Production and Purification of Membrane Scaffold Protein and Class IID bacteriocin Enterocin EJ97 from *Enterococcus faecalis* EJ97

Mahendra Chaudhary

Department of Biosciences
Faculty of Mathematics and Natural Sciences
University of Oslo

June 2017

Production and Purification of Membrane Scaffold Protein and Class IID bacteriocin Enterocin EJ97 from *Enterococcus faecalis* EJ97

Mahendra Chaudhary
Department of Biosciences
Faculty of Mathematics and Natural Sciences
University of Oslo
June 2017

June 2017

Production and purification of Membrane Scaffold Protein and Class IID bacteriocin Enterocin EJ97 from *Enterococcus faecalis* EJ97

Mahendra Chaudhary

http://www.duo.uio.no
Print: Reprosentralen, Universitetet i Oslo

Acknowledgements

The work presented in this thesis was carried out at the Department of Biosciences, Faculty of Mathematics and Natural Sciences, University of Oslo.

It gives me great pleasure in acknowledging my supervisor, Dr. Per Eugen Kristiansen for the opportunity to carry out this research in his laboratory. His support and help has remained phenomenal in accomplishing my thesis. Without his guidance and insightful advice, the completion of this work could not be possible and I appreciate the enthusiasm he has shown for my work.

I would like to thank Dr. Camilla Oppegård for her support and help to start the experiment in the laboratory. I am grateful to Bie Ekblad for her continuous help and skillful technical assistance in the laboratory. Thank you so much for always being there whenever I needed you and for always taking the time to answer my questions.

I am thankful to my dear friends and many respected teachers of UiO, who helped me maintaining the study environment sound and sharing their academic experience, that way always boosted me to seek for better performance.

Special thanks goes to all the members of Jon Nissen-meyer group and K. Kristoffer Andersson group for their precious feedback and co-operation throughout the period.

At last but not the least, I want to acknowledge my parents, wife and all family members whose love, support and perpetual inspiration have always encouraged me to pursue higher academic excellence.

Oslo, June 2017

Mahendra Chaudhary

Abstract

Enterocin EJ97 is cationic gram-positive bacteriocin, having 44 amino acid residue and molecular mass of 5.32 kDa produced by *Enterococcus faecalis* EJ97. EntEJ97 is very stable under mild heat conditions, sensitive to proteolytic enzymes and active on several gram-positive bacteria. It belongs to the LsbB-like class IId bacteriocins and contain C-terminal sequence similar to LsbB-like bacteriocins (LsbB, EntQ and EntK1) with some variations on N-terminal. LsbB and EntEJ97contains KXXXGXXPWE conserved motif. The structure of LsbB and EntK1 has been determined by NMR spectroscopy, where they show similar secondary structure. NMR spectroscopy offers lots of information on peptide/membrane systems. Membrane proteins have recently been studied in phospholipid bilayer containing nanodisc.

The optimized expression conditions can offer enough proteins required for the structural and functional analysis. In order to achieve this goal, expression conditions were optimized to increase the protein yield, which was then purified by nickel affinity chromatography.

The result shows that the expression of membrane scaffold protein (MSP) was most successful in *E.coli* BL21(DE3) strain cultured in LB medium. Induction at OD_{600} = 0.7 with 1mM IPTG at 30^{o}C for 4 hours produced high level of MSP in a soluble form. The expression of EntEJ97 fusion protein with the immunoglobulin-binding domain of streptococcal protein G (his-GB1-tev-EntEJ97) was successful in modified minimal medium induced at OD_{600} = 0.7 with 2 mM IPTG at 35^{o}C for 6 hours. However, the protease cleavage failed, leading to unsuccessful purification of EntEJ97 peptide.

Initial cleavage was performed with 1:100 and 1:50 protease to protein ratio at room temperature, though no positive hits were found. We hope that further trials would be performed for the cleavage of fusion peptide (his-GB1-tev-EntEJ97) to yield pure EntEJ97 protein with improved cleavage protocol in future.

Keywords: bacteriocin, enterocin EJ97, membrane scaffold protein (MSP), expression optimization

Table of contents

List of Tables

List of Figures

1 Introduction

Antibiotics are used for the treatment of infections, but in recent years it has experienced that pathogens are gaining resistant to commonly used antibiotics (Cotter et al., 2013). Therefore, there is a great interest in the development of alternative treatment for the infections, leading to the search for new therapeutic alternatives (Gordon et al., 2005) and new antimicrobials that can be used in clinical settings (Cotter et al., 2013). Among the most promising alternative options, ribosomally synthesized antimicrobial peptides (AMPs) have become a focus area in antibiotic research. AMPs are widely distributed in nature and are produced by a wide variety of organisms from bacteria to plants, animals and humans (Boman, 1995; Nissen-Meyer and Nes, 1997). AMPs are important for defense against microorganisms in plants and animals. They are an integral part of the innate immunity that helps in the protection against different microorganisms such as fungi and bacteria (Nissen-Meyer and Nes, 1997). AMPs have the potential for the development of useful antimicrobial additives and drugs (Lüders et al., 2003).

AMPs are often membrane permeabilizing, cationic, and amphiphilic or hydrophobic in nature (Nissen-Meyer and Nes, 1997). Different AMPs have been isolated and identified from eukaryotes and prokaryotes. In eukaryotes, AMPs constitute an important part of the host defense system (Hancock and Diamond, 2000; Hoffmann et al., 1999). In prokaryotes, AMPs are found most frequently in gram-positive bacteria (Lüders et al., 2003). Bacterial AMPs being active at pico to nanomolar concentration in contrast to the micro-molar concentration of eukaryotic AMPs (Lüders et al., 2003) and higher concentrations of antibiotics (Nes, 2011); bacterial AMPs have higher potency than eukaryotic AMPs and antibiotics. In addition, bacterial AMPs have a narrower target cell spectrum than eukaryotic AMPs (Nissen-Meyer and Nes, 1997).

Ribosomally synthesized AMPs produced by prokaryotes (bacteria) are known as bacteriocins. The Belgian scientist Andre Gratia first discovered bacteriocins in 1925 (Gratia, 1925). Bacteriocins usually contain between 25 and 70 amino acids (aa) (Diep and Nes, 2002) and are normally active against only related species or genera (Ovchinnikov et al., 2014). Both Gram-positive and Gram-negative bacteria produce bacteriocins. Bacteriocins include post-translationally modified peptides containing lanthionine and unmodified peptides (without lanthionine) (Ovchinnikov et al., 2014).

The first chapter will introduce basic knowledge on bacteriocins produced by lactic acid bacteria (LAB), its classification and their structures. Leaderless bacteriocin, Enterocin EJ97 is discussed on the second part. Subsequently, reviewed the membrane scaffold protein. Finally, the end of the chapter presents the aim of the study.

1.1 LAB bacteriocins

LAB are a group of Gram-positive, non-spore forming, nonrespiring, either coccus- or rod-shaped bacteria, which produce lactic acid as the major end product during the fermentation of carbohydrates (Schleifer and Ludwig, 1995). LAB are used in the production of milk and milk products (like cheese, yogurt, buttermilk and kefir), in food and feed additives (Nissen-Meyer et al., 1992). As LAB is a part of the natural flora in food humans have consumed for centuries, and they constitute a significant part of indigenous flora of mammals, including humans; bacteriocins produced by LAB may be considered safe for preventing growth of pathogenic/undesirable mirco-organisms (Nissen-Meyer et al., 2009). The LAB bacteriocins nisin and pediocin PA1 have been approved as preservative for use in food products (Cotter et al., 2005). The lantibiotic nisin was added to the European food additives list as number E234 in 1983 (European Economic Community Commission Directive 83/463/EEC) and was approved generally regarded as safe (GRAS) by the US Food and Drug Agency (FDA) in 1988. Among the bacteriocins produced, only nisin has been approved by the World Health Organization for use as a food preservatives and is commercialized as a dried concentrated powder (Sobrino-López and Martin-Belloso, 2008). The LAB bacteriocins have also potential use in medical application. For instance, oral intake of bacteriocins producing LAB protects mice against lethal doses of *Listeria monocytogenes* (Corr et al., 2007).

1.1.1 Classification of LAB bacteriocins

The LAB peptide bacteriocins have been classified into two main categories: The lanthionine-containing bacteriocins (Class I) and the non-lanthionine containing bacteriocins (Class II)(Cotter et al., 2005; Nissen-Meyer et al., 2009). Figure 1 shows how the peptides are divided into several Classes and sub Classes.

Class I includes post-translationally modified peptide bacteriocins, often reffered as lantibiotics in which D-alanine (derived from a serine residue) and an L-alanine (derived

from a cysteine residue) are linked by a sulphur atom forming thioether bond. Class I bacteriocins also contain methyl-lanthionine, and other non-standard residues such as dehydroalanine, dehydrobutyrine and D-alanine (Cotter et al., 2005; Pag and Sahl, 2002). Bacteriocins without modified residues (lack extensive post-translational modifications) are grouped under class II (Cotter et al., 2005). Class II bacteriocins are divided into four sub classes: class IIa, class IIb, class IIc and class IId (Cotter et al., 2005) (Figure 1).

Class IIa contains the anti-listerial one- peptide pediocin-like bacteriocins that have similar amino acid sequences (Fimland et al., 2005) whereas in class IIc N- and C-terminal ends of the peptide are covalently linked thus giving cyclic structure (Cotter et al., 2005). Class IIb contains two peptide bacteriocins, in which two different peptides required optimal antimicrobial activity (Oppegård et al., 2007). Class IId contains linear one-peptide non-pediocin-like bacteriocins (Cotter et al., 2005). Enterocin EJ97, which is the focus in this thesis, belongs to this class IId. In this thesis, the detail on class IId is the focus however; the 3D structures of representatives of sub classes will be shown.

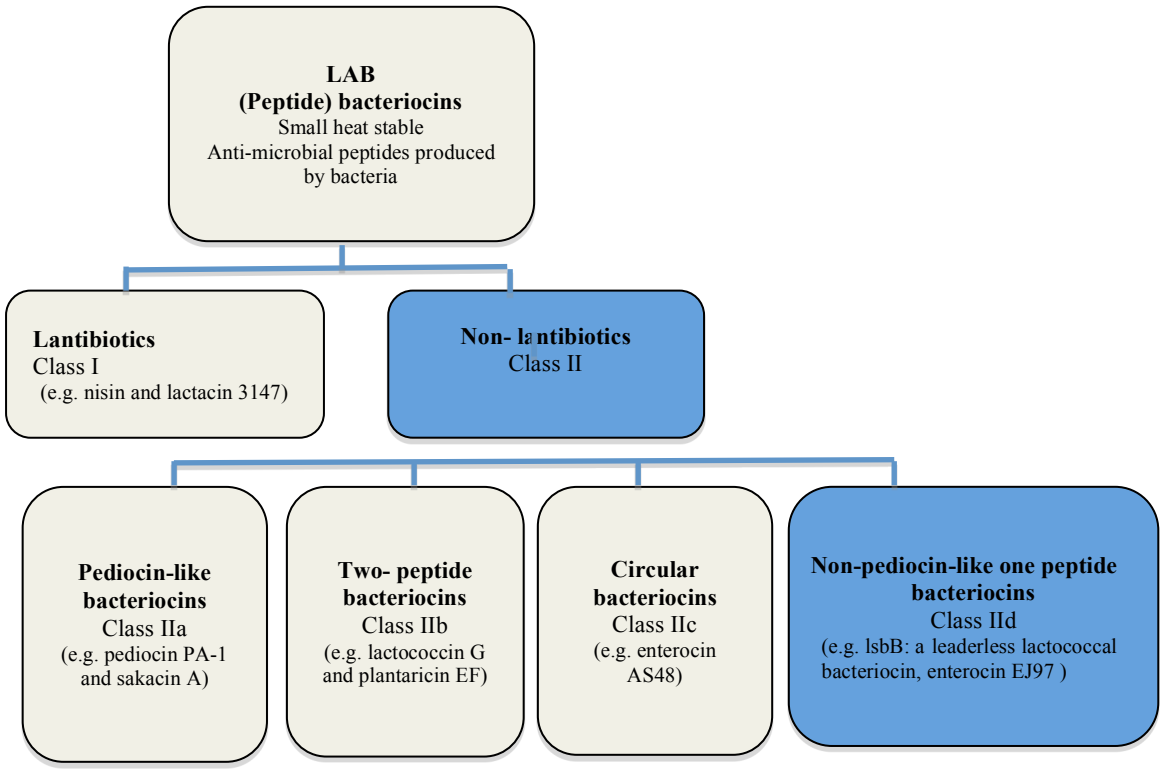


Figure 1. Classification of LAB bacteriocins.

1.1.2 Three- Dimensional Structure of bacteriocins

Circular dichroism (CD) spectroscopy and nuclear magnetic resonance (NMR) structural analysis are methods used for the determination of three-dimensional structures of bacteriocins. Most bacteriocins studied so far are unstructured in water solution but becomes structured upon interaction with membrane/membrane mimicking system (Fregeau Gallagher et al., 1997; Watson et al., 2001). However, Class IIc (circular bacteriocin) enterocin AS-48 (Cebrian et al., 2015) and Class IId (leaderless bacteriocin) enterocins 7A and 7B (Lohans et al., 2013) are structured in water.

The 3D structure of several different pediocin-like bacteriocins (class IIa) have been reported (Fimland et al., 2005). Class IIa bacteriocins are well studied and more than 20 bacteriocins have been characterized. Some of them are: leucocin A, sakacin P, curvacin A, mesentericin Y105 and pediocin PA-1. In Figure 2 the structure and cartoon representation of Class IIa bacteriocin Curvacin-A is shown (Figure 2).

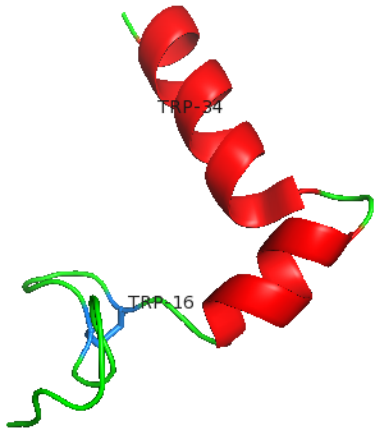


Figure 2. A Cartoon depition of pediocin-like bacteriocins (Curvacin-A, PDB: 2A2B). Disulphide bridge is shown in blue colour, helix and the loop are indicated by red and green colours respectively.

There are at least 15 two-peptide (class IIb) bacteriocins that have been isolated (Nissen-Meyer et al., 2010) and characterized since the first identification of bacteriocin lactococcin G (Nissen-Meyer et al., 1992). NMR structural studies of five Class IIb bacteriocins have been studied so far, namely lactococcin G (LcnG) (Rogne et al., 2008), plantaricin E/F (PlnEF) (Fimland et al., 2008) and plantaricin J/K (PlnJK) (Rogne et al., 2009), Carnobacteriocin X (CbnX) and Carnobacteriocin Y (CbnY) (Acedo et al., 2017). The proposed structure and membrane orientation of Lactococcin G is shown in Figure 3 (Rogne et al., 2008).

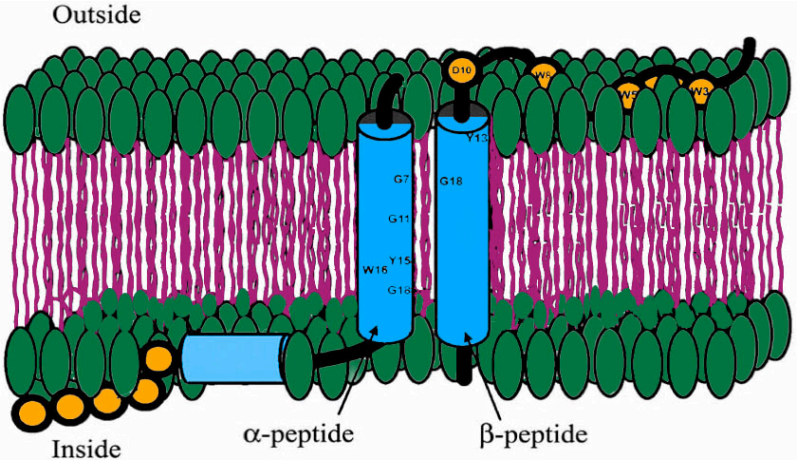


Figure 3. Proposed model of lactococcin G and its orientation in target cell membranes.
The cylinders indicate helical regions and the indicated trans-membrane helical regions in α and β peptides. The peptides (yellow) lie in a parallel orientation with their C-terminal ends on the cytoplasmic side of the membrane. Mutagenesis studies suggest that three tryptophan residues (W3, W5 and W8) in the N-terminal part of the β peptide positions themselves in the outer (interface) region, while the five basic C-terminal residues (R35, K36, K37, K38 and H39) in the α peptide are thought to position themselves on the cytoplasmic side of the membrane. The negatively charged residue (D10) in the β peptide is near the N-terminal residue in α peptide (Rogne et al., 2008).

To date, 13 different cyclic (Class IIc) bacteriocins have been characterized (Cebrian et al., 2015). The 3D structures for three circular bacteriocins, subtilosin A (Figure 4A), enterocin AS-48 and carnocyclin A have been determined.

Only few 3D NMR solution structures of leaderless bacteriocin have been reported. Some of them are enterocin 7A and 7B (Lohans et al., 2013), LsbB (Figure 4B) (Ovchinnikov et al., 2014) and Enterocin K1 (Figure 4C) (Ovchinnikov et al., 2017). The structure of LsbB (Figure 4B) shows an N-terminal α-helix whereas the C-terminal of the molecule remains unstructured.

.

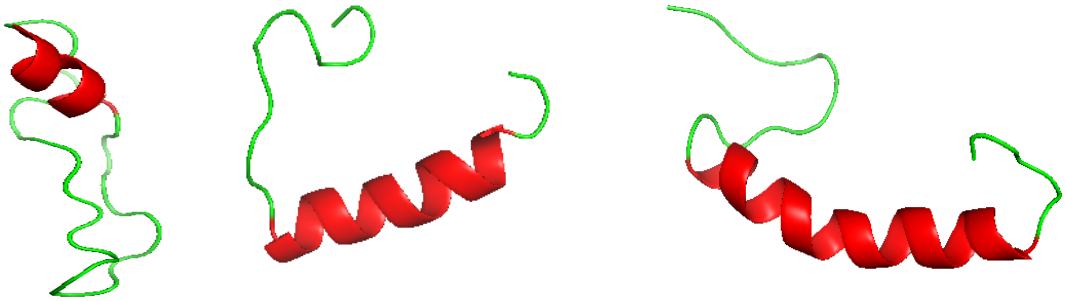


Figure 4. Cartoon representation of class IIc and class IId bacteriocins. 3D structures of A) subtilosin A (PDB:1PXQ) class IIc, B) LsBb (PDB:2MLV), and C) Ent K1 (PDB: 5L82) class IId bacteriocin. Helix and loop are represented with red and green colours respectively.

1.2 Class IId bacteriocins

The bacteriocins that have no sequence similarity to the pediocin-like bacteriocins, non-cyclic and single linear bacteriocins are under this group (class IId). This class contains leader and leaderless bacteriocins. The bacteriocins that are synthesized with an N-terminal leader and transported by a dedicated ATP-binding cassette (ABC) transporter (Cintas et al., 1997) are termed leader bacteriocins whereas leaderless bacteriocins are synthesized without an N-terminal leader sequence or signal peptide. Table 1 shows the examples of class IId bacteriocins.

Table 1. Class IId bacteriocins

Bacteriocin		Producer strain	Mass Da, (amino acids)[a]	Reference
1. Leaderless bacteriocins				
Aureocin A70	AurA	*S. aureus* A70	2,924 (31)	(Netz et al., 2001)
	AurB		2,797 (30)	
	AurC		2,955 (31)	
	AurD		3,087 (31)	
Aureocin A53		*S. aureus* A53	6,012 (51)	(Netz et al., 2002)
BHT-B		*S. rattus* strain BHT	5,195 (45)	(Hyink et al., 2005)
Enterocin K1		Chemically synthesized	(37)	(Ovchinnikov et al., 2014)
Enterocin L50A		*E. faecium* L50	5,190 (44)	(Cintas et al., 1998)
Enterocin L50B		*E. faecium* L50	5,178 (43)	(Cintas et al., 1998)
Enterocin MR10A		*E. faecalis* MRR10-3	5,202 (44)	(Martín-Platero et al., 2006)
Enterocin MR10B		*E. faecalis* MRR10-3	5,208 (43)	(Martín-Platero et al., 2006)
Enterocin Q		*E. faecium* L50	3,980 (34)	(Cintas et al., 2000)
Enterocin EJ97		*E. faecalis* EJ97	5,322 (44)	(Sánchez-Hidalgo et al., 2003)
Lacticin Q		*L. lactis* QU 5	5,926 (53)	(Fujita et al., 2007)
Lacticin Z		*L. lactis* QU 14	5,971 (53)	(Iwatani et al., 2007)
LsbB		A natural isolate of *L. lactis*	3,407 (30)	(Gajic et al., 2003)
2. Leader bacteriocins				
Lactococcin A		Strains of *L. lactis*	5,778 (54)	(Holo et al., 1991)
Lactococcin M		Strains of *L. lactis*	4,325 (48)	(Van Belkum et al., 1991a)
Lactococcin B		Strains of *L. lactis*	5,327 (47)	(Van Belkum et al., 1992)
Carnobacteriocin A		*C. piscicola* LV17A	5,052 (53)	(Worobo et al., 1994)
Piscicolin 61		*C. piscicola* LV61	5,052 (53)	(Holck et al., 1994)
Enterocin B		*E. faecium* T136	5,479 (53)	(Casaus et al., 1997)
Durancin TW-49M		*E. durans* QU 49	5,228 (54)	(Hu et al., 2008)
Divergicin 750		*C. divergens* 750	3,447 (34)	(Holck et al., 1996)
LsbA		A natural isolate of *L. lactis*	5,226 (44)	(Gajic et al., 2003)
Bovicin 255		*S. gallolyticus* LRC0255	5,967 (56)	(Whitford et al., 2001)
Thermophilin A		*S. thermophiles* ST134	5,776 (62)	(Ward and Somkuti.,1995)

[a]Theoretical molecular masses calculated from the primary structures. The numbers in the parentheses show the number of the amino-acid residues of mature peptides.

1.2.1 Leaderless bacteriocins

Leaderless bacteriocins are relatively small with 30-50 aa residues and contain no modifications (Ovchinnikov et al., 2014). The member of these bacteriocins do not involve an N-terminal leader sequence or any post translational modifications, which are responsible for the recognition process during secretion and for inactivation of bacteriocins inside producer cells (Masuda et al., 2012). However, export may be mediated by a dedicated ABC transporter (Iwatani et al., 2013; Masuda et al., 2012). 20 distinct leaderless bacteriocins have been reported to date (Masuda et al., 2012). Preliminary evaluation of the structures of aureocin A53 and lacticin Q from *S. aureus* A53 and *Lactococcus lactis* QU5 respectively have been reported with the help of CD and fluorescence experiments (Netz et al., 2002; Yoneyama et al., 2009). Three-dimensional structures of only few leaderless bacteriocins have been determined.

1.2.2 Sequence alignment of LsbB-like bacteriocins

LsbB is a hydrophilic peptide (Uzelac et al., 2015), consisting 30 aa residues (Ovchinnikov et al., 2014) produced by *Lactococcus lactis* BGMN1-5 isolated from cheese. The sequence alignment of LsbB and LsbB-like bacteriocins are shown in Figure 5.

Sequence of bacteriocins	Name
MGVYRYLELNNCFFIRVYSLWGHWISVQKLTRYEIAWFKNKHGCYPWEIPRCYHW	Duracin 41D-HP
MLAKIKAMIKKFPNPYTLAAKLTTYEINWYKQQYGRYPWERPVA	EntEJ97
MLAKIKAMIKKFPTPYTLAAKLTTYEINWYKQQYGRYPWERPVA	EntEJ97T-HP
MKFKFNPTGTIVKKLTQYEIAWFKNKHGYYPWEIPRC	EntK1
MNFLKNGIAKWMTGAELQAYKKKYGCLPWEKISC	EntQ
MKTILRFVAGYDIASHKKKTGGYPWERGKA	LsbB
KxxxGxxPWE	Consensus

Figure 5. Sequence alignment of LsbB-like bacteriocins. (Color codes pink, orange and dark blue indicate identical amino acids, light blue indicate very similar amino acids, and green indicate similar amino acids (Ovchinnikov et al., 2014).

LsbB, EntQ, EntK1 and EntEJ97 show same sequence alignment at C-terminal with some variations on N-terminal. LsbB peptide contains Trp 25 and Glu 26 as a part of conserved motif KXXXGXXPWE (Figure 5) (Ovchinnikov et al., 2014). For instance all LsbB-like bacteriocins contains such motifs.

Structure study of LsbB (Ovchinnikov et al., 2014) and EntK1 (Ovchinnikov et al., 2017) clearly revealed two distinct domains within the molecule: α-helical (N-terminal) and unstructured (C-terminal). The C-terminal of LsbB interacted with the bacteriocin receptor and Trp 25 was crucial for receptor binding (Ovchinnikov et al., 2014). The antimicrobial activity and their ability to block the activity of the full length peptide LsbB was examined and the LsbB truncated peptide (consisting of last 15 C- terminal residues) blocked the antimicrobial activity of LsbB but other truncated peptides could not (Ovchinnikov et al., 2014). To attain the blocking, the minimal molar concentration of the truncated peptide was~ 100 times higher than that of LsbB.

LsbB-like bacteriocins have high sequence similarity with LsbB at their C-terminal halves which suggests they might recognize the same receptor as LsbB (Ovchinnikov et al., 2014). The truncated peptides with sequence derived from the last 20 aa of EntEJ97 and Ent K1 (EntEJ97-C20 and EntK1-C20) were able to block the LsbB activity indicating that these peptides probably bind to the same receptor as LsbB (Ovchinnikov et al., 2014). LsbB primarily targets lactococcal cells (Gajic et al., 2003) and the enterocins with wider inhibition spectrum also includes lactococcal cells (Cintas et al., 2000; Galvez et al., 1998). EntEJ97 is one of the bacteriocins with activity against many bacterial strains (Gálvez et al., 1998) and belongs to LsbB-like bacteriocins.

1.2.3 Enterocin EJ97

Enterocins are the bacteriocins produced by many enterococcal strains of enterococci bacteria. Enterococci are gram-positive, facultative anaerobic, catalase-negative, and non-spore-forming bacteria (Fisher and Phillips, 2009) that produce lactic acid as the major end product of glucose fermentation (Neira et al., 2010). Enterococci are present in the environment and natural inhabitants of human and animal gastrointestinal tract. They are able to survive a range of stresses and hostile environment; tolerant of extreme temperature (5-65^{o}C), pH (4.5-10.0) and high sodium chloride concentrations (Fisher and Phillips, 2009). They are also found in many fermented food products made from milk and meat, especially cheeses and sausages (Giraffa, 2002). *Enterococcus avium and Enterococcus hirae,* as well as other enterococcal species are found in human stool samples (Nes et al., 2014). Among different species of *Enterococcus, Enterococcus faecalis* (90-95%) and *Enterococcus faecium* (5-10%) are two common commensal organisms in the intestine of humans. Many enterocins

are from various enterococcal species and most of the characterized enterocins are from *Enterococcus faecium* and *Enterococcus faecalis.*

Enterocin EJ97 is cationic gram-positive, low molecular mass (5.32 kDa) bacteriocin produced by *Enterococcus faecalis* EJ9. The bacteriocin, enterocin EJ97 is very stable under mild heat conditions, sensitive to proteolytic enzymes and active on several gram positive bacteria including enterococci, several species of *Bacillus, Listeria* and *Staphylococcus aureus* (Gálvez et al., 1998). In 1998, Antonio Gálvez and his co-workers isolated the strain of *Enterococcus faecalis* EJ97 from municipal wastewater. From that strain, the authors isolated the first 18 N-terminal amino acid residues of enterocin EJ97. The complete 44 amino acid sequence of enterocin EJ97 was determined in 2002 (Sanchez-Hidalgo et al., 2003).

1.2.4 Gene organization of Enterocin EJ97

The conjugative plasmid pEJ97 of 60-kb from bacteriocinogenic strain *E. faecalis* EJ97 can produce enterocin EJ97. All genes required for production and secretion of the enterocin EJ97 are located in 11.3 kb operon (Figure 6) of plasmid pEJ97. There are four clusters of genes present in the enterocin EJ97 operon: i) the enterocin EJ97 region, ii) a putative transposase-resolvase module, iii) genes involved in plasmid replication and iv) genes related to the sex pheromone response (Figure 6). Some of them correspond to homologues of genes on the pheromone-responding enterococcal plasmids pAD1, pPD1, and pAM373, as well as conserved or hypothetical proteins in *Lactoccocus* and *Streptococcus* species.

The open reading frames (ORFs) (numbered ORF1 to ORF12) found upstream and downstream of *ej97A* gene are in the same direction as *ej97A*. The enterocin EJ97 region consists of the *ej97A* structural gene (corresponding to ORF2) plus three ORFs (*ej97B*, *ej97C* and *ej97D*) and three putative ORFs transcribed in the opposite direction (orfA, orfB and orfC) to *ej97A*. The third ORF (*ej97B*) downstream from *ej97A* shows strong homology in its C-terminal domain to the superfamily of bacterial ABC-transporters.

The ORFs (ORF1 and ORF6) flanking the EJ97 module have been identified as a transposon-like structure. The ORF1 shows similarities to a transposase of the *Lactoccocus lactis* element ISS1 and is up to 50% identical to IS1216. The ORF6 (res *ej97*) shows strong

homology (97.2%) to the resolvase of plasmid pAM373 (De Boever et al., 2000) and up to 40 to 50 % similarity to a family of DNA recombinase and the recombinase of several multiresistant plasmids and transposons from *Staphylococcus aureus* and *Enterococcus faecalis*.

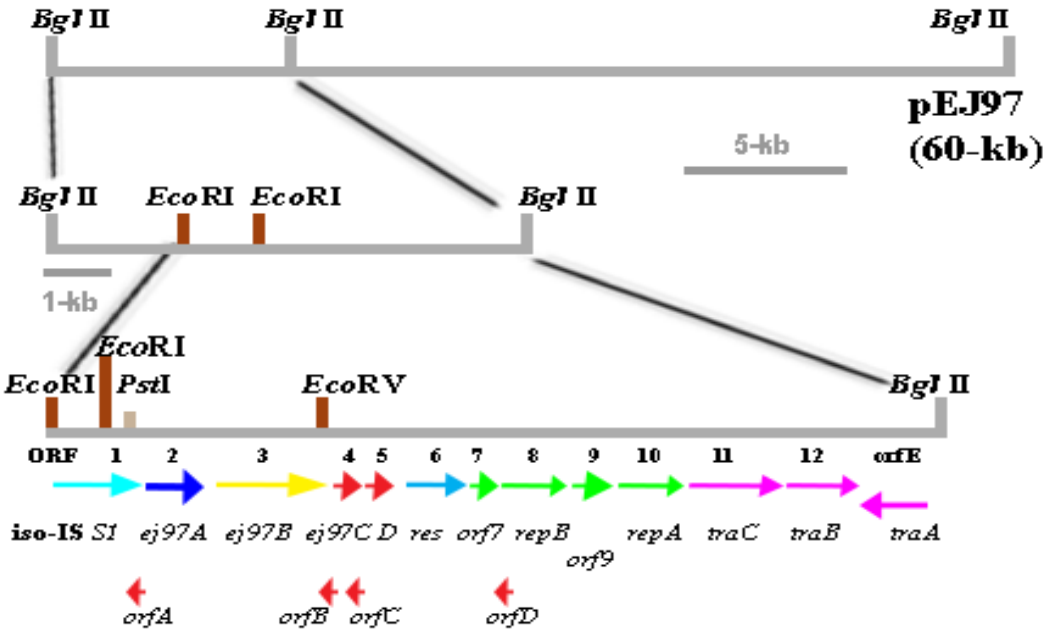


Figure 6. Genetic organization of 11.3 kb operon of enterocin EJ97 from plasmid pEJ97. Horizontal arrows with different colors indicate ORFs on plasmid pGEM-57 and the direction of transcription. The transposase gene is indicated with sky blue arrow, gene containing EJ97 is shown in dark blue, ABC transporter with yellow, gene containing hypothetical proteins in red, the putative genes involved in plasmid DNA replication in green, resolvase gene with blue and the sex pheromone response genes in pink arrow respectively (Sanchez-Hidalgo et al., 2003).

The DNA regions ORF7 to ORF10 encode the replication functions of the plasmid ej97. The ORF8 (named *repB*) predicts a gene product with high similarities to members of the ParA proteins (ParA is a family of ATPases involved in active partitioning of diverse bacterial plasmids and includes the bacterial proteins IncC, MinD, SopA and RepA). These proteins are thought to play a role in plasmid maintenance and replication (Bignell and Thomas, 2001). The ORF9 (*orf9*) could also be involved (together with the product of ORF8) in pEJ97 partitioning during cell division. The ORF10 (named *repA*) encodes a polypeptide of 317 amino acids which had significant sequence similarity to the replication initiation proteins (RepA) from several plasmids of gram positive bacteria: enterococcal plasmids pAM373 (68.8%) (De Boever et al., 2000), pPD1 (47.2%) (Nakayama et al., 1998), pCF10 (46.9%) (Ruhfel et al., 1993) and pAD1 (45.3%) (Weaver et al., 1993).

The other ORFs (ORF11, ORF12 and ORFE) from pEJ97 fragment cloned in pGEM-57 showed significant homology to genes of the pheromone response regulatory region of pPD1, pAD1, pCEF10 and the plasmid pAM373 from *Enterococcus Faecalis* (Sanchez-Hidalgo et al., 2003).

1.2.5 Structural characterization of Enterocin EJ97

Jose L. Neira and his co-workers did the structural characterization of natively unfolded enterocin EJ97 in 2010. The authors combined experimental techniques in aqueous solutions and bioinformatics approach for the structural characterization of 44 residues-long enterocin EJ97. In the experimental part, the authors used number of spectroscopic techniques, namely fluorescence, Fourier Transform Infrared (FTIR) Spectroscopy, Circular Dichroism (CD) and Nuclear Magnetic Resonance (NMR) spectroscopy to characterize the conformational properties of EJ97. Then, the authors used bioinformatics prediction programs such as PredictProtein, PONDR, RONN, IUPRED and FoldIndex as an aid to complement the experimental information about the conformational properties of EJ97.

The spectroscopic techniques showed that EJ97 is monomeric (lacks tertiary structure), natively unfolded protein in aqueous solution however, it showed some degree of α-helix or turn-like secondary structures (Table 2 and Figure 7) in water-buffer system. NMR evidences (Figure 8) also support the presence of helical- or turn-like conformation.

Table 2. Secondary structural analysis of EJ97 (in %) as determined by CD[a] and FTIR[b].

Structural assignment	Circular dichroism (k2D)	FTIR
α-helix	45	0
β-sheet	14	5.6
β-turns or loops		35.9
Random-coil	41	58.5

Table 2. [a]CD experiments were acquired at pH 7.0 (phosphate buffer), at 10μM of protein concentration and at 298 K. Deconvolution was carried out using the k2D software. [b] Experiments were carried out at 298 K. (Table taken from Neira et al., 2010).

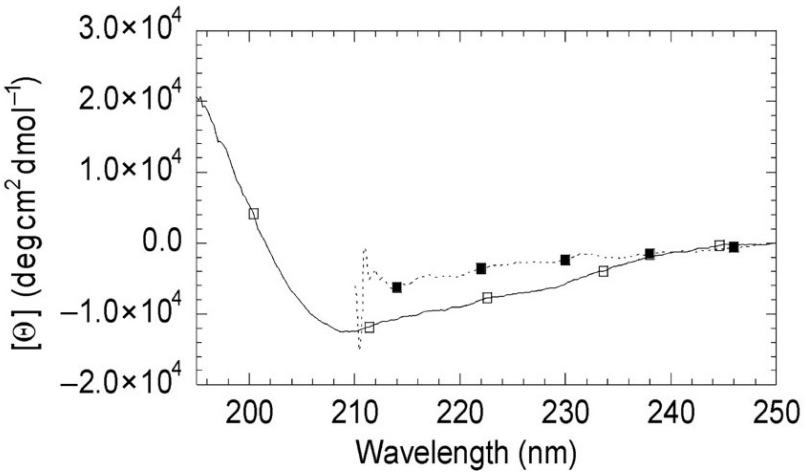


Figure 7. Far-UV CD spectra of EJ97 in aqueous solution. Far-UV CD spectra was measured in aqueous solution, phosphate buffer (blank squares, continuous line) and at 5 M GdmCl, pH 7.0 (filled squares, dotted line). 10mM buffer concentration was used in all cases. The concentration of EJ97 was 2 μM and 10 μM for fluorescence and CD spectra respectively. Spectra were acquired in either 1cm (fluorescence) or 0.1cm (CD) pathlength cells. Experiments acquired at 298 K. (Figure taken from Neira et al., 2010).

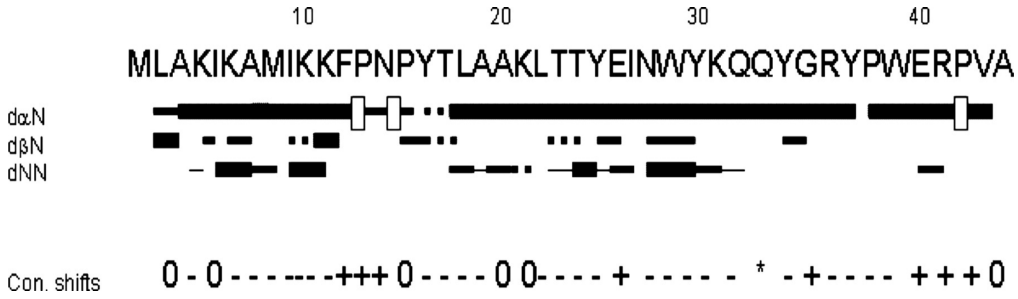


Figure 8. Summary of observed inter-residue sequential NOEs. The corresponding sequential Hα NOE with the Hδs of the adjacent proline residue are indicated by an open rectangle in the bar corresponding to the αN (i, i + 1) NOEs. Residues whose conformational shifts are larger than 0.1 ppm, positive or negative, are represented by a '+' or '−', respectively; '0' applies to those residues whose conformational shifts are lower (positive or negative) than 0.1 ppm; and '*' indicates residues which were not assigned. In Gly35, one Hα proton has a negative conformational shift. (Figure taken from Neira et al., 2010).

The theoretical analysis of the EJ97 sequence by PONDR (predictor of natural disordered regions), IUPRED, RONN and PredictProtein softwares predicted similar results that most of the polypeptide chain was ordered (Neira et al., 2010). In fact, the predicted helical region (Pro15-Gln33) by PredictProtein coincided with the polypeptides where the largest number of NN(i,i+1) NOEs were observed (Figure 8). Structure prediction servers (PONDR® VL-XT, JPred 4) used by Ovchinnikov et al. (2017) also indicated that EJ97

contains an α-helix from Lys10 to Gly35 with an unstructured C-terminal tail of 10 residues. However, Foldindex predicted that the whole EJ97 polypeptide chain was disordered (Neira et al., 2010).

The results from the application of different theoretical predictors were confusing as some algorithms predicted EJ97 a folded polypeptide chains whereas others predicted a disordered structure. Also, some predicted helical region from Pro15- Gln33 and some from Lys10- Gly33. However, the experimental results showed some degree of α-helix (Figure 7) (Neira et al., 2010).

The experimental results are unsatisfactory. The authors stated that addition of alcohols (trifluoroethanol or methanol) or other aprotic solvents (DMSO or acetonitrile) induced aggregation, which deduced signal broadening in the NMR spectra. This shows that the authors had used higher concentrations (50%) of alcohols or other polar aprotic solvents (50% DMSO or 50% acetonitrile). Since trifluoroethanol (TFE) are known for structure inducers, shares its helix-enhancing properties with other alcohols, TFE-water mixtures exhibit anomalous viscosities (suggestive of a change in water structure) (Cammers-Goodwin et al., 1996), inducing beta-sheets in some and α- helices in others when used in lower concentration. While at relatively high concentrations in water TFE disrupts hydrophobic interactions, denaturing tertiary and quaternary structure (Lau et al.,1984). And, lot more information could have been obtained in dodecylphosphocholine (DPC) micelles in addition to measuring in aqueous systems.

The structure of EntK1 and LsbB determined by CD and NMR spectroscopy showed α-helical region in N-terminal and unstructured C-terminal as that of EntEJ97. CD structures for bacteriocins EntK1 and LsbB were obtained at different (30 and 50%) TFE concentrations and the peptide to lipid ratio of 1:200 (12 mM DPC) respectively. Whilst NMR structures were obtained at 50% D3-TFE (with Milli-Q (MQ) water for both the bacteriocins (Ovchinnikov et al., 2017. This suggests that similar conditions could be used for confirming the structure of EntEJ97. Doing more experiments it should be possible to determine EntEJ97.

1.3 Membrane Scaffold Proteins

The interaction of peptides and proteins with a biological membranes is inner to a abode of biological processes for the transportation of nutrients and signalling molecules across the cell membrane; and functions (e.g. the actions of antimicrobial peptides, viral infection of cells, the function of ion channels and hormone actions) (Wong, 2017). Membrane proteins are responsible for a wide range of essential physiological processes and are difficult to study from the mechanistic perspective as many of the biophysical and chemical techniques applicable to soluble enzymes fail to deal with insoluble aggregates (Bayburt and Sligar, 2010). A membrane protein in a solubilized state allows ease in purification, functional biochemical assay, application for various spectroscopies, and crystallization for structure determination and biochemical manipulations that maintain the target protein in a stable state. Most biophysical methods for studying membrane proteins in vitro require a membrane-mimicking environment to stabilize the protein (Seddon et al., 2004). Various membrane mimicking / model membrane systems have been employed in functional studies such as vesicles, micelles, bicelles, liposomes and nanodiscs. The search for a suitable membrane model started with the use of unilamellar phospholipid vesicles that are smaller than native membranes. The morphology of this vesicle is similar to that of biomembranes. The ease in preparation has favoured the use of liposomes (or multilamellar vesicles, MLVs) (Nieh et al., 2004), which can be made more easily. Unfortunately, both phospholipid vesicles have proven inadequate for high-resolution NMR studies of membrane-associated peptides and proteins due to low reorientation rates that lead to significant line broadening, giving rise to large bandwidths (Watts et al., 1991). The most common choice is the detergent micelles that are solubilized form of membrane proteins in mixed detergent-protein-lipid and have high reorienting rates. These are significantly smaller than vesicles and have been successfully used in obtaining NMR data for membrane-associated peptides (Bechinger et al, 1998). DPC micelles are typically used for bacteriocins and peptides (Fimland et al., 2008). However, the denaturing properties and their monolayer structure (with a high curvature, a loose packing of the lipid head groups and tighter chain entanglement in the center) make them unsatisfactory for studies of many membrane proteins and their interactions. In addition, the activity of proteins and enzymes that are embedded in micelles are not always preserved (Kleinschmidt et al., 1999) and the reconstitution of transmembrane proteins can be difficult. To adopt a conformation similar to that of the native state, membrane proteins may require a bilayer. This was conceived with bilayered model

membranes. Liposomes, which are spherical membrane bilayers, are other options. These bilayers range from about 100 nm to several micrometers (Torchilin and Weissig, 2003) and provide lipid environment. However, liposomes have some disadvantages such as stability issues, compartmentalization and orientation of membrane proteins, and difficulties in controlling the oligomeric state of the membrane protein (Bayburt and Sligar, 2010).

Bicelles are discoidal lipid aggregates composed of long-chain phospholipid and either detergent or short-chain phospholipid, used for biophysical and biochemical studies of membrane associated biomolecules (Whiles et al., 2002). Phospholipid bilayers are better mimic, since their biophysical properties resemble more closely those of the native membrane and they are thus more likely to maintain membrane proteins in a stable and active state (Raschle et al., 2009). Moreover, many membrane protein systems require specific types of phospholipids to maintain active function, a requirement that is not mimicked by detergent micelles and detergents (Sanders et al., 1995). Phospholipid bicelles have been used to study membrane proteins by high-resolution biophysical methods such as NMR spectroscopy (Prosser et al., 2006) and x-ray crystallography (Caffrey, 2003). Bicelles can be prepared easily in an aqueous environment and the hydration, pH, and ionic strength can be controlled allowing biologically relevant conditions (Sizun et al., 2004). Larger bicelles can be used for solid-state NMR, while isotropic bicelles are used for solution-state NMR (Czerski and Sanders, 2000). Bicelles have shown the possibility of performing both solid and solution-state NMR investigations in a single system. However, lipid bicelles has some limitations, lipid compositions that are currently in use for isotropic bicelles are limited (e.g. mixtures of dimeyristoyl phosphatidylcholine (DMPC) and either dihexanoyl-sn-glycero-3-phosphocholine (DHPC) or cholamidopropyl-dimethylammonio-hydroxy-propanesulfonate (CHAPSO) as detergents). Moreover, DMPC/DHPC bicelles can hinder the reconstitution of membrane proteins when the bilayer thickness does not match the hydrophobic domain of proteins (Sanders and Landis, 1995). Also, during sample preparations, the lipid to protein ratio has to be kept constant in order to avoid any phase transition; detergent diffusion into lipid may cause protein instability (Triba et al., 2005).

A promising new approach for studies of membrane proteins in a phospholipid bilayer uses high-density lipoprotein nanodiscs (Bayburt et al., 2002). Phospholipid bilayer nanodiscs has been recently used for structure determination of membrane proteins (Hagn et al., 2013). Nano discs are discoidal lipid bilayer stabilized by encircling amphipathic helical

scaffold proteins and are formed upon mixing of buffered saline solutions of phospholipids, purified membrane scaffold protein (MSP), and a detergent (Bayburt et al., 2002) or by sonication of MSP and phospholipid vesicles (Jonas, 1986) in the presence or absence of, for instance, proteins. MSPs are truncated forms of apolipoprotein (apo) A-I that wrap around a patch of a liquid bilayer to form a disc-like particle or nanodisc (Figure 9) (Bayburt and Sligar, 2010). MSPs are an amphipathic peptides that provide a hydrophobic surface inside facing the lipids and outer a hydrophilic surface. This is a reason that nanodiscs are highly soluble in aqueous solutions. MSPs that are assembled into nanodiscs can be kept in solution in the absence of detergents. Nanodiscs are about 7-13 nm in diameter depending on the mutation variant of MSP used (Hagn et al., 2013). Figure 10 shows the construction of truncated MSP variants that form smaller nanodisce suitable for NMR studies of small to medium-sized membrane proteins. NMR is technique proven for obtaining high-resolution structures of smaller biological macromolecules. In NMR, as the size of the molecules increases, the recorded spectra become harder to interpret due to a low tumbling rate, broadening of the spectral peaks. When membrane proteins are reconstituted into nanodiscs, the nanodisc contribute to the measured signal as well as slow down the rotation. Nanodiscs from phospholipids are made smaller in order to increase tumbling rate, to obtain a spectrum exclusively from the protonated membrane.

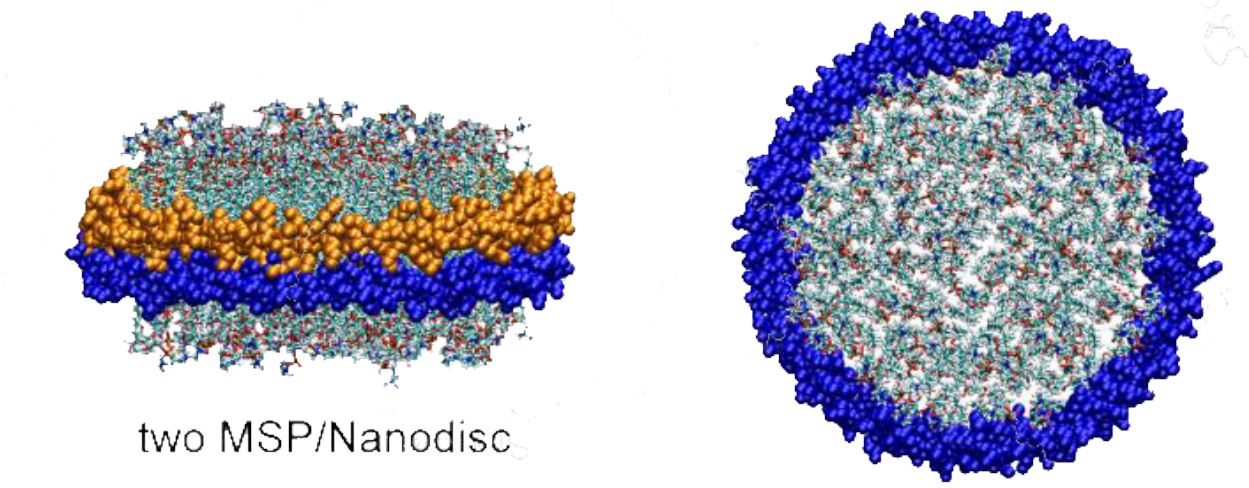


Figure 9. Nanodisc composed of MSP1D1 shown in side view and top view. The two MSPs are coloured orange and blue.

The most commonly used nanodisc has a diameter of around 10nm and a molecular weight of 150-200 kDa (Figure 10). Nanodisc has the potential to be extensively used membrane mimetic with the advantage of closely resembling a native-like lipid environment (Denisov et al., 2004). Furthermore, nanodiscs are the only available detergent-free

membrane mimetic for solution NMR spectroscopy that are suitable for studying protein-protein interactions in an unbiased lipid bilayer environment (Hagn et al., 2013). Our aim in the future is to use this system to correctly structure peptide complexes and determine membrane positioning of bacteriocins.

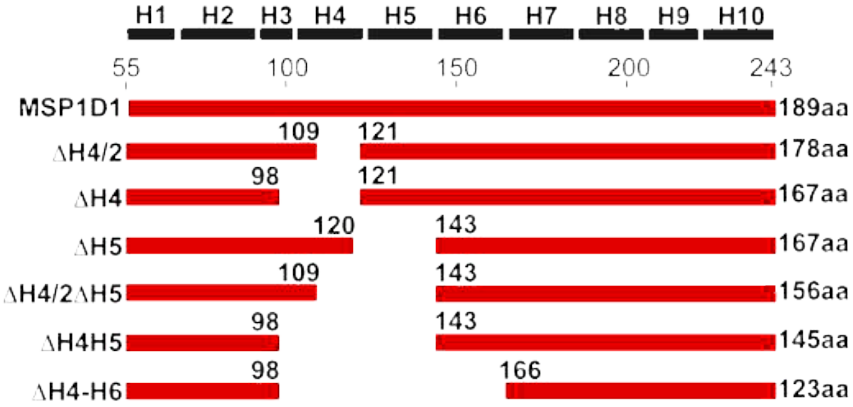


Figure 10. Construction of truncated membrane scaffold protein (MSP) variants. Deletion constructs of MSP1D1 with length. (Figure taken from Hagn et al., 2013).

2 Objective of the study

The objective of this study was to determine the optimal condition for expression and purification that will maximize the production of Enterocin EJ97. The aim was to produce the peptide Ent EJ97 in a labeled form for NMR structure investigation and to determine the structure in TFE and DPC using CD spectroscopy. Furthermore, to obtain scaffolding protein for NMR nanodisc structure determination of Ent EJ97.

The experimental work included:

I. Expression, production and purification of membrane scaffold protein
II. Optimization of expression condition for his-GB1-tev-Ent EJ97
III. Cleavage of his-GB1-tev-EntEJ97 (fusion peptide) and purification of EntEJ97

3 Materials and Methods

The chapter describes the methods applied in the study. A complete list of buffers, materials and computer software is given in the appendices.

3.1 Bacterial techniques

3.1.1 Expression vector

The enterocin EJ97 gene was previously cloned into the pUC57 vector (Appendix D) using EcoRV enzyme. pUC57 is a common used plasmid cloning vector in *E.coli*. The vector length is 2,710 bp and is isolated from *E.coli* strain DH5α by standard procedures. The vector was then inserted into the TG1 electrocompetent cells.

pCPR0012 plasmid was used as ligation independent cloning of EntEJ97 for sequencing purpose. EntEJ97 with his tag-GB1-tev gene was purchased from Genescript. The gene was initially cloned into the pET-22b(+) vector (Appendix D) and was inserted into the BL21(DE3) cells.

Similarly, MSP gene was precloned into the pET-28a(+) vector (Appendix D) using NcoI/BamHI enzymes prior to insertion into the BL21(DE3) cells.

3.1.2 Bacterial strains

The bacterial work was performed in strains of *Escherichia coli (E.coli)* cells, especially in TG1, BL21-Gold (DE3) and TOP10 cells. Each strain contains distinct characteristics making them versatile laboratory tools. TG1 Electrocompetent cells are suitable for M13 phase work, general cloning, blue/white screening and protein expression. The BL21(DE3) strain of *E.coli* is used for high-level protein expression (Jeong et al., 2015) and easy induction. This strain provides increased transformation efficiency and produce high-quality miniprep DNA. The strain is ideal for performing protein expression studies that utilize the T7 RNA polymerase promoter to direct high-level expression. Derived from *E.coli* B, the strain naturally lacks the Lon protease and the OmpT protease, which can degrade recombinant proteins during purification. E.coli and other bacterial strains used in this study are shown in Table 3.

Table 3.Bacterial strains used for expression and functional studies of MSP and Ent EJ97

	Bacterial strains	Reference/source
E.coli	K12 TB1	Jon-Nissen Mayer Lab
	TG1	Jon-Nissen Mayer Lab
	BL21-Gold (DE3)	Dirk Link' Group
	TOP10	STRATAGENE XL10-GOLD

3.1.3 Growth media

Lysogeny broth (LB) and LB agar plates supplemented with or without antibiotic were used as a general growth medium for all bacterial strains. LB and Minimal media were used for the expression of MSP and his-GB1-tev-EntEJ97. The composition of these media is shown in Appendix C.

3.1.4 Antibiotics

The antibiotics ampicillin (AppliChem) and kanamycin were used in this study. These were used as a method of selection in overnight cultures and LB agar plates. The stock concentration of ampicillin (Amp) and kanamycin (Kan) was 100 mg/ml and 50 mg/ml respectively. The working concentration was 100 µg/ml for ampicillin and 50 µg/ml for kanamycin.

3.1.5 Preparation of chemically competent *E.coli* BL21 (DE3) cells

Bacteria have the ability to take up and express exogenous DNA. When bacteria are made competent, the probability of a plasmid being transformed in a bacterial cell increases.

From the agar plate of BL21 (DE3), single colony was picked and 3 ml overnight culture was set up at 37^{o}C with shaking at 220 rpm. Next day, an overnight bacterial culture was diluted in 1:50 fresh LB and cultured to an OD_{600} of 0.3. Cells were kept on ice for 10 minutes and then collected by centrifugation at 5000 xg for 10 minutes at 4^{o}C. Then, the cells were suspended in 1/5th the original volume in ice-cold 0.1 M $CaCl_2$, followed by incubation on ice for 5 minutes, and centrifugation for 10 minutes at 4^{o}C at 5000 xg. Pelleted cells were

suspended in 1/25th the original volume in ice cold 0.1 M $CaCl_2$ containing 15 % (V/V) glycerol and incubated the cell suspension on ice for 40 minutes. 100 µl aliquots of competent cells were suspended into a pre-cooled 0.5 ml PCR tubes and frozen at -80^{o}C.

Storage and handling

Competent cells should be stored at -80°C. Storage at -20°C will result in a significant decrease in transformation efficiency. Cells lose efficiency whenever they are warmed above -80°C, even if they do not thaw.

3.1.6 Storage and growth conditions for bacterial cells

Bacterial cells can be stored for many years at -80°C. Adding some cells to LB medium supplemented with antibiotics starts a new cultivation of stock culture.

Stock culture

Bacterial cells were grown overnight in LB medium supplemented with the appropriate antibiotic at 37°C with shaking (220 rpm). 200 µl (70%) glycerol was added to 800 µl overnight culture. The solution was stored at -80°C, keeping the cells in a hibernation state and allowing them to be cultured again when needed.

Growth conditions for *E.coli* strains

E.coli strains are cultured in LB medium or LB agar plates containing the appropriate antibiotic depending in the selective marker of the plasmid. For the expression of the recombinant proteins, *E.coli* cells were transformed with the desired plasmid. The transformed cells were spread on LB agar plates with added antibiotic and cultured at 37°C for 16-20 hours.

3.1.7 Bacterial transformation by heat-shock

Bacterial transformation is a process in which foreign plasmid DNA is introduced into a bacterial host cell. The plasmid must contain a bacterial origin of replication, resulting in the replication of plasmid DNA in the host cell. Competent cells from -80^{o}C freezer are thawed on ice and the plasmid DNA is added. Low temperature causes the plasmid DNA to adhere to the bacterial cell membrane. Short duration of heat temperature introduces pores in

the bacterial cell membrane allowing the plasmid DNA enter into the host cell. The cells are spread on LB agar plates supplemented with antibiotics. The transformed plasmid DNA contains antibiotic resistance gene, ensuring that only cells containing this positive selection marker are able to grow. Single colony is selected for further use.

Transformation of plasmid DNA into host cell (competent cell)

50 µl competent cells were thawed on ice for 10 minutes and 1 µl plasmid DNA was added to the competent cell. The mixture was incubated on ice for 30 minutes. Then, cells were subjected to heat shock at 42^{o}C for 90 seconds on a water bath shaker (for BL21 (DE3) / His-GB1-tev-EntEJ97_pET-22b(+) heat shocked at 42^{o}C for 120 sec), followed by incubation on ice for 2 minutes. 500 µl LB media (pre heated at 37^{o}C) was added to the cells and were grown (300 rpm, 37 ^{o}C) on Thermomixer comfort eppendorf 1.5 ml for 45 minutes. From the mixture, 250 µl was spread on LB agar plates containing antibiotic, and incubated at 37^{o}C for 16-20 hours. Similar transformation procedure was used for all the competent cells *E.coli* TB1, TG1 and BL21 (DE3).

3.2 DNA Techniques

3.2.1 Plasmid DNA isolation

The isolation of the plasmid DNAs were achieved with the NucleoSpin® Plasmid (NoLid) and NucleoBond®Xtra Plasmid Purification kit from Macherey-Nagel for the mini prep and maxi prep method, respectively. The protocol was followed according to the manufacturers′ instructions. The protocol is based on alkaline lyses procedure, which is followed by binding of plasmid DNA to silica membrane of the NucleoSpin® Plasmid (NoLid) under high-salt conditions. The impurities are removed under high-salt or ethanolic conditions to keep the DNA bound to the membrane and the plasmid DNA is eluted in water. The plasmid DNA prepared with this kit is suitable for applications like automated fluorescent DNA sequencing, PCR, or any kind of enzymatic manipulation.

The plasmid DNA was isolated from 3 ml transformed bacterial culture (mini prep) and 100 ml (maxi prep). The growth conditions were described in section 3.1.6. The purity of the plasmid was checked by agarose gel electrophoresis, and NanoDrop spectrophotometer measured the concentration.

3.2.2 Primer design

The amplification of a segment of DNA requires primer pairs that would attach to the complementary template strand and serve as elongation starting points for the DNA polymerase. The primer sequences that were used in this study are given in Table 4.

Table 4. Primer sequences used for the amplification of EntEJ97

Primers	Sequences (5'→3')
pCPRej97 Forward	TAC-TTC-CAA-TCC-ATG-TTA-GCA-AAA-ATT-AAA-G
pCPRej97 Reverse	TAT-CCA-CCT-TTA-CTG-TTA-TGC-TAC-AGG-GCG-CTC

3.2.3 Polymerase Chain reaction (PCR)

The amplification of DNA fragment by PCR allows the use of a mixture containing a heat stable DNA polymerase, four deoxyribonucleoside tri phosphates (dATP, dCTP, dGTP, dTTP), a set of primers that are each complementary to the DNA fragment, which acts as the precursor for DNA synthesis and the DNA template. DNA polymerase needs a 3'-OH group to attach a new nucleotides to, a problem that is resolved by adding primers. Primers are short sequences of single-stranded DNA, also known as oligos, which determine the specific region of the template to be amplified.

The PCR reaction is generally composed of three steps namely denaturation, annealing and elongation or extension. The temperature and the incubation periods of the reaction are regulated according to the template DNA and the primers. The denaturation step will separate the double-stranded template, resulting in single-stranded DNA, which allows primers to bind to their recognition sequences. The primers anneal to their target sequences, and the DNA polymerase extends through the primers and attaches complementary dNTP to the template in a 5'→3' direction. Consecutive cycles of the reaction will result in an exponential amplification of the template sequences. Primer sequences are shown in 3.2.2.

The PCR reaction components were mixed in a 0.2 ml PCR tube and placed in a PTC-200 Peltier Thermal Cycler from MJ RESEARCH. When the PCR program was finished, the reaction product was purified using illustra^{TM} GFX^{TM} PCR DNA and Gel Band purification

kit from GE Healthcare, as described in the manufacturers' protocol.

Table 5. PCR reaction mixture setup

Components	Volume (μl)
MQ H_2O	38.4
Template DNA (1 ng/μl)	1
Pfu buffer (10X)	5
Forward Primer (100 ng/μl)	1.3
Reverse Primer (100 ng/μl)	1.3
dNTP (5 mM)	2
Pfu Turbo polymerase	1
Total	50

Table 6. PCR program for DNA amplification

Step	Temperature	Duration	
1	95°C	2 minutes	Denaturation (heat shock)
2	95°C	1 minute	Denaturation
3	48°C	1 minute	Annealing
4	68°C	1 minute	Elongation
5	Go to step 2	29 times	
6	68°C	7 minutes	Elongation
7	8°C	forever	Cooling

3.2.4 Agarose gel electrophoresis

The separation of macromolecules, proteins and/or DNA based on their size, electrical charge and their physical properties can be achieved by gel electrophoresis. The principle of this method lies in the different rates at which particles of different sizes and length migrate through a medium with pores (gel). Agarose gel electrophoresis separates DNA fragments according to size in a matrix of agarose polymers. Agarose is a polysaccharide typically used in a 0.5-2 % concentration in gels. Nucleic acids share a consistent negative charge imparted by their phosphate backbone, and will migrate towards a positively charged anode. Fragments of smaller size move faster and thus end up furthest in the gel whilst larger

fragment tend to take a longer time to migrate through gel. A molecular-weight size marker should always be included as a reference to help determine the size of the bands. Agarose gels are supplemented with ethidium bromide (EtBr), a carcinogenic intercalates in between the double strands of DNA. EtBr fluoresces under UV light and visualize the DNA molecules in the gel after migration. Tris-acetate-EDTA (TAE) buffer is used in the preparation of the gel and as a running buffer. TAE provides ions to carry the current.

1% agarose gels were used to verify the purity of the plasmids. The components used to make the gels are shown in Appendix C. Plasmid was loaded onto the gel with DNA loading buffer (Appendix C). GeneRuler 1 kb Plus and 0´GeneRuler 50 bp DNA ladder (Thermo scientific) were used in this study. Gels were run with 1X TAE buffer (Appendix C) with a constant voltage of 80 V for approximately 45 minutes.

Purification of DNA molecules from agarose gel

After separating DNA molecules using agarose gel electrophoresis, distinct bands can be isolated and purified for further application. The wanted DNA fragment is sliced from the gel using a scalpel blade, and subsequently dissolved in a binding buffer. The agarose should be melted at temperatures that do not denature the DNA. The solution is transferred to a column that binds DNA while other components such as salts; enzymes and agarose are washed away.

The illustra^{TM} GFX^{TM} PCR DNA and Gel Band purification kit from GE Healthcare was used to purify the restriction digested DNA molecules, according to the manufacturers' protocol.

3.2.5 Restriction digestion

The plasmid DNAs isolated with NucleoBond®Xtra plasmid purification kit was digested with restriction enzyme (BsaI) to generate fragments of varying sizes and lengths. The aim of the restriction digestion was to linearize the circular plasmid before running through the agarose gel. A reaction mixture as shown in Table 4 was made, and incubated for 1 hour at 37^{o}C in a water bath. The digested plasmids were run on 1 % agarose gel to verify the sizes of the generated fragments and low molecular ladder was used to estimate the DNA quantity.

Table 7. Restriction digestion mixture. Both enzymes and buffer came from New England Biolabs™

Components	Volume
Plasmid (pCPR0012)	1 μl (8.5 μg)
BsaI enzyme	2 μl
CutSmart®buffer	5 μl
MQ H_2O	42 μl
Total	50 μl

3.2.6 Ligation-Independent Cloning (LIC)

LIC is a form of molecular cloning developed in 1990 as an alternative to restriction enzyme/DNA ligase cloning. Inserts are usually PCR amplified and vectors are made linear either by restriction enzyme digestion or by PCR. The method uses the 3'→5' exonuclease activity of T4 DNA Polymerase (in absence of dNTPs) to generate overhangs that will anneal vector and fragment together. Incorporation of dGTP in the reaction limits the exonuclease processing to the first complementary C residue, and not present in the designed overlap, where the polymerization and exonuclease activities of T4 DNA Polymerase become "balanced". Joined fragments have 4 nicks that are repaired by *E.coli* during transformation. This technique allows efficient creation of scarless recombinant plasmids at many, but not all, positions in a vector.

Table 8. T4 DNA polymerase treatment of PCR products

Components	Volume (μl)
MQ H_2O	1.5
PCR-product	5
NEBuffer 2.1	1
dCTP (25 mM)	1
100 mM DTT	0.5
T4 DNA Polymerase 3U/μl	1
Total	10

Incubated at RT for 30 minutes and heat inactivated at 75°C for 20 minutes in PCR machine.

Table 9. T4 DNA polymerase treatment of Vector

Components	Volume (μl)
MQ H_2O	12.25
Plasmid (BsaI digested)	25
NEBuffer 2.1	5
dGTP (25 mM)	5
1 M DTT	0.25
T4 DNA Polymerase 3U/μl	2.5
Total	50
Incubated at RT for 30 minutes and then at 75^{o}C for 20 minutes in PCR machine.	

Ligation and transformation

The T4 DNA polymerase treated vector (1 μl) was dissolved in MQ H_2O (1μl) and centrifuged briefly to collect vector. 2 μl PCR products (T4 DNA Polymerase treated) was added to the vector and centrifuged briefly to ensure proper mixing. The mixture was incubated for 30 minutes at RT.

TOP10 competent cells were thawed on ice for 10 minutes and 10 μl of it was added to the reaction mixture. The mixture was incubated on ice for 30 minutes. Then, cells were subjected to heat shock at 42^{o}C for 90 seconds (Stirred waterbath WB-4MS Bioscan), followed by incubation on ice for 2 minutes. 100 μl LB media (pre heated at 37^{o}C) was added to the cells and were grown (220 rpm, 37 ^{o}C) for 1 hour. 100 μl of transformation was plated on a petri dish with LB agar (with 5 % sucrose and added 50 μg/ml Amp) and incubated overnight at 37^{o}C. Then, the isolation and purification of plasmid DNA was performed as described in 3.2.1.

3.2.7 Sequencing of DNA

The nucleotide sequence of a fragment of DNA can be determined using the DNA sequencing methods. Sequencing samples were sent to GATC Biotech in Germany. The list of sequencing primer is found in Appendix E.

Table 10. Sample requirements for shipment to GATC Biotech in Germany

Components	Volume (μl)
Template (80-100 ng/μl)	5
T7 Primer (5 μM)	5
Total	10

3.2.8 Measurement of DNA concentration

The concentration and the purity of DNA were measured using the NanoDrop 2000 UV-Vis Spectrophotometer from Thermo Scientific. The instrument measures DNA at 260 nm as well as sample purity at 260/280 nm and 260/230 nm. The nitrogenous bases in nucleotides have an absorption maximum of UV light at 260nm. Proteins have an absorption maximum of 280nm. The ratio of absorbency at 260/280 nm and 260/230 nm is a measurement of sample purity. These ratios should be between 1.65-1.85 and 1.0-2.2, respectively. An aberrant value indicates contaminants and impurities in the sample.

The background of the sample was measured by applying 2 μl of the buffer or MQ H_2O the sample was diluted. Then, 1 μl sample was applied to the machine and measured the DNA concentration.

3.3 Protein Expression and purification

3.3.1 Expression condition and optimization

A variety of expression conditions were optimized to result in a greater yield of protein MSP and, though the rate of shaking (250 rpm), and the choice and the concentration of antibiotic (50 μg/ml Kan) were kept constant. In all cases, 1 ml overnight culture was added to 100 ml of fresh LB media with appropriate antibiotic to inoculate expression cultures at 1% v/v. When the OD_{600} reached 0.7, cell culture was spilt into two equal halves and one of them was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) for a final concentration of 1 mM. In both proteins (MSP and his-GB1-tev-EntEJ97) 200 μl each sample (induced and uninduced) was taken out after 2 hours and 4 hours of incubation. Samples were immediately stored into -20^{o}C freezer.

Similarly various growth conditions for the increase in yield of protein EntEJ97 was also optimized with rate of shaking 250 rpm and the choice of antibiotic as Amp (100 μg/ml). Overnight culture was diluted with 20-fold modified minimal medium and 1:100 with LB medium; and grown to late exponential phase at 37°C until OD_{600} reached 0.7. Cell culture was split into two equal halves and IPTG was added (final conc.1 mM - 2 mM) to induce one of the cultures.

SDS-PAGE was used to compare the amount of MSP and EntEj97 under different expression conditions. SDS-PAGE reaction mixture for MSP is given in Table 11. To compare the amount of protein EntEJ97, all the stored samples were centrifuged at max. speed (Microfuge® 18 centrifuge) for 1 minute and cell pellet was dissolved in 50 μl 1x PBS buffer. Then, ≈0.5 μl DTT (final conc. 1mM) was added to the pellet solution and incubated for 1 hour at RT.

Table 11. SDS-PAGE reaction mixture setup for MSP protein

Time	OD_{600}	Dilution	Vol. for 20 μl sample (V)	Loading Dye (μl)
T_0	x	1	x*1*20	V/3
$T_{1, induced}$	x	5	x*5*20	V/3
$T_{1, uninduced}$	x	5	x*5*20	V/3

3.3.2 Large-scale protein expression

A large-scale expression of MSP was performed in 6 L of LB divided into 6 baffled flasks (2.5 L) under optimized expression condition. A single colony of freshly transformed *E.coli* BL21(DE3) was inoculated into 150 ml LB + antibiotic Kan (50 μg/ml) and incubated overnight at 37°C (220 rpm). Next morning, the overnight culture was transferred to 1 L LB medium with antibiotic in a baffled flask and incubated until OD_{600} reached 0.7. Cells were induced with IPTG (final conc. 1 mM) for 4 hours. In all cases, cells were harvested by centrifuging (Beckman coulter) at 5500 rpm for 15 minutes at 4°C using JA10 rotor. Cell pellets were transferred into a zip-locked plastic bag, weighed and stored at -80°C until further purification.

Expression of his-GB1-tev-EntEJ97 was performed in 1L of LB and 1L of modified minimal media (M9) in a baffled flask under optimized expression condition. From freshly transformed *E.coli* BL21 (DE3), a single colony was inoculated into 70 ml LB with Amp and incubated overnight at 37^{o}C (220 rpm). Next day, 10 ml overnight culture was transferred to 1000 ml LB and 50 ml overnight culture was diluted 20-fold with minimal modified media. Both cultures were incubated until OD_{600} reached 0.7. Cells were induced with IPTG (final conc. 2 mM) for 6 hours.

3.3.3 Cell lysis and extraction of crude protein

Crude MSP and EntEJ97 from pilot scale expressions were extracted by lysing the cells. Two methods were used X-press and sonication. For extraction of MSP, X-Press method was used while sonication was used for EntEJ97.

Cell lysis by X-Press method

The frozen bacterial cells were crushed under high pressure and mechanical stress by X-press equipment so that the MSP protein can be solubilized in buffer.

The frozen bacterial cells were transferred from the zip-locked plastic bag to the mounted X-press equipment, which has been precooled to -20^{o}C. Resulting paste was dissolved carefully in 100 ml 20 mM Sodium phosphate buffer pH 7.4 and 3 tablets of EDTA-free protease inhibitor cocktail was added. The cell residues were removed from the sample by centrifuging at 20,000 rpm, 4 ^{o}C for 45 minutes using JA 25.50 rotor. Supernatant was collected and stored at 4^{o}C until further use.

Cell lysis by sonication

Approximately 5.1 gm of frozen cell pellet (from -80^{o}C) obtained from 1 liter expression (BL21 (DE3)/ his-GB1-tev-EntEJ97)) was thawed on ice, and resuspended in 25 ml phosphate buffer (20 mM, pH 7.4) and 1 ml EDTA-free protease inhibitor cocktail was added (2 tablets dissolved in 3 ml MQ H_2O), and cooled on ice for 10 minutes. Then, the cell suspension was sonicated for 4 minutes (34 % amplitude, in rounds of 20 seconds /40

seconds rest for each cycle). The cell debris was removed by centrifuging (Beckman coulter) at 4^{o}C for 45 minutes at 20,000 rpm using JA 25.50 rotor.

3.3.4 Protein purification

DNA precipitation

The solution obtained from the lysed cells was treated with streptomycin. Streptomycin sulfate precipitates DNA in the solution. This process is important because solutions containing DNA tend to be viscous and may clog chromatography columns that are used to purify proteins. Addition of streptomycin sulfate to a final concentration of 2.5 % (w/v) effectively precipitates DNA, leaving protein in the solution.

Supernatant obtained from the lysed cells was transferred in a small beaker (assembled a two-beaker cooling system) on a magnetic stirrer. A volume of 10 % streptomycin corresponding to ¼ of the volume of the supernatant was added stirring carefully (e.g. 25 ml streptomycin sulfate solution if the supernatant volume is 100 ml). After centrifuging (Beckman Coulter^{TM} J-25) for 15 minutes at 15,000 rpm at 4^{o}C using JA 25.50 rotor, the supernatant was collected and measured discarding the pellet.

Preparation of protein sample for chromatography

Proper amount/conc. of NaCl and imidazole was added to the solution so that the concentration is approx. the same as in the binding buffer (otherwise, protein will probably go right through the column). The protein sample was then filtered through 0.22 μm filter (Sarstedt) prior to loading into the superloop.

Calculations for adding NaCl and imidazole:

NaCl,

$C_1V_1 = C_2V_2$

[Assume, $V_2 = V_1$ + supernatant]

$C_1V_1 = C_2V_1 + C_2$ supernatant

$V_1 = \frac{C_2 \text{ supernatant}}{C_1 - C_2}$

Then, total volume of supernatant (Ts) = V_1 + supernatant

Now, for Imidazole,

$C_3V_3 = C_4V_4$

[As, $V_4 = V_3 + Ts$]

$C_3V_3 = C_4 V_3 + C4\ Ts$

$V3 = \frac{C_4\ Ts}{C_3 - C_4}$

Where, C_1 = conc. of stock NaCl

C_2 = required conc. of NaCl

V_1= vol. of NaCl to be added

V_2 = final volume after addition of NaCl to the supernatant

C_3 = conc. of stock imidazole

C_4 = required conc. of imidazole

V_3 = vol. of imidazole to be added

V_4 = vol. after addition of imidazole and NaCl to the supernatant

Nickel affinity chromatography

As our protein consists of His_6-tag, Nickel affinity chromatography was the choice of purification. The HisTrap™ 5 ml column (GE Healthcare) was used. The column has Ni-NTA matrix charged with Ni2+, which binds to a His_6-tag of protein in the buffer. Äkta purifier system was used for the separation of proteins. The detail on concentration of buffers is given in Appendix C.

At the beginning, system wash was performed with MQ H_2O. Then, the column was washed with 5 column volume (CV) of MQ H_2O to remove preservatives. During the process, checked for air bubbles and avoided introducing air by sucking the air through the luer adapter and equilibrated with CVs of binding buffer until baseline was seen.

The test was performed with 1 ml of the protein sample using 1 ml loop. The column was first washed with MQ H_2O and then with wash buffer. The sample was then injected with 1 ml syringe (Hamilton) and a gradient elution was done at a flow rate of 1.00 ml/min. Fractions containing protein were pooled as determined by the absorbance at 280 nm. The program was run according to the parameters/variables described in Appendix F. Pressing "END" stopped the program. The fractions eluted were confirmed with SDS-PAGE.

After confirmation, loading the protein sample into a superloop and running the program according to the parameters protein was purified. 20 µl was collected and stored at -20°C for SDS-PAGE and the rest of the protein was stored at 4°C for desalting. The column was washed with 5 CV MQ H_2O then with 4 CV of 20 % ethanol, took off, screwed together and kept in shelf at 4°C. Äkta (orange) tubes were also screwed to the system and washed first with MQ H_2O then with 20 % ethanol and logged off.

Desalting of the purified fusion protein

The purified protein contains high concentration of imidazole and NaCl. Desalting procedure removes high concentration of imidazole and salt by exchanging buffer with low concentration of imidazole and salt, which is compatible with the buffer for TEV protease cleavage. Buffer exchange was done in Äkta purifier system with HiTrap™ 5 ml Desalting column from GE Healthcare.

The buffers used for desalting are given in Appendix C. The column was prepared and equilibrated with the same method as described earlier. Running the program according to the parameters/variables described in Appendix F performed the exchanging of buffers. After elution, fractions containing his tagged protein were pooled, and 20 µl aliquot was stored for SDS-PAGE, and rest was stored for next step.

Concentrating purified protein

The Amicon ® Ultra-15 centrifugal devices 3K NMWL were used for concentrating the collected purified protein. Before concentrating the sample, the devices were rinsed with Milli-Q® water.

The purified protein was loaded to the centrifugal devices and centrifuged at 5,000 xg for 60 minutes. Then the concentrated solute was pipetted out of the filter device and stored in a centrifuge tube. Absorbance of the concentrated purified protein was taken with Nanodrop spectrophotometer instrument at 280 nm and the concentration was calculated by using known protein molecular weight and extension coefficient.

Cleavage test of the pure purified his tagged protein with TEV Protease

The tobacco etch virus (TEV) protease is a protease used for the removal of fusion tags from recombinant proteins. TEV protease has a strict 7 amino acid cleavage recognition sequence of Glu-Asn-Leu-Tyr-Phe-Gln↓Gly. The cleavage of the pure his tagged protein was done according to the manufacturers` protocol. TEV protease was added to the target protein in ratio of 1:100 and 1:50 (w/w) at room temperature.

3.3.5 SDS polyacrylamide gel electrophoresis

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is a method for separating proteins according to their molecular weight. SDS is an anionic denaturing detergent. SDS-treated proteins have the same primary structure and negative charge; and are consequently migrated towards anode. Polyacrylamide is a cross-linked polymer of acrylamide and the length of the polymer chains is dictated by the concentration of acrylamide used. A higher percentage of acrylamide increases the difficulty for proteins to migrate and the proteins move slower than in a gel with a lower concentration. The stacking gel is used to align the proteins in the sample, while the separating gel is used to separate

proteins according to their molecular weight. Therefore, proteins form distinct bands of a certain size.

This study used pre-casted NuPAGE 4-12 % Bis-Tris Gels with 10 wells. Protein samples were mixed with 4X Loading Dye sample buffer at 85^{o}C for 10 minutes, and loaded into a gel. A maximum of 35 μl sample was loaded in the wells. A prestained protein ladder (Spectra ^{TM} Multicolor Low Range Protein Ladder) was also loaded to estimate the size of the protein. The polyacrylamide gel was then run in 1X MOPS SDS running buffer with a constant voltage of 200V for 40 minutes. All the chemical components for SDS-PAGE were purchased from ThermoFisher Scientific.

3.3.6 Staining Protein Bands

The separated proteins were analyzed further by staining the gel with Coomassie Brilliant Blue staining solution R250 in order to visualize the protein bands. The gels were stained for 20 seconds in a microwave. Then gels were destained with destaining solution shaking (Appendix C) until clear protein bands were visualized.

4 Results

4.1 Choice of strains and media

Strains of *E.coli* cells especially TG1, BL21(DE3) and TOP10 cells were used for the bacterial transformation. The enterocin EJ97, MSP and his-GB1-tev-EntEJ97 genes were previously cloned into pUC57, pET-28a(+) and pET22b(+) vectors respectively. The pUC57-EntEJ97 vector was then inserted into TG1 electrocompetent cells. Similarly, the MSP-pET-28a(+) vector was inserted into TB1 and BL21(DE3) competent cells. Likewise, the his-GB1-tev-EntEJ97-pET-22b(+) vector was inserted into the BL21(DE3) competent cells. Growing the cells in presence of specific antibiotics checked transformation. Strains TG1 and BL21(DE3) cells showed the successful transformation. Thus, TG1 was used for transformation of pUC57-EntEJ97. Competent cell BL21(DE3) was used for the transformation and expression of MSP-pET-28a(+) and his-GB1-tev-EntEJ97-pET22b(+) plasmid. The strain TG1 was grown in LB media and strain BL21(DE3) was grown in both LB and minimal media for selected cultivation with appropriate antibiotics.

LB medium is a staple in virtually every lab and is commonly used for maintenance and propagation of *E.coli*. Besides, it is nutrient-rich microbial broth contains peptides, amino acids and carbohydrates in a low-salt formulation. The *E.coli* cells grow faster in LB because the tryptone/peptone and yeast supply essential growth factors that *E.coli* would otherwise have to synthesize. Minimal medium comprise of minimal salts base formulation that can be supplemented with essential salts, vitamins, nitrogen and carbon sources suitable for recombinant *E.coli* strains and labelling the amino acids for ease in structure determination by NMR spectroscopy. LB media was used to inoculate all the strains for overnight culture. However, for the expression of his-GB1-tev-EntEJ97 (his-tagged protein) the media was switched to minimal medium because there was no expression or very little expression of his-tagged protein in LB media (Figure 11).

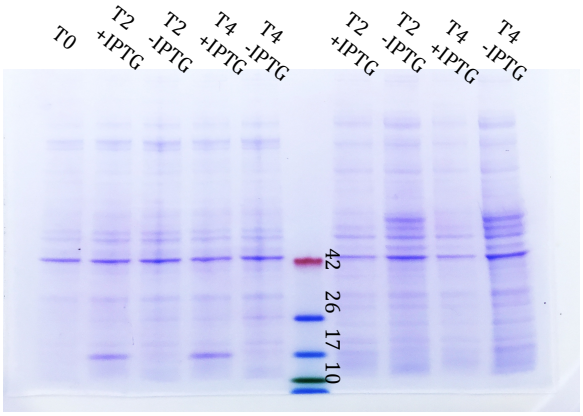


Figure 11. SDS-PAGE gel showing the protein expression in LB and minimal medium.

The cells were induced with 2 mM IPTG (final conc.) at 30°C. T0, T2 and T4 indicates time interval 0, 2 and 4 hours whereas – and + signs indicate without IPTG and with IPTG respectively. The lanes from the left of the marker indicates expression in minimal medium and the right side of marker indicates expression on LB medium.

4.2 Plasmid purification

4.2.1 Plasmid pUC57-EntEJ97

The plasmid DNAs (pUC57-EntEJ97) isolated from the overnight culture was confirmed with agarose gel electrophoresis (Figure 12A) and the concentration was measured to be 162.1 ng/µl and 195.0 ng/µl for plasmid DNA 1 and plasmid DNA 2 respectively.

Figure 12A shows the band of pUC57-EntEJ97 corresponding to the approximate size of pUC57-EntEJ97 indicating that EntEJ97 had been inserted correctly in the pUC57 vector, resulting in a size of 2.84 kb. After confirmation, EntEJ97 DNA fragment was PCR amplified and again it was confirmed with agarose gel electrophoresis (Figure 12B). Figure 12B shows the band of PCR amplified DNA EntEJ97 corresponding to the approximate size of EntEJ97 (135 bp) indicating that EntEJ97 has been correctly amplified with the help of primers. The amplified DNA fragment was ligated with pCPR0012 plasmid to produce new recombinant DNA.

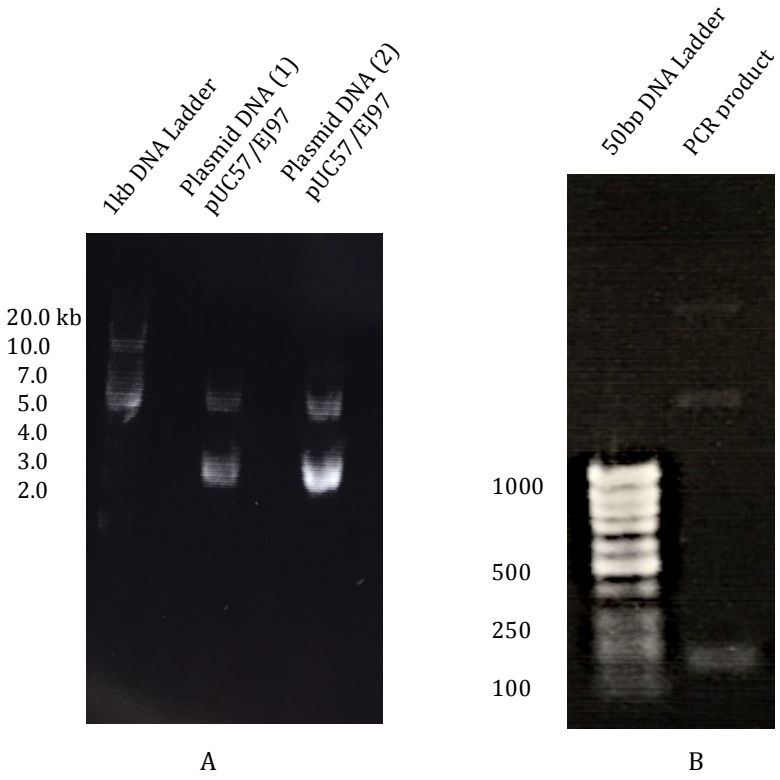


Figure 12. Agarose gels confirming the isolation of plasmid DNA pUC57/EJ97 and PCR product. An agarose gel showing the ladder, plasmid DNA pUC57/EJ97 (A) and PCR product of pUC57/EJ97 (B). In both conditions 1µl of loading dye was mixed with 5 µl of sample and loaded into the well. 1 kb and 50 bp DNA ladder was used in (A) and (B) respectively.

Purification of plasmid pCPR0012 from the overnight culture through NucleoBond®Xtra plasmid purification kit produced 8563.3 ng/µl. The purity was further verified by agarose gel electrophoresis by cutting with the BsaI enzyme and running on 1 % agarose gel (Figure 13). BsaI cutter generated bands of 2763 and 4694 bp (uncut vector). From the agarose gel, the band of around 3 kb was cut with scalpel and further purified.

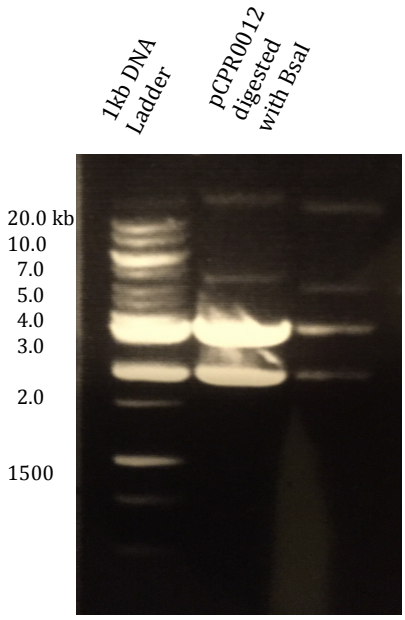


Figure 13. An agarose gel showing the digested vector. 10 µl loading dye was added to 50 µl total reaction mixture and loaded into the well. The lane containing ladder and digested vector has been indicated in the picture.

The digested purified plasmid (pCPR0012) and purified PCR plasmid (EntEJ97) were ligated together and transformed into supercompetent cell (XL10) as described in 3.2.6. The agarose gel electrophoresis confirmed the isolation of plasmid and the sequencing result from GATC proved the purity and correct sequence of plasmid pCPR0012-EntEJ97. Figure 14 shows the band of pCPR0012-EntEJ97 corresponding to the size of pCPR0012 ligated together with EntEJ97 was approximately 2.89 kb.

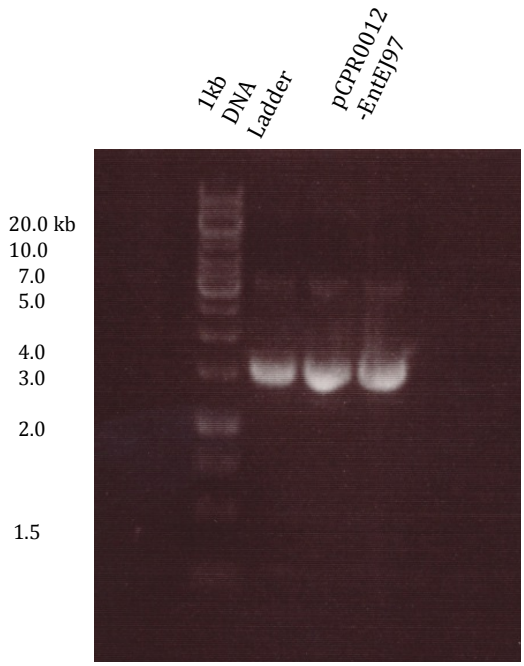


Figure 14. Agarose gel showing the purity of pCPR0012-EntEJ97 plasmids. 1 μl loading dye was mixed with 5 μl of sample and loaded into the well. The lane containing ladder and plasmids has been indicated in the picture.

The cloning in pCPR0012 produced product with N-terminal His-tag and TEV cleavage site followed by gene of interest (EntEJ97). The detail sequence information can be found in Appendix B.

atgtctggtctcatcatcatcatcattcttctggtgtagatctgggtaccgagaacctgtacttccaatcc

M S G S H H H H H H S S G V D K G T E N L Y F Q S

Gene of Interest: EntEJ97

Atgttagcaaaaattaaagcgatgattaagaagtttccgaacccttatactttagcagctaa
gctaacgacttacgaaattaattggtataaacaacaatacggtcgttatccttgggagcgcc
ctgtagcataa

4.2.2 Plasmid MSP-pET-28a(+)

The plasmid DNAs MSP-pET-28a(+) isolated from the overnight culture was confirmed by agarose gel electrophoresis (Figure 15). The size of MSP sequence was 581 bp and the vector length was 5.4 kb. Thus, the size of the pET-28a(+) encoding MSP was approximately 5.9 kb which can be identified in the Figure 15. This confirms that MSP fragment was correctly inserted into the pET-28a(+) vector. Sequencing result from GATC proved that the correct sequence of the plasmid MSP was inserted into a pET-28a(+) vector.

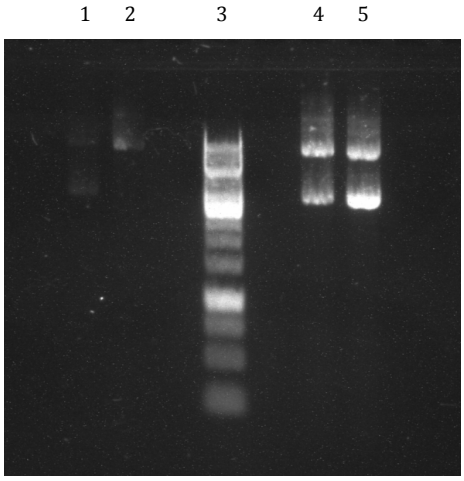


Figure 15. An agarose gel confirming the isolation of plasmid DNAs. Lane 1 and 2 indicates the isolation of plasmid MSP-pET28a(+) from TB1 bacterial cells; lane 3 indicates 1 kb ladder, lane 4 and 5 indicates isolated plasmid from BL21(DE3) bacterial cells.

4.2.3 Plasmid his-GB1-tev-EntEJ97-pET-22b(+)

Purification of Plasmid DNAs (his-GB1-tev-EntEJ97) isolated from the overnight culture was confirmed with agarose gel electrophoresis (Figure 16) and the concentration was measured as 131.0 ng/µl. The size of his-GB1-tev-EntEJ97 was 351 bp and the size of pET-22b(+) vector was 5493 bp. Figure 16 shows the band of pET-22b(+)-his-GB1-tev-EntEJ97 corresponding to the approximate size of pET-22b(+)-his-GB1-tev-EntEJ97 indicating that his-GB1-tev-EntEJ97 had been inserted correctly in the pET-22b(+) vector, resulting in a size of approximately 5.75 kb. The fragment purity was further confirmed by sequencing results from GATC. Prior to sending for sequencing, the fragment was first cut from the agarose gel band of around 3 kb and purified via Illustra^{TM} Gel Band Purification kit. Total 10 µl sample

was sent to GATC for sequencing. The result further indicates that the isolated plasmid was pure enough to perform expression experiments.

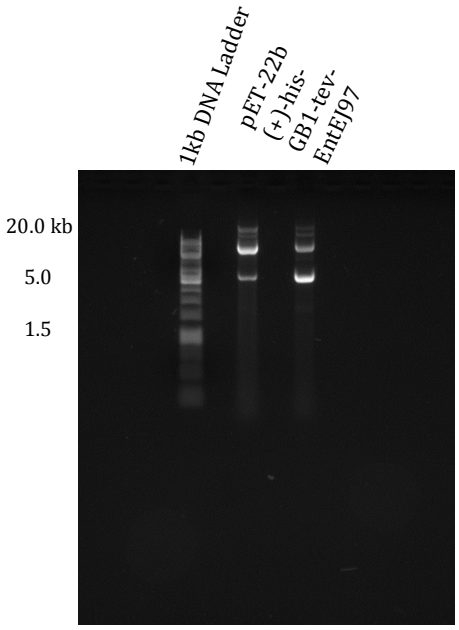


Figure 16. Agarose gel showing the purity of pET-22b(+)-his-GB1-tev-EntEj97 plasmid.
1 µl loading dye was mixed with 5 µl of sample and loaded into the well. The lane 1, 2 and 3 indicates ladder and plasmids respectively.

4.3 Expression optimization

The effect of different expression conditions on MSP-pET-28a(+) and his-GB1-tev-EntEJ97 was evaluated by SDS-PAGE assay. We examined the level of MSP-pET-28a(+) and his-GB1-tev-EntEJ97 expression under different conditions and obtained the results described below.

4.3.1 Effects of temperature

To optimize the temperature for expression, an overnight culture of freshly transformed *E.coli* BL21(DE3)-MSP was diluted 1:100 and grown in 100 ml LB medium until OD_{600} =0.7 was achieved. The culture was divided into 50 ml cultures, and marked as induced or uninduced. One of the cultures was induced with IPTG (1.0 mM final conc.) and incubated at four different temperatures (20^{o}C and 25^{o}C overnight, 30^{o}C and 35^{o}C during day for 2 hours), and then harvested. The result of temperature optimization for expression of pET-28a(+)-MSP is shown in Figure 17.

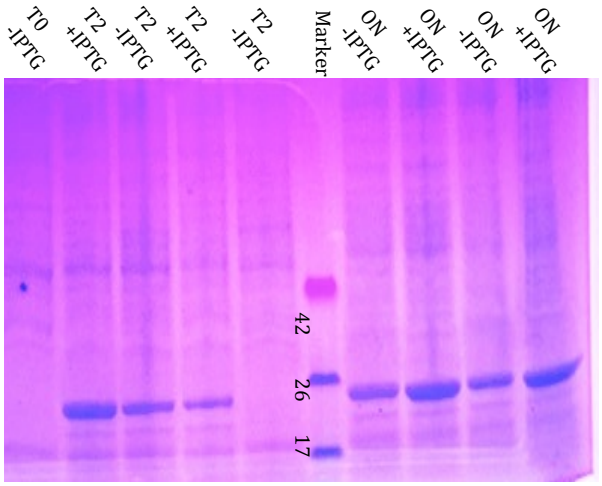
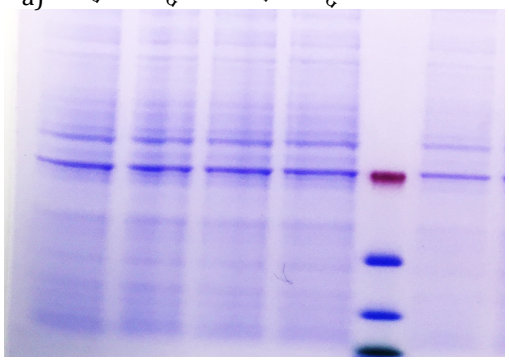


Figure 17. The SDS-PAGE gel showing the effect of temperature on the expression of MSP.

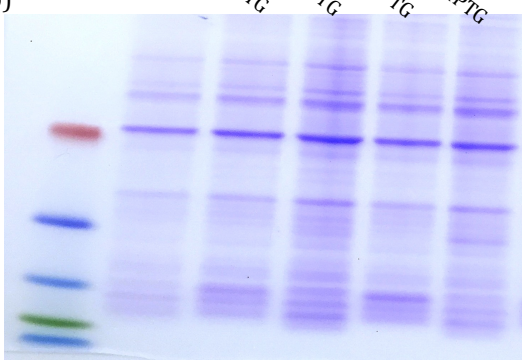
T0 and T2 indicates time interval: before induction and after 2 hours of induction. Signs " - " and " + " indicate without and with IPTG respectively. Lane 1, 2 and 3 from left shows the expression at 30°C; lane 4 and 5 shows the expression at 35°C; lane 6 indicates marker; lane 7 and 8 shows the expression at 25°C overnight; and lane 9 and 10 shows the expression at 20°C overnight. The bands around 25 kDa indicates the expression of MSP.

The temperature was optimized for expression of *E.coli* BL21(DE3)-his-GB1-TEV-EntEJ97 with the similar procedure as described above at 5 different temperatures (20°C, 25°C, 30°C, 35°C and 37°C, (Figure 18) (refers to next page). LB media was used to inoculate overnight culture and switched to minimal media for expression. Cells were induced with 1 or 2 mM IPTG.

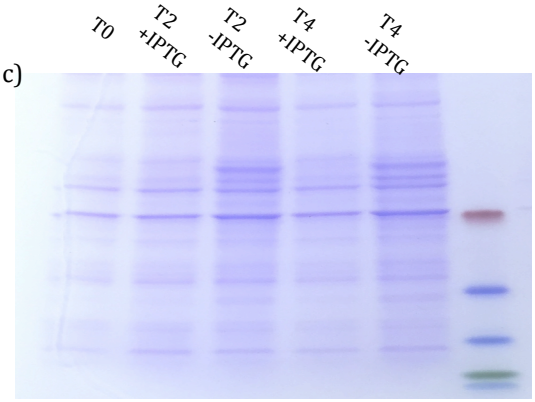
a)


T4 +IPTG
T4 -IPTG
T2 +IPTG
T2 -IPTG
T0

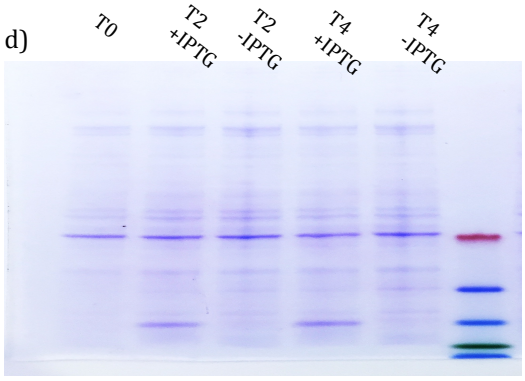
b)


T0
T2 +IPTG
T2 -IPTG
T4 +IPTG
T4 -IPTG

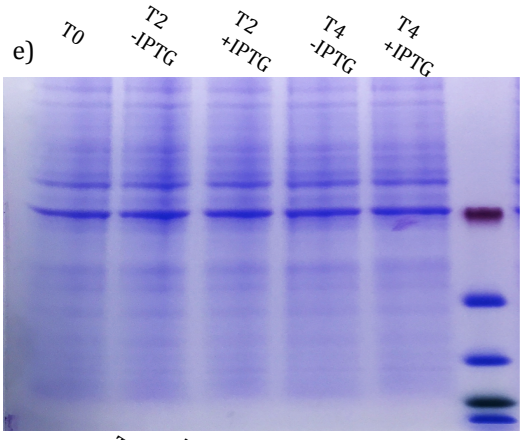
c)


T0
T2 +IPTG
T2 -IPTG
T4 +IPTG
T4 -IPTG

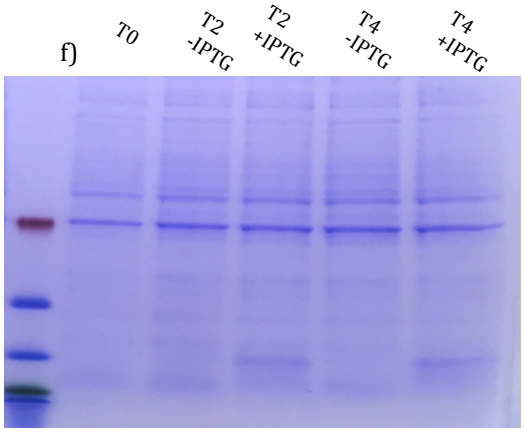
d)


T0
T2 +IPTG
T2 -IPTG
T4 +IPTG
T4 -IPTG

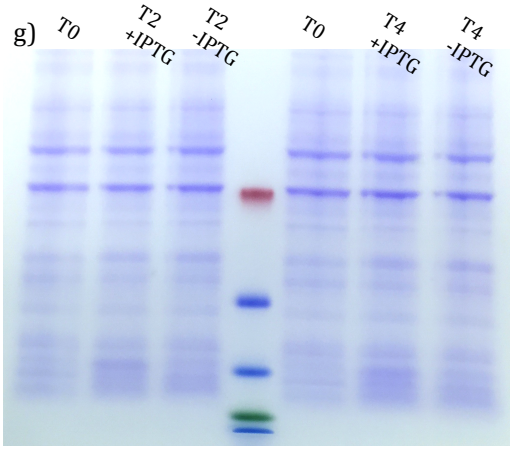
e)


T0
T2 -IPTG
T2 +IPTG
T4 -IPTG
T4 +IPTG

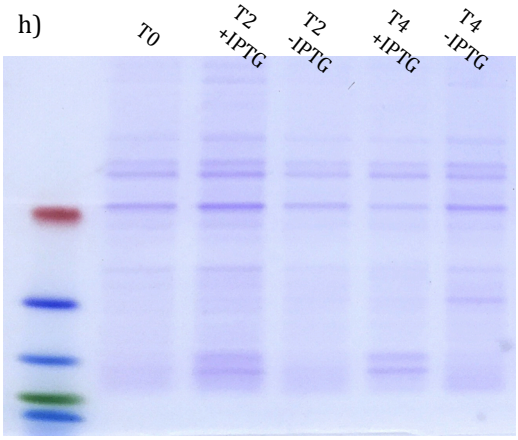
f)


T0
T2 -IPTG
T2 +IPTG
T4 -IPTG
T4 +IPTG

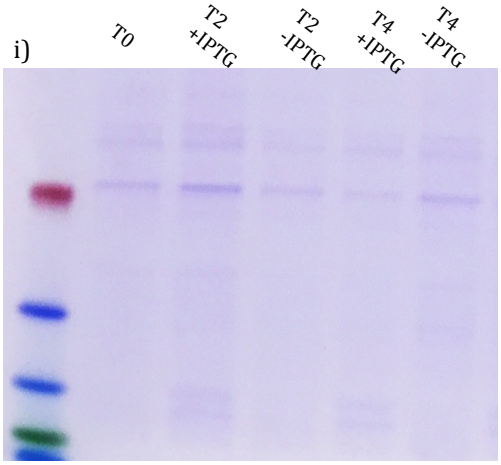
g)


T0
T2 +IPTG
T2 -IPTG
T0
T4 +IPTG
T4 -IPTG

h)


T0
T2 +IPTG
T2 -IPTG
T4 +IPTG
T4 -IPTG

i)


T0
T2 +IPTG
T2 -IPTG
T4 +IPTG
T4 -IPTG

j)

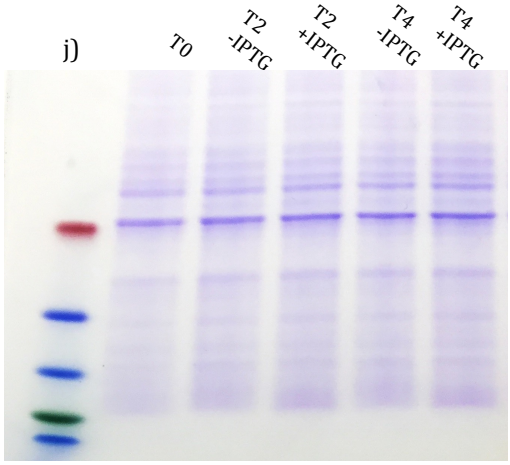

T0
T2 -IPTG
T2 +IPTG
T4 -IPTG
T4 +IPTG

Figure 18. SDS-PAGE gels showing the effect of temperature on the level of fusion protein expression at different tempearatures and IPTG concentration. The bands around 15 kDa indicate the expression of his-GB1-tev-EntEJ97. Signs " - " and " + " indicate without and with IPTG respectively. a) 37°C, induced with 1 mM IPTG; b) 37°C, induced with 2 mM IPTG; c) 35°C, induced with 1 mM IPTG; d) 35°C, induced with 2 mM IPTG; e) 30°C, induced with 1 mM IPTG; f) 30°C, induced with 2 mM IPTG; g) 25°C, induced with 1 mM IPTG; h) 25°C, induced with 2 mM IPTG; i) 20°C, induced with 1 mM IPTG; and j) 20°C, induced with 2 mM IPTG.

The effect of temperature on level of MSP and his-GB1-tev-EntEJ97 expression at different temperatures is shown in Figure 17 and Figure 18 respectively. As shown by the intensities and broad bands, the expression level of MSP and his-GB1-tev-EntEJ97 was higher at 30°C and 35°C respectively, indicating that the proteins are expressed larger at these temperatures than expressed at any other temperatures.

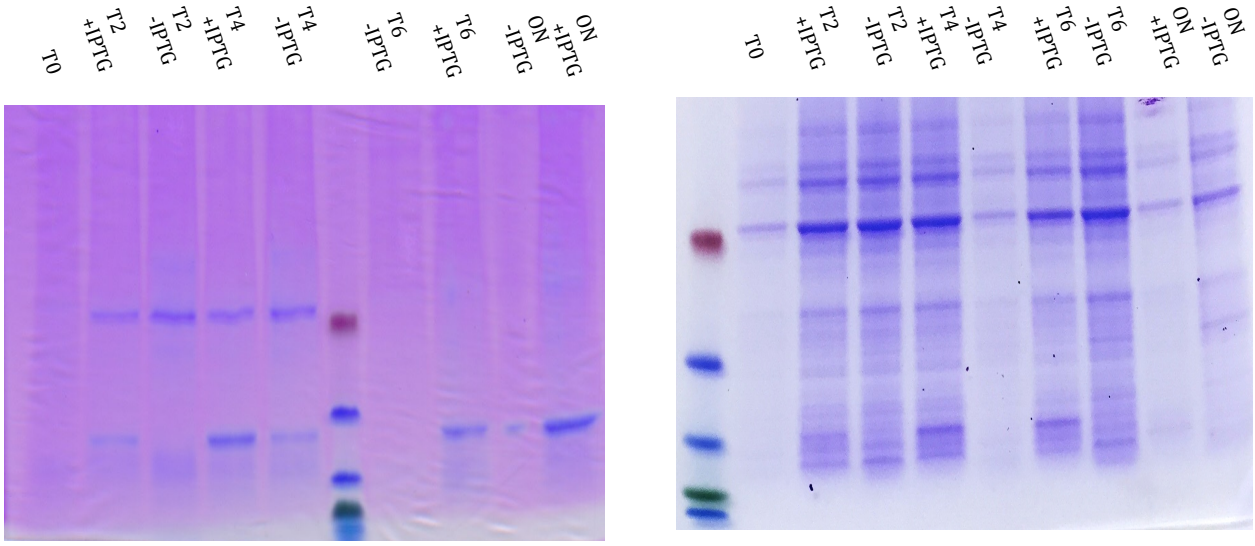
4.3.2 Effects of inducer (IPTG) concentration

To evaluate the effects of different concentrations of IPTG on the expression of fusion protein, cultures of *E.coli* BL21(DE3)-his-GB1-tev-EntEJ97 at OD_{600}= 0.7 were induced with 1-2 mM concentration of IPTG and allowed to express for 2-4 hours to check the protein yield (Figure 18).

The effect of IPTG concentration on fusion protein expression was compared by the intensity and band assay by loading equal amounts of crude extract. Both the concentrations were able to produce protein. The intensity and the band corresponding to the higher concentration was more intense and broad, indicating that higher concentrations gave the best protein yield. However, at lower concentration, the expression was reduced (Figure 18). For further optimization step, 2 mM IPTG was chosen.

4.3.3 Effects of post-induction incubation

To evaluate the maximum yield of MSP and his-GB1-tev-EntEJ97, the incubation time after induction was studied. The incubation time after induction was increased to 6 hours and overnight.



a) b)

Figure 19. The gel bands showing the effect of post-induction on the expression of MSP and his-GB1-tev-EntEJ97. a) The bands around 26 kDa indicates the expression of MSP induced with 1mM IPTG at 30°C. b) The bands around 15 kDa indicate the expression of his-GB1-tev-EntEJ97 induced with 2 mM IPTG at 35°C. Signs " - " and " + " indicate without and with IPTG respectively.

According to the gel bands in Figure 19a, the level of MSP expression started within 2 hours when incubated at 30°C, and reached its maximum level within 4 hours and remained unchanged to 6 hours. The gel bands from Figure 19b shows that his-GB1-tev-EntEJ97 has started expression within 2 hours of incubation at 35°C and increased its level within six hours. However, the expression level of his-GB1-tev-EntEJ97 reduced later and after overnight incubation. Hence, four hours and six hours was found to be the optimum post-induction time for MSP and his-GB1-tev-EntEJ97, and the rest of the experiments were carried out at this post-induction time.

4.4 Purification

4.4.1 MSP Purification

MSP expressed in large-scale culture under optimized condition was purified as described in Materials and Methods (Section 3.3.4). Affinity chromatography was used to purify protein, since MSP contained N-terminal His_6 tag, only MSP was expected to elute after increasing the imidazole concentration. A representative chromatogram for the purification step is shown in Figure 20.

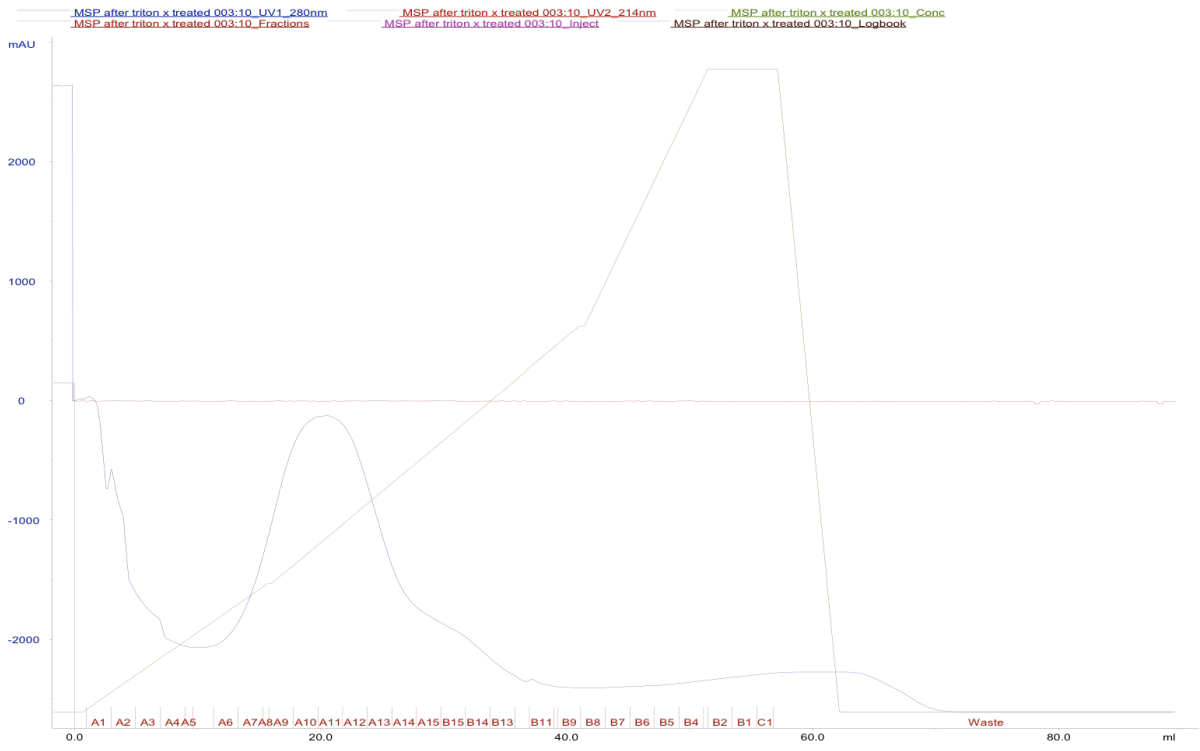


Figure 20. Chromatogram of nickel affinity purification of MSP. The protein was detected by measuring UV absorbance at 280 nm.

The molecular mass of MSP is 22.23 kDa as determined from sequence. The SDS-PAGE gel (Figure 21) shows the successful purification after affinity chromatography. The clear bands between 17 and 26 kDa shows that MSP was successfully purified. All elution fractions containing protein samples from affinity chromatography were pooled and stored at 4^{o}C in the freezer.

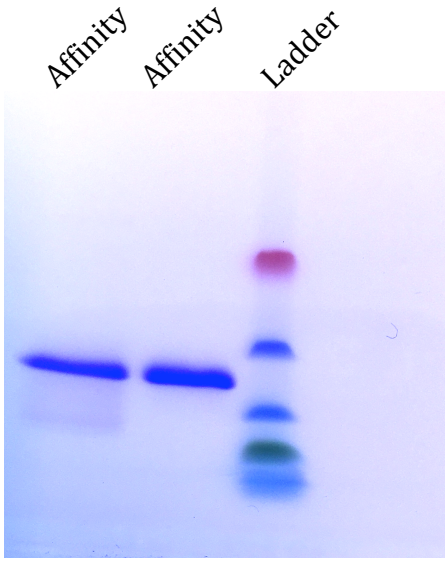


Figure 21. SDS-PAGE gel showing the purity of protein samples purified from affinity chromatography.

4.4.2 his-GB1-tev-EntEJ97 purification

Fusion peptide (his-GB1-tev-EntEJ97) expressed in large-scale culture under optimized condition was purified as described in Materials and Methods (Section 3.3.4). A two-step purification system was used, namely Affinity chromatography using HisTrap^{TM} HP column followed by buffer exchange using HiTrap^{TM} desalting column. In both steps, the protein elution was monitored by UV absorbance at 280 nm. Firstly, affinity chromatography was used to purify fusion protein. Since the target protein (his-GB1-tev-EntEJ 97) contained N-terminal His_6 tag, only fusion protein was expected to elute in nickel affinity chromatography. A representative chromatogram for both the purification step is shown in the Figure 22 and 24.

Figure 22. Chromatogram of nickel affinity purification of fusion protein. The protein was detected by measuring UV absorbance at 280 nm.

The molecular mass of fusion protein is 13.18 kDa as determined by sequence. All elution fractions containing protein samples from affinity chromatography were pooled together. SDS-PAGE confirmed the purification of fusion protein. The SDS-PAGE gel (Figure 23) shows the successful purification after affinity chromatography.

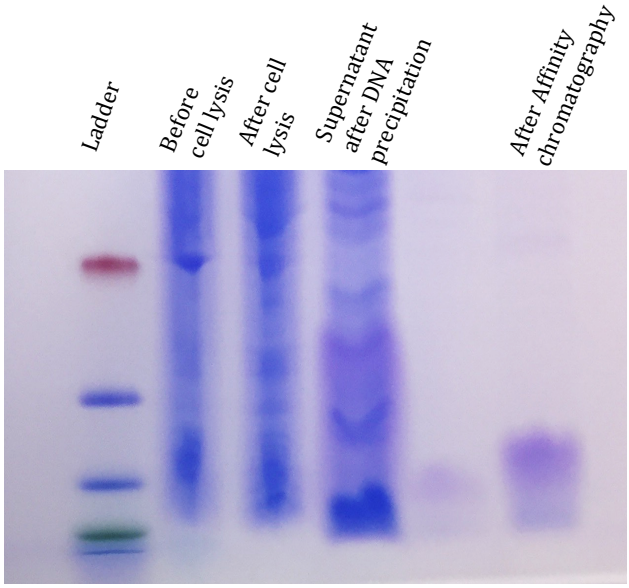


Figure 23. SDS-PAGE gel showing the purity of protein sample after affinity chromatography.

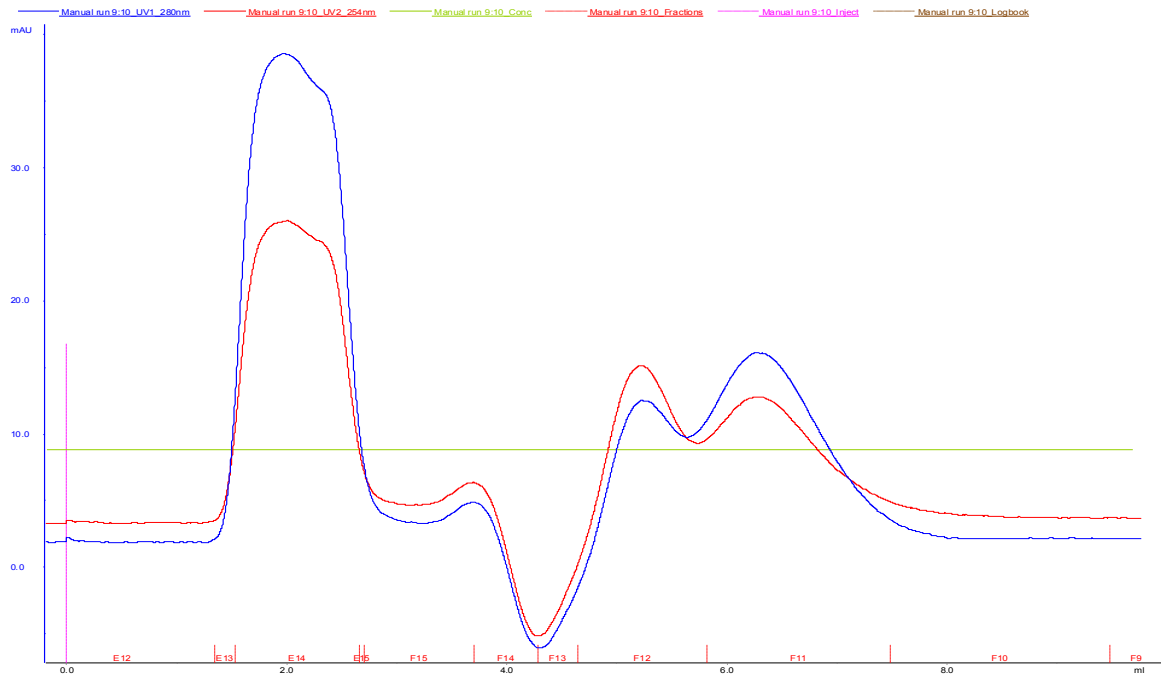


Figure 24. Chromatogram of HiTrap™ desalting of fusion protein. The protein was detected by measuring UV absorbance at 280 nm.

Secondly, the fusion protein containing high concentration of imidazole was removed using HiTrap™ desalting column. The fractions containing his-tagged bounded proteins were pooled together and concentrated with Amicon ultra centrifugal devices. A protein concentration of 1.94 mg/ml was obtained (measured by Nanodrop). The concentrated sample was used for cleavage with TEV protease.

4.4.3 Cleavage of fusion peptide

The test was performed to cleave the fusion peptide. The tev protease was added to the fusion peptide in a ratio of 1:100 and 1:50 according to manufacturers´protocol. The cleavage of his-tagged peptide could not be confirmed with SDS-PAGE because of low band intensity and was tried to purify by affinity chromatography using HisTrap ™ HP column in Äkta purifier system, the result was unsuccessful.

Several attempts were performed to obtain higher yield of fusion protein. The proteins that were pooled after affinity chromatography was collected together, buffer exchanged and then concentrated. During this process, we suspect that mislabelling of samples may have happened and impure protein was collected; samples were lost in buffer exchange process decreasing the amount and concentration. The volume and concentration of fusion protein for

cleavage may have wrongly calculated or insufficient amount was used. The trial was performed at room temperature for few hours only.

Further, research and experimental trial has to be performed in this area. Due to lack of time and insufficient samples we were unable to perform other trials.

5 Discussions

In this study, we have successfully produced and purified the protein MSP. As presented in the aim of the study in section 2.1, the starting point was to express and purify membrane scaffold protein. The next step was to optimize the expression condition for enterocin EJ97 and lastly to determine structure of EntEJ97 in TFE and DPC using CD spectroscopy. The optimization stage was successful and expression was complete. However, the structure determination of EntEJ97 could not be performed due to unsuccessful purification of fusion protein. The findings and the methods applied in our research are discussed in the following chapter.

5.1 Optimization of strain and media

The choice of strain in the expression of protein is very important. It is always recommended to use specific strains for successful expression. For the expression of protein using pET- vector, the choice of expression host is *E-Coli* BL21(DE3). The strain contains T7 RNA polymerase gene controlled by the lacUV5 promoter in its chromosal DNA (Studier and Moffatt., 1986). T7 RNA polymerase is expressed upon addition of IPTG that induces a high-level protein expression. Since, our gene was precloned into pET-vectors, the expression of protein was successful in BL21(DE3) cells.

The effect of media on MSP and his-GB1-tev-EntEJ97 expression was also examined. LB media are the common and popular choices for recombinant protein expression. The production of MSP in LB medium was well expressed (Figure 21). LB medium was also chosen as an initial expression medium for his-GB1-tev-EntEJ97. However, several initial attempts failed to express enough protein for the expression level of his-GB1-tev-EntEJ97 (Figure 11) to show up in the SDS-PAGE analysis. When the overnight cultures were switched to modified minimal media, the expressions succeed. It has been reported previously that the fusion peptide containing GB1 was successfully expressed in the same modified minimal medium induced with 2 mM IPTG (Koenig et al., 2003). We suspect that bacterial cells use the nutrients contained in LB medium only for their survival. However, the exact reason for not producing protein is unknown. The presence of trace elements, glucose, nitrogen source and vitamins in minimal media can support bacterial cells to express proteins, leading to increase protein yield.

5.2 Protein expression

5.2.1 Optimization of Temperature

The effect of different temperatures evaluated for MSP shows that, at 30^{o}C the production was highest among chosen temperatures. However, the expression for MSP on overnight at lower temperatures can be achievable. The expression of his-GB1-tev-EntEJ97 was favourable at 35^{o}C. The expression could be performed at 37^{o}C. But it is know that higher temperatures often result in protein aggregation and degradation. Therefore, higher temperatures were not tested. Moreover, the prolonged post induction growth results in significant decrease of protein yield and are susceptible to proteolysis (Ritchie et al., 2009).

5.2.2 Optimization of IPTG concentration

The inducer concentration was 1 mM IPTG for expression of MSP. This worked perfectly fine so we did not vary the concentration. The result of inducer concentration is in agreement with earlier studies, where the bacterial culture was induced with 1mM IPTG to produce higher level of MSP (Ritchie et al., 2009).

The protein (his-GB1-tev-EntEJ97) expression was optimized by varying the concentration of inducer. The amount of IPTG was varied by 2-fold range to determine the appropriate amount of IPTG needed for maximum expression of fusion protein. The clear optimum expression was found when induced with 2 mM IPTG (Figure 18d).

5.2.3 Optimization of induction parameters

The induction for the expression of protein was induced at OD_{600} =0.7, since the bacterial cultures are generally grown to the mid exponential phase before induction. It has also been reported that the bacterial culture should not be over grown in rich medium, and that the OD_{600} should be in the middle of its exponential phase to ensure higher protein expression (Sivashanmugam et al., 2009).

A time course of induction over 4 hour time period was taken to determine the change in production yield. The post induction result for MSP shows that expression yield of MSP after 4 hours increases. Figure 19a shows that expression of MSP is similar between 4 and 6

hours. However, the production yield increases drastically after overnight induction. The fusion GB1-S proteins that were expressed in modified minimal medium at OD_{600} =0.7, for 3 hours resulted in about tenfold higher yield of the S-peptide where S-peptide was expressed as a part of fusion with a 6 kDa carrier protein in *E.coli* (Koenig et al., 2003). However, the length of the S-peptide was 22 residues.

5.3 Protein purification

The purification of protein MSP was successful as shown in Figure 21 with the optimized temperature, inducer concentration, strain and media. From 6 L of large-scale expression, higher yield of MSP-protein was obtained (25 mg per liter of LB medium). The amount of protein obtained seems to be pure enough to perform further steps. As this was one of our objectives in this study, we can say that we were successful to obtain scaffolding protein for NMR nanodiscs structure determination. However, the his-tag attached to this protein needs to be cut off before producing nanodiscs.

Fusion protein (his-GB1-tev-EntEJ97) was successfully purified with affinity chromatography (Figure 23). However, several attempts were performed to obtain sample enough for cleavage of the protein. The cleavage trials were unsuccessful and we were unable to obtain pure EntEJ97 protein.

6 Conclusion

Conclusions and prespectives during the course of this thesis, we successfully optimized various parameters that influence the protein expression. Two different proteins have been expressed and purified by the optimized parameters in the course of this thesis. Plasmids encoding MSP and his-GB1-tev-EntEJ97 constructs were successfully amplified and purified. The purified plasmids were used to transform BL21(DE3) cells for protein expression. Successful expression of both MSP and his-GB1-tev-EntEJ97 were observed (after IPTG induction) by SDS gel electrophoresis. Expression of the MSP protein on large scale yielded about 25 mg/ L culture. Purification of MSP with affinity chromatography resulted pure protein.

The expression of his-GB1-tev-EntEJ97 was challenging. Though a variety of conditions were optimized, it was found that the expression of fusion EntEJ97 protein largely depended on inducer (IPTG) concentration. Changing the expression media and amount of IPTG resulted successful expression of fusion protein. However, several attempts were made to increase the amount of protein yield to be observed by protein gel electrophoresis. Temperature, inducer concentration and expression media were optimized to ensure the expression of fusion protein. The purification of fusion protein was successful.

The next step was to remove the EntEJ97 from the fusion protein. The cleavage test was performed with TEV protease. The initial trial failed and again next trial was performed, but the result was negative. Due to lack of sufficient protein sample and time duration, the cleavage test could not be completed.

Further experimental work has to be done for the cleavage of the fusion protein. The resulting pure peptide can be used for structure determination and characterization.

References

Acedo, J.Z., Towle, K.M., Lohans, C.T., Miskolzie, M., McKay, R.T., Doerksen, T.A., & Martin-Visscher, L.A. (2017). Identification and three-dimensional structure of carnobacteriocin XY, a class IIb bacteriocin produced by Carnobacteria. FEBS letters *591*(10), 1349-1359.

Bayburt, T.H., Grinkova, Y.V., and Sligar, S.G. (2002). Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins. Nano Letters *2*, 853-856.

Bayburt, T.H., and Sligar, S.G. (2010). Membrane protein assembly into Nanodiscs. FEBS letters *584*, 1721-1727.

Bechinger, B., Kinder, R., Helmle, M., Vogt, T. C., Harzer, U., & Schinzel, S. (1999). Peptide structural analysis by solid-state NMR spectroscopy. Peptide Science, *51*(3), 174-190.

Bignell, C., and Thomas, C.M. (2001). The bacterial ParA-ParB partitioning proteins. Journal of biotechnology *91*, 1-34.

Boman, H.G. (1995). Peptide antibiotics and their role in innate immunity. Annual review of immunology *13*, 61-92.

Caffrey, M. (2003). Membrane protein crystallization. Journal of structural biology *142*, 108-132.

Cammers-Goodwin, A., Allen, T. J., Oslick, S. L., McClure, K. F., Lee, J. H., & Kemp, D. S. (1996). Mechanism of stabilization of helical conformations of polypeptides by water containing trifluoroethanol. Journal of the American Chemical Society *118*(13), 3082-3090.

Cebrian, R., Martinez-Bueno, M., Valdivia, E., Albert, A., Maqueda, M., and Sanchez-Barrena, M.J. (2015). The bacteriocin AS-48 requires dimer dissociation followed by hydrophobic interactions with the membrane for antibacterial activity. Journal of structural biology *190*, 162-172.

Cintas, L. M., Casaus, P., Håvarstein, L. S., Hernandez, P. E., & Nes, I. F. (1997). Biochemical and genetic characterization of enterocin P, a novel sec-dependent bacteriocin from Enterococcus faecium P13 with a broad antimicrobial spectrum. Applied and Environmental Microbiology, *63*(11), 4321-4330.

Cintas, L.M., Casaus, P., Herranz, C., Håvarstein, L.s., Holo, H., Hernandez, P.E., and Nes, I.F. (2000). Biochemical and genetic evidence that Enterococcus faecium L50 produces enterocins L50A and L50B. the sec-dependent enterocin P, and a novel bacteriocin secreted without an N-teminal extension termed enterocin Q J Bacteriol *182*, 6806-6814.

Corr, S.C., Li, Y., Riedel, C.U., O'Toole, P.W., Hill, C., and Gahan, C.G. (2007). Bacteriocin production as a mechanism for the antiinfective activity of Lactobacillus salivarius UCC118. Proceedings of the National Academy of Sciences *104*, 7617-7621.

Cotter, P.D., Hill, C., and Ross, R.P. (2005). Bacteriocins: developing innate immunity for food. Nature Reviews Microbiology *3*, 777-788.

Cotter, P.D., Ross, R.P., and Hill, C. (2013). Bacteriocins-a viable alternative to antibiotics? Nature Reviews Microbiology *11*, 95-105.

Czerski, L., & Sanders, C. R. (2000). Functionality of a membrane protein in bicelles. Analytical biochemistry, *284*(2), 327-333.

De Boever, E.H., Clewell, D.B., and Fraser, C.M. (2000). Enterococcus faecalis conjugative plasmid pAM373: complete nucleotide sequence and genetic analyses of sex pheromone response. Molecular microbiology *37*, 1327-1341.

Denisov, I.G., Grinkova, Y.V., Lazarides, A.A., and Sligar, S.G. (2004). Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size. J Am Chem Soc *126*, 3477-3487.

Diep, D.B., and Nes, I.F. (2002). Ribosomally synthesized antibacterial peptides in Gram positive bacteria. Current drug targets *3*, 107-122.

Fimland, G., Johnsen, L., Dalhus, B., and Nissen-Meyer, J. (2005). Pediocin-like antimicrobial peptides (class IIa bacteriocins) and their immunity proteins: biosynthesis, structure, and mode of action. Journal of Peptide Science *11*, 688-696.

Fimland, N., Rogne, P., Fimland, G., Nissen-Meyer, J., and Kristiansen, P.E. (2008). Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin plantaricin EF. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics *1784*, 1711-1719.

Fisher, K., and Phillips, C. (2009). The ecology, epidemiology and virulence of Enterococcus. Microbiology *155*, 1749-1757.

Fregeau Gallagher, N.L., Sailer, M., Niemczura, W.P., Nakashima, T.T., Stiles, M.E., and Vederas, J.C. (1997). Three-dimensional structure of leucocin A in trifluoroethanol and dodecylphosphocholine micelles: spatial location of residues critical for biological activity in type IIa bacteriocins from lactic acid bacteria. Biochemistry *36*, 15062-15072.

Gajic, O., Buist, G., Kojic, M., Topisirovic, L., Kuipers, O.P., and Kok, J. (2003). Novel mechanism of baceriocin secretion and immunity carried out by lactococcal multidrug resistance proteins. Journal of Biological chemistry *278*, 34291-34298.

Galvez, A., Valdivia, E., Abriouel, H., Camafeita, E., Mendez, E., Martinez-Bueno, M., and Maqueda, M. (1998). Isolation and charaterization of enterocin EJ97, a bacteriocin produced by Enterococcus faecalis EJ97. Arch Microbiol *171*, 59-65.

Giraffa, G. (2002). Enterococci from foods. FEMS Microbiol Rev *26*, 163-171.

Gordon, Y.J., Romanowski, E.G., and McDermott, A.M. (2005). A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. Current eye research *30*, 505-515.

Gratia, A. (1925). Sur un remarquable example d´antagonisme entre deux souches de colibacille. Compt Rend Soc Biol *93*, 1040-1042.

Hagn, F., Etzkorn, M., Raschle, T., and Wagner, G. (2013). Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. J Am Chem Soc *135*, 1919-1925.

Hancock, R.E.W., and Diamond, G. (2000). The role of cationic antimicrobial peptides in innate host defences. Trends Microbiology *8*, 402-410.

Hoffmann, J.A., Kafatos, F.C., Janeway, C.A.J.a., and Ezekowitz, R.A.N.B. (1999). Phylogenetic perspectives in innate immunity. Science *284*, 1313-1318.

Iwatani, S., Horikiri, Y., Zendo, T., Nakayama, J., and Sonomoto, K. (2013). Bifunctional Gene Cluster lnqBCDEF Mediates Bacteriocin Production and Immunity with Differential Genetic Requirements. Applied and environmental microbiology *79*, 2446-2449.

Jeong, H., Kim, H. J., & Lee, S. J. (2015). Complete genome sequence of Escherichia coli strain BL21. Genome announcements *3*(2), e00134-15.

Kleinschmidt, J. H., Wiener, M. C., & Tamm, L. K. (1999). Outer membrane protein A of E. coli folds into detergent micelles, but not in the presence of monomeric detergent. Protein Science *8*(10), 2065-2071.

Koenig, B. W., Rogowski, M., & Louis, J. M. (2003). A rapid method to attain isotope labeled small soluble peptides for NMR studies. Journal of biomolecular NMR, *26*(3), 193-202.

Lau, S. Y., Taneja, A. K., & Hodges, R. S. (1984). Effects of high-performance liquid chromatographic solvents and hydrophobic matrices on the secondary and quarternary structure of a model protein: Reversed-phase and size-exclusion high-performance liquid chromatography. Journal of Chromatography A, *317*, 129-140.

Jonas, A. (1986). Reconstitution of high-density lipoproteins. Methods in enzymology *128*, 553-582.

Lohans, C.T., Towle, K.M., Miskolzie, M., McKay, R.T., Van Belkum, M.J., McMullen, L.M., and Vederas, J. (2013). Solution structures of the linear leaderless bacteriocins enterocin 7A and 7B resemble carnocyclin A, a antimicrobial peptide. Biochemistry *52*, 3987-3994.

Lüders, T., Birkemo, G.A., Fimland, G., Nissen-Meyer, J., and Nes, I.F. (2003). Strong synergy between a eukaryotic antimicrobial peptide and bacteriocins from lactic acid bacteria. Applied and environmental microbiology *69*, 1797-1799.

Marcotte, I., & Auger, M. (2005). Bicelles as model membranes for solid-and solution-state NMR studies of membrane peptides and proteins. Concepts in Magnetic Resonance Part A *24*(1), 17-37.

Masuda, Y., Zendo, T., and Sonomoto, K. (2012). New type non-lantibiotic bacteriocins: circular and leaderless bacteriocins. Beneficial Microbes *3*, 3-12.

Nakayama, J., Takanami, Y., Horii, T., Sakuda, S., and Suzuki, A. (1998). Molecular mechanism of peptide-specific pheromone signaling in Enterococcus faecalis: functions of pheromone receptor TraA and pheromone-binding protein TraC encoded by plasmid pPD1. J Bacteriol *180*, 449-456.

Neira, J.L., Contreras, L.M., de los Panos, O.R., Sanchez-Hidalgo, M., Martinez-Bueno, M., Maqueda, M., and Rico, M. (2010). Structural characterisation of the natively unfolded enterocin EJ97. Protein engineering, design & selection : PEDS *23*, 507-518.

Nes, I.F. (2011). History, current knowledge, and future directions on bacteriocin research in lactic acid bacteria. In Prokaryotic Antimicrobial Peptides, D. Drider, and S. Rebuffat, eds. (Springer New York Dordrecht Heidelberg London: Springer Science+ Business Media).

Nes, I.F., Diep, D.B., and Ike, Y. (2014). Enterococcal Bacteriocins and Antimicrobial Proteins that Contribute to Niche Control. In Enterococci: From Commensals to Leading Causes of Drug Resistant Infection, M.S. Gilmore, D.B. Clewell, Y. Ike, and N. Shankar, eds. (Boston: Massachusetts Eye and Ear Infirmary).

Netz, D.-J.A., Pohl, R., Beck-Sickinger, A.G., Selmer, T., Pierik, A.J., Bastos, M.C.F., and Sahl, H.-G. (2002). Biochemical characterization and genetic analysis of Aureocin A53, a new, atypical bacteriocin from *Staphylococcus aureus*. Journal of molecular biology *319*, 745-756.

Nieh, M. P., Harroun, T. A., Raghunathan, V. A., Glinka, C. J., & Katsaras, J. (2004). Spontaneously formed monodisperse biomimetic unilamellar vesicles: the effect of charge, dilution, and time. *Biophysical journal*, *86*(4), 2615-2629.

Nissen-Meyer, J., Holo, H., Håvarstein, L., Sletten, K., and Nes, I. (1992). A novel lactococcal bacteriocin whose activity depends on the complementary action of two peptides. Journal of bacteriology *174*, 5686-5692.

Nissen-Meyer, J., and Nes, I.F. (1997). Ribosomally synthesized antimicrobial peptides: their function, structure, biogenesis, and mechanism of action. Archives of microbiology *167*, 67-77.

Nissen-Meyer, J., Oppegard, C., Rogne, P., Haugen, H.S., and Kristiansen, P.E. (2010). Structure and Mode-of-Action of the Two-Peptide (Class-IIb) Bacteriocins. Probiotics and antimicrobial proteins *2*, 52-60.

Nissen-Meyer, J., Rogne, P., Oppegard, C., Haugen, H., and Kristiansen, P. (2009). Structure-function relationships of the non-lanthionine-containing peptide (class II) bacteriocins produced by gram-positive bacteria. Current pharmaceutical biotechnology *10*, 19-37.

Oppegård, C., Rogne, P., Emanuelsen, L., Kristiansen, P. E., Fimland, G., & Nissen-Meyer, J. (2007). The two-peptide class II bacteriocins: structure, production, and mode of action. Journal of molecular microbiology and biotechnology, *13*(4), 210-219.

Ovchinnikov, K.V., Kristiansen, P.E., Straume, D., Jensen, M.S., Aleksandrzak-Piekarczyk, T., Nes, I.F., and Diep, D.B. (2017). The Leaderless Bacteriocin Enterocin K1 Is Highly Potent against Enterococcus faecium: A Study on Structure, Target Spectrum and Receptor. Frontiers in Microbiology *8*.

Ovchinnikov, K.V., Kristiansen, P.E., Uzelac, G., Topisirovic, L., Kojic, M., Nissen-Meyer, J., Nes, I.F., and Diep, D.B. (2014). Defining the structure and receptor binding domain of the leaderless bacteriocin LsbB. Journal of Biological Chemistry *289*, 23838-23845.

Pag, U., and Sahl, H.-G. (2002). Multiple activities in lantibiotics-models for the design of novel antibiotics? Current pharmaceutical design *8*, 815-833.

Prosser, R.S., Evanics, F., Kitevski, J.L., and Al-Abdul-Wahid, M.S. (2006). Current applications of bicelles in NMR studies of membrane-associated amphiphiles and proteins. Biochemistry *45*, 8453-8465.

Raschle, T., Hiller, S., Yu, T.Y., Rice, A.J., Walz, T., and Wagner, G. (2009). Structural and functional characterization of the integral membrane protein VDAC-1 in lipid bilayer nanodiscs. J Am Chem Soc *131*, 17777-17779.

Ritchie, T. K., Grinkova, Y. V., Bayburt, T. H., Denisov, I. G., Zolnerciks, J. K., Atkins, W. M., & Sligar, S. G. (2009). Chapter eleven-reconstitution of membrane proteins in phospholipid bilayer nanodiscs. Methods in enzymology, *464*, 211-231.

Rogne, P., Fimland, G., Nissen-Meyer, J., and Kristiansen, P.E. (2008). Three-dimensional structure of the two peptides that constitute the two-peptide bacteriocin lactococcin G. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics *1784*, 543-554.

Rogne, P., Haugen, C., Fimland, G., Nissen-Meyer, J., and Kristiansen, P.E. (2009). Three-dimensional structure of the two-peptide bacteriocin plantaricin JK. Peptides *30*, 1613-1621.

Ruhfel, R.E., Manias, D.A., and Dunny, G.M. (1993). Cloning and characterization of a region of the Enterococcus faecalis conjugative plasmid, pCF10, encoding a sex pheromone-binding function. J Bacteriol *175*, 5253-5259.

Sanchez-Hidalgo, M., Maqueda, M., Galvez, A., Abriouel, H., Valdivia, E., and Martinez-Bueno, M. (2003). The genes coding for enterocin EJ97 production by Enterococcus faecalis EJ97 are located on a conjugative plasmid. Appl Environ Microbiol *69*, 1633-1641.

Sanders, C. R., & Landis, G. C. (1995). Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. Biochemistry *34*(12), 4030-4040.

Schleifer, K.H., and Ludwig, W. (1995). Phylogenetic relationships of lactic acid bacteria. In The Genera of Lactic Acid Bacteria, B.J.B. Wood, and W.H. Holzapfel, eds. (Boston, MA: Springer US), pp. 7-18.

Seddon, A.M., Curnow, P., and Booth, P.J. (2004). Membrane proteins, lipids and detergents: not just a soap opera. Biochimica et biophysica acta *1666*, 105-117.

Sizun, C., Aussenac, F., Grelard, A., & Dufourc, E. J. (2004). NMR methods for studying the structure and dynamics of oncogenic and antihistaminic peptides in biomembranes. Magnetic Resonance in Chemistry *42*(2), 180-186.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. Protein Science, *18*(5), 936-948.

Sobrino-López, A., and Martin-Belloso, O. (2008). Use of nisin and other bacteriocins for preservation of dairy products. International Dairy Journal *18*, 329-343.

Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. Journal of molecular biology *189*(1), 113-130.

Torchillin, V. P., & Weissig, V. (2003). Liposomes. New York: Oxford University Press, *2*, 5-40.

Triba MN, Warschawski DE, Devaux PF (2005) Reinvestigation by phosphorus NMR of lipid distribution in bicelles. Biophys Journal *88*:1887–1901

Uzelac, G., Miljkovic, M., Lozo, J., Radulovic, Z., Tosic, N., and Kojic, M. (2015). Expression of bacteriocin LsbB is dependent on a transcription terminator. Microbiological research *179*, 45-53.

Watts, A., & Spooner, P. J. (1991). Phospholipid phase transitions as revealed by NMR. Chemistry and physics of lipids, *57*(2-3), 195-211.

Watson, R.M., Woody, R.W., Lewis, R.V., Bohle, D.S., Andreotti, A.H., Ray, B., and Miller, K.W. (2001). Conformational changes in pediocin AcH upon vesicle binding and approximation of the membrane-bound structure in detergent micelles. Biochemistry *40*, 14037-14046.

Weaver, K.E., Clewell, D.B., and An, F. (1993). Identification, characterization, and nucleotide sequence of a region of Enterococcus faecalis pheromone-responsive plasmid pAD1 capable of autonomous replication. J Bacteriol *175*, 1900-1909.

Whiles, J. A., Deems, R., Vold, R. R., & Dennis, E. A. (2002). Bicelles in structure–function studies of membrane-associated proteins. Bioorganic chemistry, *30*(6), 431-442.

Wong, T.C. (2017). Micelles and bicelles as membrane mimics for Nuclear Magnetic Resonance studies of peptides and proteins. In Encyclopedia of Surface and Colloid Science (Taylor & Francis).

Yoneyama, F., Imura, Y., Ichimasa, S., Fujita, K., Zendo, T., Nakayama, J., Matsuzaki, K., and Sonomoto, K. (2009). Lacticin Q, a lactococcal bacteriocin, causes high-level membrane permeability in the absence of specific receptors. Applied and environmental microbiology *75*, 538-541.

Appendix

A. Abbreviations

µ	Micro
ABC- transporter	ATP binding cassette - transporters
Amp	Ampicillin
AMPs	Antimicrobial peptides
ATP	Adenosine Tri-phosphate
CD	Circular dichroism spectroscopy
CEX	Cation exchange chromatography
C-terminal	Carboxyl-terminal
bp	Base pairs
CV	Column Volume
Da	Dalton
dH_2O	Distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide Tri-phosphate
DPC	Dodecylphosphocholine
DTT	Dithiothreitol
E.coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetate
EEC	European Economic Community Commission Directive
et al	And others (Latin: et alibi)
EtBr	Ethidium bromide
FPLC	Fast-protein liquid chromatography
FTIR	Fourier Transform Infrared spectroscopy
gm	Gram
GdmCl	Guanidinium chloride
GPCR	G-protein-coupled receptor
GRAS	Generally recognized as safe
GST	Glutathione S-transferase
IPTG	Isopropyl β-D-1-thiogalactopyranoside

Kan	Kanamycin
Kb	Kilo base pair
kDa	Kilo Dalton
L	Liter
LAB	Lactic acid bacteria
LB medium	Lysogeny broth
LsbB	Leaderless lactococcal bacteriocin
m	mili
M	Molar
MQ H_2O	Milli-Q water
Mw	Molecular weight
MWCO	Molecular weight cut-off
N-terminal	Amino-terminal
NEB	New England Biolabs
nm	Nanometer
NMR	Nuclear magnetic resonance spectroscopy
OD	Optical density
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
pI :	Isoelectric point
PMSF	Phenylmethylsulfonyl fluoride
rpm	Revolutions per minute
SDS	Sodium dodecyl sulfate
SDS- PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
Sec	Secretory
TAE buffer	Tris-acetate-EDTA buffer
TE buffer	Tris-EDTA buffer
TFE	Trifluoroethanol
UV light	Ultraviolet light
vs	Versus
xg	Centrifugal force

Amino acid	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic Acid	Asp	D
Cysteine	Cys	C
Glutamic Acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Nucleotide	one letter
Adenine	A
Cytosine	C
Guanine	G
Thymine	T
Purine (A/G)	R
Pyrimidine (C/T)	Y
Any nucleotide	N
Not G (A, C, or T)	H

B. DNA and protein sequences

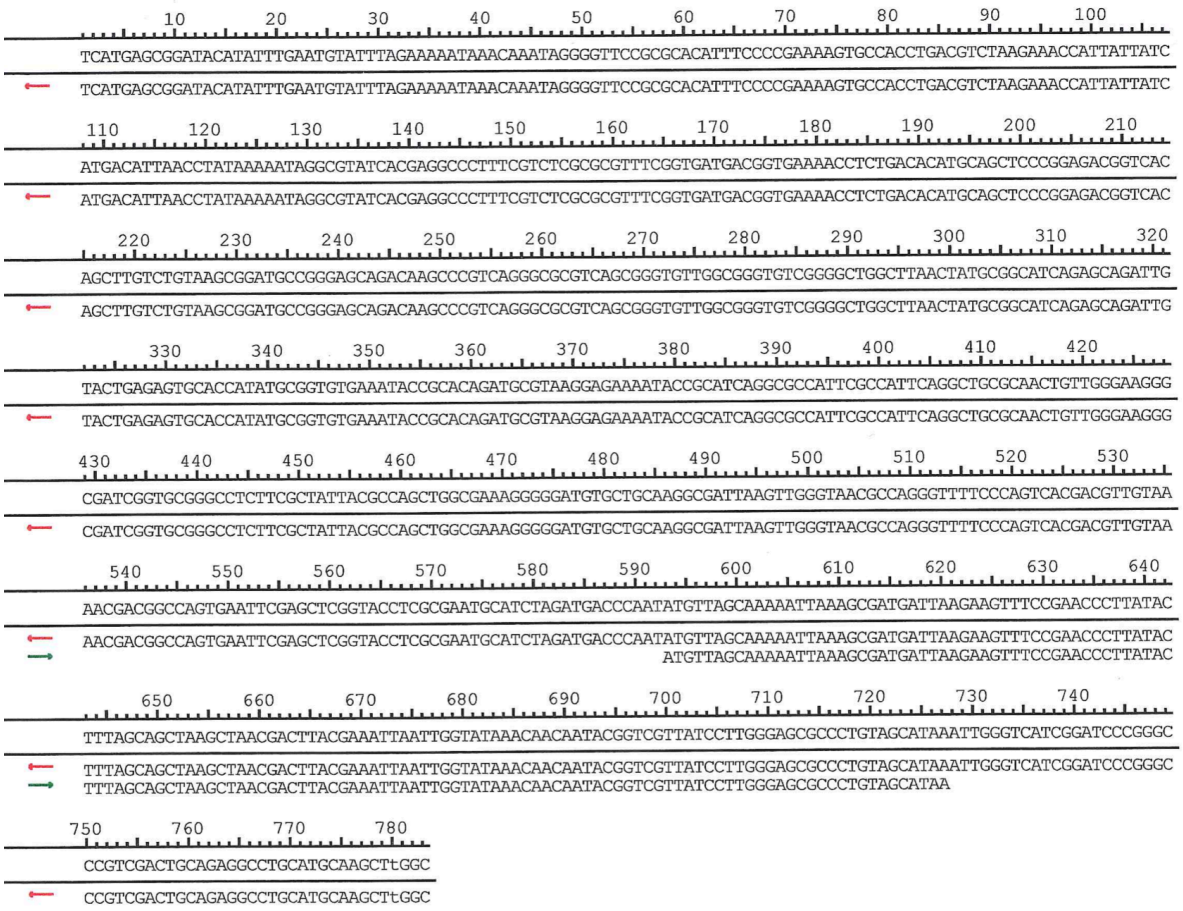
Gene* name: EntEJ97

Gene length: 135 bp

Original gene sequence

M	L	A	K	I	K	A	M	I	K	K	F	P	N	P	15
ATG	TTA	GCA	AAA	ATT	AAA	GCG	ATG	ATT	AAG	AAG	TTT	CCG	AAC	CCT	45
Y	T	L	A	A	K	L	T	T	Y	E	I	N	W	Y	30
TAT	ACT	TTA	GCA	GCT	AAG	CTA	ACG	ACT	TAC	GAA	ATT	AAT	TGG	TAT	90
k	Q	Q	Y	G	R	Y	P	W	E	R	P	V	A	-	45
AAA	CAA	CAA	TAC	GGT	CGT	TAT	CCT	TGG	GAG	CGC	CCT	GTA	GCA	TAA	135

The gene EntEJ97 was cloned into pUC57 vector (shown below):



pCPR0012

tgatccggctgctaacaaagcccgaaaggaagctgagttggctgctgccaccgctgagca
ataactagcataaccccttggggcctctaaacgggtcttgaggggtttttgctgaaagg
aggaactatatccggataattcttgaagacgaaagggcctcgtgatacgcctatttttat
aggttaatgtcatgataataatggtttcttagacgtcaggtggcacttttcggggaaatg
tgcgcggaacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatga
gacaataaccctgataaatgcttcaataatattgaaaaaggaagagtatgagtattcaac
atttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacc
cagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttaca
tcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttc
caatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtgttgacgccg
ggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactcac
cagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgcca
taaccatgagtgataacactgcggccaacttacttctgacaacgatcggaggaccgaagg
agctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaac
cggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgtagcaatgg
caacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaat
taatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccgg
ctggctggtttattgctgataaatctggagccggtgagcgtggctcacgcggtatcattg
cagcactggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtc
aggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagc
attggtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatt
tttaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcccttt
aacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttctt
gagatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccag
cggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttca
gcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttca
agaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctg
ccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataagg
cgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacct
acaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaaggga
gaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagc
ttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttg
agcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacg
cggccttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgt
tatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgcc
gcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcctgatgc
ggtattttctccttacgcatctgtgcggtatttcacaccgcaatggtgcactctcagtac
aatctgctctgatgccgcatagttaagccagtatacactccgctatcgctacgcttcgtt
aatacagatgtaggtgttccacagggtagccagcagcatcctgcgatgcagatccggaac
ataatggtgcagggcgctgacttccgcgtttccagactttacgaaacacggaaaccgaag
accattcatgttgttgctcaggtcgcagacgttttgcagcagcagtcgcttcacgttcgc
tcgcgtatcggtgattcattctgctaaccagtaaggcaaccccgccagcctagccgggtc
ctcaacgacaggagcacgatcatgcgcacccgtggccaggacccaacgctgcccgagatc
tcgatcccgcgaaattaatacgactcactataggaggaccacaacggtttccctctagaa
ataatttgtttaactttaagaaggagatataccatgtctggttctcatcatcatcatca
tcattcttctggtgtagatctgggtaccgagaacctgtacttccaatccatggagaccga
cgtccacatatacctgccgttcactattatttagtgaaatgagatattatgatattttct
gaattgtgattaaaaaggcaactttatgcccatgcaacagaaactataaaaaatacagag
aatgaaaagaaacagatagatttttagttctttaggcccgtagtctgcaaatcctttta
tgattttctatcaaacaaaagaggaaaatagaccagttgcaatccaaacgagagtctaat

agaatgaggtcgaaaagtaaatcgcgcgggtttgttactgataaagcaggcaagacctaa
aatgtgtaaagggcaaagtgtatactttggcgtcacccttacatattttaggtcttttt
ttattgtgcgtaactaacttgccatcttcaaacaggagggctggaagaagcagaccgcta
acacagtacataaaaaaggagacatgaacgatgaacatcaaaaagtttgcaaaacaagca
acagtattaacctttactaccgcactgctggcaggaggcgcaactcaagcgtttgcgaaa
gaaacgaaccaaaagccatataaggaaacatacggcatttcccatattacacgccatgat
atgctgcaaatccctgaacagcaaaaaaatgaaaaatataaagttcctgagttcgattcg
tccacaattaaaaatatctcttctgcaaaaggcctggacgtttgggacagctggccatta
caaaacactgacggcactgtcgcaaactatcacggctaccacatcgtctttgcattagcc
ggagatcctaaaaatgcggatgacacatcgatttacatgttctatcaaaaagtcggcgaa
acttctattgacagctggaaaaacgctggccgcgtctttaaagacagcgacaaattcgat
gcaaatgattctatcctaaaagaccaaacacaagaatggtcaggttcagccacatttaca
tctgacggaaaaatccgtttattctacactgatttctccggtaaacattacggcaaacaa
acactgacaactgcacaagttaacgtatcagcatcagacagctctttgaacatcaacggt
gtagaggattataaatcaatctttgacggtgacggaaaaacgtatcaaaatgtacagcag
ttcatcgatgaaggcaactacagctcaggcgacaaccatacgctgagagatcctcactac
gtagaagataaaggccacaaatacttagtatttgaagcaaacactggaactgaagatggc
taccaaggcgaagaatctttatttaacaaagcatactatggcaaaagcacatcattcttc
cgtcaagaaagtcaaaaacttctgcaaagcgataaaaaacgcacggctgagttagcaaac
ggcgctctcggtatgattgagctaaacgatgattacacactgaaaaaagtgatgaaaccg
ctgattgcatctaacacagtaacagatgaaattgaacgcgcgaacgtctttaaaatgaac
ggcaaatggtacctgttcactgactcccgcggatcaaaaatgacgattgacggcattacg
tctaacgatatttacatgcttggttatgtttctaattctttaactggcccatacaagccg
ctgaacaaaactggccttgtgttaaaaatggatcttgatcctaacgatgtaacctttact
tactcacacttcgctgtacctcaagcgaaaggaaacaatgtcgtgattacaagctatatg
acaaacagaggattctacgcagacaaacaatcaacgtttgcgcctagcttcctgctgaac
atcaaaggcaagaaaacatctgttgtcaaagacagcatccttgaacaaggacaattaaca
gttaacaaataaaaacgcaaaagaaaatgccgatatcctattggcattgacggtctccag
taaaggtggatacg

Kuttesete for BsaI (ggtctc)
sacB
T7primer
Start kodon with His-tag
TEV site
LIR sequence
Sequence complementary to rev primer

Protein Sequence of his-GB1-tev-EntEj97

HMHHHHHHQYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYDD
ATKTYTVTEENLYFQGILAKIKAIIKKFPNPYTLAAKLTTYEINWYKQQYGRYPWER
PVAStopGS

Gene name: his-GB1-tev-EntEJ97

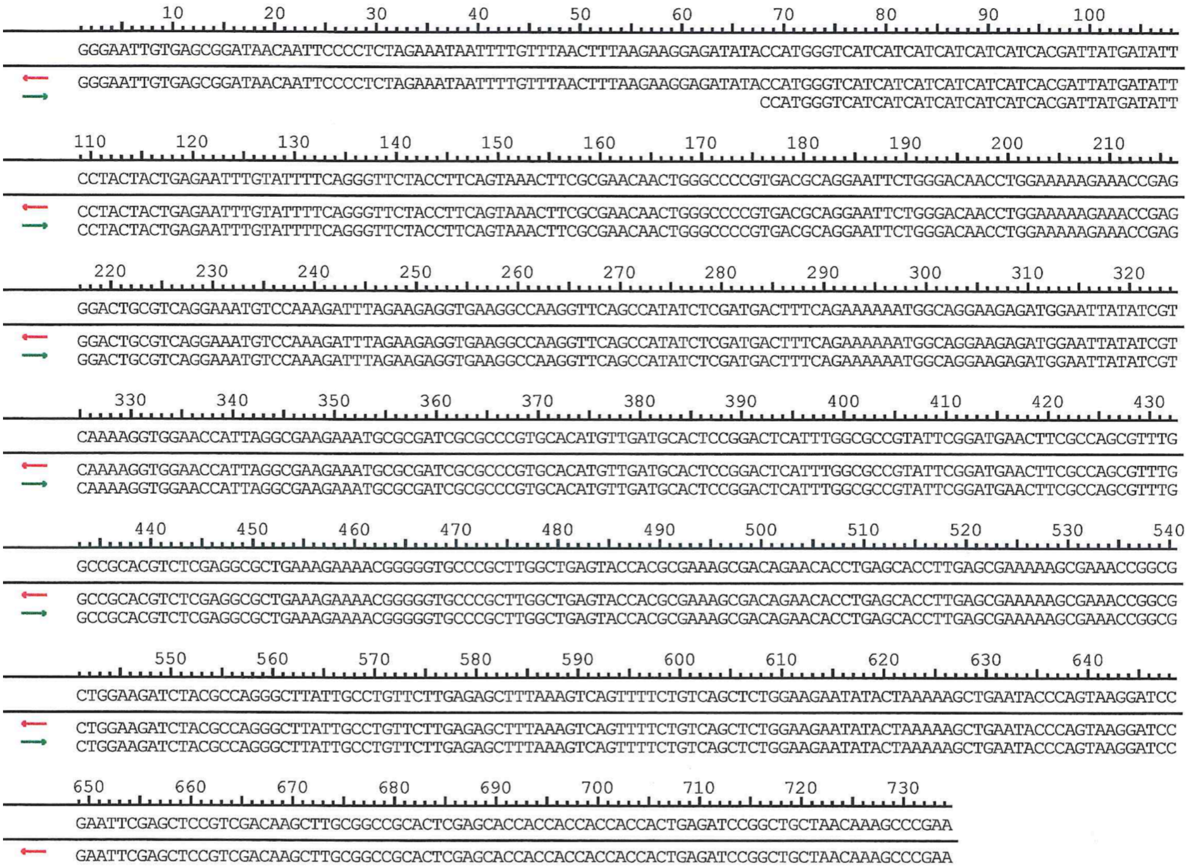
Gene length: 351 pb

The gene his-GB1-tev-EntEJ7 was cloned in pET-22b(+) vector by NdeI and BamHI.

Gene name: MSP

Gene length: Actual MSP that was inserted 573 bp

The gene MSP was cloned in pET-28a(+) NcoI and BamHI.



C. Buffers and media

Media, buffers, and solutions used in this study for various purposes as cited in materials and method section is shown here.

Lysogeny Broth medium

Components	Company
10 g NaCl	VWR Chemicals
10 g Peptone	Sigma-Aldrich
5 g Yeast extract granulated	Merck Millipore
MQ H_2O to bring final volume 1L	
Final medium was autoclaved and stored at room temperature	

LB agar medium

Components	Company
10 g NaCl	VWR Chemicals
10 g Peptone	Sigma-Aldrich
5 g Yeast extract granulated	Merck Millipore
20 g Agar	VWR Chemicals
MQ H_2O to bring final volume 1L	
Final media was autoclaved and when the media cooled to around 55^oC then antibiotic was added and poured on petri plates.	

LB agar with 5 % sucrose

Components	Company
1 g NaCl	VWR Chemicals
1 g Peptone	Sigma-Aldrich
0.5 g Yeast extract granulated	Merck Millipore
2 g Agar	VWR Chemicals
5 gm Sucrose	VWR Chemicals
MQ H_2O to bring final volume 100 ml	
Final media was autoclaved and when the media cooled to around 55^oC then antibiotic was added and poured on petri plates.	

Modified Minimal Media (pH 7.0-8.0)

Components	Company
1.2 gm NH_4Cl	Sigma-Aldrich
5 gm D (+)-Glucose	Merck Millipore
0.2 gm Yeast extract granulated	Merck Millipore
Adjusted pH to 7.4 and MQ H_2O to a final volume of 993.8 ml.	
Autoclaved at 121^oC for 20 minutes. When the solution cooled down then following chemical solutions were added to above medium.	
100 µl Biotin (from stock)	Sigma-Aldrich
2 ml Thiamine (from stock)	Sigma-Aldrich
2 ml $MgSO_4$ (from stock)	VWR Chemicals
100 µl $CaCl_2.2H_2O$ (from stock)	Merck Millipore
1 ml trace elements (from stock)	
1 ml Ampicillin (from stock)	Calibochem

Stock Solutions

Antibiotics

1) **10 mg/ml Ampicillin**

Components	
1 g Ampicillin Sodium Salt (Mw. 371.40 gm/mol)	Calibochem
MQ H_2O to bring final volume 10 ml.	
Final volume was sterile filtrate.	

2) **50 mg/ml Kanamycin**

Components	
0.5 gm Kanamycin disulfate salt (Mw. 483.51 gm/mol)	Sigma-Aldrich
MQ H_2O to bring final volume 10 ml.	
Final volume was sterile filtrate.	

1 mg/ml Biotin

Components	
50 mg Biotin (Mw. 244.31 g/mol)	Sigma-Aldrich
MQ H_2O to bring final volume 50 ml.	
Final volume was sterile filtrate.	
1 ml aliquots was prepared and stored at -20^oC.	

1 M $CaCl_2.2H_2O$

Components	
14.7 gm $CaCl_2.2H_2O$	Merck Millipore
MQ H_2O to bring final volume 100 ml.	
Final volume was autoclaved at 121^oC for 20 minutes.	

1 M Dithiothreitol (DTT)

Components	
1.54 g DTT (Mw. 154.24 g/mol)	VWR Chemicals
MQ H_2O to bring final volume 10 ml.	
Final volume was sterile filtrate.	

1 M Imidazole

Components	
34.04 g Imidazole (Mw. 68.08 g/mol)	USB Corporation
MQ H_2O to bring final volume 500 ml	
Final volume was sterile filtrate.	

1 M Isopropyl-1-thio-β-D-galactopyranoside (IPTG)

Components	
2.383 gm IPTG (Mw: 238.3 g/mol)	VWR Chemicals
Dissolved in 10 ml MQ H_2O.	
Final volume was sterile filtrate.	

1 M $MgSO_4.7H_2O$

Components	
24.65 g $MgSO_4.7H_2O$	VWR Chemicals
MQ H_2O to bring final volume 100 ml.	
Final Volume was autoclaved at 121^oC for 20 minutes.	

2 M NaCl

Components	
58.44 g NaCl (Mw. 58.44 g/mol)	VWR Chemicals
MQ H_2O to bring final volume to 500 ml	
Final solution was sterile filtrate.	

Phosphate Buffer Saline (PBS)

Components	
1.16 g Na_2HPO_4	Merck Millipore
0.1 g KCl	Merck Millipore
0.1 g K_3PO_4	Merck Millipore
4 g NaCl	VWR Chemicals
MQ H2O to bring final volume 500 ml (pH 7.4)	

1M Phenylmethylsufonyl fluoride

Components	
1.742 gm PMSF (Mw. 174.19 g/mol)	Thermo Scientific
Dissolved in Ethanol to a final volume of 10 ml.	

0.2 M Sodium phosphate buffer

Components	Company
15.602 g NaH_2PO_4 (Mw. 156.02 g/mol)	Merck Millipore
17.799 g Na_2HPO_4 (Mw. 177.99 g/mol)	Merck Millipore
Dissolved in 400 ml MQ H_2O, pH was adjusted to 7.4 and final volume was brought to 500 ml.	
The final solution was sterile filtrate.	

20mM Sodium phosphate buffer

Components	Volume
0.2 M Sodiumphosphate buffer (from stock)	100 ml
MQ H_2O	900 ml
Total	1000 ml

1 mg/ml Thiamin

Components	
50 mg Thiamin-HCl (Mw. 337.27 g/mol)	Sigma-Aldrich
MQ H_2O to bring 50 ml.	
Final volume was sterile filtrate.	
1 ml aliquots was prepared and stored at -20^{o}C.	

100 ml Trace Elements solution contained per litre

Components	Company
600 mg $CaCl_2.2H_2O$	Merck Millipore
600 mg $FeSO_4.7H_2O$	Sigma-Aldrich
500 mg EDTA	Applichem
115 mg $MnCl_2.4H_2O$	Merck Millipore
80 mg $CoCl_2.6H_2O$	Merck Millipore
70 mg $ZnSO_4.7H_2O$	Merck Millipore
30 mg $CuCl_2.2H_2O$	Sigma-Aldrich
25 mg $(NH_4)_6Mo_7O_{24}.4H_2O$	Merck Millipore
2 mg H_3BO_3	Merck Millipore

Agarose gels and corresponding buffers

Components of 1% agarose gel

Components	Company
SeaKem®LE Agarose	LONZA
SYBR® Safe DNA gel stain	Invitrogen™

50X TAE buffer

Components	Company
242 gm Tris base	Angus
57.1 ml glacial acetic acid	Sigma-Aldrich
20.81 gm EDTA	AppliChem
MQ H_2O to bring final volume 1 L	

1X TAE buffer

20 ml TAE buffer taken from 50X stock solution
MQ H_2O to bring final volume 1 L

HisTrap™ HP 5 ml (Ni-column) Buffers

Binding Buffer

Component	Volume
20 mM Phosphate buffer (pH 7.4)	100 ml
0.5 M NaCl	250 ml
20 mM Imidazole	20 ml
MQ H_2O	630 ml
Total	1000 ml

Elution Buffer

Component	Volume
20 mM Phosphate buffer (pH 7.4)	100 ml
0.3 M NaCl	150 ml
0.3 M Imidazole	300 ml
MQ H_2O	450 ml
Total	1000 ml

20 % Ethanol

Components	Volume
Ethanol	200 ml
MQ H_2O	800 ml
Total	1000 ml

Degassed MQ H_2O

1 liter MQ H_2O
To be degassed the same day as purification.

HiTrap™ 5 ml Desalting Buffers

Binding Buffer

Component	Volume
20 mM Phosphate buffer (pH 7.4)	100 ml
0.3 M NaCl	150 ml
0.3 M Imidazole	300 ml
MQ H_2O	450 ml
Total	1000 ml

Elution Buffer

Component	Volume
20 mM Phosphate buffer (pH 7.4)	100 ml
0.3 M NaCl	150 ml
20 mM Imidazole	20 ml
MQ H_2O	730 ml
Total	1000 ml

SDS-PAGE corresponding buffers and stains

1X MOPS SDS Running Buffer

Components	
50 ml 20X NuPAGE® MOPS SDS Running Buffer	Thermo Scientific
950 ml MQ H_2O to bring final volume 1 L.	

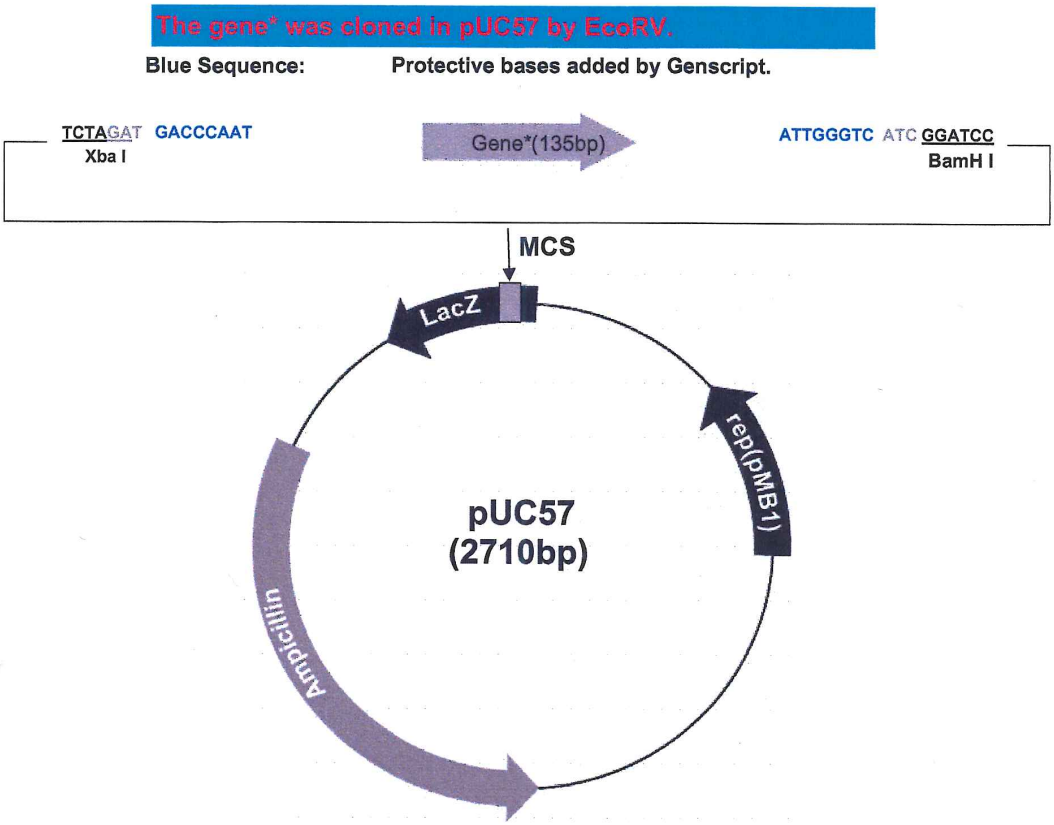
Composition of Coomassie Brilliant Blue R250 staining solution

Components	Company
1 g Brilliant Blue R250	Sigma-Aldrich
10 % acetic acid	Sigma-Aldrich
40 % ethanol	Arcus
MQ H_2O to bring final volume 1L	

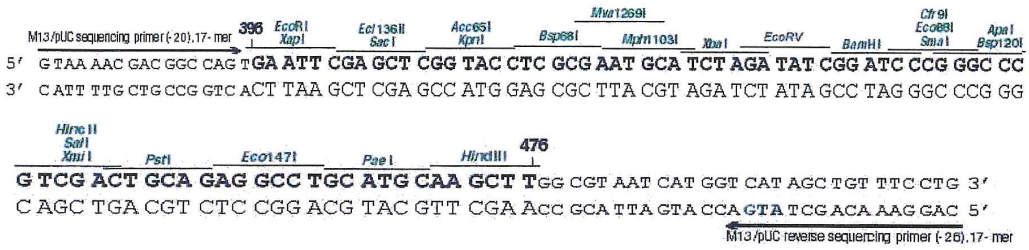
Composition of Coomasie Brilliant Blue destaining solution

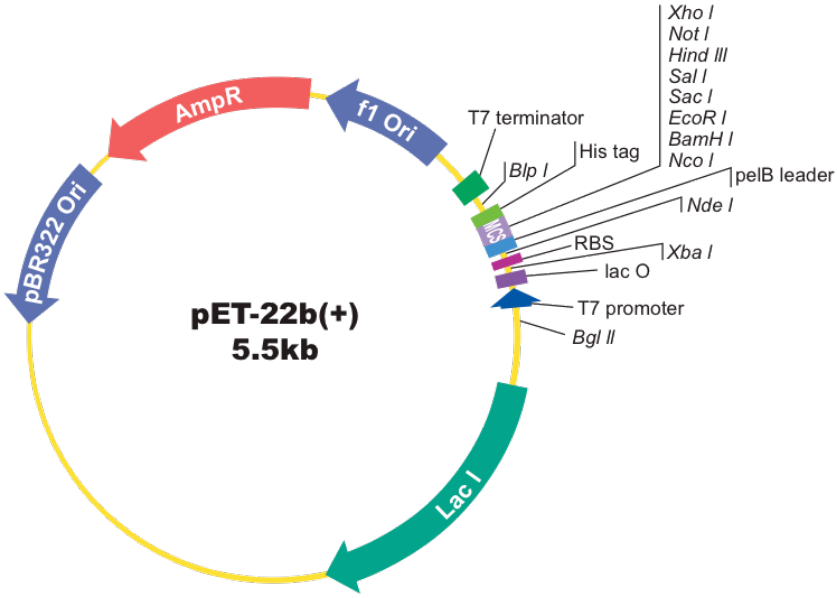
Components	Company
10 % ethanol	Arcus
7.5 % acetic acid	Sigma-Aldrich
MQ H_2O to bring final volume 1L	

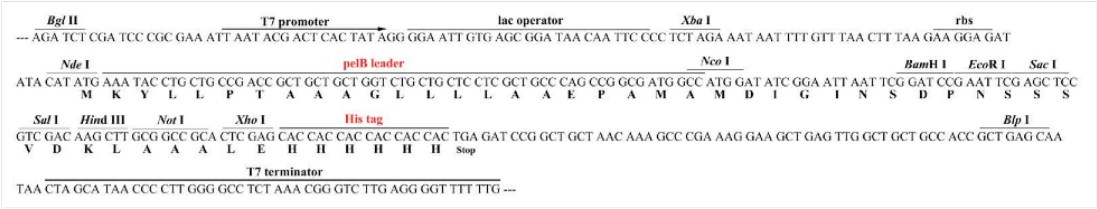
D. Plasmid construct map

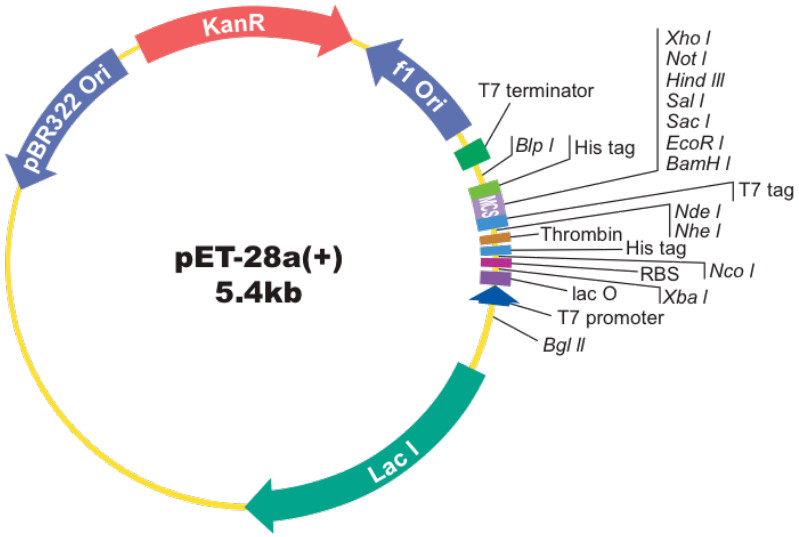


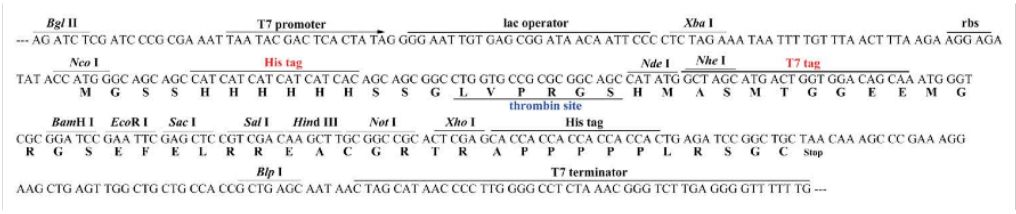
MCS:




Xho I
Not I
Hind III
Sal I
Sac I
EcoR I
BamH I
Nco I
pelB leader
Nde I
Xba I
T7 terminator
His tag
Blp I
RBS
lac O
T7 promoter
Bgl II
AmpR
f1 Ori
pBR322 Ori
pET-22b(+)
5.5kb
Lac I


Bgl II
T7 promoter
lac operator
Xba I
rbs
--- AGA TCT CGA TCC CGC GAA ATT AAT ACG ACT CAC TAT AGG GGA ATT GTG AGC GGA TAA CAA TTC CCC TCT AGA AAT AAT TTT GTT TAA CTT TAA GAA GGA GAT
Nde I
pelB leader
Nco I
BamH I
EcoR I
Sac I
ATA CAT ATG AAA TAC CTG CTG CCG ACC GCT GCT GCT GGT CTG CTG CTC CTC GCT GCC CAG CCG GCG ATG GCC ATG GAT ATC GGA ATT AAT TCG GAT CCG AAT TCG AGC TCC
M K Y L L P T A A A G L L L A A E P A M A M D I G I N S D P N S S S
Sal I
Hind III
Not I
Xho I
His tag
Blp I
GTC GAC AAG CTT GCG GCC GCA CTC GAG CAC CAC CAC CAC CAC CAC TGA GAT CCG GCT GCT AAC AAA GCC CGA AAG GAA GCT GAG TTG GCT GCT GCC ACC GCT GAG CAA
V D K L A A A A L E H H H H H H Stop
T7 terminator
TAA CTA GCA TAA CCC CTT GGG GCC TCT AAA CGG GTC TTG AGG GGT TTT TTG ---


KanR
f1 Ori
pBR322 Ori
pET-28a(+)
5.4kb
Lac I
T7 terminator
Blp I
His tag
Xho I
Not I
Hind III
Sal I
Sac I
EcoR I
BamH I
T7 tag
MCS
Nde I
Nhe I
Thrombin
His tag
RBS
Nco I
lac O
Xba I
T7 promoter
Bgl II


Bgl II
T7 promoter
lac operator
Xba I
rbs
--- AG ATC TCG ATC CCG CGA AAT TAA TAC GAC TCA CTA TAG GGG AAT TGT GAG CGG ATA ACA ATT CCC CTC TAG AAA TAA TTT TGT TTA ACT TTA AGA AGG AGA
Nco I
His tag
Nde I
Nhe I
T7 tag
TAT ACC ATG GGC AGC AGC CAT CAT CAT CAT CAT CAC AGC AGC GGC CTG GTG CCG CGC GGC AGC CAT ATG GCT AGC ATG ACT GGT GGA CAG CAA ATG GGT
M G S S H H H H H H S S G L V P R G S H M A S M T G G E E M G
thrombin site
BamH I
EcoR I
Sac I
Sal I
Hind III
Not I
Xho I
His tag
CGC GGA TCC GAA TTC GAG CTC CGT CGA CAA GCT TGC GGC CGC ACT CGA GCA CCA CCA CCA CCA CCA CTG AGA TCC GGC TGC TAA CAA AGC CC GAA AGG
R G S E F E L R R E A C G R T R A P P P P P P L R S G C Stop
Blp I
T7 terminator
AAG CTG AGT TGG CTG CTG CCA CCG CTG AGC AAT AAC TAG CAT AAC CCC TTG GGG CCT CTA AAC GGG TCT TGA GGG GTT TTT TG ---

E. Primer sequences

T7 Sequencing Primer

5´ TAA-TAC-GAC-TCA-CTA-TAG-GG-3´

Primers for subcloning

pCPRej97 Forward	TAC-TTC-CAA-TCC-ATG-TTA-GCA-AAA-ATT-AAA-G
pCPRej97 Reverse	TAT-CCA-CCT-TTA-CTG-TTA-TGC-TAC-AGG-GCG-CTC

F. Variables/parameters Äkta Purifier system

HisTrap™ HP column Variables/parameters

Block	Variable	Value
Main	Column	HisTrap_HP_5_ml
Flow rate	ml/min	5.00
Column Pressure limit	Mpa	0.9
Wavelength	Wavelength 1 (nm)	280
	Wavelength 2 (nm)	214
Eluent A_Inlet	Pump A_Inlet	A1
Eluent B_Inlet	Pump B_inlet	B1
Start_conc_B	Start_concB (%B)	0.0
Column Equilibration	Equilibrate with (CV)	2
Flowthrough Fractionation	Flowthrough Tube	12mm
	Flowthrough Frac Size (ml)	2.00
Sample injection	Empty loop with (ml)	10
PeakFrac_parameters_UV	UV Mode	Slope
	Meanpeak width (min)	0.250
	Peak start slope (mAU/min)	100.00
	Peak End slope (mAU/min)	75.00
Fractionation 1	Tube type	12mm
	Frac size (ml)	2.00
	Peak Frac size (ml	2.00
	Peak Frac start at	Next tube
Gradient Segment 1	Target concB 1 (%B)	20
	Length of gradient (ml)	3.00
Fractionation 2	Tube type	12mm
	Frac size (ml)	2.00
	Peak Frac size (ml	2.00
	Peak Frac start at	Next tube
Gradient Segment 2	Target concB 2 (%B)	60
	Length of gradient (ml)	5.00

Fractionation 3	**Tube type**	**12mm**
	Frac size (ml)	**2.00**
	Peak Frac size (ml	**2.00**
	Peak Frac start at	**Next tube**
Gradient Segment 3	**Target concB 3 (%B)**	**100**
	Length of gradient (ml)	**2.00**
Fractionation 4	**Tube type**	**12mm**
	Frac size (ml)	**2.00**
	Peak Frac size (ml	**2.00**
	Peak Frac start at	**Next tube**
Gradient Segment 4	**Target concB 4 (%B)**	**100**
	Length of gradient (ml)	**1.00**
Fractionation 5	**Tube type**	**12mm**
	Frac size (ml)	**2.00**
	Peak Frac size (ml	**2.00**
	Peak Frac start at	**Next tube**
Gradient Segment 5	**Target concB 5 (%B)**	**0.00**
	Length of gradient (ml)	**5.00**

HiTrap™ Desalting column variables/parameters

Block	Variable	Value
Main	Column	HiTrap_Desalting
Flow rate	ml/min	5.00
Column Pressure limit	Mpa	0.9
Wavelength	Wavelength 1 (nm)	280
	Wavelength 2 (nm)	214
Eluent A_Inlet	Pump A_Inlet	A1
Eluent B_Inlet	Pump B_inlet	B1
Start_conc_B	Start_concB (%B)	0.0
Sample injection	Empty loop with (ml)	1.0
PeakFrac_parameters_UV	UV Mode	Slope
	Meanpeak width (min)	0.1
	Peak start slope (mAU/min)	100.00
	Peak End slope (mAU/min)	75.00
Fractionation	Tube type	12mm
	Eluate Fra size (ml)	2.00
	Eluate Frac Start at	Next tube
	Peak Frac size (ml	2.00
	Peak Frac start at	Next tube
Length of Elution	Length of Elution (CV)	1

G. Materials, equipment, and computer software

Material	Producer	Product no.
Buffers		
CutSmart	New England Biolabs®	B7204S
NuPAGE LDS Sample buffer (4x)	Thermo Scientific	NP0007
NEB 2.1	New England Biolabs®	B7202S
Pfu Buffer	Aligent	Supplied
Enzymes		
BsaI	New England Biolabs®	Supplied
EDTA-free protease Cocktail	Sigma	Supplied
Pfu Turbo DNA polymerase	Aligent	Supplied
TEV protease	Sigma	T4455-1MG
T4 DNA polymerase	New England Biolabs®	M0203S
Loading dye		
6X DNA loading dye	Fermentas	R0611
Gel loading dye (6X)	New England Biolabs®	B7024S
Ladders		
GeneRuler 1 kb Plus DNA ladder	Thermo Scientific	SM0311
O´Generuler 50 bp DNA ladder	Thermo Scientific	SM1138
Spectra multicolor Low range Protein Ladder	Thermo Scientific	26628
Protein and DNA techniques		
GFX PCR and DNA purification kit	GE Healthcare	Supplied
NucleoSpin® Plasmid (NoLid)	Macherey-Nagel	Supplied
NucleoBond®Xtra	Macherey-Nagel	Supplied

Equipment

Product	Producer	Model/Product no.
Centrifuge	Beckman	Avanti™ J-25
Cuvettes	VWR	634-0676
Digital Dri-block heater	Techne	DB-20
Electrophoresis power supply	Pharmacia Biotech	EPS 600
Fixed-Angle platform rocker	Grant-bio	PMR-30
Illuminated Refrigerated Incubator Shaker	New Brunswick Scientific	innova™4340
Incubator Shaker	New Brunswick Scientific	Series 25
Magnetic stirrer	Snijders	34521
Microcentrifuge	Beckman Coulter	Microfuge®18
Microcentrifuge	Orto Alresa	Biocen 21
Microwave	Samsung	T.D.S.
NanoDrop UV-Vis Spectrophotometer	Thermo Scientific	2000
Peltier Thermal Cycler	MJ Research	PTC-200
pH/ION Meter	MeterLab	PHM240
SDS-PAGE power supply	Bio-rad	Power Pac 300
Stirred Water bath	BioSan	WB-4MS
Tabletop Laboratory centrifuge	Alfa Lab	
Thermomixer Comfort 1.5 ml	Effendorf	MTP
Ultrasonic Processor	Sonics	Vibra-Cell™
Ultraschallprozessor	dr. hielscher	UP 400 s
Ultra-pure water system	Millipore	Milli-Q_{PLUS}

Computer software

Program	Version	Company
Excel	14.0.0 (100825)	Microsoft®Excel® for Mac 2011
NanoDrop 2000	1.6	Thermo Scientific
Word	14.0.0 (100825)	Microsoft®word for Mac 2011
Unicorn	5.3	
MacPyMOL	1.7.4.5	SchrÖdinger LLCResults