specific and tight binding efficiency of the tags to the resins available. An example is the recovery of His-tagged proteins using Ni^{2+} loaded nitrilotriacetic acid-agarose resins in the process of immobilized metal ion affinity chromatography. Adding on, the main problem with this fusion approach is the removal of the tag from the target peptide without changing the sequence of the latter. The removal of fusion tags from the peptide also needs chemical or enzymatic cleavage which adds to the expense and complexity of the system. Moreover, the peptide toxicity for the host cell itself and proteolysis of the final product also make the process unsuitable (Wang, et al., 2018).

Self-cleaving tags are a special group of fusion tags, serve as an alternative to the existing affinity strategy for simple and easy protein purification and tag removal. The tags undergo cleavage when triggered by low molecular weight compound or due to change of conformation (Costa, et al., 2014). Several AMPs were previously expressed in by fusing with intein tag. Inteins are part on the introns of a protein that can cleave themselves by undergoing autocatalytic cleavages from the precursor protein and rejoin the two surrounding flanking areas using a self-catalyzed protein splicing reaction. The advantage of using self-cleaving inteins is that they have an enzymatic nature, which omits the use of additional chemicals/enzymes to achieve cleavage. Inteins also play a vital role to form a construct that is expressed in a stable form as well as is not toxic to the host cells.

They are used with the MBP, CBD affinity tag or elastin – like peptide (ELP), which enables the combined proteins to be isolated by using chitin resin beads. The target peptide with either N-terminus or C-terminus can be generated upon excision (Li, 2011). Figure 1.3 shows example of a

fusion of the intein tag to the N-terminus of the target peptide. This method does not require enzymatic cleavage and goes through just one step purification to enhance the final yield of the peptide.

One such example of using intein mediated purification of protein was shown by Warren, et al., (2013) where ligation-independent cloning (LIC) cloning technique was used with the ΔI-CM intein pET vector system for successful expression and purification of proteins. The LIC vectors encoding the ΔI-CM intein which was fused to an affinity tag showed successful gene insertion without the use of ligation based cloning methods. Intein mediated expressions systems have also been used to express AMPs like bacteriocins, cecropins, β-defensins and moricin CM4 (Diaz, et al., 2012). However recombinant expression of TMPG peptide in *E.coli* cells have not been reported before. SMAP-29 cathelicidin-like peptide is an example of another AMP that have been expressed in E.coli cells using intein mediated system (Morassutti, et al., 2005).

1.6 IMPACT-TWIN system

Intein based technology is commercially available through the IMPACT-SYSTEM (New England Biolabs), in order to purify proteins and peptides by fusing the affinity tag, the CBD, C or N-terminal to the intein sequence. Therefore, in this study we used the IMPACT-TWIN (Intein Mediated Purification with an Affinity Chitin-binding Tag-Two Intein) system which compromises the use of commercial vectors such as pTWIN1 (7375 bp) and pTWIN2 (7192 bp) for expression of recombinant protein in a soluble form and for purification purpose. Both the vectors contain Sap 1 sites where the gene of interest is cloned between the intein tags. They both also make use of the modified Ssp DnaB intein (Intein 1) which is fused at its N-terminus to CBD and differ only in case of intein 2 where modified *Mxe* GyrA intein in pTWIN1, is replaced by *Mth* RIR1 intein in pTWIN2. However, the similarity among these two vectors is that they both contain a common multiple cloning sites to facilitate insertion and cloning as shown in Figure 1.2.

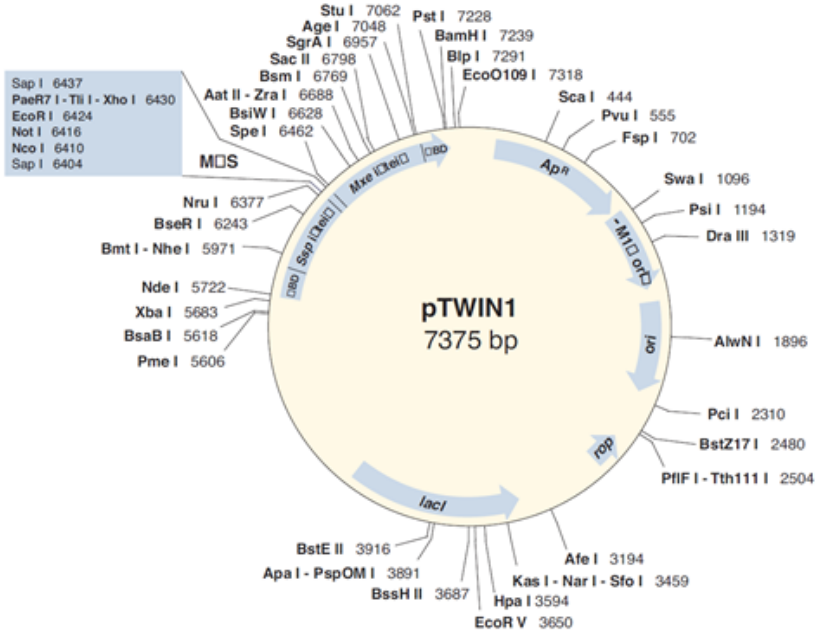


Figure 1.2: IMPACT pTWIN1 vector map, showing the multiple cloning site, the origin of replication and the restriction enzyme sites.

In some cases , thiol containing compounds such as dichlorodiphenyltrichloroethane (DDT) or cysteine is required to cleave the N-terminal sequence from the thioester intermediate that is formed within the intein sequence. For example, the intein in pTYB11 and pTYB12 vector undergoes cleavage at both N-terminal and C-terminal junctions with the use of DTT or cysteine. However, some proteins are sensitive to DDT and so its use as a cleavage reagent should be avoided (Mitchell & Lorsch, 2015). pTWIN1 plasmid catalyzes cleavage reactions based on pH and temperature shifts without addition of any thiol containing compound. There are many advantages of using the IMPACT-TWIN system compared to other purification strategies. The specific connection between the chitin resins and the CBD tag isolates highly purified peptides and utilizing the intein mediated purification system generates native peptide sequences without requiring the use of proteases or chemicals to cleave the affinity tag. Using IMPACT-TWIN system also produces native peptides without vector borne sequences or non-specific residues. The use of proteases would lead to the risk of proteolysis along with the need of higher temperature for incubation at which some peptides might not be stable. Additional chromatographic steps are also required to remove the protease. Whereas using intein mediated purification system, native recombinant proteins either containing N-terminal cysteine or a reactive C-terminal thioester can easily be purified by undergoing one-step affinity purification (Mitchell & Lorsch, 2015).

Affinity purification is widely used to simplify purification steps and has several advantages over normal chromatographic methodology such as the target proteins does not come into direct contact with the chitin resin, they can easily be obtained through one step purification, different proteins can be purified and is less time consuming procedure. The advantage of using this fusion system for protein purification is that peptide can be expressed and purified without an N-terminal

methionine residue or other non-specific residues. Adding on, using N-terminal fusion, no thiol reagent is required for cleavage which is important for DTT sensitive proteins. For cloning 5' end of the target gene, the use of SapI site allows isolation of protein without addition of extra N-terminal amino acids. However, if SapI site in pTWIN1 vector is chosen for cloning purpose, SpeI, PstI or BamHI sites, downstream of the polylinker should be used as 3' cloning site (Xu & Jr, 2003).

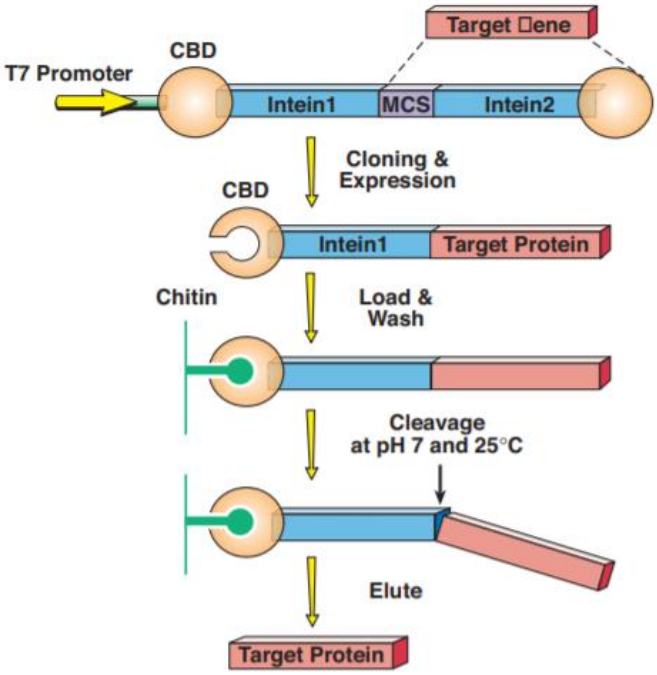


Figure 1.3: Fusion of the intein tag to the N-terminus of the target peptide in pTWIN1 vector. The release of the peptide was induced by the change in pH and temperature.

1.7 General aim and specific objectives

The aim of this study was to clone the TMPG-encoding gene fragment into the pTWIN1 expression vector followed by optimization of recombinant expression of the CBD-Int-TMPG protein construct from the expression host *E. coli* BL21 (DE3). Note that the TMPG peptide will first be expressed as a protein construct (N-terminally fused to the CBD and intein tags) from the vector which was the focus of the current study. Following that, the native TMPG can be excised from the protein construct through intein-mediated self-cleavage in the future. Therefore, the specific objectives of this project were:

1. To design and construct recombinant pTWIN1 plasmid carrying TMPG-encoding gene
2. To express gene encoding for TMPG peptide in a suitable bacterial host *E. coli* BL21 (DE3) using IPTG

3. To enhance the expression of CBD-Int-TMPG protein construct by optimizing the IPTG concentration, induction temperature, induction duration, and incorporation of agitation.

Chapter 2: Materials and Methods

2.1 Equipment, Software and kits

Table 2.1: Lists the equipment, software, detection and purification kit used to carry out this study.

Equipment	Company	Under Section
Minispin	Eppendorf	2.2
Water bath	Biocote	2.4
Shaker incubator	Lab companion	2.6
37°C incubator	Memmert	2.6
Centrifuge 5810R	Eppendorf	2.7
Take 3 Micro-volume plate	Biotek	2.5
Chemi Doc imaging system	Bio-Rad	2.8
Pierce G2 fast blotter	Thermoscientific	2.9
Mini rocker shaker	Biosan	2.4
T100 thermal cycler	Bio-Rad	2.5
Thermomixer	Eppendorf	2.4
SDS-PAGE electrophoresis tank	GE healthcare Biosciences	2.8
Electrophoresis power supply	GE healthcare Bioscience	2.8
White light Conversion screen (imaging)	Bio-Rad	2.8
Sonicator (Ultrasonic homogenizer)	Interscience	2.6

Software	Company
Gen5 2.04 software	Biotek
ImageLab version 5.2.1	Bio-Rad
Blast software	-
Image J	-

Kit	Brand
DNA clean and Concentrator™ -5	Zymo Research
NucleoSpin Plasmid EasyPure	Macherey-Nagel
Pierce ECL Western blotting substrate detection kit	Thermo-scientific

2.2 Bacterial strains and plasmid expression vector

E. coli DH5α was used as the host cell for propagation of pTWIN1 plasmid expression vector (New England Biolabs, UK) and *E. coli* BL21 (DE3) was used as the host cell for recombinant expression of CBP-Int-TMPG protein construct from the pTWIN1-TMPG vector. Both the *E. coli* strains were obtained from Dr. Fang Chee Mun (University of Nottingham Malaysia) and freeze-stored in -80°C for future use. For freeze stocking, 500µl of overnight grown cultures were added to 500µl of Brain heart infusion (BHI) medium and 15% (V/V) glycerol mixture. To revive the culture, the freeze-stock strains were passaged on nutrient agar at 37°C overnight before use.

2.3 Preparation of chemically competent *E. coli* cells

To prepare the chemically competent host cells, a total of 300µl of overnight grown *E. coli* culture was inoculated into 30ml of fresh LB broth (ratio 1:100) and cultivated at 37°C with agitation speed of 300rpm to OD_{600} 0.5-0.7. The cells were then pelleted at 4000rcf for 10 mins at 4°C and was resuspended in 25ml of ice-cold 0.1M $CaCl_2$ followed by incubation on ice for 30 mins. The cells were pelleted using the same conditions in the previous step and resuspended in 500µl of ice-cold 0.1M $CaCl_2$ with 15% glycerol. The cells were distributed into 50µl aliquots and stored at -80°C until use (Chang, et al., 2017).

2.4 Gene design and oligonucleotide synthesis

TMPG, a 13-amino acid AMP with peptide sequence FFPVIGRILNGIL was used in this study. The sequence was reverse-translated and codon-optimized for *E. coli* using the online tool Sequence Manipulation Suite for reverse translation (http://www.bioinformatics.org/sms2/rev_trans.html). The predicted TMPG-encoding gene sequence is 5'-TTTTTTCCGGTGATTGGCCGCATTCTGAACGGCATTCTG-3'. To facilitate cloning into pTWIN1 based on the N-terminal fusion approach, the TMPG-encoding gene was flanked with SapI and PstI sites and designed in such a way to produce the corresponding cohesive ends following annealing of oligonucleotides and to ease direct insertion of the annealed gene fragment into the double digested pTWIN1 backbone. The final oligonucleotide sequences generated and the two PCR primer sets are shown in Table 2.2. The forward primer PTW1_seq-F (5'-ACTGGGACTCCATCGTTTCT-3') is

equivalent to primer S1269S (NEB, UK) and PTW1_seq-R is equivalent to S1271S (NEB, UK) which is also known as T7 terminator reverse primer (5'-TATGCTAGTTATTGCTCAG-3'). Native pTWIN1 without insert is expected to produce a band size of 1015bp while pTWIN1 with TMPG-encoding gene fragment insert is expected to produce a band size of 277bp. The smaller band size expected for pTWIN1 with insert as compared to the native pTWIN1 plasmid is due to the removal of a much larger 738bp fragment from the multiple cloning site (MCS) following double digestion. Visual validation of the annealed oligonucleotides using gel electrophoresis were not performed due to the small size (50bp) of the assembled fragment and was therefore directly proceed with ligation.

Table 2.2: List of oligonucleotides used in this study.

Oligonucleotide	**Sequence (5' to 3')**	**Expected band size**
PTW1_seq-F (S1269S)	ACTGGGACTCCATCGTTTCT	1015bp
PTW1_seq-R (S1271S)	TATGCTAGTTATTGCTCAG	1015bp
TMPG-F	AACTTTTTTCCGGTGATTGGCCGCATTCTGAACGGCATTCTGTAACTGCA	277bp
TMPG-R	GTTACAGAATGCCGTTCAGAATGCGGCCAATCACCGGAAAAAA	277bp

2.5 Oligonucleotide assembly and construction of pTWIN1-TMPG vector

Chemically competent *E. coli* DH5α cells were first transformed with pTWIN1 vector using heat shock method in order to isolate and propagate the plasmid. Nucleospin plasmid easypure extraction kit was used to extract the pTWIN1 plasmid from *E. coli* DH5α cells following the manufacturer guidelines. Firstly, the transformed *E. coli* cells were cultivated in LB/Amp broth by incubating at 37°C with shaking (200rpm) until OD_{600} was 3. The culture was pelleted by centrifugation at 12000 rcf and resuspended in Resuspension buffer A1 (kit component). Lysis buffer A2 was used to liberate the plasmid from the *E. coli* host cells followed by using Buffer A3 for neutralizing the lysate to create suitable conditions for the plasmid to bind to the silica membrane of the Nucleospin plasmid easypure column. Another round of centrifugation was carried out to pellet the cell debris and precipitated proteins. The supernatant was loaded into the column and using Buffer AQ the contaminants were removed in a single washing step. Purified plasmid was eluted in ddH_2O. The concentration of the extracted plasmid was measured on a Take3 micro-volume plate using Gen5 2.04 software for obtaining high purity plasmids used for cloning.

The extracted pTWIN1 plasmid was double digested with SapI and PstI (NEB, UK). Insertion of gene of interest into the MCS between Sap 1 to Pst 1 would result in N-terminal fusion of TMPG

adjacent to the intein tag thus producing peptide with no vector-derived amino acid residues. For double digestion, 5µl of 10X restriction buffer (NEB, UK), 1µl SapI and 0.5µl PstI enzyme, 10µl of extracted plasmid (100ng/µl) and ddH_2O were added to a final reaction mixture of 50µl and incubated at 37°C for 4h. The restriction enzymes were heat inactivated at 65°C for 20 mins after double digestion.

To assemble the TMPG-encoding gene fragment, 5µl (ratio 1:1) of both TMPG-F and TMPG-R were incubated at room temperature for 15mins. Subsequently, insertion of the assembled gene fragment was performed by carrying out ligation of the digested plasmid (vector) and TMPG-encoding gene fragment (insert), using 1µl T4 DNA Ligase (NEB), 2µl of 10X T4 ligation buffer, insert (5µl of TMPG-encoding gene) to vector(5µl of digested pTWIN1 plasmid) ratio of 1:1, and 7µl of ddH_2O to a final reaction volume of 20µl. Fifty microliter of *E. coli* BL21 (DE3) cells were transformed with 2µl of pTWIN1-TMPG vector using heat shock method. Following transformation, selected colonies growing on the LB/amp agar plate were picked and colony PCR was carried out to confirm the insertion of TMPG-encoding gene in pTWIN1 vector. Colony PCR followed by gel electrophoresis, which was carried out to confirm the size of the gene and the cloned pTWIN1 vector. Colony PCR was performed in a 25µl reaction mix containing 12.5µl of Mytaq mix polymerase enzyme (Bioline, UK), 1.25µl of PTW1_seq-F primer (10µM), 1.25µl of PTW1_seq-R primer (10µM) and 1µl of DNA sample. DNA sample was prepared by picking one large transformed colony using a sterile metallic loop and mixing it with 1µl ddH_2O. The thermal cycling condition was as follow: denaturation for 5 min at 95°C, 29 cycles at 94°C for 45s, annealing for 30s at 54°C and extension for 45s at 72°C, then followed by a final extension step for 5 min at 72°C. Presence of the target PCR product was confirmed on a 1.5% agarose (W/V) gel stained with 1X SYBR safe dye (Invitrogen, USA) at 80V for 50 mins. The 100bp DNA ladder (Bioline, UK) was prepared by mixing the ladder with 5X blue dye and water at the ratio 1:1:3 respectively and samples were prepared by adding the blue dye at a ratio of 4:1, before loading into the well. Gel image was acquired under UV using the ImageLab version 5.2.1 software (Bio-Rad image detecting system).

After electrophoresis, DNA purification was carried out to remove the unwanted materials using purification kit. Positive clones containing the gene of interest were isolated using the same Nucleospin plasmid easypure extraction kit (Macherey-Nagel, Germany) and the concentration of was measured using Gen5 2.04 software. PCR sequencing using Sanger sequencing method (Apical Scientific) was performed to further confirm the positive recombinant plasmid – pTWIN1-TMPG before proceeding with peptide expression and optimization.

2.6 Expression of CBP-Int-TMPG protein construct

E. coli BL21 (DE3) harboring the pTWIN1-TMPG vector was cultivated in 200ml of LB/Amp broth at 37°C with agitation at 220rpm until OD_{600} 0.5-0.7. Induction of peptide expression was initiated by addition of 0.3mM IPTG into the culture and incubating them at two different baseline setting using 15°C for 24h and 37°C for 2 hours recommended by the manufacturer's guidelines (https://www.neb.com/-/media/catalog/datacards-or-manuals/impact-twin-manual-e6950-708.pdf). An uninduced control without the addition of IPTG was included. Following induction, the cells were pelleted at 5000rcf for 10 mins at 4°C and the supernatant was discarded for cell lysis purpose.

2.7 Cell Lysis

Harvested cell pellet from induction was resuspended in ice cold lysis buffer B1 containing 20mM Tris-HCl, 500mM NaCl, 1mM EDTA, and 20mM PMSF protease inhibitor. The suspension was mixed by gentle pipetting before subjected to cell lysis using physical disruption methods combining rapid freeze-thaw cycle at -80°C (10 mins) and 80°C (10 mins) with sonication was carried out to maximize cell lysis. Sonication was carried out using a thin probe and the suspension was kept chilled on ice during the whole process. The sonication protocol was tested at 10/20s (where 10/20s means 10 seconds of pulse with 20 seconds off) using 35% amplitude and 3 cycles. Subsequently, the lysates were clarified by centrifugation at 18,000rcf for 30 mins at 4°C. The clarified cell lysate (supernatant) was separated from the pellet fraction into individual tubes. The pellet fraction was resuspended in 200µl of ddH_2O for subsequent tests. In addition, 50µl of the clarified cell lysate and the bacterial culture were plated using L-shaped glass rod on LB/Amp to assess the extent of cell lysis before and after the physical disruption procedures. Lysis of cells was confirmed by observing the growth of colonies on the agar plates following overnight incubation at 37°C. Twenty microlitre of supernatant and pellet that was resuspended and diluted in 100µl of ddH_2o were aliquoted to be analyzed by SDS-PAGE and WIA.

2.8 Optimization of CBP-Int-TMPG protein construct expression

Optimization of expression was performed following the one-factor-at-a-time approach by optimizing a single parameter at one time in sequential stages while keeping the other parameters constant (Sankaran, et al., 2018). The optimized settings will be accumulated by stages to achieve maximum expression of CBP-Int-TMPG protein construct as defined in this study. The key parameters influencing protein expression were tested: (i) IPTG concentration, (ii) IPTG induction temperature, (iii) IPTG induction duration, and (iv) agitation. To begin with, the first parameter tested was concentrations of IPTG using the range of 0.1mM, 0.3mM, 0.6mM, 0.9mM and 1.2mM. In the second stage induction temperature was tested which included inducing the culture with 0.3mM

IPTG with induction temperature of 37°C for 2h, 15°C overnight (24h), and 25°C overnight (24h). In this stage the temperature and duration are coupled together. The temperature used followed the recommended duration and IPTG concentration based on the manual. The third stage involved altering the duration of induction, while the selected temperature from stage two and optimized IPTG concentration from stage one were kept constant. The duration of induction period was changed using the intervals 2h, 8h, 24h and 48h, respectively. In the last stage, the effect of introducing agitation during induction was determined. Appropriate controls were included for each parameter. For example, a control run without the addition of IPTG (uninduced cells) was included while optimizing the effect of IPTG concentrations. After completing the OFAT, a final run was conducted to compare the overall expression enhancement before and after optimization.

2.9 SDS-PAGE gel electrophoresis for protein separation

The presence and expression level of TMPG was determined based on WIA by targeting the CBD co-expressed on the CBD-Int-TMPG protein construct. All protein samples were electrophoresed on a discontinuous SDS-PAGE gel. The composition of the stacking and resolving gels are listed in table 2.3. The two gels vary from each other according to their differences in pore size, ionic strength and different pH value. The size of the pores in the gel was controlled by changing the concentration of acrylamide as well as the ratio of acrylamide to bis-acrylamide. Polymerization of the gel takes place when 10% (W/V) ammonium persulfate (APS) and Tetramethyl ethylenediamine (TEMED) are added into the gel solution. In this experiment, 30% (W/V) (37.5:1) acrylamide Bis- acrylamide ready to use solution (Bio-Rad) was used. The samples were added with 10µl SDS sample buffer with 5% (V/V) β-2-marcaptoethanol and heated to 95°C for 10 mins before loading sample loading at 10ul per well. The protein marker (polypeptide SDS-PAGE standards, Bio-rad) used was pre-diluted with ddH_2O at the ratio of 1:4 before addition of SDS sample buffer as abovementioned. An additional pre-stained 3-Color high range protein marker (SMOBIO) was also included alongside the polypeptide SDS-PAGE standards as a blot transfer indicator (the polypeptide SDS-PAGE standards was not visible after transfer). The SDS sample buffer contains 0.01% (V/V) of the tracking dye bromophenol blue which migrates with the samples being separated. It also contains 30% (V/V) glycerol which helps the samples to sink to the bottom of the well after loading. After the gel is casted it is placed vertically into the electrophoresis chamber and filled with running buffer. According to the methods described by Schagger and Jagow (with some modifications), 16% (W/V) SDS was carried out using tricine cathode running buffer at 110v for 1h 10 mins (Schagger & Jagow, 1987). Following electrophoresis, the gel was stained with staining solution (composed of 50% (V/V) Methanol, 10% (V/V) Acetic acid, ddH_2O and 0.1% (W/V) Coomassie brilliant blue (CBB) G250 staining dye on a shaker for 24h and destained with destaining solution which has the same

composition as the staining solution except the CBB dye. The gel was incubated in destaining solution for 25 mins for 3 cycles on a shaker before proceeding with western immunoblot assay (WIA).

Table 2.3: Composition of stacking and resolving gels.

Solution (for 2 gels)	Stacking (4%)	Resolving (16%)
30% Acrylamide/Bis	0.75 ml	5.3 ml
0.5 M Tris-HCl, pH 6.8	1.26 ml	------
1.5 M Tris-HCl, pH 8.8	--------	2.5 ml
10% SDS	60 µl	100 µl
Distilled deionized water	3.5 ml	2.7 ml
10% APS	30 µl	100 µl
TEMED	8.4 µl	8.4 µl
Total volume	**5 ml**	**10 ml**

2.10 Western immunoblot assay

PVDF membrane and filter papers were first equilibrated in Bjerrum Schafer-Nielsen transfer buffer (48mM Tris base, 39mM glycine, 20% (W/V) methanol, 500ml water, and 1.3mM SDS) for 5 mins and a transfer sandwich was assembled before semi-dry blotting (Thermo semi-dry transfer system) at 25V 1-3A for 5 min. Following blotting, the membrane was incubated in blocking solution containing phosphate buffered saline with 0.1% (V/V) tween 20 (PBST) and 5% (W/V) skim milk for 1h. The membrane was rinsed with PBST for 5 mins and then incubated with mouse anti-CBD monoclonal antibody (NEB, UK) in PBST with 1% skim milk (dilution factor 1:1000, antibody:1% skim milk in PBST) . The membrane was washed with PBST for 10 mins twice before incubated with secondary goat anti-mouse antibody (R and D systems) conjugated with horseradish peroxidase (HRP) (1:10,000) in PBST with 1% skim milk. Detection of bands was performed using the chemilumi ECL detection kit (Thermo-scientific). The ECL substrate was prepared according to the manufacturer's instructions with 500ml of each reagent A and B were mixed (1:1) and subsequently incubated onto the membrane for 5 mins in dark. The chemiluminescence signal was captured using the Bio-Rad ChemiDoc imaging system at an exposure time of 5 mins. The densitometric program Image J (National Institutes of Health, USA) was used for WIA image analysis and quantitation of the relative abundance (%) of the sample signal intensity.

2.11 Bradford assay

Protein concentration was measured using the Bradford assay with bovine serum albumin (BSA) as a standard to generate the standard curve according to the manufacturer's instructions (sigma-aldrich.com). A serial dilution of stock BSA at 2mg/ml was prepared in duplicates in 96-well plate (2, 1.75, 1.5, 1.25, 1.0, 0.75, 0.5, and 0.1 mg/ml). Sterilized distilled water was included as the blank control. After loading BSA, the bradford reagent (Sigma Aldrich, Germany) required no dilution and therefore 100µl was directly added to each well containing 10µl of sample and pipetted up and down to mix the regents properly. The plate was incubated at room temperature for 5 mins for colour change. After incubation, the absorbance was measured at OD of 595nm. The brilliant blue G dye present in the Bradford regent will form a complex with the proteins which causes a shift in the absorption. The concentration of proteins was calculated based on the standard curve generated.

Chapter 3: Results

3.1 Construction of recombinant pTWIN1-TMPG expression plasmid

Following assembly of TMPG-encoding oligonucleotides (TMPG-F and TMPG-R), the TMPG gene fragment was inserted into the SapI and PstI double digested pTWIN1 vector. Following ligation, the pTWIN1-TMPG recombinant plasmid was transformed into *E. coli* DH5α for propagation. The recombinant plasmid was extracted and transformed into *E. coli* BL21 (DE3) for expression (expression host). Transformants obtained were confirmed with colony PCR using PTW1_seq-F and PTW1_seq-R (Figure 3.1). Transformants carrying pTWIN1-TMPG recombinant plasmid produced a band size of 277bp (Figure 3.1B) while transformants carrying pTWIN1 only (without insert) produced a band size of 1015bp (Figure 3.1A). DNA sequencing of the pTWIN1-TMPG amplicon further confirmed the orientation and sequence identity of the cloned TMPG gene, which matched precisely to the sequence of oligonucleotides used (Figure 3.2).

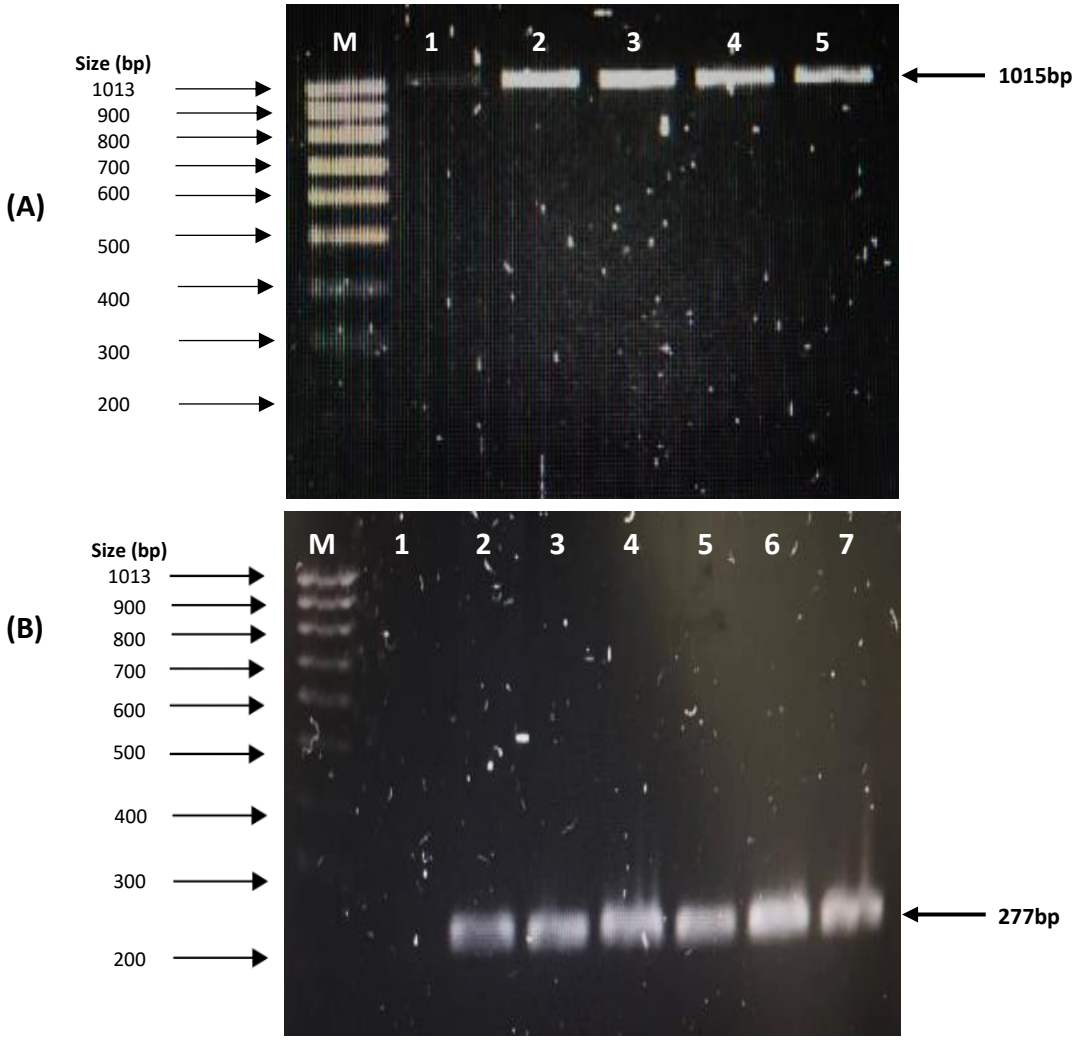


Figure 3.1: Colony PCR screening of pTWIN1 vector and pTWIN1-TMPG transformants. (A) *E. coli* BL21 (DE3) carrying the non-recombinant pTWIN1 vector (no insert) produced a band size of 1015bp. (B) *E. coli* BL21 (DE3) carrying the recombinant plasmid pTWIN1-TMPG produced a smaller band size of 277bp. The PCR products were analyzed on a 1.5% agarose gel and stained with 1X SYBR safe dye. M, 100bp DNA ladder (Bioline, UK); Lane 1, non-template control; other lanes represent the selected colonies for screening.

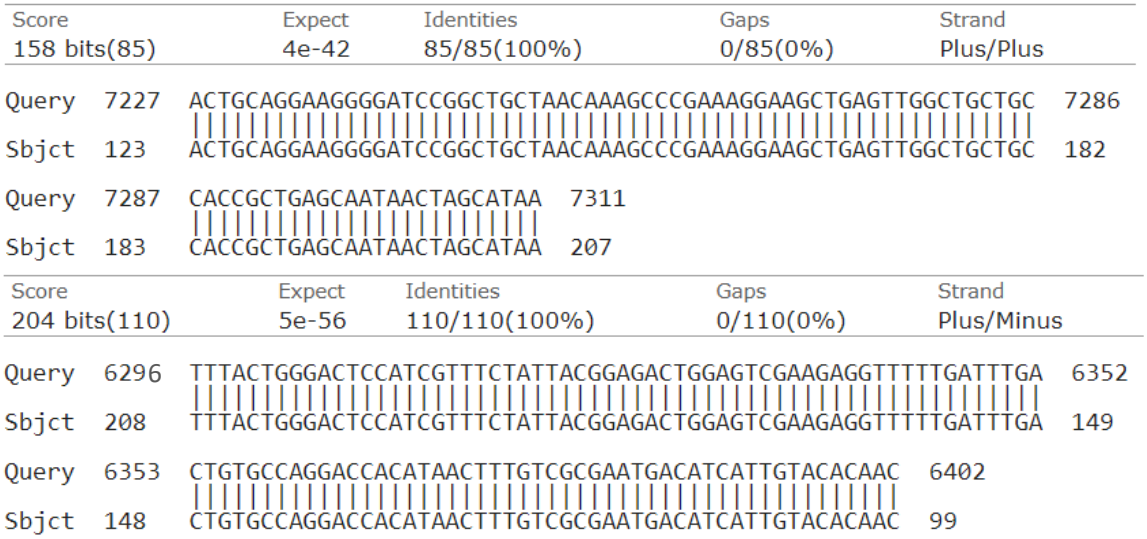


Figure 3.2: Confirmation of sequence identity of the inserted TMPG-encoding gene using NCBI BLASTn .

3.2 Efficiency of cell lysis by sonication

To ensure complete cell lysis and thus release of total protein and the expressed CBD-Intein-TMPG protein construct, cells were lysed by probe sonication and were kept chilled on ice during the process. The endpoint was to achieve maximal cell lysis and therefore release of cellular content based on the absent of viable colonies after overnight plating of the clarified cell lysate on LB/Amp agar. The first sonication protocol tested was at pulse time 10/20s, 15/20s and 20/20s (where 5/20s means 5 seconds of pulse with 20 seconds off) using 35% amplitude and 6 cycles for each condition. From the results in Figure 3.3(A, I), after carrying out cell lysis by sonication and plating on LB/Amp agar plates, colonies overgrowth was observed in the control plate (without sonication). Exposing the bacterial culture to longer pulse time (20/20s) decreased the number of colonies and established an inverse relationship to the number of colonies grown on agar plate. Fewer colonies were observed on the agar plate with pulse time 20/20s. This setting was not satisfactory considering the abundance of surviving colonies observed (endpoint not achieved) and the harsh sonication protocol used generated substantial heat in the solution which accumulated after each cycle of sonication. Such drastic conditions would have negative impact on the quality of proteins in the clarified cell lysates.

To test the effect of sonication cycle on cell lysis, another test was conducted by setting the sonication time constant at 35% amplitude and pulse time 15/20s while increasing the number of sonication cycles. The protocol tested was no sonication, 2 cycles, 4 cycles, and 6 cycles at pulse time 15/20s. From the results (figure 3.3B), gradual reduction in viable colonies was observed, however, the presence of viable colonies was still observable even at the highest 6 cycles tested.

To further enhance the cell lysis efficiency, an additional freeze-thaw step at -80°C for 10 mins and 80°C for 10 mins, prior to sonication (3 cycles using 35% amplitude at pulse time 15/20s) was included. The additional freeze-thaw cycle significantly increased the efficiency of cell lysis and achieved 100% lysis of the cells. The results as shown in Figure 3.3(C) compare the efficiency of cell lysis when sonication was carried out, to the extent of lysis when an additional of freeze-thaw step in combination with sonication was used.

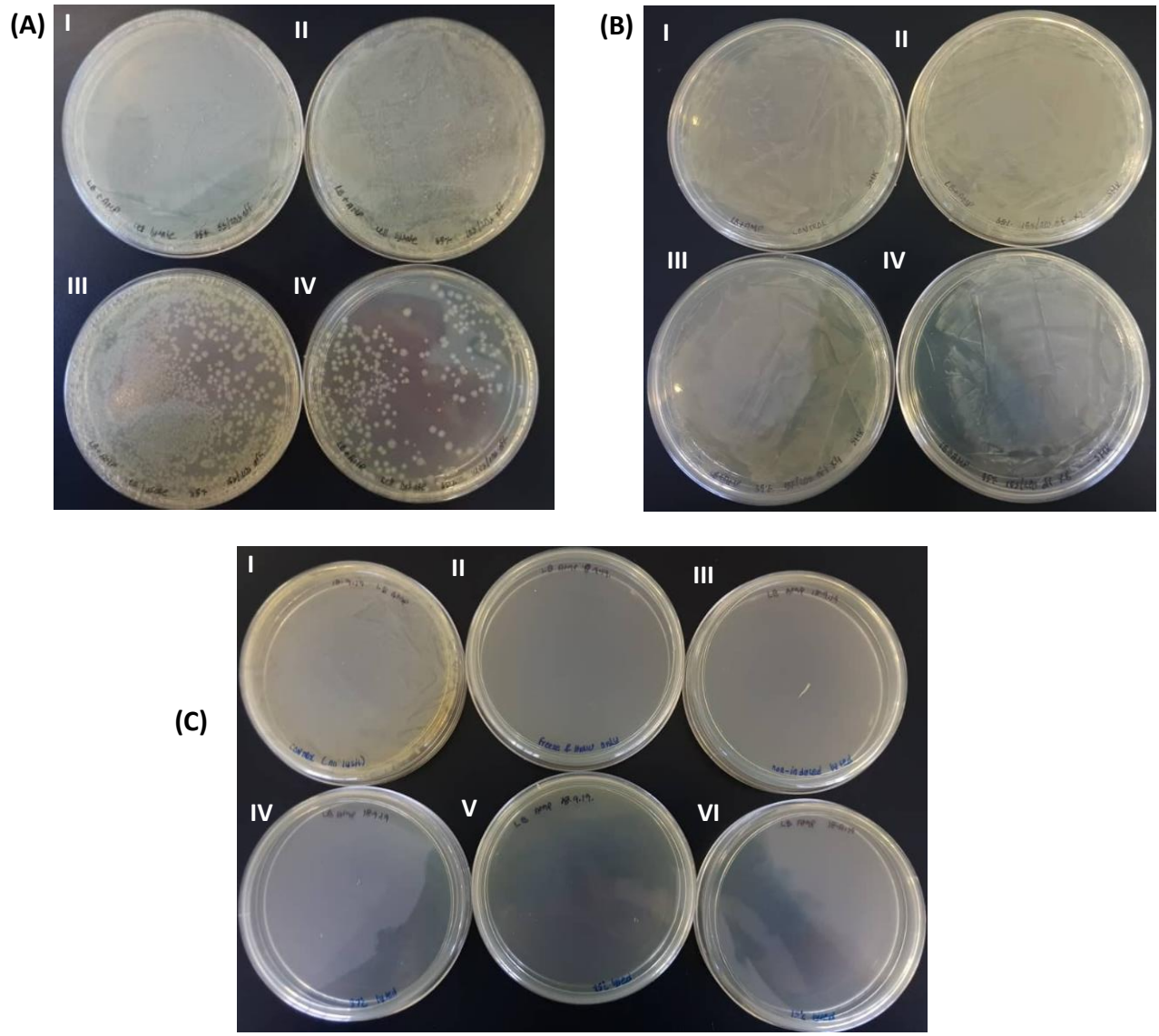


Figure 3.3: Optimization of cell lysis using probe sonication at varying pulse time, sonication cycle, and incorporation of free thaw step. (A) Effect of sonication at 35% amplitude for 6 cycles with increasing pulse time at 10s/20s, 15s/20s, 20s/20s and no sonication (control). (B) Effect of sonication at 35% amplitude and pulse time 15s/20s with increasing number of cycles at 2 cycles, 4 cycles, 6 cycles, and no sonication (control). (C) Effect of an additional freeze-thaw cycle prior to sonication resulted in complete cell lysis as observed where (I) control plate (without sonication and freeze-thaw cycle) showing overgrowth (no lysis), (II) plate showing the effect of only freeze and thaw (III) effect of sonication using 35% amplitude and pulse time 15s/20s for 3 cycles without freeze-thaw and (IV,V and VI) effect of freeze-thaw prior to sonication (35% amplitude pulse time 15s/20s for 3 cycles) resulting in 100% lysed cells.

3.3 Optimization of recombinant TMPG-Int-CBD protein construct expression

3.3.1 Effect of IPTG concentration on protein expression

To determine the optimum concentration of IPTG for the induction of protein expression from pTWIN1-TMPG in *E. coli* BL21 (DE3), IPTG at 0.1mM, 0.3mM, 0.6mM, 0.9mM and 1.2mM were tested while keeping the induction temperature and duration constant at 15°C for 24h without agitation (baseline setting). The presence of an approximately 26.6 kDa band in WIA analysis in both clarified cell lysate and pellet fraction indicated the CBP-Int-TMPG protein construct. Both the clarified cell lysate and cell pellet fractions of each sample were compared. As shown in Figure 3.4(A), the highest expression for clarified cell lysate was observed with 0.9mM IPTG concentration (36.14%), followed by 1.2mM (17.39%) and 0.6mM (17.37%) IPTG concentration. The band size by inducing with 0.6mM and 1.2mM of IPTG concentration gave approximately similar intensity for clarified cell lysate (Figure 3.4A). In contrast, for the pellet fraction highest expression was achieved under 1.2mM IPTG induction followed by 0.9mM, 0.6mM and 0.3mM, respectively. The expression of CBP-Int-TMPG protein construct was directly proportional to the IPTG concentration used for the pellet fraction (Figure 3.4B). Induction using 0.1mM IPTG produced a band with the lowest intensity as compared to the higher IPTG concentrations for both clarified cell lysate and pellet fraction. From the results of the change in concentration, 0.9mM was selected for the next optimization test.

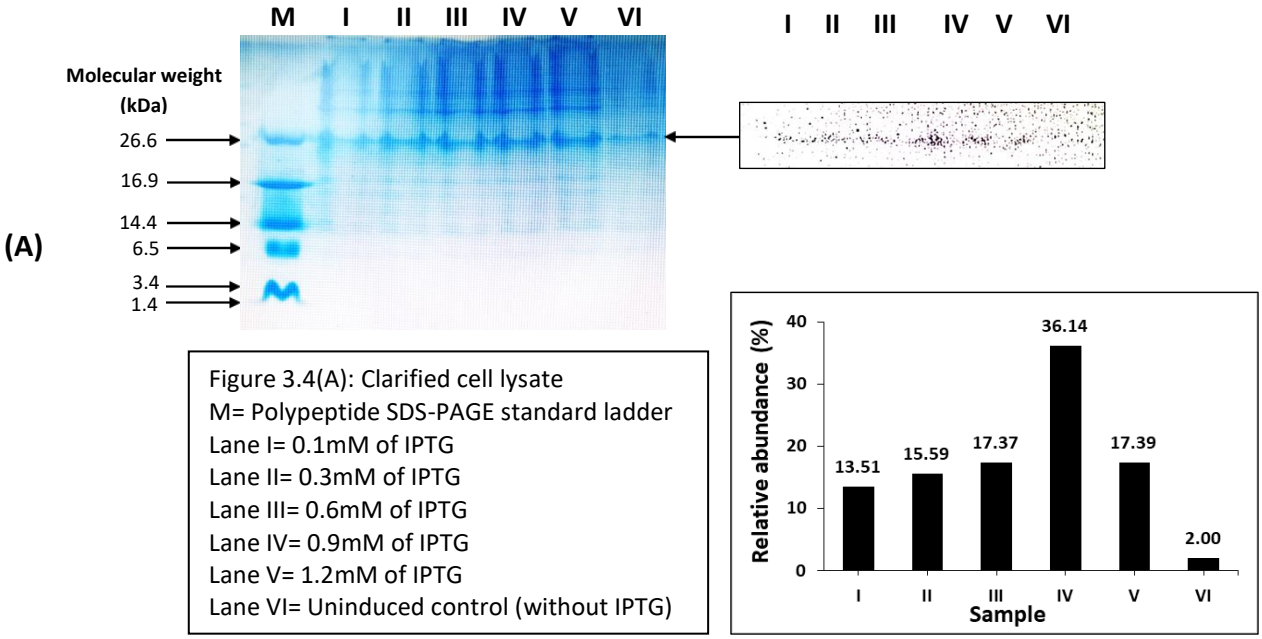


Figure 3.4(A): Clarified cell lysate
M= Polypeptide SDS-PAGE standard ladder
Lane I= 0.1mM of IPTG
Lane II= 0.3mM of IPTG
Lane III= 0.6mM of IPTG
Lane IV= 0.9mM of IPTG
Lane V= 1.2mM of IPTG
Lane VI= Uninduced control (without IPTG)

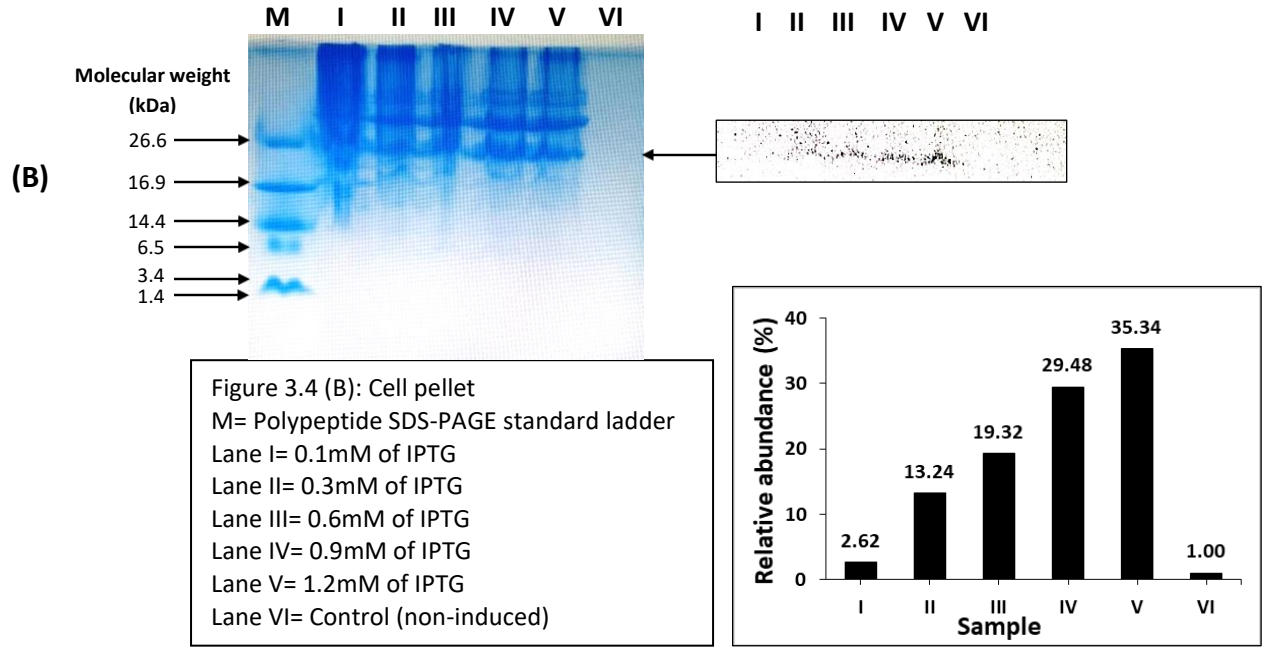


Figure 3.4 (B): Cell pellet
M= Polypeptide SDS-PAGE standard ladder
Lane I= 0.1mM of IPTG
Lane II= 0.3mM of IPTG
Lane III= 0.6mM of IPTG
Lane IV= 0.9mM of IPTG
Lane V= 1.2mM of IPTG
Lane VI= Control (non-induced)

Figure 3.4: Effect of varying IPTG concentrations on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction. (A) SDS-PAGE and WIA results of clarified cell lysate when induced using different IPTG concentrations; lane I: 0.1mM, lane II: 0.3mM, lane III: 0.6mM, lane IV: 0.9mM and lane V: 1.2mM of IPTG and lane VI: control (non-induced cells). (B) SDS-PAGE and WIA result of pellet fraction; lane I: 0.1mM, lane II: 0.3mM, lane III: 0.6mM, lane IV: 0.9mM and lane V: 1.2mM of IPTG and lane VI: control (non-induced cells). The detection of peptide expression is based on presence of CBD protein using anti-CBD antibody. (M : Polypeptide SDS-PAGE standard ladder)

3.3.2 Effect of induction temperature on protein expression

The effect of varying temperature on the expression of CBP-Int-TMPG was determined at 37°C for 2h, 25°C for 24h, and 15°C for 24h at 0.3mM IPTG without agitation. The temperature settings tested at this stage were coupled with varying durations using 0.3mM IPTG as recommended by the manufacturer, and a separate test on the effect of induction duration alone was carried out with the optimized induction temperature and IPTG concentration in the next stage.

The highest density band was observed for cells induced at 25°C for 24h, for both clarified cell lysate (68.87%) and pellet fraction (70.38%). No bands were observed for the uninduced control cells (Figure 3.5A and B, lane IV) indicating no protein expression. Therefore 25°C temperature was chosen as the optimized temperature to carry out further tests.

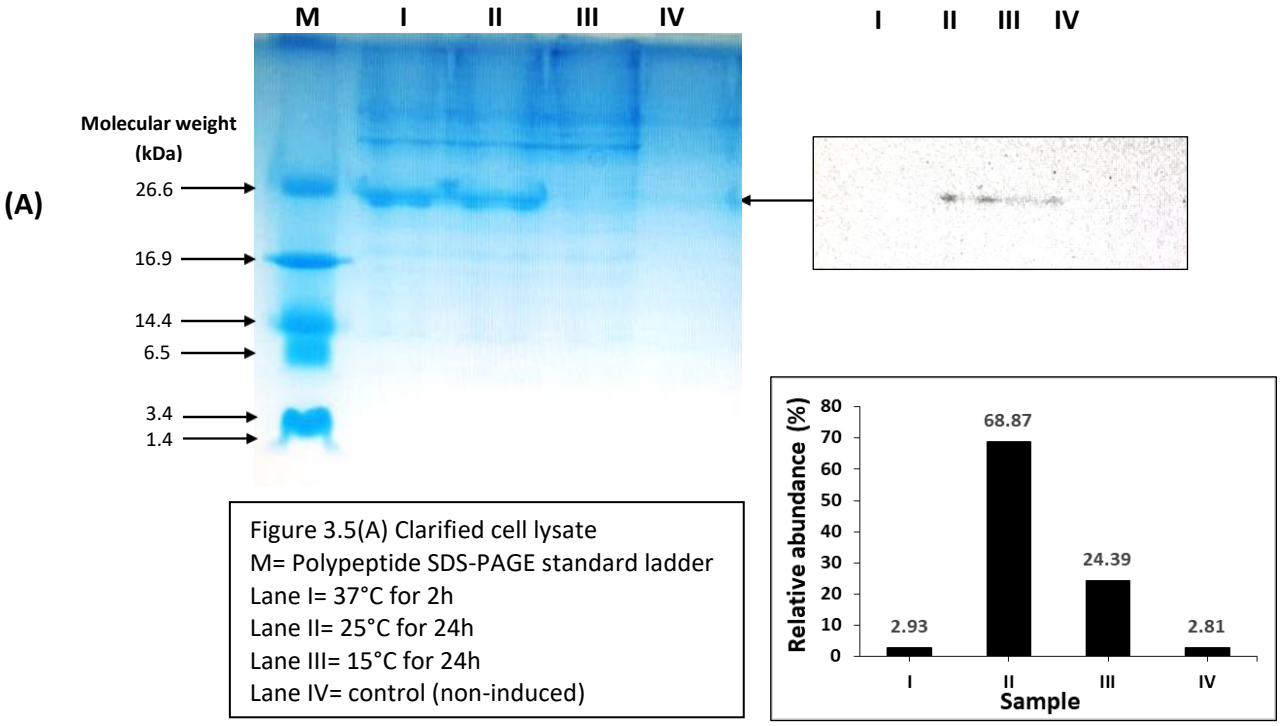


Figure 3.5(A) Clarified cell lysate
M= Polypeptide SDS-PAGE standard ladder
Lane I= 37°C for 2h
Lane II= 25°C for 24h
Lane III= 15°C for 24h
Lane IV= control (non-induced)

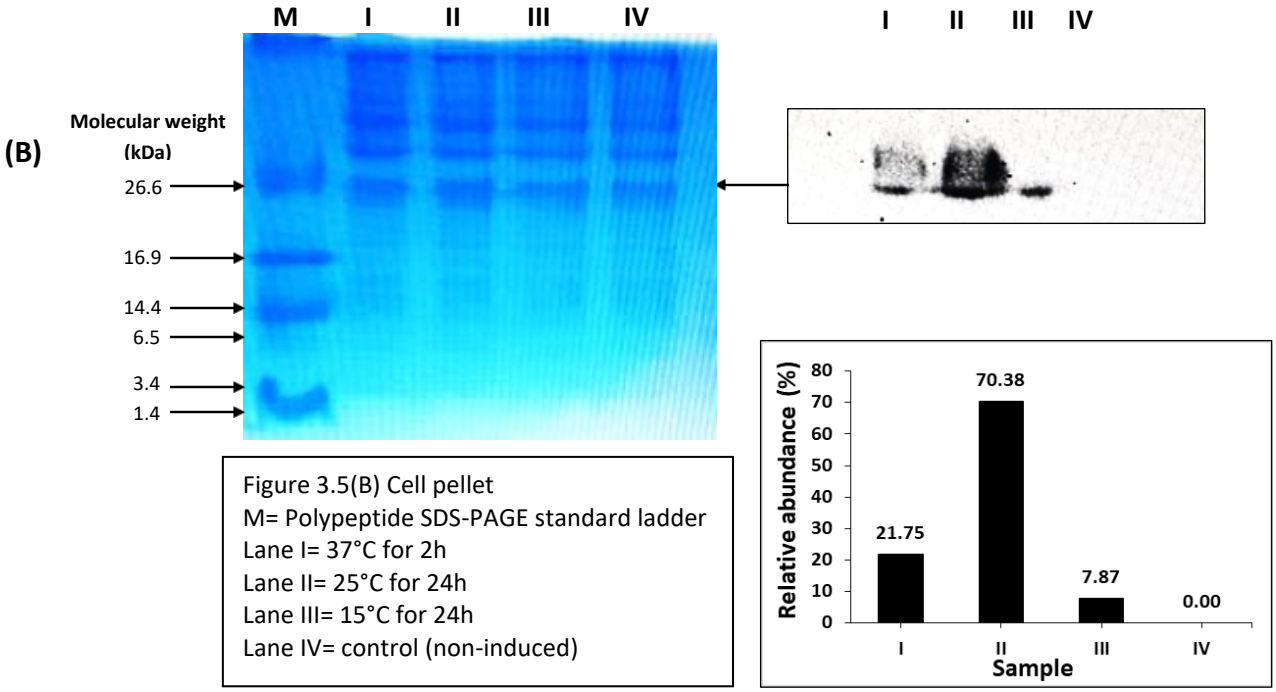


Figure 3.5(B) Cell pellet
M= Polypeptide SDS-PAGE standard ladder
Lane I= 37°C for 2h
Lane II= 25°C for 24h
Lane III= 15°C for 24h
Lane IV= control (non-induced)

Figure 3.5: Effect of varying temperature on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction. (A) SDS-PAGE and WIA results of clarified cell lysate; lane I: 37°C for 2h, lane II: 25°C for 24h, lane III: 15°C for 24h and lane IV: control (no induction carried out) using Coomassie-stained 16% tricine polyacrylamide gels. (B) SDS-PAGE and WIA analysis of the pellet fraction from the same conditions mentioned in (A). (M: Polypeptide SDS-PAGE standard ladder)

3.3.3 Effect of induction duration on protein expression

The induction duration was tested at 2h, 8h, 24h, and 48h at 25°C using 0.9mM IPTG. From the results, increasing the duration of induction enhanced the expression of CBP-Int-TMPG protein construct proportionally for both clarified cell lysate and pellet fraction (Figure 3.6A and B). At 24h and 48h, the expression levels were strongest as seen with the presence of very dense bands in WIA for both fractions. In contrast, 2h induction was the least ineffective with negligible protein expression while a slightly detectable expression was observed for 8h induction. Following the results, 48h duration was chosen as the optimized duration.

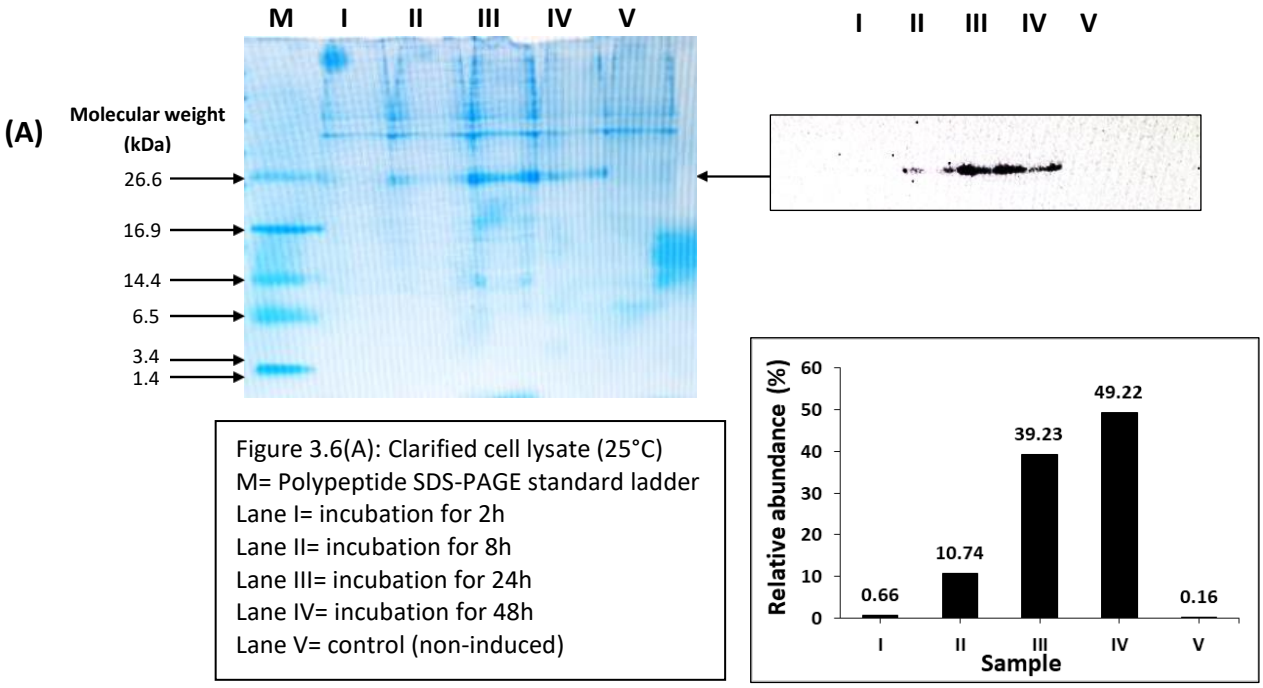


Figure 3.6(A): Clarified cell lysate (25°C)
M= Polypeptide SDS-PAGE standard ladder
Lane I= incubation for 2h
Lane II= incubation for 8h
Lane III= incubation for 24h
Lane IV= incubation for 48h
Lane V= control (non-induced)

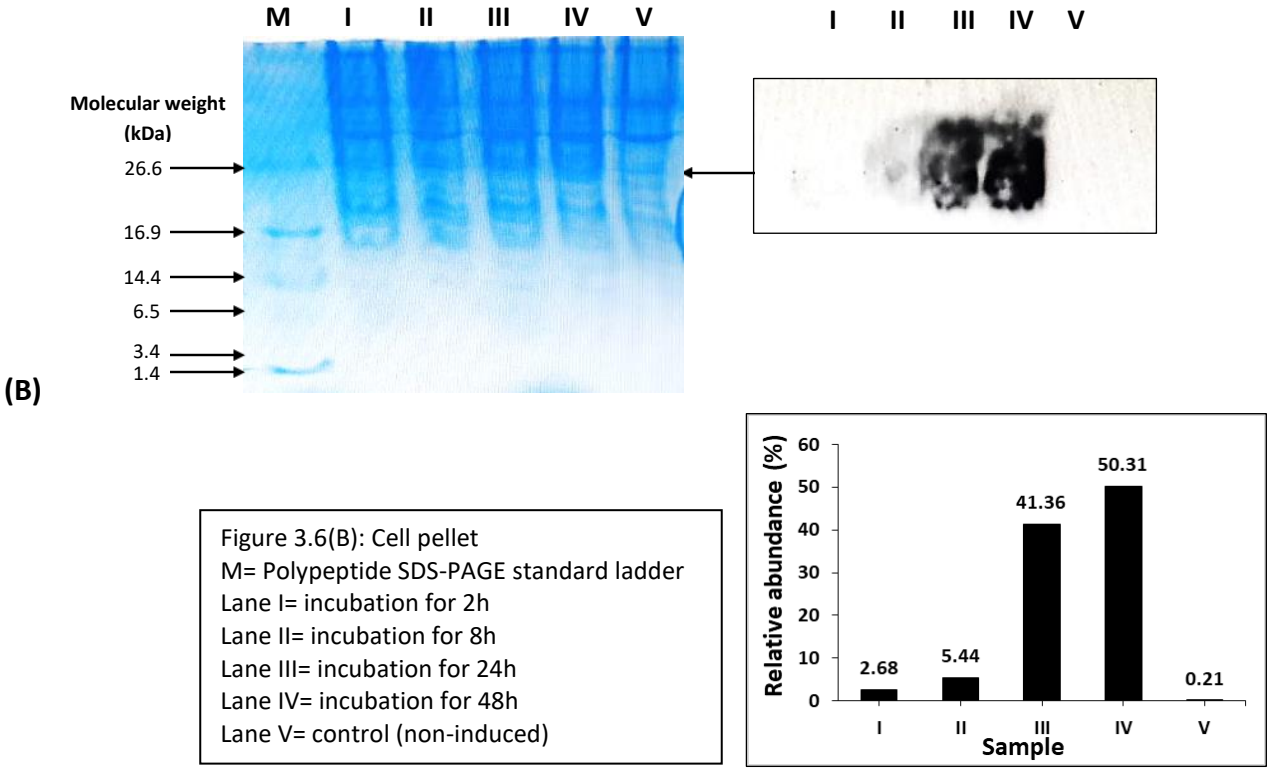


Figure 3.6(B): Cell pellet
M= Polypeptide SDS-PAGE standard ladder
Lane I= incubation for 2h
Lane II= incubation for 8h
Lane III= incubation for 24h
Lane IV= incubation for 48h
Lane V= control (non-induced)

Figure 3.6: Effect of different induction duration on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction. (A) SDS-PAGE and WIA analysis of clarified cell lysate when induced with 0.9mM IPTG at different incubation duration lengths but keeping the temperature constant at 25°C; lane I: incubation for 2h, lane II: incubation for 8h, lane III: incubation for 24h, lane IV: Incubation for 48h, lane V: uninduced cells (control). (B) SDS-PAGE and WIA analysis of cell pellet; lane I: incubation for 2h, lane II: incubation for 8h, lane III: incubation for 24h, lane IV: Incubation for 48h, lane V: uninduced cells (control) (respectively. (M: Polypeptide SDS-PAGE standard ladder)

3.3.4 Effect of agitation on protein expression

Based on the optimized settings determined to this stage (IPTG concentration 0.9 mM, 25°C, 48h), next the effect of agitation at 180rpm on inducing CBP-Int-TMPG expression was tested. From the results, expression level was clearly higher in cells induced under shaking as compared to without shaking (Figure 3.7). Both the clarified cell lysate and pellet fractions showed similar observation and the expression level is significantly higher in the pellet fraction. The presence of the comparatively thicker bands from WIA also showed that the recombinant protein was primarily distributed in the pellet fraction. The final induction condition determined based on the optimized settings was 0.9mM IPTG at 25°C for 48h with shaking. Figure 3.8 shows the overall flow of optimization protocol used in this study.

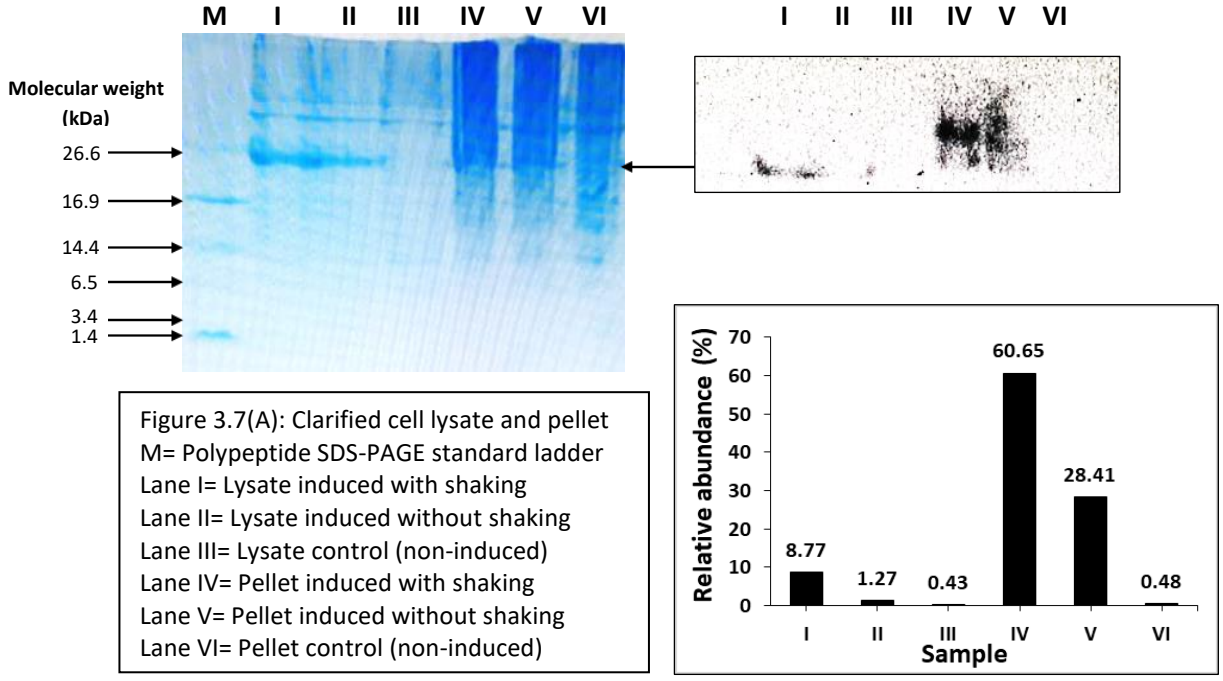


Figure 3.7(A): Clarified cell lysate and pellet
M= Polypeptide SDS-PAGE standard ladder
Lane I= Lysate induced with shaking
Lane II= Lysate induced without shaking
Lane III= Lysate control (non-induced)
Lane IV= Pellet induced with shaking
Lane V= Pellet induced without shaking
Lane VI= Pellet control (non-induced)

Figure 3.7: Effect of shaking on the expression of TMPG-carrying protein construct from E. coli BL21 (DE3) for both clarified cell lysate and pellet fraction. SDS-PAGE and WIA analysis with and without shaking using optimized conditions (25°C for 48h using 0.9mM IPTG); lane I: clarified cell lysate fraction induced with shaking, lane II: clarified cell lysate induced without shaking and lane III: clarified cell lysate of the control (no induction), lane IV: cell pellet fraction induced with shaking, lane V: cell pellet induced without shaking and lane VI: cell pellet from control (non-induced cells). (M: Polypeptide SDS-PAGE standard ladder)

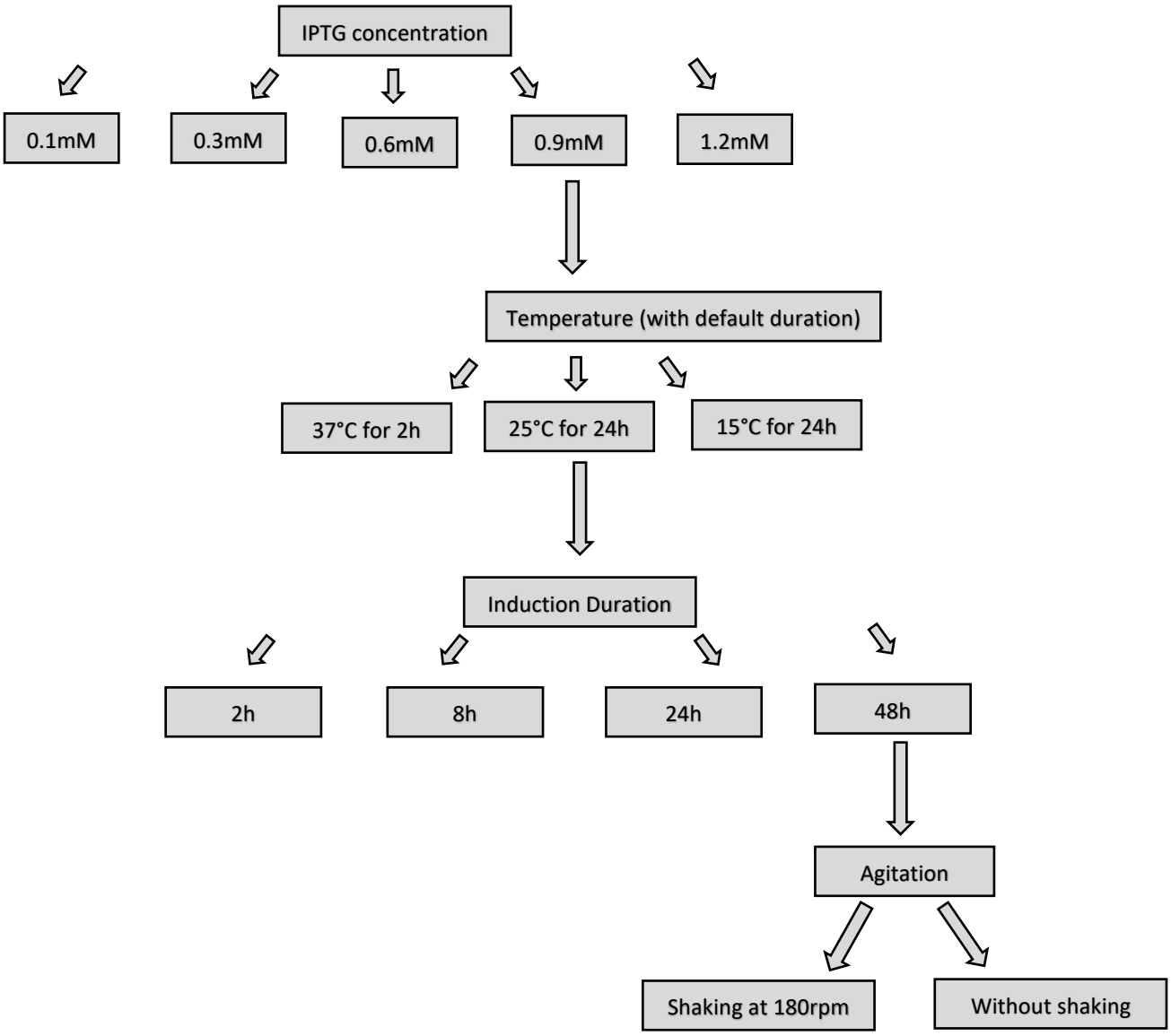


Figure 3.8: The flow chart of the overall optimization process tested in this study. The final optimized conditions for inducing the cell culture were using 0.9mM IPTG at 25°C for 48h with 180rpm shaking.

3.4 Comparison of recombinant protein expression before and after optimization

To evaluate the overall enhancement in expression level of CBP-Int-TMPG from pTWIN1-TMPG vector in *E. coli* BL21 (DE3) before and after optimization, cells induced under both optimized (IPTG concentration 0.9mM, 25°C, 48h, shaking at 180rpm) and pre-optimized baseline setting (IPTG concentration 0.3mM, 15°C, 24h, without shaking) were compared. The total protein yields were 3.3-fold and 4.2-fold higher for the optimized induction setting (Table 3.1). As shown in Figure 3.9, the optimized induction produced a significantly higher protein expression (8-fold) than the pre-

optimized conditions in the target pellet fraction. Similar observation was noted for clarified cell lysate which showed 7-fold increase in expression after optimization although the protein yield in clarified cell lysate fraction was minimal and in general not the target fraction to harvest the expressed protein. The higher IPTG concentration and temperature used for an extended duration with shaking appeared to contribute to the higher expression level in the host cell.

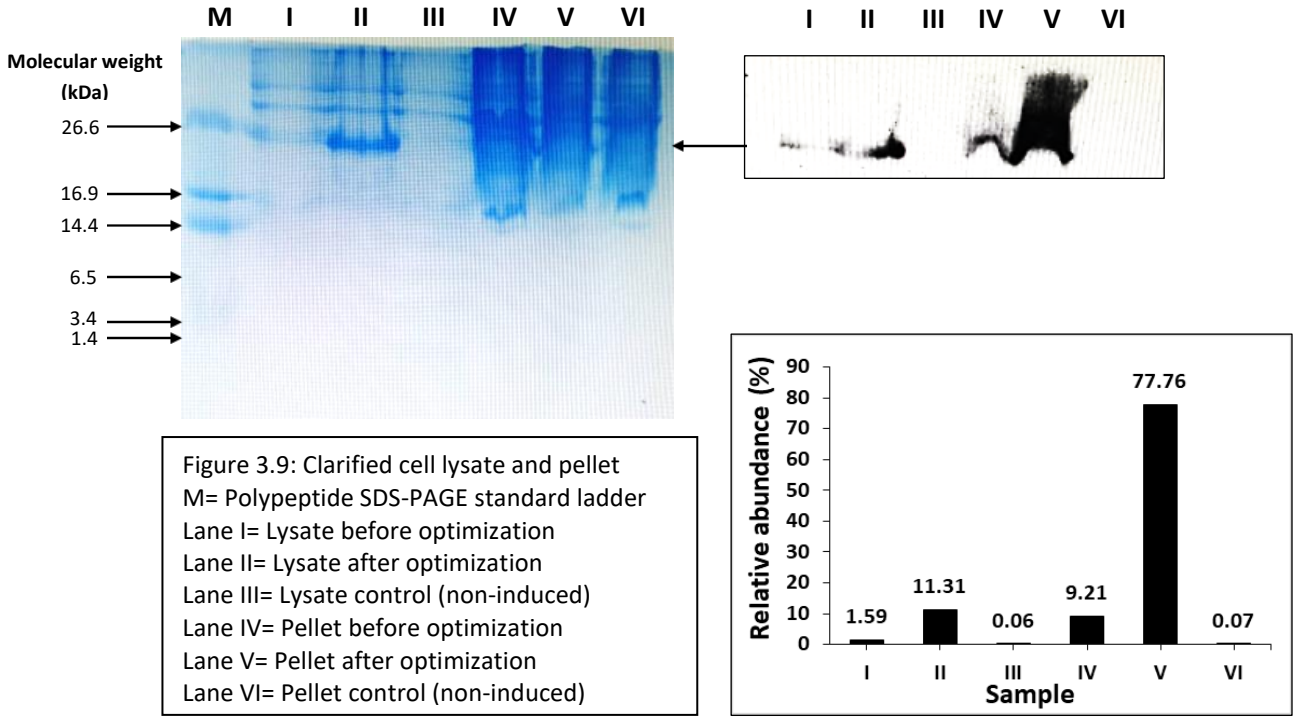


Figure 3.9: Clarified cell lysate and pellet
M= Polypeptide SDS-PAGE standard ladder
Lane I= Lysate before optimization
Lane II= Lysate after optimization
Lane III= Lysate control (non-induced)
Lane IV= Pellet before optimization
Lane V= Pellet after optimization
Lane VI= Pellet control (non-induced)

Figure 3.9: Comparison of recombinant protein expression profiles before and after optimization. SDS-PAGE analysis of the optimization results; lane I: clarified cell lysate without optimization, lane II: clarified cell lysate after optimization, lane III: clarified cell lysate from control (non-induced), lane IV: cell pellet before optimization, lane V: cell pellet after optimization and lane VI: cell pellet for control (non-induced) using Coomassie-stained 18% tricine polyacrylamide gel. WIA results for the changes in peptide production before and after optimization by detection of peptide expression based on the presence of CBD protein using anti-CBD antibody. ImageJ analysis of the relative abundance for both the clarified cell lysate and pellet fraction before and after optimization. (M: Polypeptide SDS-PAGE standard ladder)

Table 3.1: Absorbance and concentration values of the protein before and after optimization.

Sample	Protein concentration (mg/ml)	
	pre-optimized	Optimized
Cell lysate	0.29	1.23
Cell pellet	1.07	3.53

Chapter 4 : Discussion

Antibiotic resistance has been a major issue worldwide since the past decades. This situation not only limit the choice of empirical antibiotic treatment in cases involving ARB, but ineffective empirical treatment would also severely predispose patients to greater risks of morbidity and mortality. ARB associated infections are highly critical particularly in patients with life-threatening infections such as meningitis, pneumonia, and bacteremia. Ineffective first antibiotic treatment shall mean a second (or more) antibiotic regimen would be needed immediately to avoid further delay in treatment and recovery. Continued proliferation and disease establishment in the patients would incur severe damages to the infected sites causing poorer prognosis and long term post-treatment sequelae. For example, bacterial meningitis would result in learning disability, neuronal damage, motor functions impairment, and others (Hoffman & Weber, 2009). *Streptococcus pneumoniae* and *Neisseria meningitidis* are two of the common pathogens causing meningitis. Another major issue is that most often the pathogens surrounding us in the environment may cause only mild or non-lethal infections such as strep throat, pneumonia, sepsis and otitis media (Weiser, et al., 2018). However, ineffective antibiotic treatment would allow more time and space for these pathogens to continue establishing wider and deeper infections to sterile sites including bloodstream and the vital organs. Therefore, of utmost concern is that the common and non-lethal diseases could now turn into life-threatening infections. Therefore, the continued development of new classes of antimicrobial agents is in great demand and AMPs have been recognized as one of the most promising alternative antimicrobials to conventional antibiotics.

The discovery and bioproduction of AMPs have become a demanding need. Recombinant expression of AMPs is a potential approach for the low cost and scalable production of peptides for various applications. A number of studies have documented the recombinant expression of AMPs using various fusion partners such as TrxA (Xie, et al., 2013), SUMO, maltose-binding protein and intein (Costa, et al., 2014). *E.coli* is one of the most used microorganism for the production of AMP (Wong, et al., 2018). Some examples of previously expressed AMPs in *E.coli* are adenoregilin, cecropin, defensin, OG2, indolicidin, and more (Parachin, et al., 2012). Even though all these AMPs were produced in *E.coli* cells, they did not use the intein mediated pathway for expression of the peptide. In this study, recombinant expression of TMPG peptide was successfully carried out in *E.coli* BL21 (DE3) using intein system. As compared to other fusion proteins, one of the unique advantages of using intein tag is the self-cleaving property which allows the release of native peptides/proteins without vector-borne residues to which the biological activities might be highly sensitive to the sequence alteration (Li, 2011). This is particularly crucial for small compounds and compounds for

medical and therapeutic applications and thus the reason we chose this expression system in the current study. Previously in a study carried out by Wang, et al., the AMP cecropin was expressed and purified using IMPACT-TWIN system in *E.coli* cells. This AMP was shown to be toxic to the host cells at high temperatures like 37°C and hence needed to be stored at low temperature (Wang, et al., 2012). In this study, TMPG-encoding gene fragment was designed and successfully cloned by N-terminal fusion to the Ssp DnaB intein tag flanked by CBD on pTWIN1. The TMPG-encoding gene did not show any toxicity against the host cell at higher temperatures and was successfully expressed when induced with IPTG which was confirmed by the presence of protein bands in SDS-PAGE gel and WIA.

With the construction of pTWIN1-TMPG recombinant expression vector, we focused on optimization of the key induction parameters including IPTG concentration, induction temperature, induction duration, and incorporation of agitation in order to maximize the expression and thus the protein yield. The recombinant protein construct was present in both the clarified cell lysate and pellet fraction but the concentration of protein in pellet fraction was more compared to the lysate which was confirmed by WIA analysis. The recombinant protein construct was also seen to be soluble in water and this was confirmed by the absence of insoluble particle when tested with water, and from the detection of high concentration of protein using WIA analysis. In the same study by Wang, et al., the expression was observed in the supernant fraction only which was also confirmed by SDS-PAGE analysis.

Prior to recombinant protein expression and optimization, we tested on the physical disruption methods by probe sonication to ensure an efficient and complete cell lysis to release the intracellular content and thus maximizing the protein yield. Our results showed that sonication alone was not efficient and failed to achieve complete cell lysis up to the 20s/20s and 6 cycles tested despite the observation that the duration and number of cycles of sonication were proportionate to cell lysis. Results from Figure 3.3(A and B) showed that with a lower sonication pulse time (less sonication energy), the efficiency of cell lysis is poorer and thus the higher number of surviving colonies after overnight incubation. Increasing total sonication time with continuous sonication burst-cooling cycles was also shown to increase the yield of protein (more lysed cells) (Shrestha, et al., 2018). However, more sonication energy input also generated excessive flow of heat during cell disruption. Excessive heat generated especially with extended sonication cycles and higher pulse time could negatively impact on the quality of proteins (Kwon & Jewett, 2015). Moreover, it had been demonstrated that sonication could produce free radicals which would further affect the quality of proteins released (Santos, et al., 1996).

Freeze-thaw is a physical method that can be applied to disrupt the bacterial cell and releasing the total cell content. When cells are frozen, the ice crystals formed inside the cells could damage the cellular structures such as cell membrane and intracellular organelles. The mechanism increases water stress inside the cells causing the cell membrane and cell wall to be sheared causing release of proteins (and other components) from the cells. Notably, we incorporated a freeze-thaw cycle at -80^{o}C and 80^{o}C prior to sonication and the lysis effect was dramatically enhanced to 100%. Therefore, complete cell lysis was achieved with the use of sonication at pulse 15s/20s (35% amplitude) with 3 cycles and 10 mins of freeze-thawing at -80°C and 80°C, respectively. In addition, using both freeze-thaw and sonication which are physical disruption methods means we could avoid the use of chemicals/additives that may affect the stability and quality of the proteins produced. Furthermore, both techniques can be applied and upscaled conveniently in large-scale biosynthesis in the future.

Following that, optimization of the recombinant expression of CBD-Int-TMPG protein construct from the host *E. coli* BL21 (DE3) was conducted using the OFAT approach by targeting the key induction parameters which were IPTG induction concentration, induction temperature, induction duration, and incorporation of agitation. Besides increasing the yield, optimization of IPTG induction concentration can be applied to reduce the formation of inclusion bodies during recombinant protein production (Fathi-Roudsari, et al., 2016). Low concentration of inducer would result in inefficient protein expression while presence of inducer in excess may be toxic and cause impairment to cell growth thereby reducing overall protein yield (Shafiee, et al., 2017).

IPTG is similar to a molecular mimicry of lactose called allolactose that activates the *lac* promoter thus triggering transcription of the *lac* operon to induce protein expression. Like lactose, IPTG binds to the *lac* repressor thus preventing the *lac* repressor from binding to the promoter-operator region. However, IPTG is not a natural substrate of any metabolic pathways and hence its concentration will remain constant and not be used up by the cell (Fernandez-Castane, et al., 2012). From the results in figure 3.4, using 0.9mM IPTG concentration gave the thickest band for clarified cell lysate whereas for the pellet fraction the thickest band observed was using 1.2mM IPTG concentration followed by 0.9mM. According to Shafiee, et al., (2017), the use of too high concentration of IPTG can also be shown to be toxic to the cells and hence in return lower the expression level. However, in our case none of the IPTG concentration was shown to be toxic to the cell and the relative difference in expression level between induction at 1.2mM and 0.9mM IPTG was insignificant (for the pellet fraction) and therefore the lower concentration of 0.9mM IPTG was chosen. This was considered based on the potential toxicity against the host cells and cost-

effectiveness especially for future upscaling. To improve on T7 promoter-based expression system as in pTWIN1 and other similar expression systems, the addition of low concentration of glucose prior to IPTG induction could limit the basal expression of gene products particularly peptides/proteins which are potentially toxic to the host cells (Angov, et al., 2011). However, a disadvantage associated with the use of glucose is that the metabolic byproducts of glucose could acidify the culture condition thereby reducing cell fitness and overall density. Hence, glucose was not supplemented with IPTG for the purpose of expression induction in this study.

Induction temperature is another important physical factor that affect bacterial growth and protein expression. Higher temperature for induction promotes cell growth thereby increasing the total protein yield but may on the other hand causing plasmid loss (Shafiee, et al., 2017). An increase in temperature (20°C to 37°C) increases the growth rate of *E.coli*, which results in higher protein synthesis (Malik, et al., 2016). However, current results demonstrated that the optimal temperature for expression of the recombinant protein construct CBD-Int-TMPG was 25°C (Figure 3.5). This setting produced the highest protein yield when incubated for an extended period of 24h for both clarified cell lysate and cell pellet. This indicated that inducing at lower temperature (25°C) with prolonged induction duration (24h) time might result in higher expression level. The hydrophobic interactions that determine the formation of inclusion bodies are temperature dependent and hence, lower temperature enhances protein stability and leads to correct folding of the proteins (Sahdev, et al., 2008). Therefore 25°C was found to be the optimum temperature and further optimization were performed at this temperature.

Duration of induction also affects the yield of the peptide (Boshtam, et al., 2018). However in contrast, in one of the reports of Malik et al it was shown that after 1h of induction duration, the protein expression reached at its maximal level and the expression level remain unchanged with increased duration (Malik, et al., 2016). Whereas, from the results in figure 3.6, it was observed that the expression of the recombinant protein was elevated with increase in duration time (from 2h to 48h) and the highest expression was observed when induced for 48h. The involvement of agitation for inducing the culture has been shown useful for better expression of the peptide according to figure 3.7. During batch culture, with the increase in cell density, oxygen may limit the growth of the bacterial cell and protein expression. Therefore, inducing the culture with shaking leads to increased aeration and in return increased bacterial growth rate which leads to higher expression of recombinant protein (Sina, et al., 2015) .

This research has potential limitations. Some of the SDS PAGE gels were seen to have poor resolution which could be due to the incorrect running buffer pH or impurities present in the

reagents. The amount of β-mercaptoethanol (reducing agent) could have been less which led to distorted M-shaped bands in some cases. Overloading of the samples in the well could have caused the formation of thick bands.

Chapter 5: Conclusions and Future Recommendations

In our current study, the TMPG-encoding gene was successfully cloned into pTWIN1 bacterial expression vector and transformed into *E. coli* BL21 (DE3) as the host cell for recombinant protein expression. The optimized condition selected (0.9mM IPTG, 24°C for 48h with agitation) following a systematic OFAT approach significantly maximized the recombinant protein expression of CBP-Int-TMPG protein construct as compared to the baseline setting without optimization (0.3mM IPTG, 15°C for 24h without agitation). Therefore, the efficiency of the current expression system for AMP biosynthesis had been greatly enhanced.

Further studies focusing on CBD affinity purification of the recombinant protein using chitin resins followed by a temperature- and pH-shift intein-mediated self-cleavage of the native AMP (TMPG in this study) from the protein construct are recommended. The peptides yielded shall be sequence-verified and tested using suitable antimicrobial assays to validate its biological activities in comparison to the chemically synthesized peptides as the reference standard. Optimization of the culture media can also be carried out to enhance the expression of the peptide. As CBD affinity purification using the anti-CBD resins is rather costly despite convenient, an alternative cost-effective and scalable purification approach such as the Fast Protein Liquid Chromatography (FPLC) method could be applied. FPLC utilizes the medium pressure chromatography to separate peptides with high resolution according to size and charge (Pontis, 2017). The use of optimization of the peptide induction as outlined in this study can also be used for scaling up the recombinant protein production. We envision that our bacterial expression system can be applied for both a cost-effective and up-scalable approach at the larger AMP bioproduction scale which would benefit the research and development on AMPs and its potential applications in the medical and pharmaceutical industries.

References

1. Morassutti, C., Amicis, F. D., Bandiera, A. & Marchetti, s., 2005. Expression of SMAP-29 cathelicidin-like peptide in bacterial cells by intein-mediated system. *Protein expression and purification,* 39(2).

2. Andersson, D., Hughes, D. & Kubicek-Sutherland, J., 2016. Mechanisms and consequences of bacterial resistance to antimicrobial peptides. *Drug Resistance Updates,* Volume 26, pp. 4-57.

3. Angov, E., Legler, P. M. & Mease, R. M., 2011. Adjustment of Codon Usage Frequencies by Codon Harmonization Improves Protein Expression and Folding. In: *Heterologous Gene Expression in E.coli Methods and protocols.* Hertfordshire: Humana press.

4. Bahar, A. A. & Ren, D., 2013. Antimicrobial peptides. *Pharmaceuticals (Basel),* 6(12), pp. 1543-1575.

5. Boshtam, M., Shahreza, H. K., Feizollahzadeh, S., Rahimmanesh, I. & Asgary, S., 2018. Expression and purification of biologically active recombinant rabbit monocyte chemoattractant protein1 in Escherichia coli. *FEMS Microbiology Letters,* 365(9).

6. Chang, A. Y., Chau, V. W., Landas, J. A. & Pang, Y., 2017. Preparation of calcium competent Escherichia coli and heat-shock transformation. *JEMI methods,* Volume 1, pp. 22-25.

7. Costa, S., Almeida, A., Castro, A. & Domingues, L., 2014. Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in Microbiology,* Volume 5.

8. Costa, S., Almeida, A., Castro, A. & Domingues, L., 2014. Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: the novel Fh8 system. *Recombinant protein expression in microbial systems.*

9. Diaz, M., Venturini, E., Marchetti, S., Arenas, G. & Marshall, S. H., 2012. Intein-mediated expression of cecropin in Escherichia coli. *Electronic Journal of Biotechnology,* 15(2).

10. Fathi-Roudsari, M., Akhavian-Tehrani, A. & Maghsoudi, N., 2016. Comparison of Three Escherichia coli Strains in Recombinant Production of Reteplase. *Avicenna Journal of medical biotechnology,* 8(1), pp. 16-22.

11. Fernandez-Castane, A., Caminal, G. & Lopez-Santin, J., 2012. Direct measurements of IPTG enable analysis of the induction behavior of E. coli in high cell density cultures. *Microbial cell factories,* Volume 11.

12. Gaiser, R. A., Rivas, L. & Lopez, P., 2011. Production of eukaryotic antimicrobial peptides by bacteria - A review.. *Science against microbial pathogens: communicating current research and technological advances,* pp. 992-1002.

13. Geldart, K., Forkus, B., McChesney, E., McCue, M. & Kaznessis, Y. N., 2016. pMPES: A modular peptide expression system for the delivery of antimicrobial peptides to the site of Gastrointestinal infections using probiotics. *Pharmaceuticals (Basel),* 9(4).

14. Hoffman, O. & Weber, R. J., 2009. Pathophysiology and Treatment of Bacterial Meningitis. *Ther Adv Neurol Disord,* 2(6), pp. 1-7.

15. Kim, D. S., Kim, S. W., Song, J. M., Kim, S. Y. & Kwon, K., 2019. A new prokaryotic expression vector for the expression of antimicrobial peptide abaecin using SUMO fusion tag. *BMC Biotechnology.*

16. Kwon, Y.-C. & Jewett, M. C., 2015. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific reports,* Volume 5.

17. Lewies, A., Plessis, L. H. D. & Wentzel, J. F., 2018. Antimicrobial Peptides: the Achilles'Heel of Antibiotic Resistance?. *Probiotics and Antimicrobial Proteins,* 11(8).

18. Li, C., Blencke, H., Paulsen, V., Haug, T. & Stensvag, K., 2010. Powerful workhorses for antimicrobial peptide expression and characterization. *Bioengineered Bugs,* 1(3), pp. 217-220.

19. Li, Y., 2011. Recombinant production of antimicrobial peptides in Escherichias colu: A review. *Protein Expression and Purification,* 80(2), pp. 260-267.

20. Li, Y., 2011. Self-cleaving fusion tags for recombinant protein production. *Biotechnology Letters,* Volume 33, pp. 869-881.

21. Made, V., Els-Heindl, S. & Beck-Sickinger, A. G., 2014. Automated solid-phase peptide synthesis to obtain therapeutic peptides. *Beilstein J Org Chem,* Volume 10, pp. 1197-1212.

22. Makarova, O. V., Makarov, E. M., Sousa, R. & Dreyfus, M., 1995. Transcribing of Escherichia coli genes with mutant T7 RNA polymerases: Stability of lacZ mRNA inversely correlates with polymerase speed. *Microbiology,* Volume 92, pp. 12250-12254.

23. Malik, A., Alsenaidy, A. M., Elrobh, M., Khan, W., Alanazi, M. S. & Bazzi, M. D., 2016. Optimization of expression and purification of HSPA6 protein from Camelus dromedarius in E. coli. *Saudi journal of biological sciences,* 23(3), pp. 410-419.

24. Mangoni, M. L. & Shai, Y., 2009. Temporins and their synergism against Gram-negative bacteria and in lipopolysaccharide detoxification. *Biochimica et Biophysica Acta (BBA) - Biomembranes,* 1788(8), pp. 1610-1619.

25. Mitchell, S. F. & Lorsch, J. R., 2015. Protein Affinity Purification using Intein/Chitin Binding Potein Tags. In: *Methods in Enzymology.* s.l.:Elsevier Inc., pp. 111-125.

26. Narayana, J. L. & Chen, J.-Y., 2015. Antimicrobial peptides: Possible anti-infective agents. *Peptides,* Volume 72, pp. 88-94.

27. Overton, T. W., 2014. Recombinant protein production in bacterial hosts. *Drug discovery today,* 19(5), pp. 590-601.

28. Papaneophytou, C. P. & Kontopidis, G., 2014. Statistical approaches to maximize recombinant protein expression in Escherichia coli: A general review. *Protein Expression and Purification,* Volume 94, pp. 22-32.

29. Patocka, J., Nepovimova, E., Klimova, B., Wu, Q. & Kuca, K., 2018. Antimicrobial Peptides: Amphibian Host Defense Peptides. *Current Medicinal Chemistry,* Volume 25, pp. 1-21.

30. Pontis, H. G., 2017. Chapter 3 - Protein and Carbohydrate Separation and Purification. In: *Methods for Analysis of Carbohydrate Metabolism in Photosynthetic Organisms.* s.l.:Academic press, pp. 45-63.

31. Rinaldi, A. C. & Conlon, J. M., 2013. Temporins. In: *Handbook of Biologically Active Peptides: Amphibian/Skin Peptides .* s.l.:Elsevier Inc., pp. 400-406.

32. Rizkia, P. R., Silaban, S., Hasan, K., Kamara, D. S., Subroto, T., & Maksum, S. I. P., 2015. Effect of isopropyl-B-D-thiogalactopyranoside concentration on prethrombin-2 recombinan gene expression in Escherichia coli ER2566. *Procedia Chemistry,* pp. 118-124.

33. Rosano, G. L. & Ceccarelli, E. A., 2014. Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology,* Volume 5.

34. Sahdev, S., Khattar, S. K. & Saini, K. S., 2008. Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Molecular and Cellular Biochemistry,* Volume 307, pp. 249-264.

35. Sang, M., Wu, Q., Xi, X., Ma, C., Wang, L., Zhou, M., Burrows, J. F. & Chen, T., 2018. Identification and target-modifications of temporin-PE: A novel antimicrobial peptide in the defensive skin secretions of the edible frog, Pelophylax kl. esculentus. *Biochemical and Biophysical Research Communications,* pp. 2539-2546.

36. Sankaran, R., Show, P. L., Cheng, Y.-S., Tao, Y., Ao, X., Nguyen, T. D. P. & Quyen, D. V., 2018. Integration Process for Protein Extraction from Microalgae Using Liquid Biphasic Electric Flotation (LBEF) System. *Molecular Biotechnology,* pp. 749-761.

37. Santos, J., Belo, I., Mota, M. & Cabral, J., 1996. Freeze/thawing and sonication of Escherichia coli TBl cells for cytochrome b5 recovery. *Bioseperation,* pp. 81-89.

38. Schagger, H. & Jagow, G. V., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytic Biochemistry,* 166(2), pp. 368-379.

39. Shafiee, F., Rabbani, M. & Jahanian-Najafabadi, A., 2017. Optimization of the Expression of DT386-BR2 Fusion Protein in Escherichia coli using Response Surface Methodology. *Advanced Biomedical Research,* 6(22).

40. Shrestha, P., Holland, T. M. & Bundy, B. C., 2018. Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing. *Biotechniques,* 53(3).

41. Simmaco, M., Mignogna, G., Canofeni, S., Miele, R., Mangoni, M. L. & Barra, D., 1996. Temporins, antimicrobial peptides from the European red frog Rana temporaria. *Eur. J. Biochem,* Volume 242, pp. 788-792.

42. Sina, M., Farajzadeh, D. & Dastmalchi, S., 2015. Effects of Environmental Factors on Soluble Expression of a Humanized Anti-TNF-α scFv Antibody in Escherichia coli. *Advanced Pharmaceutical Bulletin,* 5(4), pp. 455-461.

43. Stawikowski, M. & Fields, G. B., 2002. Introduction to Peptide Synthesis. *Cuu Protoc Protein Sci.*

44. Subbalakshmi, C. & Sitaram, N., 1998. Mechanism of antimicrobial action of indolicidin. *FEMS Microbiology letters,* 160(1), pp. 91-96.

45. Wang, G., Mishra, B., Lau, K., Lushnikova, T., Golla, R. & Wang, X., 2015. Antimicrobial Peptides in 2014. *pharmaceuticals,* pp. 123-150.

46. Wang, H., Meng, X.-L., Xu, J.-P., Wang, J., Wang, H. & Ma, C.-W., 2012. Production, purification, and characterization of the cecropin from Plutella xylostella, pxCECA1, using an intein-induced self-cleavable system in Escherichia coli. *Applied genetics and molecular biotechnology,* Volume 94, pp. 1031-1039.

47. Wang, M., Zheng, K., Lin, J., Huang, M., Ma, Y., Li, S., Luo, X. & Wang, J., 2018. Rapid and efficient production of cecropin A antibacterial peptide in Escherichia coli by fusion with a self-aggregating protein. *BMC Biotechnology,* 18(62).

48. Warren, T. D., Coolbaugh, M. J. & Wood, D. W., 2013. Ligation-independent cloning and self-cleaving intein as a tool for high-throughput protein purification. *Protein expression and Purification,* 91(2), pp. 169-174.

49. Weiser, J. N., Ferreira, D. M. & Paton, J. C., 2018. Streptococcus pneumoniae: transmission, colonization and invasion. *Nat Rev Microbial,* 16(6), pp. 355-367.

50. Wong, W.-K.R., Ng, K.-L., Sivakumar, T., Hu, X.-H., Wang, H. & Lai, L.-C. N., 2018. *Escherichia coli: A Versatile Platform for Recombinant Protein Expression.* s.l.:IntechOpen.

51. Wu, Y., Qiao, R., Chen, T., Wu, J., & Du, S., 2017. Identification and molecular cloning of novel antimicrobial peptides from skin secretions of the Chinese bamboo leaf odorous frog (Odorrana versabilis) and the North American pickerel frog (Rana palustris). *Journal of Traditional Chinese Medical Sciences,* Volume 4, pp. 297-305.

52. Xie, Y.-G., Han, F.-F, Luan, C., Zhang, H.-W., Feng, J., Choi, Y.-J., Groleau, D. & Wang, Y.-Z., 2013. High-Yield Soluble Expression and Simple Purification of the Antimicrobial Ppetide OG2 Using the Intein System in Escherichia coli. *Biomed Research International,* pp. 1-6.

53. Xu, M.-Q. & Jr, T. C. E., 2003. Purification of Recombinant Proteins from E. coli by Engineered Inteins. *E.coli gene expression protocols,* Volume 205, pp. 43-68.

54. Parachin, N. S., Mulder, K. C., Viana, A. A. B., Dias, S. C. & Franco, O. L., 2012. Expression systems for heterologous production of antimicrobial peptides. *Peptides,* pp. 446-456.

Appendices

Reagents and Buffer:

Solution/Buffers	Composition
PBS buffer	1X stock solution: 137mM NaCl 2.7mM KCl 10mM Na_2HPO_4 1.8mM KH_2PO_4 Adjust the pH to 7.4
TBE buffer	10X stock solution: 1M Tris base 1M Boric acid 0.02M disodium salt (EDTA) Adjust pH to 8
Tris HCL pH 6.8	0.5M Tris-Cl
Tris HCL pH 8.8	1.5M Tris-Cl
SDS sample running buffer	3X solution: 188mM Tris-Cl 3%SDS 30% Glycerol 0.001% Bromophenol Blue 5% B-marcaptoethanol
Staining solution	50% Methanol 10% Acetic acid ddH_2O 0.1% Coomassie brilliant blue G250
Destaining solution	50% Methanol 10% Acetic acid ddH_2O
Ampicillin stock	50mg/ml stock solution: 250mg of Ampicillin powder Dissolved in 5ml of ddH_2O

	Stored in -20°C
SOC medium	SOB medium dissolve in appropriate volume of ddH_2O (autoclaving) 1M filtered sterilized glucose solution (0.22um filter) added to SOB
10% APS	1g APS in 10ml water multiple small aliquots and store in -20°C
TE buffer	1mM EDTA 10mM Tris-Cl ddH_2O
Bjerrum Transfer buffer	48mM Tris base 39mM Glycine 20% Methanol ddH_2O 10% SDS
Tricine SDS running buffer	10X stock solution: 1M Tris base 1M Tricine 1% SDS ddH_2O
Blocking solution	5% Skim milk 0.1% Tween-20