# Overproduction of recombinant VirG from *Shigella flexneri*.

**Yumiko Tashiro**

**Supervisor**

**Prof. Richard Pickersgill**

## Abstract

The ability of *Shigella flexineri* to spread within and between epithelial cells is essential for *Shigella* infection causing bacillary dysentery. This is a particular problem in the developing world. The movement of *Shigella* within the host cell requires the accumulation of actin at one pole of the bacterium and the protein VirG is responsible for this function. While the C-terminal domain (β–domain) of VirG is integrated into the outer membrane of *Shigella*, the N-Terminal domain (α–domain) is exposed on the surface of the bacterium. The β-domain acts as autotransporter of the α–domain. The exposed α–domain has multiple binding partners including N-WASP, Vinculin and IcsB that are required for infection in man and cell to cell spread. To understand the molecular basis of VirG's activity, it is first necessary to produce the protein in quantity; this study investigates the expression of VirG α–domain in *E.coli*. The optimum construct corresponded to residues 58-506 of VirG expressed in Rossetta-gami cells.

## Acknowledgments

I have many thanks to say.

First of all, I would like to thank Prof. Richard Pickersgill for the supervision and giving me an opportunity to study such an interesting project.

Many thanks to the member of Pickersgill and Viles groups and lovely colleagues in the department, especially, Teng, Arefeh, and Salyha for their supports.

Finally, thank you to the Queen Mary University of London, School of Biological and Chemical Science for giving me such a grateful opportunity to do study.

# Table of contents

## Chapter 3. Over-production of the alpha-domain of VirG using pGEX2T vector



## Chapter 4. Overproduction of the alpha domain of VirG using pET system



## Chapter 5. Over-production and purification of WIP-WASP

## Chapter 6. Conclusion and Future work

## List of figures

## List of tables

# Abbreviations

| | |
|---|---|
| Amp | Ampicillin |
| b.p | Base pair |
| Cam | Chloramphenicol |
| Da | Dalton |
| DMF | Dimethylformamide |
| DMSO | Dimethyl sulfoxide |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxyribonucleotide triphosphates |
| DTT | Dithiothreitol |
| *E.coli* | *Escherichia coli* |
| EDTA | Ethylenediaminetetraacetic acid |
| FPLC | Fast protein liquid chromatography |
| g | Gram |
| Gly | Glycine |
| GST | Glutathione S-transferase |
| HCl | Hydrochloric acid |
| His | Histidine |
| IPTG | Isopropyl-1-thio-β-D-galactopyranoside |
| Kan | Kanamycin |
| kb | Kilo base |
| Kb.p | Kilo base pair |
| kDa | Kilo Dalton |
| L | Litre |
| LB broth | Luria-Bertani broth |
| M | Molar |
| mg | Milligram |
| $MgCl_2$ | Magnesium Chloride |
| $MgSO_4$ | Magnesium sulfate |
| ml | Millilitre |
| mM | Millimolar |
| MME | Monomethylether |
| MPD | 2-methyl-2, 4- pentanediol |
| MWCO | Molecular weight cut off |
| N-WASP | Neuronal Wiskott–Aldrich Syndrome protein |
| NaCl | Sodium chloride |
| NMR | Nuclear magnetic resonance |
| OD | Optical density |
| O/N | Over night |
| PBS | Phosphate buffer saline |
| PCR | Polymerase chain reaction |
| PET | Polyethylene glycol |
| rcf | Relative centrifugal force |

| | |
|---|---|
| rpm | Revolutions per minute |
| sec | Seconds |
| Ser | Serine |
| SDS | Sodium dodecyl sulfate |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| TAE | Tris-acetate-EDTA |
| Tet | Tetracycline |
| V | Voltage |
| WASP | Wiskott–Aldrich Syndrome protein |
| WIP | WASP-interacting protein |
| x g | Times gravity |
| x-gal | 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside |
| % | Persentage sign |
| $^{o}$C | Degrees celsius |
| µg | Microgram |
| µl | Microlitre |
| µM | Micromolar |

# Chapter 1

# Introduction

## 1.1 **A brief introduction to bacterial infection**

Some bacteria, such as Lactic acid bacteria and Bifidobacteria, are not harmful but beneficial to humans. However, quite a few bacteria are responsible for diseases in higher plants, animals and humans. For instance in humans, tuberculosis, pneumonia and foodborne illnesses are caused by pathogenic bacteria. Diseases from bacterial infections include tetanus, typhoid fever, diphtheria, syphilis and leprosy.

Common pathogenic bacteria are *Streptococcus*, *Pseudomonas*, *Shigella*, *Campylobacter*, *Salmonella*, enteropathogenic *Escherichia coli* and mycobacteria. All cause potentially lethal infections. For example, the mycobacterium, *Mycobacterium tuberculosis* causes tuberculosis. *M. bovis bacilli* Calmette–Guérin (BCG) vaccine is available for vaccination against tuberculosis and continues to be administered to infants at birth in most regions where tuberculosis is endemic. However the BCG vaccine can cause fatal disseminated infection in immunosuppressed patients, such as HIV-infected newborns. Unlike tuberculosis, the vaccine for shigellosis that is caused by *Shigella* is still under development.

Bacteria are classified into two main groups of bacteria, Gram-positive and Gram-negative. Gram-negative bacteria can cause many types of infections. The tool that is used to defferencite the bacteria group is Gram–staining. Gram-staining has four basic steps, primery staing with crytal violet, fixing the cell wirh iodine as mordant, decolourization with mixture of ethanol and aceton and finaly, counterstaining with safranin. Both Gram-positive and –negative bacteria are stained deep purple colour when the staining substance, crystal violet and iodine, penetrate into the cell wall and form crystal violet-iodine complex during the staining process. Bacteria reveal its differenciatation after decolourization and counterstining treatment. Gram-positive bacteria and Gram-negative bacteria are stained purple and pink respectively (Figure 1.1 and Figure 1.2) after completion of stining process. The difference in

colour of the stained bacteria is depending on the organization and the structure of the bacterial cell wall. Cell wall is located outside of bacterial plasma membrane (cell membrane). The Gram-positive bacterial cell wall consist of the thick proteoglycan layer (20-80 nm), while the Gram-negative bacterial cell wall are made of two layers, outer membrane and thin proteoglycan layer (2-3 nm) (Figure 1.2). The cell wall of Gram-positive bacteria shrinks due to dethydration of cell wall during the decoloriaion step. Therefore the staining substances are trapped in cell and the bacteria are stained purple. Gram-negative bacteria, on the other hand, have a thin layer of proteoglycan that is located between plasma membrane and outer membranes. The decolorizer (the mixture of ethanol and acetone) during declorization step dissolves the lipid layer of outer membrane of the Gram-negative bacteria. The removal of lipid layer expose thin peptidoglycan layer and causes the crystal violet-iodine complex wash away from the cell. As a result, the cell loses the colour. After counterstaining with safranin, the Gram-negative bacteria cell remains pink colour. The decolorization step is critical. The crystal violet stain can be removed from both gram-positive and negative cells if the decolorizing agent is left on too long.

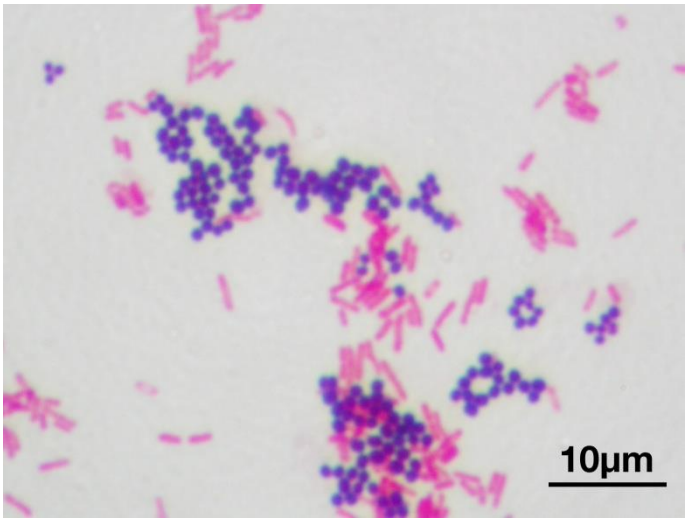



**Figure 1.1 Image of a Gram stain of mixed Gram-positive cocci (*Staphylococcus aureus* ATCC 25923, purple) and Gram-negative bacilli (*Escherichia coli* ATCC 11775, red). The image was recorded using an optical microscope at a magnification of 1000 times. The image is taken from http://commons.wikimedia.org/wiki/File:Gram_stain_01.jpg**

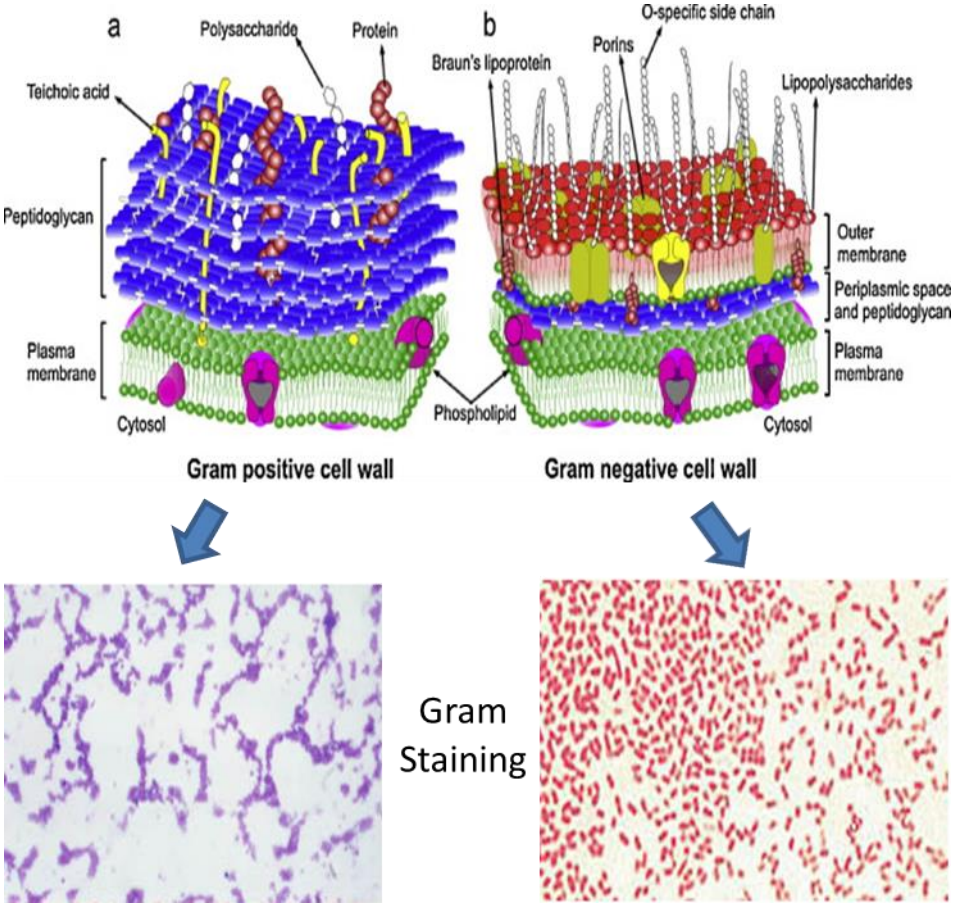



**Figure 1.2. The schematic overview of Gram-positive and -negative cell walls and the result of the Gram-staining. (Tripathi *et al.* 2012 and http://www.nmpdr.org/FIG/wiki/view.cgi/FIG/GramStain)**

## 1.2 Gram negative bacteria

Many antibiotics inhibit Gram-negative bacteria less effectively than Gram-positive bacteria. This is because of the low permeability of the outer membrane barrier that prevents ingress of the antibiotic. Gram-negative bacterial also have outer membrane pumps that pump out antibiotics and confer drug resistance.

Unlike Gram-positive bacteria, Gram-negative bacteria have a thin layer of proteoglycan and has extra layer of outer membrane for protection. The outer membrane consists of a phospholipid membrane, similar to the cell or plasma membrane. The major comportment of the bacterial cell wall is Lipopolysaccharide (LPS). LPS is a highly-branched fatty sugar that is attached to the outer membrane and acts as endotoxin that induces fever and shock in the

human host. Selective permeability of the outer membrane is retained by transporter protein called porins. Porins are embedded in the outer membrane of Gram-negative bacteria and control the flow of small moleculars across the inner membrane (IM) into the cytoplasm. The small number of drug molecules that traverse the outer membrane through transporters similar to porins in antibiotic resistant Gram-negative bacteria can be efficiently inactivated by the enzyme such as β-lactamases or is transported to extracellular space. The process of translocation of drug out of the bacterial cell is often catalyzed by widely distributed multidrug efflux pumps. Transport of much larger molecules, such as extracellular enzymes, toxins and bacterial effectors, use dedicated secretory pathways such as the Type Three Secretion System.

**1.3 Secretion systems in Gram negative bacteria**

Protein secretion plays a central role in modulating the interactions of bacteria with their environments. Seven types of secretion systems in gram negative bacteria have been described so far. There are named Type One Secretion system (T1SS), Type Two Secretion system (T2SS), Type Three secretion system (T3SS), Type Four Secretion System (T4SS), Type Five secretion system (T5SS), Type Six secretion system (T6SS) and Type Seven Secretion system (T7SS). Figure 1.3 shows the schematic overview of the major protein secretion systems in Gram-negative bacteria. Only T4SS is found in both Gram-negative and positive bacteria, the other five secretion systems that exclude T7SS are present only in Gram-negative bacteria. The T7SS is found in Mycobacteria but this system is not yet well-characterised. All secretion systems are dedicated to the specific secretion of exoproteins. Such proteins are transported from the cell's interior across the cell envelope to extracellular space or the host cell cytosol and exhibit various activities useful for bacterial adaptation to the environment or for bacterial pathogenicity.

The secretion systems can be subdivided into Sec-independent and Sec-dependent pathways.

The Sec dependent secretion systems, T2SS and T5SS, translocate proteins to extracelluar space in two steps. First step is the translocation of the precursor protein to periplasm by crossing the inner membrane. This process is achieved by the Tat pathway and Sec pathway in T2SS and T5SS respectively. Sec or Tat apparatus are targeted by N-terminus signal peptide of the expressed precursor protein. The distinctive difference between two pathways is that Sec pathway transports unfolded proteins while Tat system delivers folded protein. Sec apparatus is comprised of 4 proteins, SecYEG, SecDF, SecA, and YidC are found. Further translocation across the outer membrane is accomplished by a dedicated system- T2SS or T5SS. T2SS was found in the *Legionella pneumophila* or *Yersinia Entrocolitica* and composed of about 16 different proteins. The complex of protein is called secreton that embedded in the outermembrane of bacteria. Secreton like complex are also found in the T3SS.

Unlike the Sec dependent secretion systems, T1SS, T3SS, T4SS and T6SS export protein to extracellular space by bypassing the periplasm. Three secretion systems (T3SS, T4SS and T6SS) have needle-like structures that can penetrate the targeted host cell membrane and transports bacterial protein for infection (Figure 1.3). T1SS is depending on the subclass of efflux pumps, either a proton motive force or ATP hydrolysis that is utilized as energy source.

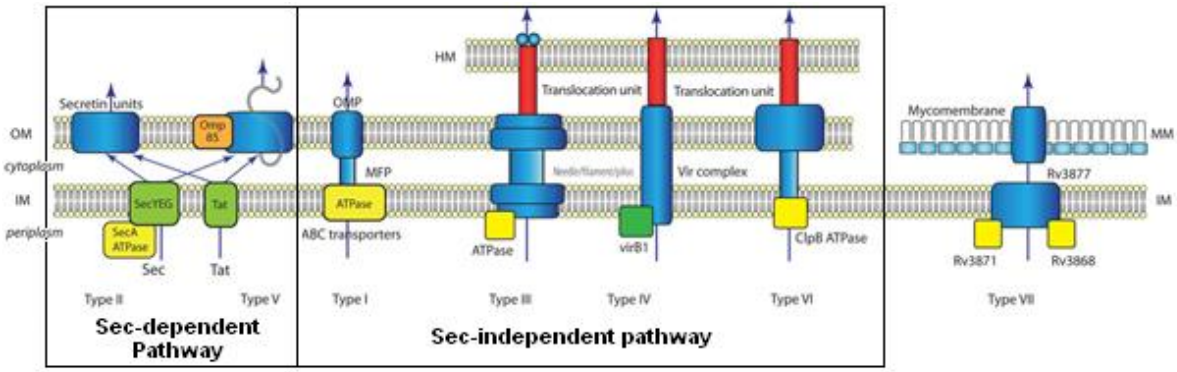



**Figure 1.3 The schematic overview of the six major protein secretion systems (Type I ~ Type VI) in Gram-negative bacteria and Type VII secretion system in Mycobacteria. (Figure taken from Tyler *et al*. 2009)**

### 1.4 Type Three Secretion System (T3SS)

The Type III secretion System (T3SS) is an essential determinant of the interaction of many different Gram-negative bacteria, such as animal pathogens *Salmonella*, *Shigella* and *Vibrio*, plant pathogens *Pseudomonas*, *Xanthomonas*, and *Ralstonia*. The T3SS is characterised by host contact–mediated induction, ATPase energy requirement for protein secretion and translocation or the protein effectors into host cells. Although the structure of whole T3SS has been determined by electron microscopy (Blocker *et al.,* 2001), not all of the individual proteins have had their structures solved. T3SS seems has sight variations in structure. In T3SS of *Shigella flexineri*, part of T3SS is embedded in the membrane. The outer membrane region (OMR) of T3SS is composed of three rings, OMR1–3. OMR1-2 is similar to that which is seen in secretins of the type II secretion system. The C-terminal of OMR3 is known to be embedded in the outer membrane and its N-terminal domain is periplasmic and soluble. The inner membrane region (IMR) is formed of one thick ring with 24-fold symmetry which create thinner than that of *Salmonella tryphimurium*. The structure of T3SS is similar to the flagellum that is responsible for the bacterial motility (Figure 1.4). There is an evolutional relationship between T3SS and the flagellum, although it is not clear if the ancestor was a flagellum or a secretion system, perhaps a secretion system is more

probable as the flagellum assembles via a mechanism similar to secretion. (Sophie S. Abby and Eduardo P. C. Rocha, 2012)

The T3SS of *Shigella flexneri* is composed of an external needle, a transmembrane domain and a cytoplasmic bulb. *Shigella flexneri* MxiH assembles to form the helical needle structure (Cordes *et al.*, 2003). The study showed that the structure of helical packing of needles has similar parameters to those of the flagella rod, hook and filament, but the needle monomer protein displays no primary sequence homology with any flagella axial components. MxiH mutants showed several alterations in needle structure and function of the TTSS, morphologically (defect of needle length) and physically (reduction of interaction with the host cells and non-induced continuous secretion of molecules, Kenjale *et al.*, 2005). These results suggest the needle may control the activity of the TTSS directly.

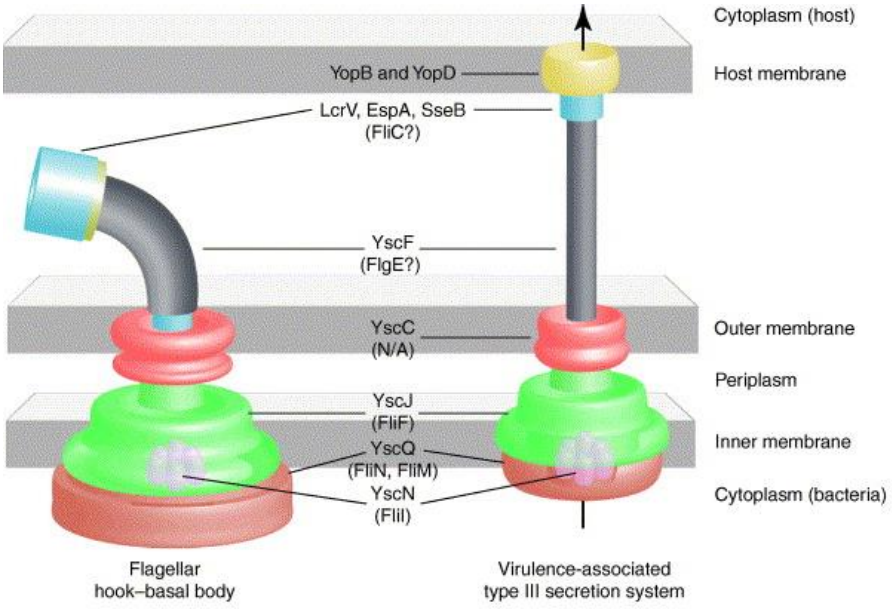



**Figure 1.4 The schematic image of flagella and type III secretion system (Yip and Strynadka 2006).**

### 1.5 *Shigella flexneri*

The Gram-negative bacterium *Shigella* comprises 4 species: *S. dysenteriae*, *S. flexneri*, *S. boydii* and *S. sonnei*. *Shigella* is probably most well known as the cause of

bacillary dysentery or shigellosis as it is otherwise known. Over 20 million people, mostly young children in developing countries, are suffering from the disease and one million people are killed by the infection every year. Most cases of infection in the developed countries are due to *S. sonne,* whereas *S. dysenteriae* and *S. flexneri* are common species in developing countries. Transmission of bacteria to humans occurs when people consume water contaminated by faeces from infected individuals. Some symptoms of illness are vomiting and diarrhea, which may contain mucus and blood, and the conditions may occur within 2-3 days after exposure. The ingestion of 100-200 *S. flexneri* can cause the disease. People infected by *S. flexneri* may develop the condition called Reiter's syndrome that shows pain in their joints, irritation of the eyes, and painful urination. It can last for months or years, and can lead to chronic arthritis that is difficult to treat.

*Shigella* and *E. coli* share many characteristics. The only reason they were originally separated into two different genera was that most of the strains of *Shigella* can cause bacillary dysentery while most of the strains of *E. coli* do not. This difference between *Shigella* and *E. coli* was revealed by the recent study of *S. flexneri* that has been the subject of extensive molecular pathogenesis and genetics research. A major difference is that *Shigella flexneri* (*S. flexneri*) has a large plasmid that contains many virulence associated genes that is not present in *E. coli*. *S. flexneri* serotype 5a virulence plasmid (Venkatesa *et al*., 2001), genome sequence of *S. flexneri serotype* 2a strain 301 (Jin *et al*., 2002), and *S. flexneri* serotype 2a strain 2457T (Wei *et al*., 2003) were completed relatively recently. Comparison of these genomes to that of *E. coli.* showed that *Shigella* has specific regions containing many genes that encode proteins with roles in virulence and the proteins of the type III secretion system to deliver these virulence proteins. Analysis of these virulence genes and their protein products will reveal the genetic basis for this pathogenic organism's distinctive lifestyle.

### 1.6 *Shigella* infection

*Shigella* enters the human body via the ingestion of contaminated food or water. Once *Shigella* has accessed the large intestine in the human body it is able to establish an infection. Once in the large intestine *Shigella* crosses the epithelial barrier via the microfold cell (M-cells, Figure 1.5). *Shigella* is released from M-cells and encounters resident macrophages that engulf and degrade foreign bodies. After bacterium-induced phagocytosis *Shigella* induces apoptosis of the macrophages in order to escape phagocytosis. Apoptosis induced by *Shigella* involves activation of enzyme caspase 1 by *Shigella* effector protein, IpaB (Hilbi, Moss *et al*. 1998), which causes maturation of two inflammatory cytokines, interleukin 1β and interleukin 18. The mature peptide of cytokines initiate intestinal inflammation characteristic of Shigellosis (Fantuzzi and Dinarello 1999).

After the release of *Shigella* from macrophages, the bacterum contacts the basolateral side of epithelial cell where it secretes effectors, such as VirA, IpaA, IpaB, IpaC and IcsB, by mean of the Type Three Secretion System (TTSS). The translocated effector, VirA is involved in rearrangement of the host cell cytoskeleton, resulting in membrane ruffling (Figure 1.6). This causes the uptake of *Shigella* into the epithelial cell by phagocytosis. When the bacteria invade the cell an autophagosome is formed which usually results in destruction of the invading bacteria. *Shigella* can escape from this digestive process before the formation of autophagolysosome. In the cell, the formation of autophagosome is initiated by Atg proteins. The interaction of the two proteins, VirG and Atg5, trigger the digestive process of invaded object in host cell. *Shigella* survives this digestive process in the autophagasome as IcsB competes with the binding site on VirG with Atg5. A recent study showed that the IcsB deletion mutant in *Shigella flexneri* was trapped by autophagy within the host cell (Ogawa *et al*., 2005, Ogawa, 2006). This result provided evidence that the effector IcsB is involved in escape of *Shigella* from autophagy. The phagocytic vesicle lyses immediately and the

bacteria are released into the cytosol where they multiply. *Shigella* then initiates actin polymerisation within the host cell, which leads to inter- and intracellular movement of *Shigella* and results in dissemination of the bacteria into neighbouring cells (Figure 1.5).

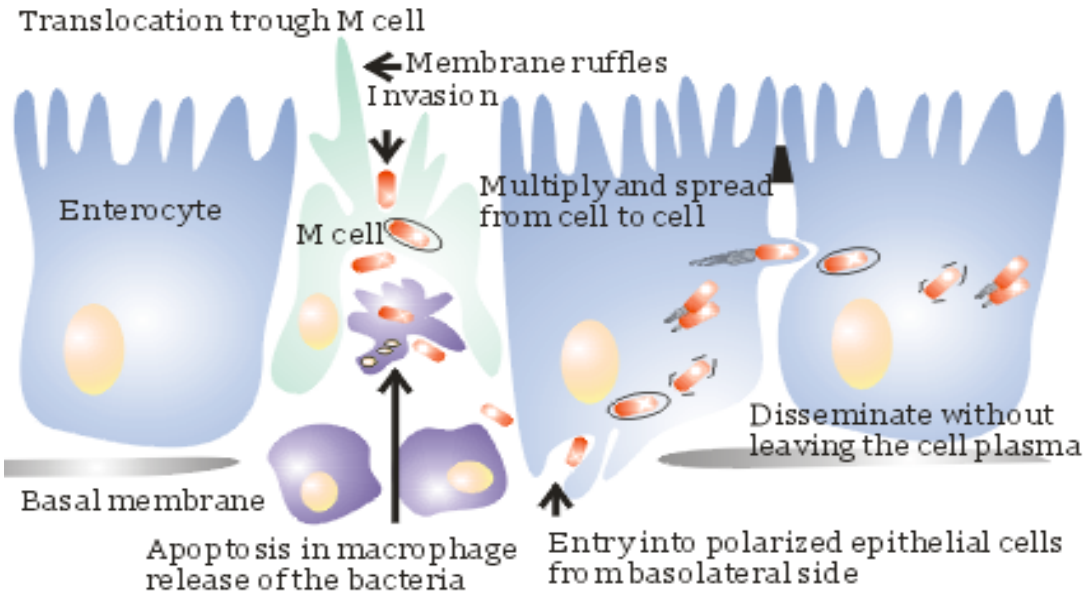



**Figure 1.5 *Shigella* invasion strategies. (Figure taken from Sansonetti *et al.*, 2001)**

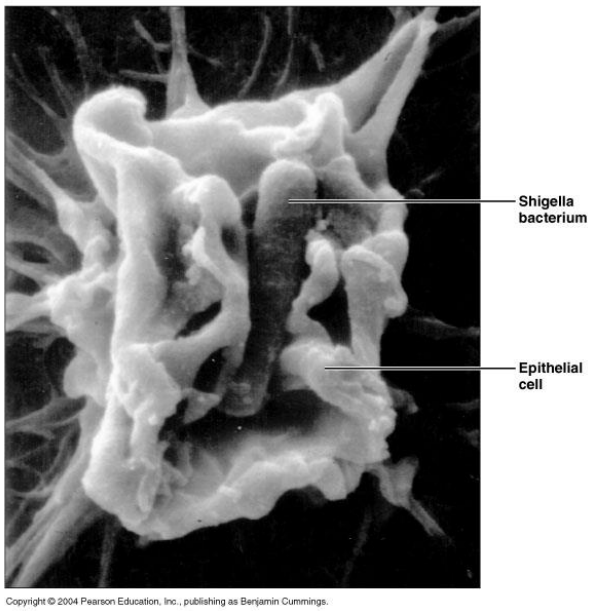



**Figure 1.6 Membrane ruffling on the basolateral side of epithelial cell by effector VirA that is inserted into the host by *Shigella*.(Picture was taken from Copyright©2004 Pearson Education Inc., publishing as Benjamin Cummings)**

Most of the genes required for invasion of epithelial cells and intercellular spread are located on the 220kb virulence plasmid, although full virulence also requires the expression of chromosomally encoded genes. A study confirmed that a 31kb region of the *S. flexneri* virulence plasmid is responsible for invasive characteristic of the bacterium as transformation of the 31kb region of the virulence plasmid conferred the invasive phenotype on *E. coli* K12 (Maurelli *et al*., 1985).

### 1.7 The *Shigella flexneri* effector VirG

A locus of *virG* was first identified on the *Shigella flexneri* virulence plasmid pWR100 by Bernardini (Bernaridini *et al*., 1989). VirG is expressed at one pole of the bacteria. It is a 120kDa surface-exposed outer membrane protein composed of 1,102 amino acid residues which form three domains, the N-terminal signal sequence (residues 1-52), α–domain (residues 53-758) and the β–core (residues 759-1102,Figure 1.7).

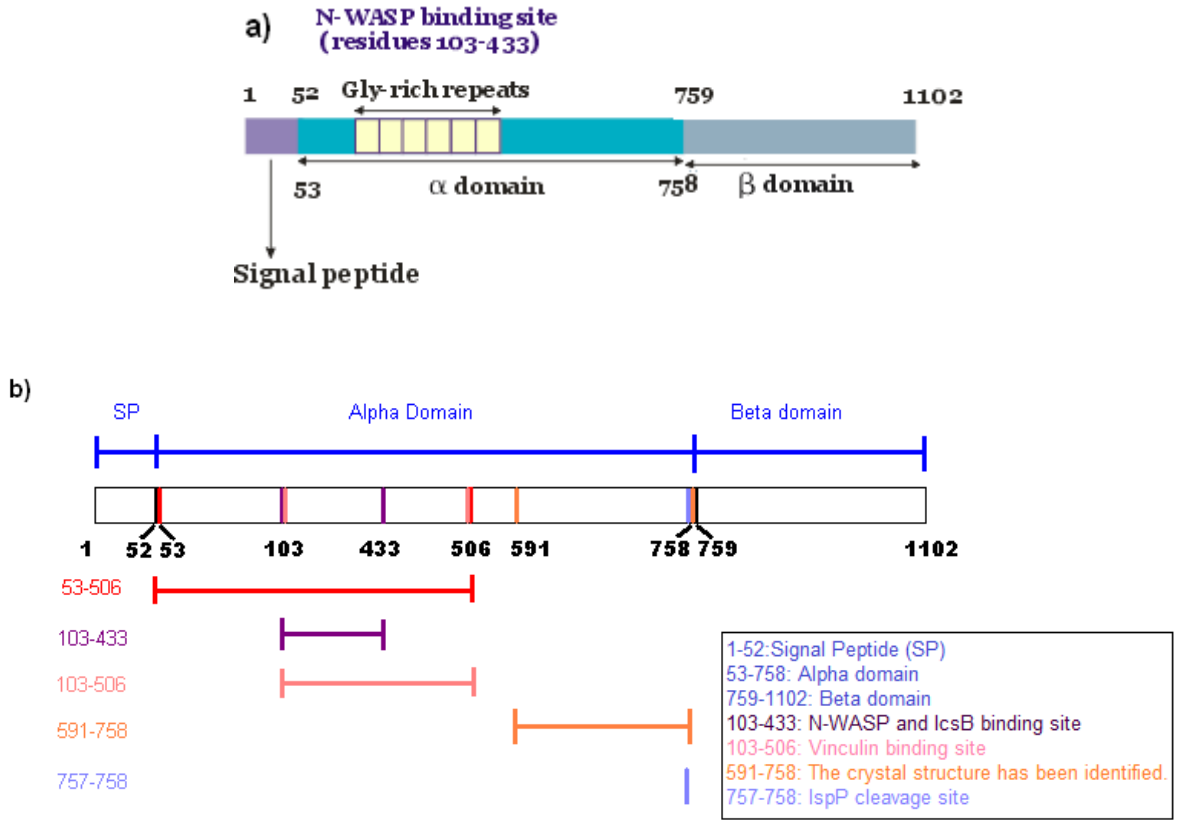



**Figure 1.7. a) Domain structure of VirG showing the α and β domains and the glycine-rich repeats within the α–domain. b) The binding sites for IcSb, vinculin and the cleavage site are also indicated.**

The α–domain of VirG is exposed on surface of the bacterial membrane. This domain is translocated by the β-domain which acts as autotransporter. The glycine rich repeats within the α–domain have multiple binding domains for N-WASP (Suzuki *et al*., 1998) and vinculin (Suzuki *et al*., 1996). The recruited N-WASP domains promote assembly of F-actin tails via recruitment of the ARP2/3 complex. There is another binding site on the α–domain. Residues 320-433 were determined as binding the *Shigella* effector IcsB and autophagy protein Atg5 by the study of VirG α–domain derivatives, which have deletion of sequences in α–domain (Ogawa *et al*., 2005). This region is responsible for escape of *Shigella* from autophagy as the IcsB mutant assay indicated that binding of autophagy protein Atg5 to VirG induced autophagy but induction of autophagy was inhibited by competition of IcsB binding to VirG.

The α–domain of VirG has an IcsP protease cleavage site at Arg758-Arg759, and cleavage releases a 95 kDa fragment of VirG. VirG and IcsP activities are regulated by the regulators VirF and VirB. This study shows that the mutation of VirB decreases the activity of IcsP and increases the proportion of uncleaved VirG on the surface of the bacterium. Therefore actin-based motility rises. The overexpression of IcsP results in more VirG being removed. The other study suggested that the VirG in the wild type bacterium is located across the entire surface of bacterium by moving VirG from the particular pole, where VirG are delivered, to the side part of bacteria on the surface of bacteria body. As VirG were inserted at the particular pole, the VirG concentration on the surface of bacterium seems higher at the pole than the other part of bacteria even if IcsP cleve the VirG anywhere on the surface of bacteria (Steinhauer *et al*., 1999). IcsP is not required for polar localization of VirG however; it contributes to maintaining a tight polar cap of VirG on the bacterial surface. (Wing *et al*., 2004)

The β-domain is embedded in the outer membrane and a structure model of the β–barrel channel consists of amphipathic antiparallel transmembrane β–strands with interspersed hairpin turns and loops (Suzuki *et al*., 1995). The β-domain is capable of translocation of passenger polypeptides in the same manner as secretion of the 80kDa α–domain from the periplasmic side of the outer membrane to the external side. Therefore the β-domain is demonstrated to be the autotransporter for the α–domain of VirG (Suzuki *et al*., 1995).

### 1.8 Localisation of VirG is essential for *Shigella* invasion system

Localisation of VirG is crucial for motility in infected cells and defects in the polar localisation of VirG reduce the efficiency of bacterial movement to adjacent cells.

The polar localization of VirG does not require signal peptide-mediated location or the β-domain as VirG-GFP fusion protein without signal peptide or the β-domain localized to the

pole. Absence of amino acid residues 58-103 and 507-729 decreases the polar localization of VirG. These results suggest that these two regions are required for the localization of VirG (Charles *et al*., 2001).

VirG crosses the cytoplasmic membrane using the Sec apparatus (Brandon *et al*., 2003). When VirG is translocated from the cytoplasm to the outer membrane, it is present transiently in the periplasm during transportation (Brandon and Goldberg, 2001). DegP acts as a chaperone at low temperature and a protease at high temperature. A reduction in the quantity of VirG on the surface of the bacteria was observed in DegP deletion mutant in both *Shigella* and *E.coli* without affecting the overall amount of VirG produced. These results suggest the folding of VirG in the periplasm or the rapid transit of VirG to the outer membrane may be facilitated by DegP (Purdy *et al*., 2002).

### 1.9 VirG mediated actin polymerisation during infection by *S.flexineri*

The actin-based motility of *Shigella* is dependent on VirG. The loss of VirG causes a reduction of actin assembly and reduction of spread of the bacterium to adjacent cells. The initiation of actin polymerisation is by binding to the Neural Wiskott-Aldrich syndrome protein (N-WASP).

N-WASP accumulates at the front of the actin tail and is capable of interaction with VirG both *in vitro* and *in vivo*. N-WASP is homologous to WASP and has several distinct domains. These are Enabled /VASP homology domain (EVH1) containing a pleckstrin homology (PH) domain and Calmodulin binding domain (IQ), proline rich region, GTPase binding domain (GBD) and Verprolin-homology region, a Cofilin-homology sequence and an Acidic terminal segment domain (VCA domain). These domains are required for the actin tail assembly.

Binding of N-WASP, VirG and Arp2/3 complex initiates actin polymerisation. The binding assay using N-WASP chimeras containing PH domain or IQ domain or GBD domain

suggested these domains were sufficient for interaction with VirG (Suzuki *et al*., 2002). After binding of N-WASP and VirG, Arp2/3 complex bind to N-WASP. The activation of N-WASP-Arp2/3 complex requires Cdc42 and this interaction leads to nucleation of actin and initiates actin polymerisation (Egile *et al*., 1999). This model of interaction of N-WASP by VirG is shown in Figure 1.8.

Profilin I also binds to VirG through binding to N-WASP. GST pull-down assay showed wild type recombinant Profilin I and wild type recombinant N-WASP binds to VirG. However, VirG failed to bind wild type Profilin I when N-WASP mutant (Proline-rich region deletion) failed to bind Profilin I and caused loss of the actin tail formation. Also wild type Profilin I alone was added to see if cell motility was elevated, however, only the addition of two molecules, wild type Profilin I and G-actin, can restore the original motility rate of VirG expressed *E.coli*. As actin nucleation was restored by both Arp2/3 complex and wild type or mutant N-WASP, this result suggests that profilin is associated with the elongation of actin filaments (Mimuro *et al*., 2000).

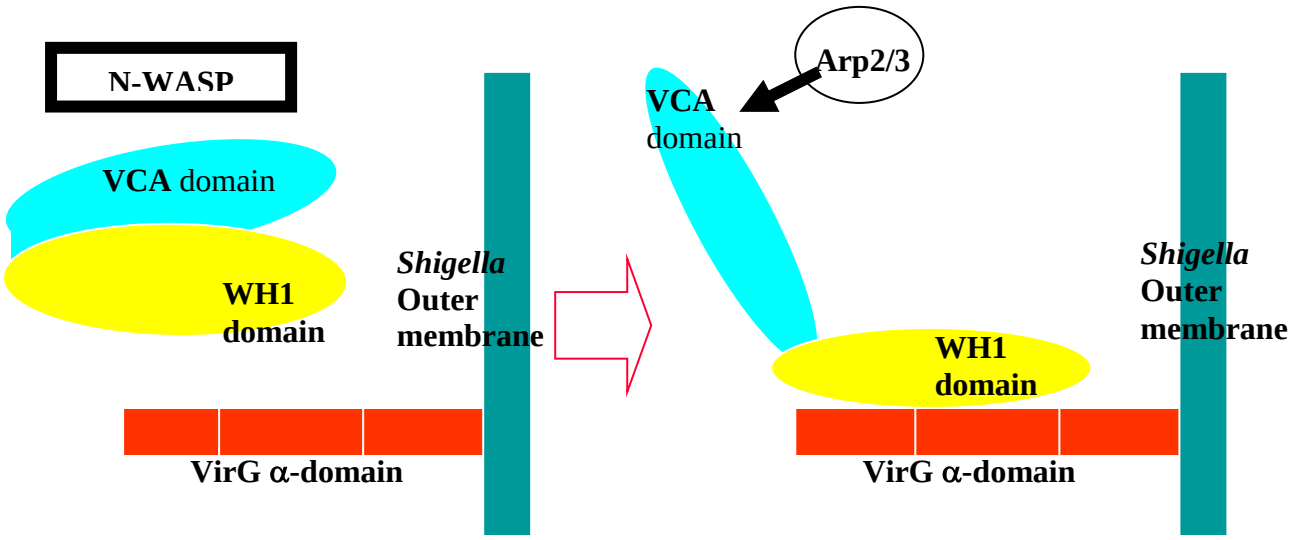



**Figure 1.8 Simple model of the interaction of VirG and N-WASP. N-WASP comprises VCA and WH1 domains and it is the WH1 domain of N-WASP that interacts with the alpha domain of VirG enabling the VCA domain to bind the Arp2/3 complex leading to actin polymerisation. The model shows how the binding of WH1 to the alpha domain might reveal the binding surface of the VCA domain via a conformational change in the N-WASP protein.**

N-WASP also has a WIP (WASP Interaction Protein) binding domain. WIP is a proline-rich protein of 503 amino acids that has homology in its amino-terminal end to the yeast polarity–development protein verprolin and contains two regions, actin-and profiling-binding motifs (Ramesh *et al*., 1997). Interaction of N-WASP EVH1 (residues 26-147) and WIP (residues 461-485) was confirmed by NMR study (Volkman *et al*., 2002).

### 1.10 Aims and Objectives

The aim of this study was to discover more about *Shigella flexneri* VirG by overproducing the protein and studying its properties. The objective was to overproduce the soluble (external) α-domain or a sub-domain of the α-domain of VirG, in *E. coil* for further studies. These further studies would include crystallization trials, various biophysical measurements, and interaction studies using a number of biochemical and biophysical techniques.

# Chapter 2

# Materials and Methods

## 2.1 Materials

### 2.1.1 Water

15.0MΩcm purity water (Distilled water = $dH_2O$) by produced by the ELGA Purelab water purifier was used to make up all buffers. Autoclaved $dH_2O$ was used for the molecular biology experiments.

### 2.1.2 Restriction Enzymes

| Restriction Enzymes | Restriction site |
|---|---|
| BamHI | 5'G/GATCC 3' |
| NcoI | 5'C/CATGG 3' |
| NotI | 5'GC/GGCCGC3' |
| XhoI | 5' C/TCGCG 3' |
| DpnI | 5'-$Gm^6A$/TC-3' |

**Table 2.1 List of restriction enzymes and restriction site that were used for cloning. "/" represents the cleavage site (New England BioLabs).**

### 2.1.3 Polymerase

Table 2.2 showed the list of the polymerase that were used for amplification of DNA in cloning and site directed mutagenesis.

| Polymerase |
|---|
| KOD Hot start DNA polymerase (Novagen) |
| Pfu turbo DNA polymerase (Stratagene) |
| Hot start Taq (QIAGEN) |
| DNA polymerase (New England BioLab) |
| Go Taq DNA polymerase (Promega) |

**Table 2.2 List of polymerases used for cloning and site directed mutagenesis.**

### 2.1.4 Primers

The table shows the primer sequence for different constructs. Sequences highlighted in red represent restriction enzyme binding sites.

| Construct | Gene encoded | Direction of primer | Primer sequence |
|---|---|---|---|
| Single DNA cloning | | | |
| pET32a VirG53 | VirG Residues Ala53-Val506 | Forward | 5'- CAT CCA TGG CTA CTC TTT CGG GTA CTC C-3' (NcoI) |
| | | Reverse | 5'- ATG GGA TCC TCA AAC AGT AAG TTC AGC GTT TTC TTT CAG -3' (BamHI) |
| pET32a VirG53 | VirG Residue Ala53-Val506 | Forward | 5'-CAT CCA TGG CTA CTC CTC TTT CGG GTA CTC AAG AAC TTC ATT TTT CAG AGG-3' (NcoI) |
| | | Reverse | 5'-ATG GGA TCC TCA AAC AGT AAG TTC AGC GTT TTC TTT CAG AT T GAA CTC TCC-3' (BamHI) |
| pET32a VirG103 | VirG Residue Arg103-Ala433 | Forward | 5'- CC ATG GCT AGA ACT AAG CTA CGG ATT AAC -3' (NcoI) |
| | | Reverse | 5'- CTA CTC GAG CTA CGC CCA GTT ATT TAT AAT AGC -3' (XhoI) |
| pET20b VirG53 | VirG Residues Ala53-Val506 | Forward | 5'-CCA TGG CTA CTC CTC TTT CGG G-3' (NcoI) |
| | | Reverse | 5'- CTA CTC GAG CTA AAC AGT AAG TTC AGC-3' (XhoI) |
| pET20b VirG103 | VirG Residue Arg103-Ala433 | Forward | 5'- CC ATG GCT AGA ACT AAG CTA CGG ATT AAC -3' (NcoI) |
| | | Reverse | 5'- CTA CTC GAG CTA CGC CCA GTT ATT TAT AAT AGC -3' (XhoI) |
| Mutagenesis | | | |
| pGEM2T VirG53 C130S | VirG Residue Ala53-Val506 | Forward | 5'-TGA TAA TAA TGA TGG TAA TAG TGG CGG TAA TGG TGG TG-3' (Point mutation) |
| | | Reverse | 5'-CAC CAC CAT TAC CGC CAC TAC TAT TAC CAT CAT TAT TAT CA-3' (Point mutation) |

**Table 2.3 List of primers used for DNA amplification**

### 2.1.5 Vector

The Table below (Table 2.4) shows the list of vectors that were used in this study. Most of the vectors shown below were protein expression vectors, apart from pGEM®-T Easy (Promega) which is a cloning vector.

| Vector | Antibiotics | Tags |
|---|---|---|
| pGEM®-T Easy (Promega) | Ampicillin | -- |
| pGEX2T(GE Healthcare) | Ampicillin | N-terminus glutathione S-transferase (GST) tag |
| pET14b (Novagen) | Ampicillin | N-terminus 6x Histidine tag |
| pET20b (Novagen) | Ampicillin | C-terminus 6x Histidine tag. |
| pET32a (Novagen) | Ampicillin | C and N-terminus 6x Histidine tag. |
| pACYC184 * | Chloramphenicol Tetracycline | No Tag |

***** This vector was kindly donated by a previous lab member

**Table 2.4 List of Vectors used for cloning and protein expression.**

### 2.1.6 Medium

Luria Bertani medium (LB medium) was used for bacterial cell growth. LB liquid media and LB-agar medium were made up to 1 liter of distilled water (see Tables 2.5 and 2.6 for details) and sterilised by autoclave. All chemical were from Fisher Scientific.

| Amount | Chemicals |
|---|---|
| 5g | NaCl |
| 10g | Tryptone |
| 10g | Yeast extract |

**Table 2.5 Component of LB medium (per 1 liter)**

| Amount | Chemicala |
|---|---|
| 5g | NaCl |
| 10g | Tryptone |
| 10g | Yeast extract |
| 15g | Agar |

**Table 2.6 Components of 1.5% LB-agar medium (per 1 liter)**

### 2.1.7 Competent cells

The competent cells segregate into two groups (Table 2.7), the first for plasmid applications and the second group for protein production.

| Application | Competent Cell | Selection |
|---|---|---|
| DNA replication | DH 5 alpha (Invitrogen)<br>Alpha select (Bioline)<br>XL1-Blue (Stratagene) | Ampicillin<br>Ampicillin<br>Tetracycline |
| Protein Production | BL21(DE3) (Bioline)<br>BL21(DE3)pLysS (Bioline)<br>Origami2 (DE3) pLysS<br>Rossetta-gami2(DE3)pLysS | Ampicillin<br>Ampicillin<br>Chloramphenicol/Tetracycline/Kanamicin<br>Chloramphenicol/Tetracycline |

**Table 2.7 List of competent cells for cloning and protein expression and the antibiotics used for their selection.**

### 2.1.8 Antibiotics

Antibotics were used for selection of competent cells. Apart from Tetracycline and Chloramphenicol, antibiotics were made up with distilled water (15.0MΩcm purity) and filter sterilised using aseptic technique. Stocks were kept at -20$^{o}$C for long term storage. All antibiotics were purchased from Fisher Scientific.

| Antibiotics | Stock concentration (mg/ml) | Working concentration (µg/ml) |
|---|---|---|
| Ampicillin | **100** | **100** |
| Chloramphenicol | **34** | **35** |
| Kanamycin | **30** | **30** |
| Tetracycline | **12.5** | **12.5** |

.

**Table 2.8 List of antibiotics used for selection of competent cells.**

### 2.1.9 Constructs

A list shows the different constructs that were produced and used in this study (Table 2.9)**.** pGEX2TVirG53 construct was provided by Prof. Chihiro Sasakawa and colleagues (University of Tokyo, Japan).

| Plasmid name | Inserts | Vector | Tag | Description of vector | Location of expressed protein |
|---|---|---|---|---|---|
| pGEMTEasyVirG53 | VirG T53-V506 | pGEMTEasy | non | Cloning vector | - |
| pGEMTEasyVirG103 | VirG R103-A433 | pGEMTEasy | non | Cloning vector | - |
| pET32aVirG53 | VirG T53-V506 | pET32a | His-tagged at N-terminus | Protein Expression vector | Cytoplasm |
| pET32aVirG103 | VirG R103-A433 | pET32a | His-tagged at N-terminus | Protein Expression vector | Cytoplasm |
| pET20bVirG53 | VirG T53-V506 | pET20b | His-tagged at C-terminus and Signal peptide at N-terminus | Protein Expression vector | Periplasm |
| pET20bVirG103 | VirG R103-A433 | pET20b | His-tagged at C-terminus and Signal peptide at N-terminus | Protein Expression vector | Periplasm |
| pGEX2T VirG53 (Provided by Prof. Chihiro Sasakawa) | VirG T53-V506 | pGEX2T | GST tagged at N-terminus | Protein Expression vector | Cytoplasm |
| pGEX2TVirG53C130S | VirG T53-V506 Cysteine(C) was mutated to Serine(S) | pGEX2T | GST tagged at N-terminus | Protein Expression vector | Cytoplasm |

**Table 2.9 List of plasmids that were constructed in this study**

### 2.1.10 Columns

Columns used are listed below in Table 2.10

| Separation method | Columns | Material | Molecular weights separation range |
|---|---|---|---|
| Affinity | 5ml prepacked HiTrap chelating column | Chelating Sepharose™ High Performance | -- |
| | Empty column | Glutathione Sepharose | -- |
| Size exclusion | Superdex 200 10/300 GL | Spherical composite of cross-linked agarose and dextran | between 10,000 and 600, 000 |

**Table 2.10 List of columns used for protein purification. All columns were purchased from GE Healthcare.**

#### 2.1.10.1 Chelating Sepharose

Immobilized-metal affinity chromatography (IMAC) is used for separation of proteins by using the characteristic of amino acid residues such as histidine and cysteine that form complexes with the chelated metals. Chelating Sepharose is made by coupling a metal chelate forming ligand to Sepharose (Figure 2.1). Chelating Sepharose immobilises the metal by chelation. Nickel was used in this study. His-tagged proteins were trapped on the chelated Nickel in the column. Any proteins which do not have His-tag will tend to pass through the column. The binding capacity of 5ml HiTrap Chelating column is 60mg/column.

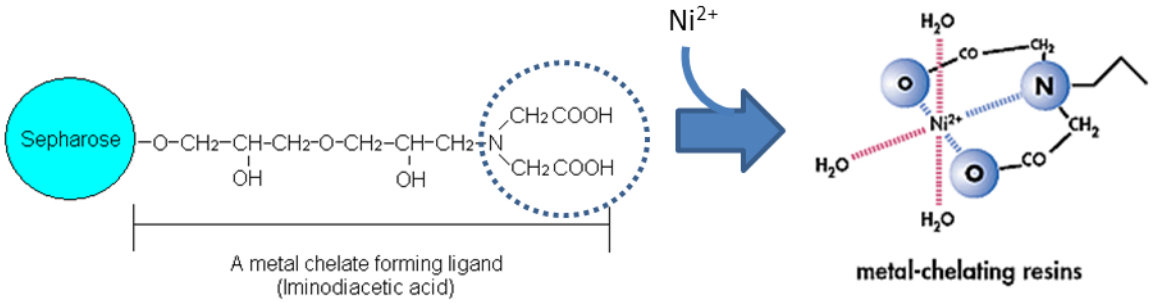



**Figure 2.1 Structure of Chelating Sepharose High Performance and the complex of Ni ion and ligand. Image of Chelating Sepharose structure were obtained from GE Healthcare and metal-chelating resin structure from QIAGEN.**

### 2.1.10.2 Glutathione Sepharose 4 Fast flow

Glutathione Sepharose is made of the coupling of highly cross linked agarose and glutathione that is linked to agarose by 10-carbon linker arm. GST has an affinity toward to glutathione. The total binding capacity is approximately 10 mg recombinant GST protein per ml of beads. The Glutahione column is not recommended for use with strong denaturing agents, such as urea, as they reduce the binding capacity of GST tag due to its denaturation.

### 2.1.10.3 Spherical composite of cross-linked agarose and dextran

Unlike affinity column, size exclusion columns like Superdex 75 10/300 or 200 10/300, separate molecules according to the size. A bead of Superdex is composed of agarose and dextran that produce various pores for the separation. The applications of size exclusion column are, for example, buffer exchange, a polishing step in purification, determining molecular weight or separation of monomer and dimer. Superdex 200 separates the molecules between 10,000 and 600,000 Da. Therefore, any molecules more than 600,000 Da will be eluted in the void fraction.

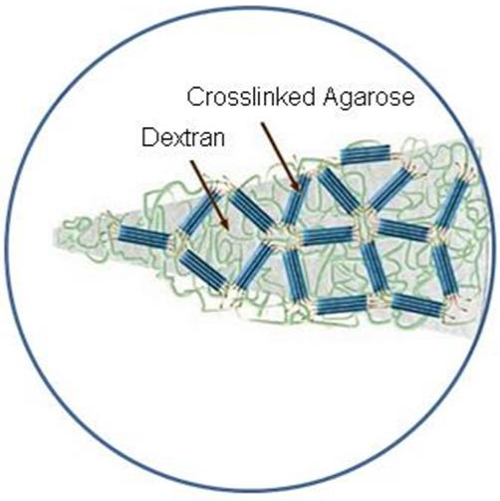



**Figure 2.2 The structure of Superdex beads. The circle represents a bead. The dextran chains are covalently linked to a highly cross-linked agarose matrix (GE Healthcare).**

### 2.1.10.4 Buffer for protein purification column

Buffers were made up with $dH_2O$ (ELGA) and adjusted by concentrated HCl to the (Table 2.11) appropriate pH. Phosphate buffered saline (PBS) tablets were purchased from Fisher Scientific. One tablet was dissolved in 200mL $dH_2O$ and produce 0.01M phosphate buffer, 0.0027M KCl, and 0.137M NaCl.

| HiTrap chelating column | | | Size exclusion |
|---|---|---|---|
| Purified sample | Binding buffer | Elution buffer | |
| His-tagged VirG | 20mM TrisHCl pH7.5, 20mM Immidazole, 500mM NaCl. | 20mM TrisHCl pH7.5, 500mM Immidazole, 500mM NaCl. | 20mM TrisHCl pH7.5, 500mM NaCl. |
| His-tagged VirG | 20mM TrisHCl pH7.5, 500mM NaCl, 20mM Immidazole, 0.1mM n-Dodecyl β-D-maltoside | 20mM TrisHCl pH7.5, 500mM NaCl, 500mM Immidazole. 0.1mM n-Dodecyl β-D-maltoside | 20mM TrisHCl pH7.5, 500mM NaCl, 0.1mM *n*-Dodecyl β-D-maltoside |
| His-tagged VirG | 20mM TrisHCl pH8.0, 10mM Imidazole, 500mM NaCl | 20mM TrisHCl pH8.0, 500mM Imidazole, 500mM NaCl | 20mM TrisHCl pH8.0, 500mM NaCl |
| Glutathione sepharose column | | | Size exclusion |
| Purified sample | Binding buffer | Elution buffer | |
| GST-tagged VirG | Phosphate Buffered Saline (PBS) | Reduced glutathione in PBS | 20mM TrisHCl pH8.0, 500mM NaCl,10% glycerol |
| GST-tagged VirG | Phosphate Buffered Saline (PBS) | Reduced glutathione in PBS | 20mM Tris-HCl pH 8.0, 500mM NaCl, 10% Glycerol, 10mM DTT. |

**Table 2.11 List of buffers used with the column**

## 2.2 Methods

### 2.2.1 Polymerase chain reaction

The polymerase chain reaction (PCR) was used to amplify the regions of the gene of interest. The reaction requires primers that recognise the gene or part of gene of interest on the template DNA. Primers were designed with 5'-and 3'- restriction enzyme sites (Table 2.1) which were be used for insertion of the DNA fragment to the cloning or expression vectors. DNA amplification was achieved using primers, deoxynucleotide triphosphates (dNTPs), template DNA and DNA polymerase (Table 2.2) and the buffer that was provided.

The chemicals and DNA were added to 0.5ml sterile PCR tubes. The components of PCR mixture are shown in Table 2.12. The mixture was located in the PCR thermocycler to facilitate the PCR. The program of PCR cycle is shown in Table 2.13. Condition of PCR reaction mixture and PCR cycle for amplify *virG* (residues 52-506) for producing pET32a construct were shown in Chapter 4.

| Components of reaction mixture for PCR | |
|---|---|
| Volume | Components |
| 10µl | 10× reaction buffer |
| 2µl | dsDNA template pGEX2TVirG53 |
| 2.5µl | oligonucleotide primer Forward |
| 2.5µl | oligonucleotide primer Reverse |
| 1.6µl | 10mM dNTP mix |
| 2.0µl | Pfu DNA polymerase |
| 79.4µl | $ddH_2O$ to a final volume of 100µl |
| Total 100µl | |

**Table 2.12 The PCR reaction mixture**

**a)**

| Segment | Cycles | Temperature | Time | Steps |
|---|---|---|---|---|
| 1 | 1 | 95°C | 45seconds | Activation |
| 2 | 30 | 95°C | 45 seconds | Denature |
| | | 60.3°C | 45 seconds | Annealing |
| | | 72°C | 2 minute | Extention |
| 3 | 1 | 72°C | 10 minute | Final extension |

**b)**

| Segment | Cycles | Temperature | Time | Steps |
|---|---|---|---|---|
| 1 | 1 | 95°C | 45 second | Activation |
| 2 | 30 | 95°C | 45 second | Denature |
| | | 58.4°C | 45 second | Annealing |
| | | 72°C | 2 minute 30 sec | Extention |
| 3 | 1 | 72°C | 10 minute | Final extension |

**Table 2.13 PCR cycle a) VirG Arg103-Ala433 and b) VirG Ala53-Val506**

The PCR products were analyzed by agarose gel electrophoresis (see Agarose gel electrophoresis later in this Chapter). PCR products were purified using a gel extraction kit or PCR purification kit (QIAGEN, PEQlab) prior to enzyme digestion.

### 2.2.2 Enzyme digestion of PCR products and vectors

The amplified PCR product was directly cloned into the appropriate vector (Table 2.4). The PCR product and vector were treated with appropriate restriction enzymes (Table 2.1) to create unique restriction sites at both sides of gene before the ligation. The digested PCR products and vector were purified by PCR purification kit and Gel purification kit (QIAGEN, PEQlab) respectively prior to the ligation. Ligated circular DNA was transfered to appropriate competent cells (Table 2.7) depending on the application, either DNA replication or protein production, for transformation of cells.

The conditions used for double digestion and ligation are shown below.

All chemicals and DNA material (Table 2.14) were mixed in the sterile 1.5ml microcentrifuge tubes. The digestion mixture was left to incubate for one hour.

| Double Digestion | |
|---|---|
| Amount | Reagent and DNA sample |
| 10µl | vector or PCR protduct of gene |
| 2 µl | 10x buffer |
| 1 µl* | restriction enzyme #1 |
| 1 µl* | restriction enzyme #2 |
| 0.2 µl | mM BSA |
| 5.8 µl | Sterile ddH2O |
| 20µl | Total volume of reaction mixture |

* 10 units

**Table 2.14 Double digestion mixture**

After the completion of enzyme digestion of PCR products and vector, the reaction mixture were purified by agarose gel electropshesis. Gel extraction kit (QIAGEN) was used to extract DNA fragment from agarose gel.

### 2.2.3 Ligation of DNA fragments and vector

Clean digested PCR product and vector were mixed with Takara ligation mixture (TAKARA) that contains T4 ligase. Details of the ligation mixture are shown in the table below. The mixture was left to incubate at room temperature for 30 minutes then the mixture was incubated at 4ºC for 2-3 days.

| Ligation | |
|---|---|
| Amount | Reagent and DNA sample |
| 2 µl | digested PCR product |
| 6 µl | digested vector |
| 8 µl | Takara ligation mixture |
| Total 16µl | |

**Table2.15 Ligation mixture**

.

| Ligation for using pGEM-T Easy vector | |
|---|---|
| Volume | Components |
| 5µl | 2xRapid ligation buffer for T4DNA Ligase |
| 1 µl | pGEM-T easy vector(50ng) |
| 2 µl | PCR products |
| 2 µl | T4 DNA Ligase (U) |
| 0 µl | $ddH_2O$ |
| Total 10µl | |

**Table2.16 Ligaton mixture for pGEM-T easy**

### 2.2.4 Site-Directed Mutagenesis

QuikChange®Site-Directed Mutagenesis Kit (Promega) was used for mutageneisis of the *virG* alpha domain. All PCR reaction components (Table 2.17) were mixed in a sterile 0.5 ml PCR tube. The thermalcycler was programmed using recommended cycling parameter for PfuTurbo DNA polymerase (Table 2.18).

| Components of reaction mixture for PCR | |
|---|---|
| Volume | Components |
| 5 µl | 10× reaction buffer |
| 4 µl | 2.5ng/ µl pGEX2TVirG53 dsDNA template |
| 1 µl (125 ng) | oligonucleotide primer #1 |
| 1 µl (125 ng) | oligonucleotide primer #2 |
| 1 µl | 10 mM dNTP mix |
| 37 µl | $ddH_2O$ to a final volume of 49 µl |
| Then add 1 µl | PfuTurboDNA polymerase (2.5 U/µl) |
| Total 50 µl | |

**Table 2.17 PCR reaction mixture**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 95°C | 30 seconds |
| 2 | 16 | 95°C | 30 seconds |
| | | 55°C | 1 minute |
| | | 68°C | 7 minute * |

*1minute/kb: The product was 6.1kbp therefore set up for 7 minutes.

**Table 2.18 Cycling program for the mutagenesis method**

Once DNA amplification was completed, the amplification product was cleaved by adding 1 µl DpnI restriction enzyme. The mixture of DpnI and DNA product was gently mixed by pipetting and centrifuged at 14,000g for 1 mintue in a microcentrifuge (Eppendorf). The reaction mixture was then incubated at 37°C for 1 hour to allow complete cleavage of the methylated DNA. 1 µl of digested DNA was transfere to the XL1-Blue competent cells for transformation. 100 µl of 2% of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and 100 µl of 10 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) were spread on LB agar plate prior to plating the transformants for white and blue colour selection. X-gal was prepared in dimethylformamide (DMF). IPTG was prepared in $dH_2O$ and then filter sterilized.

### 2.2.5 Agarose Gel Electrophoresis for the Separation of DNA Fragments

Agarose gel electrophoresis has proven to be an efficient and effective way of separating digested plasmid and PCR products from the low molecular mass fragments. Molecular sieving is determined by the size of pores generated by the bundles of agarose in the gel matrix. In general, the higher the concentration of agarose the smaller the resulting pore size. Therefore low percentage gels are used for the separation of large DNA fragments. Table 2.19 showed the separation size of DNA fragment according to the concentration of agarose gel.

#### 2.2.5.1 Agarose gel

An agarose gel was made up with running Tris Acetate EDTA (TAE) Buffer (40 mM Tris-acetate, 1 mM EDTA). The agarose/buffer mixture was melted by microwave. 5µl of liquid ethidium bromide stock (10 mg/ml) were added to 100ml melted the agarose/buffer mixture to make final concentration of 0.5 µg/ml. The molten agarose was poured into the gel mold to

make gel for DNA separation by gel electrophoresis. The agarose was allowed to set at room temperature and then the comb was removed. The gels can be stored at 4 °C until use.

| Percentage of agarose gel | Agarose | TAE | Ethidium bromide final concentriaton | Separation size |
|---|---|---|---|---|
| 0.7% | 0.8g | 100ml | 0.5 µg/ml | 800bp-10kb |
| 0.9% | 1g | 100ml | 0.5 µg/ml | 500bp-7kb |
| 1.2% | 1.2g | 100ml | 0.5 µg/ml | 400bp-6kb |
| 1.5% | 1.5g | 100ml | 0.5 µg/ml | 200bp-3kb |

**Table 2.19 Components of agarose gel for different agarose concentrations.**

### 2.2.5.2 Running agarose gel by gel electrophoresis

DNA samples were prepared with 6x loading dye (New England BioLabs). DNA sample mixtures (10-30 µl) were loaded to the each well on the agarose gel. DNA fragments were separation by gel electrophoresis apparatus (Biorad) at 100 V for 1 hour. The gel was exposed to UV light and an image of a gel was taken with a gel documentation system. DNA bands showed up as orange

fluorescent bands on the gel. The DNA ladder was run alongside of DNA sample to determine the sizes of sample bands.

### 2.2.6 Transformation of competent cells

100 µl of competent cells (Table 2.7) was transformed with 1 µl of plasmid (Table 2.9) and incubated for 30 min on ice followed by heat-shock (see the table 2.20 for the detail of heat-shock conditions for each competent cell). 800 µl LB liquid media was added to the cell suspension which was then incubated for 1 hour at 37°C. Transformed cell were centrifuged at 130,000 rpm (Eppendorf) at room temperature for 1 minute and pellets were resuspended with 200µl LB liquid media. Resuspended cells were plated on LB-plates containing appropriate antibiotics (Table 2.21). Plates were then incubated at 37°C overnight.

| Competent Cell | Heat-shock temperature (°C) | Duration (sec) |
|---|---|---|
| DH 5 alpha (Invitrogen) | 42 | 45 |
| Alpha select (Bioline) | 42 | 45 |
| XL1-Blue (Stratagene) | 42 | 45 |
| BL21(DE3) (Bioline) | 42 | 30 |
| BL21(DE3) plysS (Bioline) | 42 | 30 |
| Rosetta-gami™2 (DE3) pLysS | 42 | 30 |

**Table 2.20 The condition of heat-shock treatment for transformation of competent cells**

### 2.2.7 Plasmid DNA purification

After transformation for DNA amplification, each single colony was inoculated into 5ml of LB media with appropriate antibiotics (Table 2.21). The cell culture was incubated at 37ºC for overnight on the shaker. The overnight cell culture was then centrifuged to separate the cell culture from medium. The collected pellet was used to purify DNA by using a DNA purification kit (QIAGEN, PEQlab). Purification was carried out by following the procedure in the kits. Purified DNA was extracted with 50µl of sterilised $dH_2O$ and stored at -20ºC. Alternatively pelleted cell culture was frozen at -20ºC for long term storage. These plasmids were used for transformation.

### 2.2.8 Protein production

After the transformation of cells for protein production, a single colony was selected and inoculated in 10ml of LB liquid medium which had appropriate antibiotics (Table 2.21) for overnight at 180rpm at 37ºC. 750ml of overnight culture were mixed with 250ml of sterilised glycerol to make glycerol stocks that were kept in -80ºC. 10ml of overnight culture was transferred to 1L of LB liquid media with appropriate antibiotics (Table 2.21). 1L of cell culture was grown at 37ºC for 3-4 hours until the optical density was approximately ~0.6 at 600nm. Final concentration of 0.5~1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was

added to the cell culture for induction of recombinant protein production. The cells were spun down at 8,000 rpm for 15minutes at 4$^{o}$C using a Avanti centrifuge (Beckman using a JA10 rotor) after further 3 hours or overnight of incubation at 18-37$^{o}$C. The pellet was resuspended in 10-30ml of binding buffer (Table 2.24). Pellets were immediately used for the sonication step or kept in -20$^{o}$C for storage for future use.

| Vector / Cell | pET32a | pET20b | pET14b |
|---|---|---|---|
| *E.coli* BL21 DE3 | Amp (100µg/ml)* | Amp (100µg/ml)* | Amp (100µg/ml)* |
| *E.coli* BL21 DE3 plysS | Cam (25µg/ml)*<br>Amp (100µg/ml)* | Cam (25µg/ml)*<br>Amp (100µg/ml)* | Cam (25µg/ml)*<br>Amp (100µg/ml)* |
| *E.coli* Rosetta-gami 2 DE3 plysS | Cam (25µg/ml)*<br>Tet (12.5µg/ml)*<br>Amp (100µg/ml)* | Cam (25µg/ml)*<br>Tet (12.5µg/ml)*<br>Amp (100µg/ml)* | Cam (25µg/ml)*<br>Tet (12.5µg/ml)*<br>Amp (100µg/ml)* |

Amp: Ampicillin, Cam: Chloramphenicol, Tet: Tetracycline *(Final concentration)

**Table 2.21 Antibiotics used for selection**

### 2.2.9 Protein purification

Protein purification protocols contained one or more chromatographic steps. The basic procedure in chromatography is to pass the solution containing the protein through a column packed with various beads. The various proteins interact differently with the column material, and can thus be separated by size or affinity of the protein towards the beads. Most proteins are detected by their absorbance at 280 nm or the Bio-Rad Protein Assay that is based on the method of Bradford when they are coming off the column.

#### 2.2.9.1 Column preparation for affinity column and size exclusion column

The table shows the conditions used to run columns. (See Table 2.11 for list of the Buffers for each column.)

| Column | Condition | |
|---|---|---|
| | Flow rate | Fraction (ml) |
| 5ml prepacked HisTrap chelating column | 1-5ml per min | 5ml |
| Superdex 200 10/300GL | 0.5-1ml per min | 1ml |
| Superdex 75 10/300GL | 0.5-1ml per min | 1ml |

**Table 2.22 FPLC parameters used for protein purification**

### 2.2.9.2 Sample preparation and purification by HiTrap chelating column

Hexahistidine ($His_6$) is a popular tag for recombinant protein due to its small size and affinity for $Ni^{2+}$ ions. Unlike other commonly used tags, Glutathione s Sepharose, $His_6$ tag can be used under the denaturing conditions.

HiTrap Chelating column (GE Healthcare) was used for His-Tagged protein purification. HiTrap Chelating column was regenerated with 2.5 ml of 0.1 M Nickel sulphate salts ($NiSO_4$) in distilled water and then washed with 50ml of $dH_2O$ prior to loading supernatant. The regenerated columns were then equilibrated with binding buffer (Table 2.22).

The pellets were kept on ice to prevent degradation of the protein by heat. 20ml of induced cell culture pellet were sonicated (VibraCell) for 30 sec and rest for 1 minute. The sonication interval were repeated for 17 times and then centrifuged by 18,000 rpm for 30 minutes at $4^oC$ by Avanti centrifuge (Beckman). The supernatant and pellet were applied to 12% SDS-PAGE gel to determine the extent of expression of recombinant protein.

20ml of supernatant was loaded on to pre-packed 5ml HisTrap column. Following purification was performed on an AKTA system (GE Healthcare). The column was washed with 25ml binding buffer (Table 2.11) then eluted against a gradient from 0%-100% of elution buffer (Table 2.11). 5ml of sample were collected per faction. Each fractions of protein sample were checked by 12% SDSPAGE gel. High concentration of protein sample

were pooled and concentrated by MWCO 10,000 vivaspin (Generon) concentrator for further purification.

#### 2.2.9.3 Sample preparation and purification by Glutathione Sepharose column

Glutathione Sepharose column was used for purification of Glutatione S-transferase (GST)-tagged protein. 3ml of Glutathione Sepharose 4 fast flow bead (GE Healthcare) were packed into a clean plastic column (PD-10, GE Healthcare) and washed with 30ml of PBS.

The pellets were kept on ice to prevent degradation of the protein by heat. Induced cell culture pellet were resuspended with 20ml of PBS and then sonicated by sonicatior (VibraCell) for 30 sec and rest for 1 minute. The sonication were repeated for 17 times and then centrifuged by 18,000 rpm for 30 minutes at 4$^{o}$C by Avanti centrifuge (Beckman). The supernatant and pellet were applied to 12% SDS-PAGE gel to determine the extent of expression of recombinant protein.

10ml of supernatant were loaded into the column and incubated at 4 $^{o}$C for one hour. After the incubation, applied supernatant were passed through the column. Incubation and elution were repeated until all supernatant were loaded to the column. The column was then washed with 30ml of PBS in order to wash out contaminants. The protein concentrations of flow through fraction were checked by BioRad protein assay solution. Mixture of 10µl (10 units) of thrombin (1u/µl) and 3ml of PBS were added to the glutathione Sepharose column to cleave the GST tag from fusion protein. The column was incubated at 4$^{o}$C overnight. Protein sample were eluted at 4$^{o}$C. An extra 3ml PBS were added and the sample was eluted after one hour of incubation at 4$^{o}$C. Each fraction of the protein sample was checked by 12% SDS-PAGE. Eluted fractions were pooled and concentrated by MWCO 10,000 vivaspin (Generon) concentrator for further purification.

### 2.2.9.4 Purification by size exclusion chromatography

Unlike affinity column, proteins do not bind to the beads in a size exclusion column. The size exclusion column separates molecules according to differences in size as they pass through a packed gel filtration medium in a column. Superdex 200 10/300 GL and Superdex 75 10/300 GL column (GE Healthcare) were used in this study. Each size exclusion column has different fractionation range (see table 2.22).

Size exclusion columns were washed with 36ml (1.5 column volumes) of buffer (Table 2.11) prior to loading the sample.

Concentrated protein was centrifuged at 13,000rpm (Eppendorf) for 10 minutes at 4ºC to separate any precipitated material in the solution. Up to 1ml of supernatant was loaded to the column. Protein purification was performed by AKTA (GE Healthcare). Separation was run for 1.5 column volumes. Table 2.22 shows the purification running conditions for the AKTA. Each fraction of protein was checked by 12% SDS-PAGE. Eluted fractions containing a high concentration of protein of interest were pooled and concentrated by MWCO 10,000 vivaspin (Generon) concentrator for further experiments.

### 2.2.9.5 Removal of the tags from recombinant protein with thrombin protease

The tags were removed by incubating tagged recombinant protein with thrombin protease (Novagen) in a solution overnight at 4°C. Cleaved protein was purified by running the sample through HisTrap or Glutathione Sepharose column using AKTA.

### 2.2.10 SDS-PAGE Gel Electrophoresis for protein

#### 2.2.10.1 SDS-PAGE gel

SDS-PAGE was used to separate the protein by size. The SDS-PAGE was made using the various components listed in the Table 2.23. The resolving gel was poured between thin glass plates by using BioRad SDS-PAGE gel casting system. A layer of ethanol was added to top of dried resolving gel to retain straight surface on the gel. When the gel set, the ethanol was discarded and the spacer between the glasses rinsed with $dH_2O$. The stacking gel solution then was then added and the desired comb (10-15 wells) was inserted between the spacers at the top of the set gel to create the well on the gel.

| Gel Type / Solution Components | 5% stacking gel (ml) | 12% resolving gel (ml) | 15% resolving gel (ml) |
|---|---|---|---|
| dH2O | 2.7 | 1.5 | 1.1 |
| 1.5M Tris-HCl(pH8.8) | - | 1.3 | 1.3 |
| 1.0M Tris-HCl(pH6.8) | 0.5 | - | - |
| 10% SDS | 0.04 | 0.05 | 0.05 |
| 10% Ammonium persulfate | 0.04 | 0.05 | 0.05 |
| 30% acrylamide | 0.67 | 2.0 | 2.5 |
| TEMED | 0.004 | 0.002 | 0.002 |

**Table 2.23 List of solution components and volumes for the stacking gel and resolving gel used to produce the SDS-PAGE gel.**

#### 2.2.10.2 Running SDS-PAGE gel

30µl of protein sample and 30µl of 1x SDS loading dye were mixed in the 1.5ml microcentrifuge tube and then boiled at 100$^{o}$C in the heat block for 10 minutes. The mixtures were spun down by microcentrifuge (1 mins, 13,000 rpm, Eppendorf) and 30µl of sample were loaded on to the 12% SDS PAGE gel. Protein markers were loaded alongside samples to guide the determination of the size of protein bands.

The Mini-PROTEAN 3 system (Biorad) was used for running SDS PAGE gels. The prepared SDS PAGE gel was placed in the electrophoresis tank and run at 100V-200V for 30-60 minutes until the dye front reach to the bottom of the gel. The gel was then stained using staining dye (see the Table 2.24 for details) for 1 minite by microwave then placed on the shaker (scientific) for another 15 minutes at 70 rpm. After staining, the gel was destained with destaining solution for 1 minute in a microwave (760 V) and then placed on the shaker until the protein bands were visible on the gel.

| 1x SDS running buffer | | 1x SDS loading dye | | Staining dye | | Destaining solution | |
|---|---|---|---|---|---|---|---|
| 30.2g | Tris-HCl pH8.3 | 100mM | Tris | 1g | Comassie blue | 30% | Acetic acid |
| 188g | Glycine | 200mM | DTT | 30%(v/v) | Methanol | 40% | Methanol |
| 10g | SDS | 4%(w/v) | SDS | 10%(v/v) | Acetic acid | ~1L | dH2O |
| ~1L | dH2O | 0.2%(w/v) | Bromophenol blue | ~1L | dH2O | | |
| | | 20%(v/v) | Glycerol | | | | |

**Table 2.24 List of solutions used for SDS-PAGE experiments.**

### 2.2.11 Protein Crystallisation

The hanging drop method was used for protein crystallization. 500 µl reagents were added to each well on the 24 well plates. 1µl of protein sample and 1µl of screening reagent were mixed on the siliconised glass. Automated liquid handling equipment, Mosquito (TTP), was also used to dispense protein samples and screening reagents for setting up drops on the 96 well plates.

Screening reagent for protein used for the experiments are Structure Screen 1, Structure Screen 2 and Morpheus® (Molecular dimensions).

# Chapter 3

# Over-production of the alpha-domain of VirG using pGEX2T vector

### 3.1 Over-production and purification of the alpha-domain of VirG with pGEX2T VirG

Suzuki *et al.* (1998; 2002) reported that the alpha domain of VirG, amino acid residues 53-506, was sufficient to bind N-WASP and cause the polymerisation of actin. The pGEX2T plasmid harbouring the DNA corresponding to the alpha-domain was kindly provided by Prof. Chihiro Sasakawa and colleagues (University of Tokyo, Japan). The DNA corresponding to the alpha-domain of VirG was amplified using primers with appropriate restriction sites and subsequently ligated into the pET14b expression plasmid; this was the work of Dr Syeed Hussain. The advantage of the pET14b expression plasmid is that it encodes a hexa histidine-tag which produces a His-tagged protein that can be readily purified using a nickel column.

The quantity of the alpha domain produced by the two expression systems, pET14b and pGEX2T, and importantly the solubility of the alpha domain were investigated. The pET14b expression system produced only small quantities of the alpha domain of VirG and that which was produced was mostly insoluble (results not shown). However, the pGEX2T expression system produced a greater quantity of the alpha domain of VirG and a greater quantity of the protein was in the soluble fraction. The GST domain fused to the N-terminal end of the alpha domain of VirG improved the solubility of the protein compared to that of the His-tagged protein but still most of the produced VirG was insoluble. The soluble VirG produced was purified using a glutathione affinity column. Protein was eluted using reduced glutathione and the GST-tag was then cleaved overnight using thrombin. A size-exclusion column was then used to separate the GST-domain and the thrombin from the alpha domain of VirG. After removal of the GST-tag the alpha domain remained soluble; Fig. 3.1 shows the protein before (1) and after (2) cleavage of the GST tag from the alpha domain as analysed by 12% SDS-PAGE. VirG remained soluble after cleavage but the result of subsequent gel chromatography revealed that VirG has a molecular mass in excess of 200,000 Daltons.

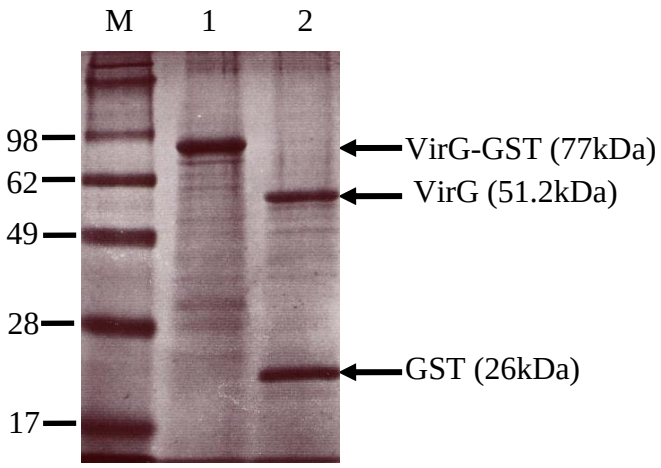



**Figure 3.1 Overproduction of the alpha domain of VirG in pGEX2T analyzed using 12% SDS-PAGE. Lanes: M-markers, 1-before and 2-after cleavage of the GST tag using thrombin. The alpha domain of VirG remained in the soluble fraction, but the protein has a high mass as judged by size exclusion chromatography. (Marker: SeeBlue® Plus2 Pre-Stained Standard, Life technology)**

## 3.2 Production of VirG 53 protein with pGEX2TVirG53 in Rosetta-gami™2 (DE3) pLysS

### 3.2.1 Rare codon analysis

Analysis of amino acid of VirG53-506 revealed there were total of 19 rare codons (Rare Codon Calculator (RaCC) by NIH MBI Laboratory for Structural Genomics and Proteomics. These rare codon included ATA, AGA, CTA and CCC and also double and triple rare codon were indicated within the gene of interest (see also Appendix). To supplement these codons, *E.coli* Rosetta-gami™2 (DE3) pLysS (Novagen) was used as host cell for protein production.

| | Rare codon | | | | | | |
|---|---|---|---|---|---|---|---|
| **VirG - alpha domain** | **ATA (11)*** | **-----** | **AGA (1)*** | **CTA (4)*** | **CCC (3)*** | **-----** | **-----** |
| **Rosetta-gami 2 (DE3) pLysS** | **AUA** | **AGG** | **AGA** | **CUA** | **CCC** | **GGA** | **CGG** |

***The number of codon in the VirG DNA sequence.**

**Table 3.1 Number of rare codon identified in VirG53-506 by RaCC and rare codons that supplied by rare tRNAs in Rosetta-gami™2 (DE3)pLysS .**

### 3.2.2 Protein production

Transformation and protein production were followed by the methods described in Chapter 2. The cell culture was incubated at 37 ºC for 7 hours which was long induction duration than BL21 (DE3) cell. Final concentration of 1mM IPTG was used at OD of 0.6 for protein production and incubated at 30 ºC for overnight. Protein was successfully produced but result showed that pGEX2TVirG53 produced in Rosetta-gami™2 (DE3) pLysS cell was detected in the insoluble pellets (Figure 3.2).

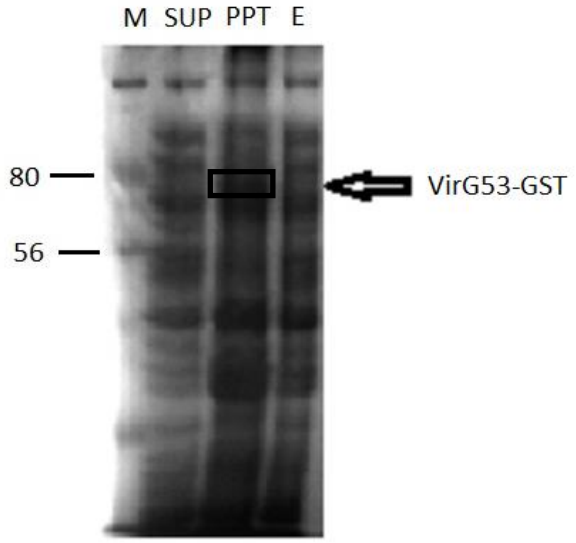



**Figure 3.2 Overproduction of VirG residues 53-506 pGEX2T by Rosetta-gami™2 (DE3) pLysS. Lines M: Marker, SUP: supernatant and PPT: pellet.**

## 3.3 Production of VirG 53 C130S mutant protein with pGEX2TVirG53C130S in BL21 (DE3) pLysS

### 3.3.1 Mutagenesis of alpha domain of VirG

The results from the size exclusion column showed that VirG aggregates. This might be due to the formation of inappropriate disulphide bonds between VirG molecules. There are three cysteine residues on the alpha domain of VirG. Substitution of cysteine at position at 130 was attempted by mutagenesis. Cysteine 130 was successfully mutated to serine in alpha domain of VirG. The result was confirmed by DNA sequencing (Eurofins MWG Operon). The plasmid that produced the mutated alpha domain of VirG53-506 was named pGEX2TVirG53C130S.

### 3.3.2 Optimisation of IPTG concentration for protein production.

Previous studies with VirG53-506 protein expressed using pET14b and pGEX2T vectors revealed that the protein produced was at least partially insoluble and aggregated in the solution. To improve the solubility of protein, Cysteine 130 was mutated to Serine within the alpha domain of VirG53-506. In this part, plasmid containing the mutated alpha domain of VirG53-506 was used. VirG protein was produced by using pGEX32TVirG53C130S plasmid transformed into BL21 (DE3) pLysS competent cells with two different IPTG concentrations, 0.5mM and 1mM. Produced proteins were tagged with the glutatione S-transferase (GST) at the N-terminus of VirG protein (VirGC130S-GST).

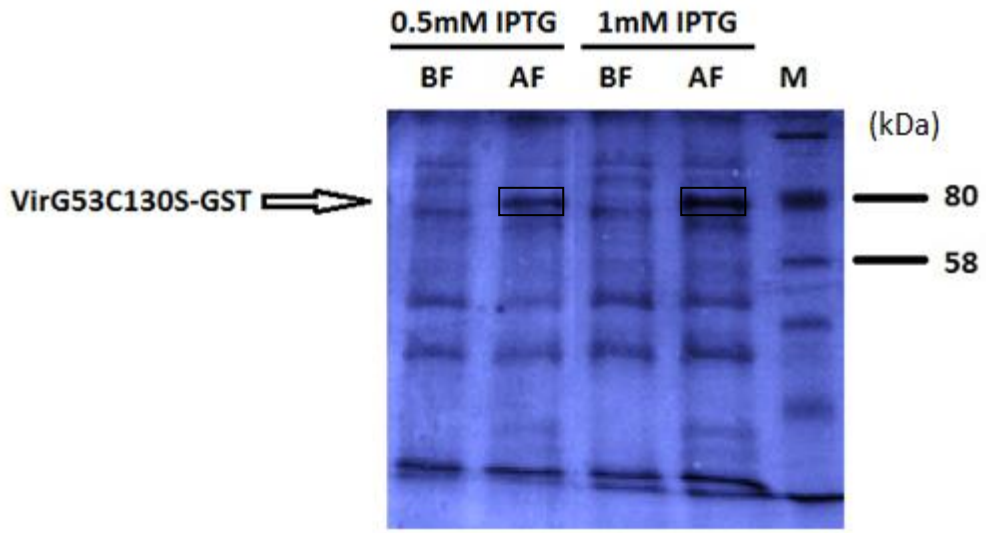



**Figure 3.3 Optimization of IPTG concentration for protein induction. Overproduction of 53-506 residues of mutated VirG alpha domain with GST tag. Each cell culture was induced protein production by 0.5 mM and 1mM IPTG. Lines BF: protein sample before IPTG induction. AF: protein sample after IPTG induction. M: Prestain Protein Marker (Invitrogen) P7703S**

Although, the final concentration of 1mM IPTG produced an apparently better yield of VirGC130S-GST (Figure 3.3), the final amount of purified protein produced was not greater. Therefore, optimisation of the Optical Density (OD) of the cell for protein induction was explored. Cell culture was induced to an OD = 0.8 and 1.0 and protein production was determined by SDS PAGE. 1mM IPTG was used for protein induction in this experiment (Figure 3.3).

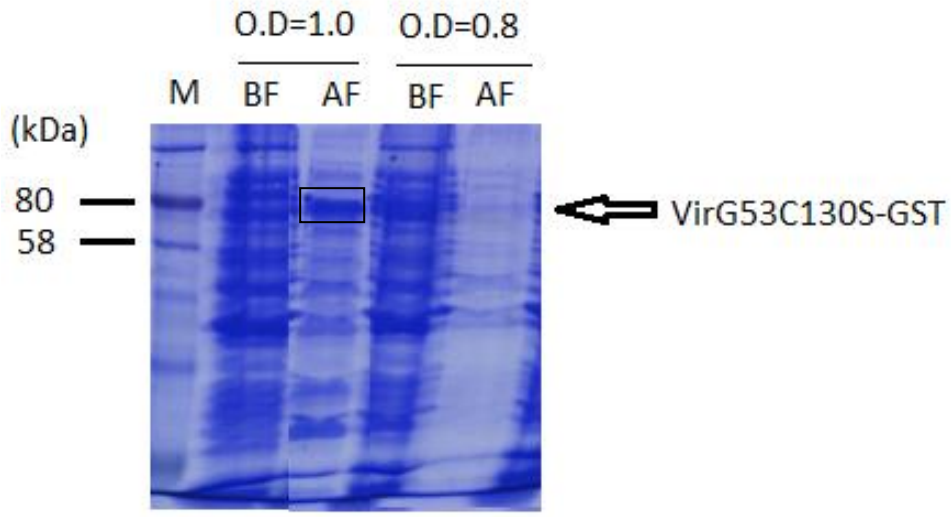



**Figure 3.4 Overproduction of residues 53-506 of mutated VirG alpha domain with GST tag. The cell culture was induced at an optical density of 0.8 and 1.0. Lines BF: protein samplebefore IPTG induction. AF: protein sample after IPTG induction. M: Prestain Protein Marker (Invitrogen) P7703S**

This experiment clearly suggests that more protein was produced better when the cells were grown to an OD of 1.0.

These conditions were used for protein production of pGEX2TVirG53C130S in *E. coli* BL21 (DE3) pLysS cells.

### 3.3.3 Protein purification

A GST column was used for first step of VirG53C130S purification. The GST tag was cleaved by thrombin from the VirG protein on the column. 15µl of Glutathione Sepharose beads were loaded onto a 12% SDS PAGE gel to check if the cleaved GST tag was bound to the beads. VirG was confirmed to be in the eluate. Figure 3.5 shows that the eluted fraction contained VirG53C130S and cleaved tags were indeed trapped on the Glutathione Sepharose beads. VirG also appeared in the pellet fraction which shows that some of the protein was insoluble.

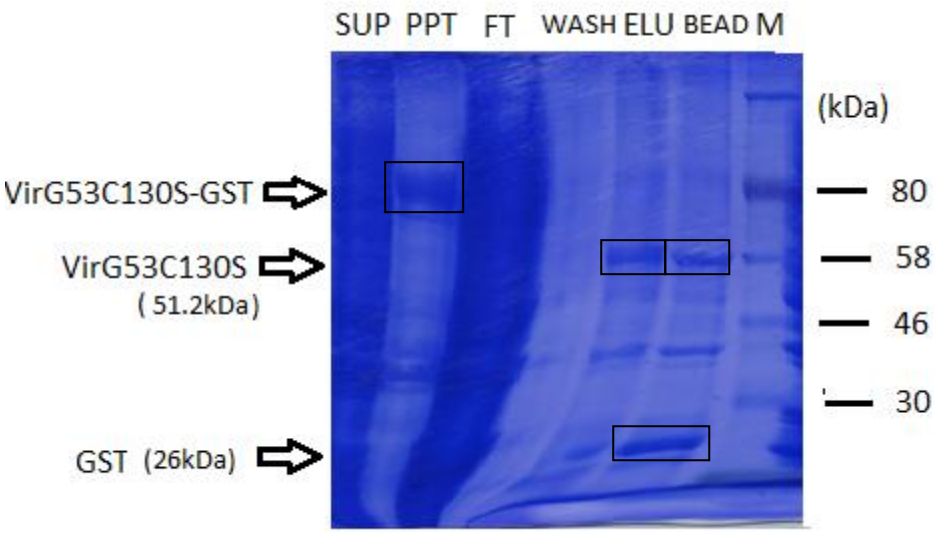



**Figure 3.5 VirG53C130S purification using Glutathione sepharose after thrombin digestion. 12% SDS PAGE was used to separate protein sample. Lines SUP: supernatant, PPT: pellet, FT: Flowthrough fraction, WASH: Wash fraction, ELU: Elution fraction and BEAD: Glutathione Sepharose beads. M: NEBPrestainProteinMarker P7703S**

The PBS buffer of the supernatant was exchanged with buffer comprising 20mM TrisHCl pH 8.0, with various NaCl concentrations (250mM, 400mM, 500mM) by the dialysis overnight at 4$^{o}$C. The buffer containing 250mM and 400mM caused precipitation of protein during the dialysis. Buffer with 500mM NaCl also produced white precipitate during concentration of protein solution after dialysis. Therefore 10% glycerol was added to the buffer to aid solubility of the protein for further purification by size exclusion column. The protein had a calculated pI of 4.29 (protein without GST tag), PBS has a pH of 7.3; therefore Tris-HCl pH 8.0 was used as the buffer.

The eluted protein sample was however slightly cloudy. The precipitation was apparent after thrombin digestion and dialysis. The sample was therefore spun down using a microcentrifuge at 13,000 rpm for 5mins at 4$^{o}$C. Separated supernatant and pellet were loaded onto a SDS-PAGE gel to identify if the precipitate contains VirG53C130S.

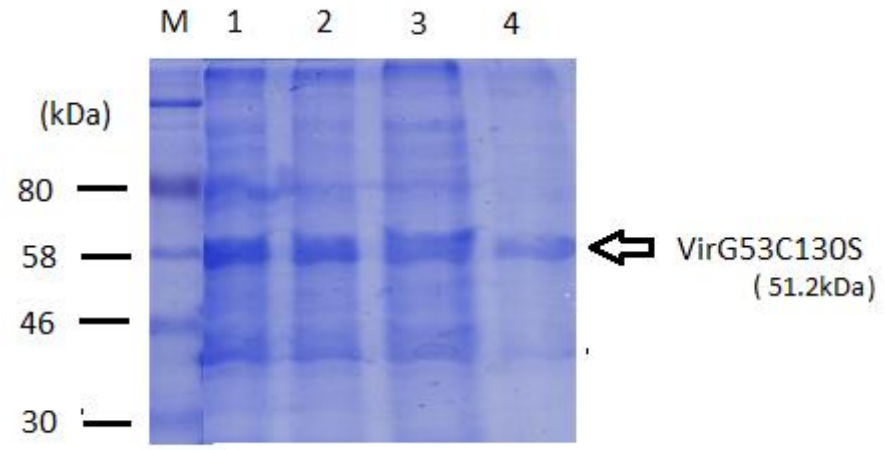



**Figure 3.6 Analysis of VirG53C130S by 12% SDS PAGE. Lines 1: Supernatant from elution fraction after thrombin digestion. 2: Pellet from elution fraction after thrombin digestion, 3: Supernatant from protein sample after dialysis 4: Pellet from elution fraction after dialysis M: Marker (NEB Prestain Protein Marker).**

The VirG53C130S protein was present in all samples (Fig. 3.6). This result showed that at least some of the protein is insoluble. The protein in the soluble fraction was thought to be soluble at this stage.

The elution fractions still contained contamination. In order to further purify the sample, the protein sample was concentrated to 1ml (20.5mg/ml) using 10,000 MWCO Viva spin concentrator (Generon) at 3,000 rpm for 20min. 1ml of concentrated protein sample was loaded into a Superdex 200 10/300 GL column (GE Heathcare). Two buffers containing two different detergents were used to optimise protein solubility for the further protein purification steps. The column was equilibrated with 20mM Tris-HCl pH 8.0, 500mM NaCl, 10% Glycerol or 20mM Tris-HCl pH 8.0, 500mM NaCl, 10% Glycerol, 10mM DTT. Fractions were collected 1ml per tube and the elution volume was 1.5 column volumes. Protein sample were centrifuged for 10min at 4$^{o}$C to separate any precipitate present in the solution. Eluted protein was checked by SDS-PAGE. 30µl of sample was loaded into the 12% SDS PAGE gel.

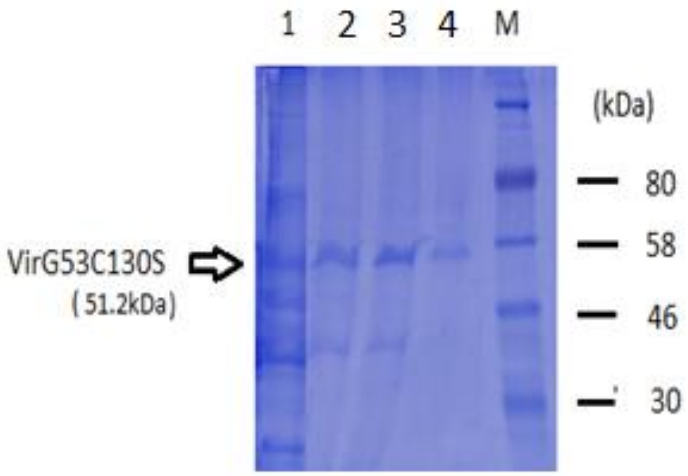



**Figure 3.7 Protein purification of VirG35C130S using a Superdex 200 10/300 GL column.Lines 1: Protein sample before the purification. 2-4: collected fraction.**

Some of the contaminating protein was separated by size exclusion column chromatography but some contamination remained in the same fraction as VirGC130S. VirG remained soluble after purification but the result of subsequent gel chromatography revealed that VirG has a molecular mass in excess of 200,000 Daltons revealing that it was aggregated.

# Chapter 4

# Overproduction of the alpha domain of VirG using pET system

### 4.1 Overproduction of the alpha domain of VirG using pET32a vector

VirG protein was found to be either partially insoluble or aggregates in the solution (Chapter 3). VirG aggregates might be due to the formation of inappropriate disulfide bonds within alpha domain of VirG that has three cysteine residues. pET32a vector has thioreodoxin encoded gene which reduces disulfide bonds to enable their rearrangement and facilitate folding of the protein. Therefore, expression of the alpha domain of VirG in pET32a may help disulfide exchange and the correct folding of the protein. In this chapter, two different length of alpha domain VirG (residues 53-506 and 103-433) were amplified and cloned into pET32a vector in order to achieve the protein solubility.

#### 4.1.1 *virG* Polymerase Chain Reaction (PCR) amplification

The alpha domain of VirG was amplified by PCR using the following primers.

| Short | Sequences | Length | Tm* |
|---|---|---|---|
| **Primer Forward** | 5'-CATCCATGGCTACTCCTCTTTCGGGTACTC-3' *Nco*I | 30 | 59 ºC |
| **primer Reverse** | 5'-ATGGGATCCTCAAACAGTAAGTTCAGCGTTTTC -3' *BamH*I | 33 | 57 ºC |
| **Long** | | | |
| **Primer Forward** | 5-CATCCATGGCTACTCCTCTTTCGGGTACTCAAGAACTTCATT *Nco*I TTTCAGAGG-3’ | 51 | 65ºC |
| **Primer Reverse** | 5’-ATGGGATCCTCAAACAGTAAGTTCAGCGTTTTCTTTCAGAT *BamH*I TGAACTCTCC-3’ | 51 | 64 ºC |

*Tm = Melting temperature

**Table 4. 1 Primers used for amplification of the alpha domain of *virG* encoded gene for VirG53**

In order to optimise the PCR amplification of *virG*, different polymerases, reaction mixtures and templates were used.

Betaine, Dimethyl sulfoxide (DMSO) and $MgCl_2$ were used to optimize the PCR reaction mixture. Betaine and DMSO facilitate strand separation as Betaine equalizes the contribution of GC- and AT base to the stability of the DNA duplex and DMSO disrupts base pairing. Hence, these agents help to increase the primer binding to the single strand template DNA

and help to produce a high yield of PCR products. $MgCl_2$ increases specificity of primer binding to template DNA. Therefore, this reduces contamination and increases the purity of the PCR products. Annealing temperature of the PCR cycle was also optimized depending on the primers used (short and long primers are shown in Table 4.1.).

Table 4.2 shows how the PCR was optimized.

(a)

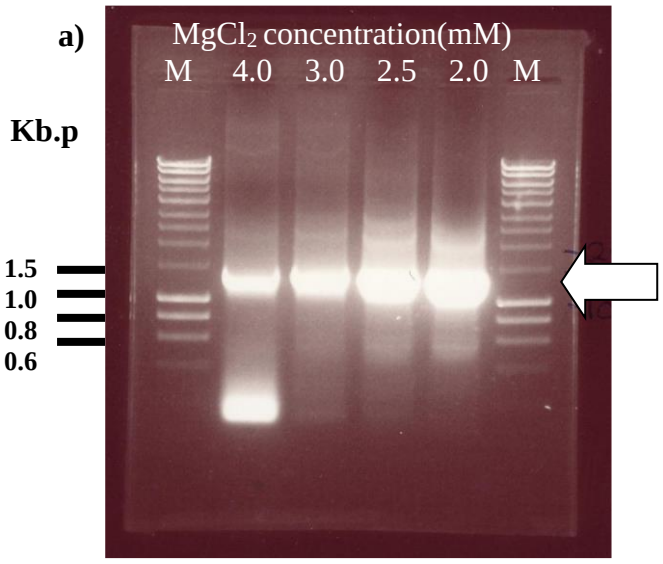
| **Polymerases** | **Template** | **Optimising the PCR reaction mixture** | **Primers** |
|---|---|---|---|
| Pfu polymerase | pET14b::*virG* | ±Betaine(1.2M), ±DMSO(10%)<br>Annealing template:<br>See Table 4.2 (b) | Short |
| | pGEX2T::*virG* | N/A | Short |
| Hot start polymerase | pET14b::*virG* | $MgCl_2$ concentration:<br>1.5mM,2.5mM,3.5mM<br>Annealing template:<br>See Table 4.2 (b) | Short |
| Hot start pfu polymerase | pET14b::*virG* | N/A | Short |
| Taq polymerase | pET14b::*virG* | ±Betaine(1.2M), ±DMSO(10%) | Short |
| Go Taq polymerase | pET14b::*virG* | Annealing template:<br>See Table 4.2 (b) | Short/<br>Long |
| | pGEX2T::*virG* | $MgCl_2$ concentration<br>2.0mM,2.5mM,3.0mM and 4.0mM<br>Annealing template:<br>See Table 4.2 (b) | Long |

(b)

| Polymerase / Cycle | **Pfu polymerase** | | **Hot start polymerase** | | **Hot start Pfu polymerase** | | **Taq polymerase** | | **Go Taq polymerase** | |
|---|---|---|---|---|---|---|---|---|---|---|
| Initial Denature | 95ºC | 1 min | 95ºC | 15min | 95ºC | 2 min | 95ºC | 2min | 95ºC | 2 min |
| Denature | 95ºC | 1 min | 95ºC | 1 min | 95ºC | 30 sec | 95ºC | 30 sec | 95ºC | 1 min |
| Annealing | 52ºC | 30sec | 52ºC or 55Cº | 1 min | 52ºC | 30 sec | 55ºC | 30 sec | 52ºC or 59ºC | 1min |
| Extension | 72ºC | 3 min | 72ºC | 1 min 30sec | 72ºC | 1min 30sec | 72ºC | 2 min | 72ºC | 1min 30sec |
| Number of Cycle | 30 | | 30 | | 30 | | 35 | | 35 | |
| Extra extension | 72ºC | 5 min | 72ºC | 10 min | 72ºC | 10 min | 72ºC | 7 min | 74ºC | 5min |

**Table 4.2 Optimisation of PCR reactions mixture for amplification of *virG* alpha domain. (a) Table for optimising PCR reaction mixture. (b) Table of PCR cycles for each polymerase used.**

a)




b)

| Finalized components of reaction mixture for PCR with 2mM $MgCl_2$ | |
|---|---|
| Volume | Components |
| 10µl | Reaction buffer |
| 1µl | dsDNA template pGEX2TVirG53 |
| 5µl | Long primer Forward |
| 5µl | Long primer Reverse |
| 0.8µl | 25mM dNTP mix |
| 0.25µl | GoTaq DNA polymerase |
| 3.75µl | $ddH_2O$ to a final volume of 50µl |
| 3.2 µl | $MgCl_2$ |
| Total 50µl | |

**Figure 4.1 a) *virG* (residues 53-506) PCR products for construct by using GoTaq polymerase and long PCR primers. b) Optimized PCR reaction mixture used in this study.**

Table 4.2 shows the final optimized condition used to amplify the *virG* alpha domain. PCR amplified better using Go Taq polymerase and long primers with annealing temperature at 59 $^{o}$C. Amplified PCR products from the reaction mixture containing 3.0 mM and 4.0 mM of $MgCl_2$ showed less contaminant compared with the other two conditions (the reaction mixture containing 2.0 mM and 2.5 mM of $MgCl_2$). However, latter conditions gave high yield the PCR products. The PCR product was extracted by gel extraction kit (QIAGEN). pET32a vector (Novagen) and the *virG* PCR products were digested by two restriction enzymes, *BamHI* and *NcoI*. Therefore the *virG* alpha domain amplified with optimized condition of 2.0mM $MgCl_2$ was cloned into the pET32a vector (Novagen) at the site of *BamHI* and *NcoI* to generate pET32aVirG53construct. For the pETVirG103 construct, different primers were used (see table 2.3 for detail.). The *virG* (Residues 103-433) PCR products were digested by two restriction enzymes, *XhoI* and *NcoI* was used and successfully ligated into pET32a vector.

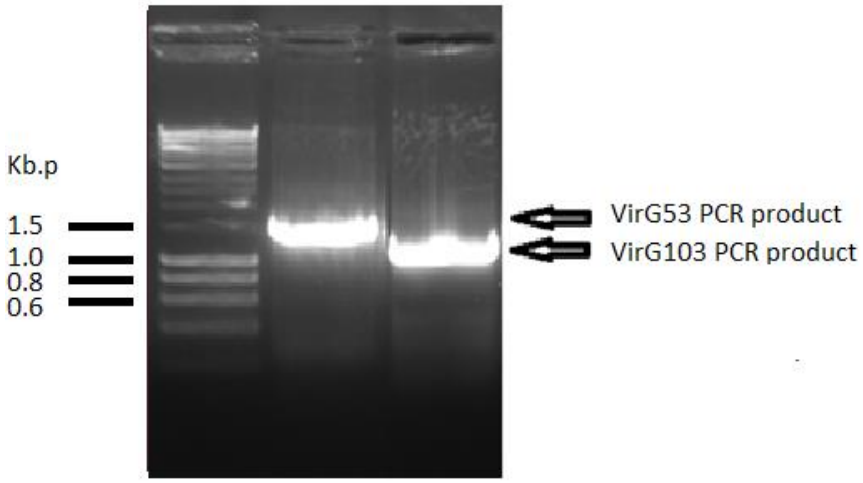



**Figure 4.2 Agarose gel electrophoresis of PCR products.**

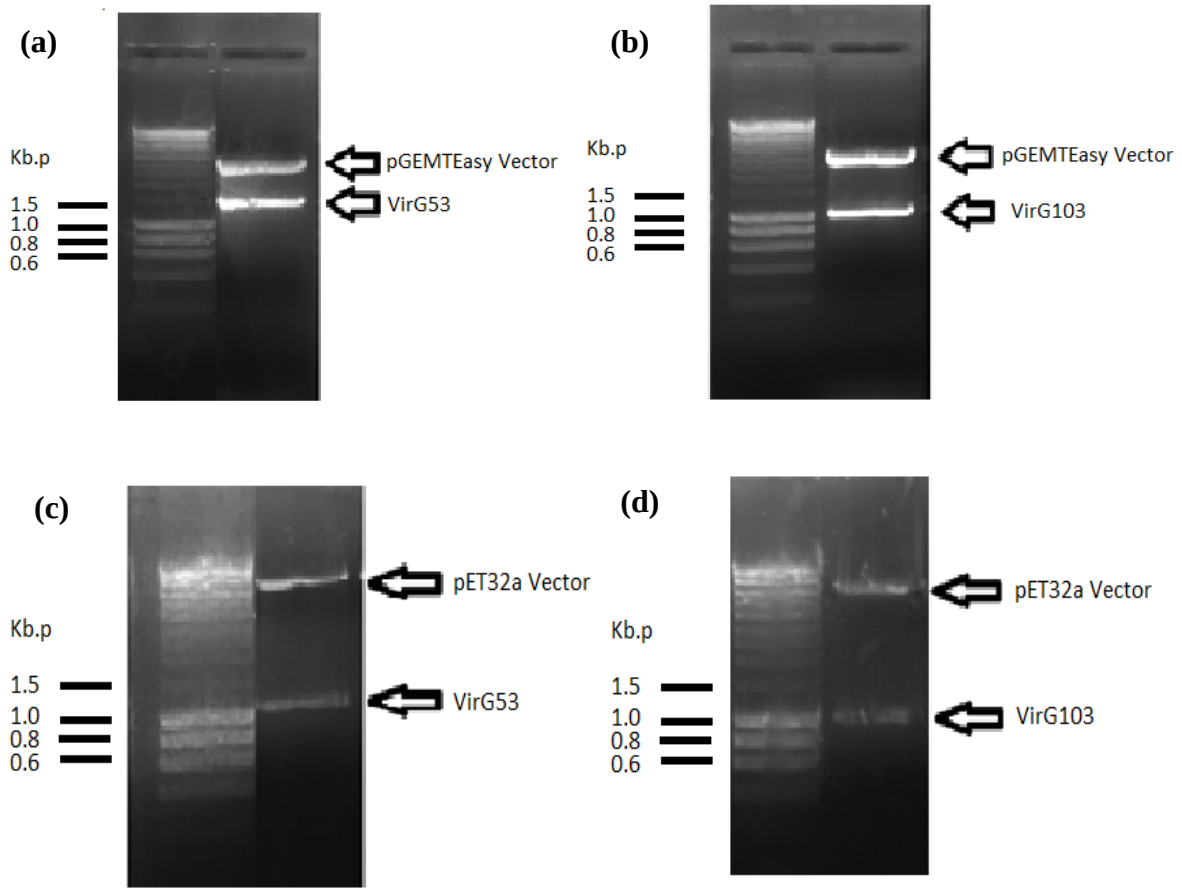



**Figure 4.3 An image of a gel electrophoresis. The image showed the inserted DNA corresponding to the alpha-domain VirG103 and VirG53 into pGEMEasy (a and b) and pET32a vector (c and d) by agarose gel electrophoresis.**

### 4.1.2 Production and purification of VirG by pET32aVirG construct

pET32aVirG53 and pET32aVirG103 were transformed into *E. coli* BL21 (DE3) competent cells. A single colony was inoculated into Luria–Bertani (LB) media containing Amp and Cam for growth overnight at 37ºC. The proteins produced were His-tag at the N terminus of VirG protein.

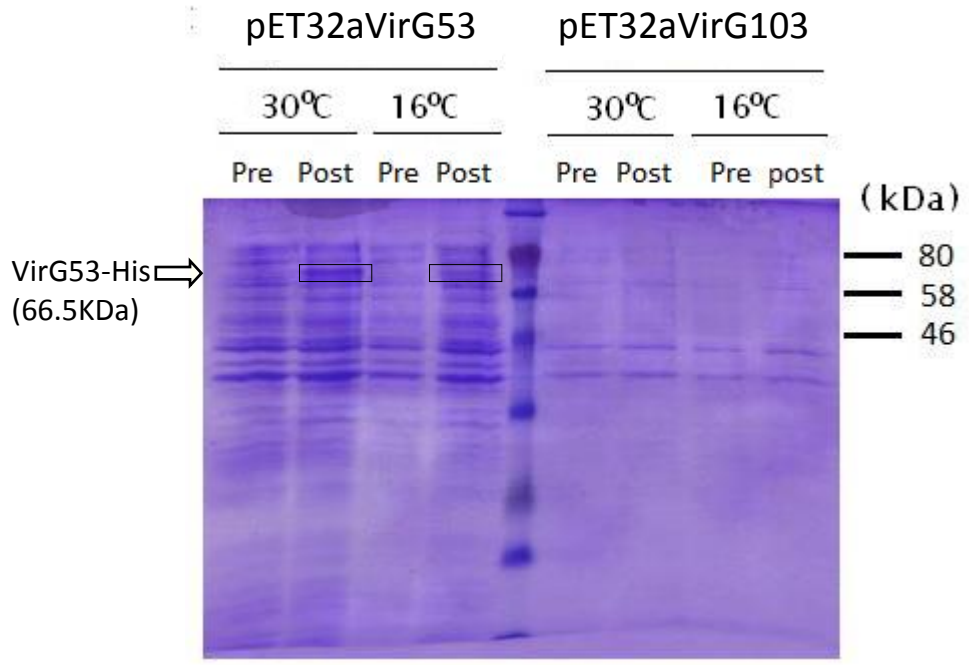



**Figure 4.4 Overexpression of VirG residues 53- and 103- with pET32a in BL21 (DE3). Black box on the gel shows band of produced VirG53 protein. Pre: Pre-induction sample, Post: Post-induction sample**

The result from SDS-PAGE showed that VirG53 was produced at both induction temperatures (30ºC and 16ºC) while VirG103 was not produced. Therefore, further purification was carried by using pET32aVirG53 in BL21(DE3).

### 4.1.3 Protein purification of VirG53 produced by *E.coli* BL21 (DE3).

Protein samples were prepared by following the method in Chapter 3. The cells were lysed by sonication and centrifuged. The supernatant was loaded on to pre-packed 5ml HisTrap column (GE Healthcare). The column was washed with 25ml wash buffer (20mM TrisHCl

pH7.5, 20mM Immidazole, 500mM NaCl.) then eluted against a gradient from 0%-100% of elution buffer (20mM TrisHCl pH7.5, 500mM Immidazole, 500mM NaCl.) by automated protein purification equipment AKTA FPLC (GE Healthcare). 5ml of elution fractions were collected per tube. Each fraction of protein samples were checked by 12% SDSPAGE gel. High concentration of protein sample were pooled and concentrated from 40ml to 2.0ml by MWCO 10,000 vivaspin (Generon) concentrator.

Concentrated protein samples were then centrifuge at 13,000 rpm for 10 minutes at 4$^{o}$C to extract any aggregated protein in the solution prior to run the column. Protein samples were loaded on to a Superdex 200 10/300 column (GE Healthcare) for further purification. The column was equilibrated using buffer (20mM Tris HCl pH7.5, 500mM NaCl) before running the column. 1.5ml of protein sample was collected per tube. Protein concentration was checked by chromatogram and by SDS-PAGE.

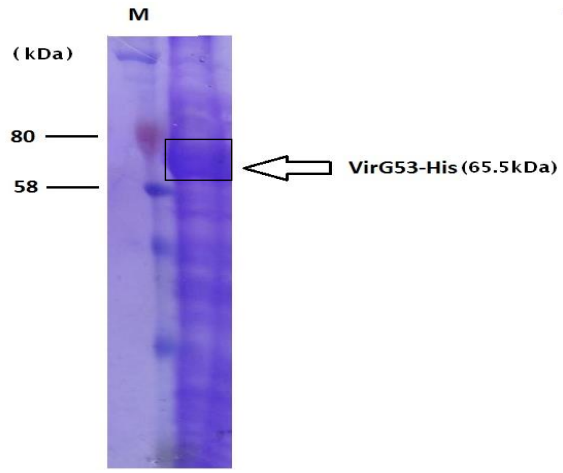



**Figure 4.5 Concentrated sample VirG. M:NEB colour plus protein marker**

Eluted VirG53 protein was observed in the void fraction. Therefore VirG has a molecular mass in excess of 200,000 Daltons (Da).

To improve solubility of protein in solution, final concentration of 0.1mM n-Dodecyl β-D-maltoside was added to the buffer for all purification method. However, the solubility of VirG protein produced was not enhanced.

### 4.1.4 Protein induction of VirG53 and VirG103 at high *E. coli* cell density.

Although protein was produced in the previous experiment, the amount was not sufficient. In this study, the *E. coli* cells were grown to a high density (high O.D) before protein was induced. Two different time duration (3 hours and overnight) of protein induction were also tested. *E. coli* harbouring pET32aVirG53 and pET32aVirG103 plasmids were used in this study. pET32aVirG53 and pET32aVirG103 were transformed into *E. coli* BL21 (DE3) competent cells. Proteins were produced with His-tag at N terminus of VirG protein.

150µl of overnight culture was added to 15ml of LB-Amp media and the cell culture was grown at 37$^{o}$C for 3hours until the optical density was approximately 1.3. 1M IPTG was then added to the culture at final concentration of 1mM for induction of recombinant protein production. The cells were grown overnight at 30$^{o}$C and then were harvested at 7000g for 20 minutes at 4 $^{o}$C. The cells were spun down after overnight incubation. The pellet was resuspended in sonication buffer, and then the cells were lysed by sonication and centrifuged. The supernatant and pellet were applied to 12% SDS-PAGE gel to determine the extent of over-expression of VirG. Cell pellet of pre-induction sample and post-induction sample were resuspended using 30µl of 20mM Tris-HCl pH7.5, 500mM NaCl, 20mM Immidazole and mixed with 30µl of loading dye. SDS-PAGE protein samples were boiled at 100 $^{o}$C for 10 minutes. 50µl of each protein sample were loaded into the 12% gel. Gel was run for 30 minutes, 100V. Figure 4.6 shows the protein production in different time schedule. Protein production was greater using overnight induction in both constructs.

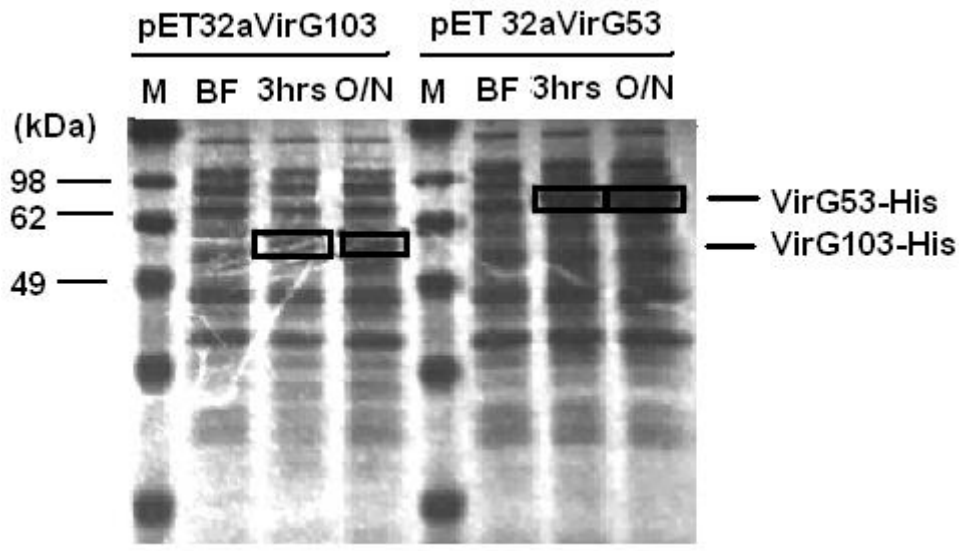



**Figure 4.6 Overproduction of VirG53 and VirG103 in pET32a with E. coli BL21 (DE3) cells in different induction times. (BF: Pre-induction, 3hrs: 3 hours, O/N: overnight).**

Overnight induction protein samples were analysed further. Pellets were sonicated and centrifuged to separate the supernatants and pellets. Both samples were prepared and loaded to 12% SDS-PAGE gel. Most of VirG103 was found in the insoluble fraction while some of the VirG53 were found in the soluble fraction (Figure. 4.7).

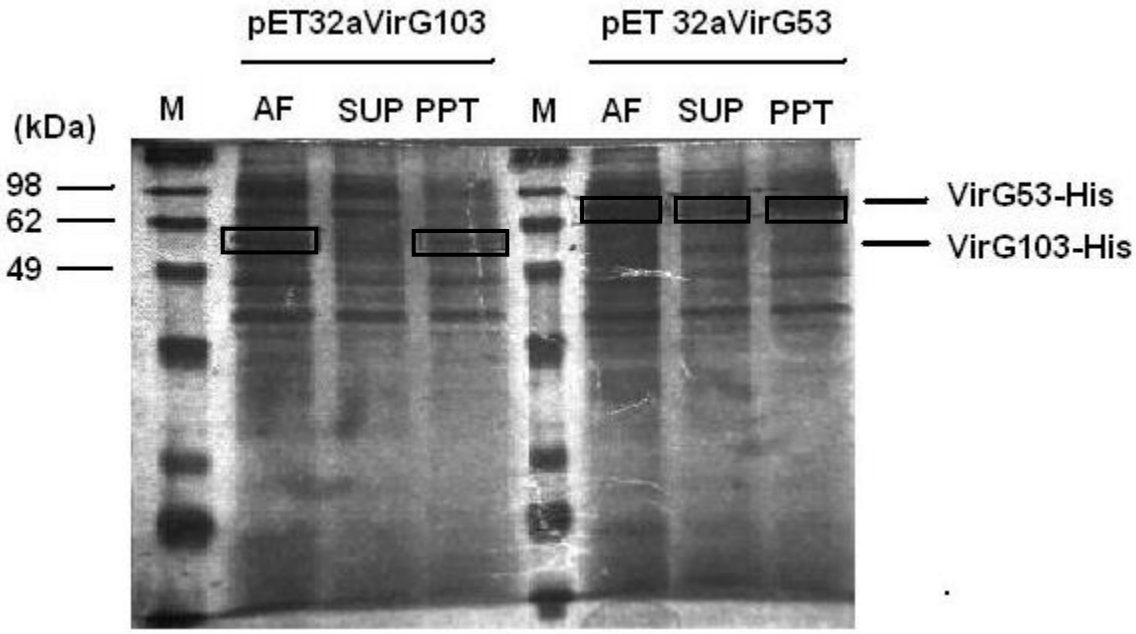



**Figure 4.7 Overproduction of VirG53 and VirG103 in pET32a with BL21 (DE3) in overnight induction. (AF: Post-induction, SUP: Supernatant, PPT: Pellet).**

### 4.2 Overproduction of the alpha domain of VirG using pET20b

The pET-20b(+) vector carries an N-terminal *pelB* signal sequence for potential periplasmic localization and adds 6x Histidine sequence at the C-terminal of protein during expression process. The DNA fragments that encod VirG 53 and VirG103 alpha domain of VirG were successfully cloned into pET20b. Figure 4.8 shows digestion of pET20bVirG53 and pET20bVirG103 constructs. Insert were identified using agarose gel electrophoresis.

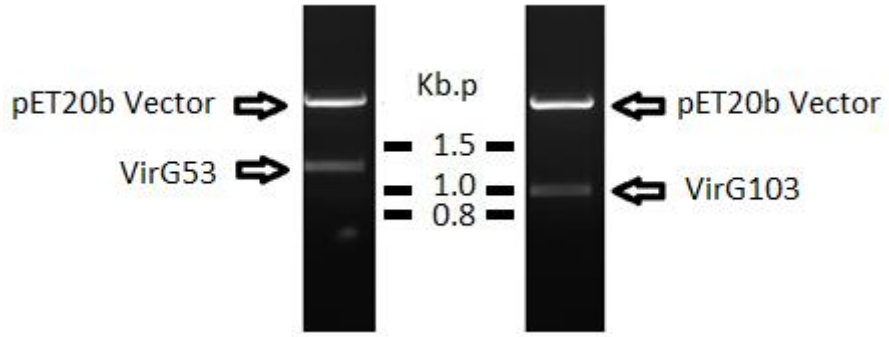



**Figure 4.8 Confirmation that alpha-domain VirG103 and VirG53 were successfully ligated into pET20b vector by agarose gel electrophoresis.**

pET20bVirG53 was transformed into Rosetta-gami™2 (DE3) pLysS competent cell. A single colony was inoculated into Luria–Bertani (LB) media containing Cam (12.5 μg /ml), Tet (30 μg /ml), Amp (100 μg /ml) overnight at 37$^{o}$C.

10ml of overnight culture was added to 1L of LB media and the cell culture was grown at 37$^{o}$C for 5 hours until the optical density was approximately ~0.6. 1L of cell culture was induced the protein with final concentration of 1mM IPTG for overnight at 30 $^{o}$C.

The extent of over-expression of VirG was determined by 12% SDS-PAGE gel in post-induction cell culture sample. Induced cell were resuspened by 20mM Tris-HCl pH8.0, 500mM NaCl, 10mM Imidazole. Cell culture were sonicated and centrifuged. Supernatant was loaded to HiTrap column (GE Healthcare) for protein purification.

Protein purification by HiTrap column were showed in Chaptor 2. The elution buffer in this study was 20mM TrisHCl pH8.0, 500mM Immidazole, 500mM NaCl. 5ml of elution fraction

were collected per tube. A high concentration of protein in fractions A9 and A10 was indicated by SDS PAGE gel (Figure 4.9). These two fractions contained contaminant therefore, these fractions were pooled and concentrated from 10ml to 1.0ml by MWCO 10,000 vivaspin (Generon) concentrator for further purification by size exclusion column.

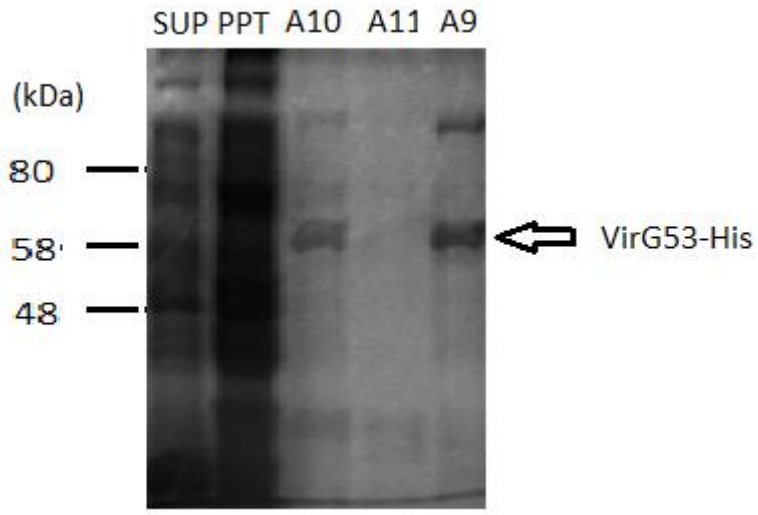



**Figure 4.9 Protein has been expressed by pET20bVirG53 with Rosetta2 (DE3) pLysS competent cells. The protein has been purified with HiTrap chelating column and selected fractions (A9, A10 and A11) were concentrated with MWCO 10,000 vivaspin. The histidine tag is still attached to the protein.**

The concentrated protein sample was then centrifuged at 13,000 rpm for 10 minutes at 4$^{o}$C to remove any aggregated protein prior to running the size exclusion column. Protein sample were loaded on to a Superdex 200 10/300 column (GE Healthcare). The column was equilibrated using buffer (20mM Tris-HCl pH8.0, 500mM NaCl) before running the column. 1.0ml of protein sample was collected per tube.

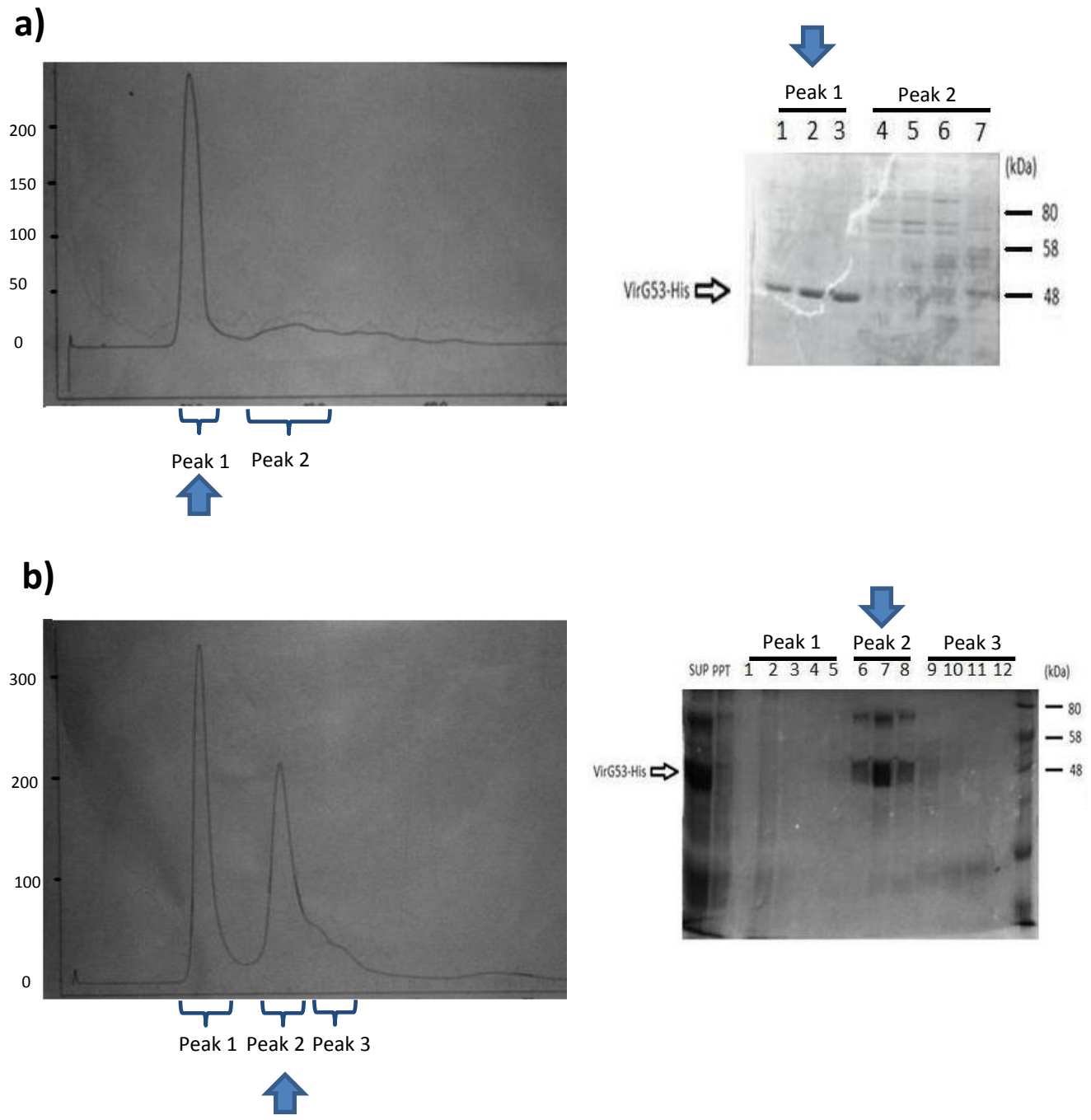



**Figure 4.10 Comparison of location of the eluted VirG53 protein with Histidine tag that was produced in different conditions. Locations of the eluted protein were determined by chromatogram (left) and SDS PAGE (right). Protein has been expressed by a) pET32aVirG53 with BL21 (DE3) competent cell and by b) pET20bVirG53 with Rosetta2 (DE3) pLysS competent cell. The protein has been purified using Superdex 200 10/300 size exclusion column and results were showen in the chromatogram. Blue arrow specifiy the peak and the fructions that contain protein of interest. a) Protein was indicated in fraction 1, 2 and 3 of peak 1 which were void fractions. On the other hand, b) Protein was indicated in fraction 6, 7 and 8, that were between the fractionation ranges, of peak 2. Fraction 6, 7 and 8 from b) condition were selected and were concentrated with MWC10, 000 vivaspin (Generon) for further experiment. Histidine tag is still attached to the protein (white arrow).**

Figure 4.10 shows chromatogram and the protein fraction from the size exclusion column. VirG proteins were eluted in the fraction 6, 7 and 8 at the secondpeak. The first and third peaks did not contain VirG.

The Superdex 200 10/300 column separates the protein size between 10,000 and 600, 000 (the fractionation range). Proteins that were more than 600kDa were too large to be fractionated by the column and eluted in the void volume. The void volume of the Superdex 200 10/300 column was 8ml (1/3 of volume in total bed volume of column 24ml). Fractions 6-8 were eluted at 18-21ml (fraction: 1.5 ml/tube). This suggested that VirG protein was no longer producing high molecular weight protein complex in the solution as has been seen in the experiment (Figure 4.10 a). VirG fraction 6, 7 and 8 were collected and concentrated by using a 10,000 MWCO concentrator.

Protein crystallisation screens were setup using VirG35-His (3.2mg/ml) using Morpheus® Additive OptiMax Kit (See apendex for the compostion of screen) Molecular dimentions) and Crystal screens 1 and 2 (See apendex for the compostion of screen, Hampton Reserch). VirG protein was introduced into hanging-drop vapor diffusion crystallization trials in 96 well plates at 19ºC. Sample (500nl) and reagents (500nl) were dispensed and mixed in the 96 well plates by the automated liquid dispenser, Mosquito (TTP Labtech).

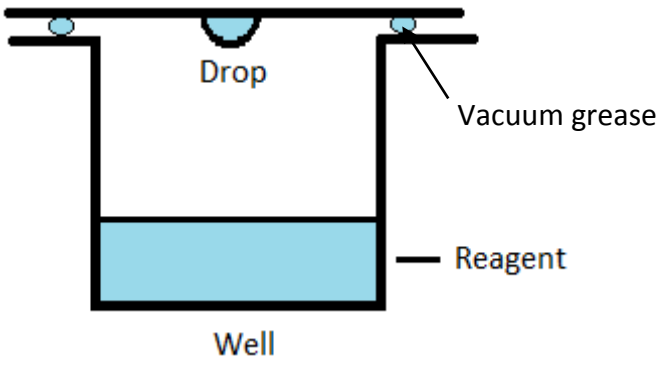



**Figure 4.11 Hanging drop vapour diffusion crystallization experiment**

Some of the trials showed precipitation in the protein drop, but no crystals were produced.

# Chapter 5

# Over-production and purification of WIP-WASP

### 5.1 Overexpression and production of N- Wiskott–Aldrich syndrome protein (WASP)

One method of enhancing the solubility of the alpha domain of VirG may have been to co-express the alpha domain together with its binding partner N-Wiskott–Aldrich syndrome protein (WASP). Therefore, WASP-interacting protein (WIP)-WASP was overproduced in order to check its solubility prior to co-expression with VirG. The structure of WIP-WASP structure has been characterised by NMR (Volkman *et al*., 2002), but not by X-ray crystallography. WIP-WASP was also used in crystallization trials.

The pBH-WIP-WASP construct was kindly provided by Prof. Wendell Lim from University of California. pBH is derived from the pET19b vector and contains a TEV protease site and a His-tag. The solubility of the WH1 domain of N-WASP was greatly enhanced by its interaction with WIP and the fusion protein WIP-WASP comprises a fragment of WIP (residues 461-485) and the WH1 domain of N-WASP (residues 26-147) linked by a Gly-Ser-Gly-Ser-Gly linker. The study showed the N-terminal fusion of WH1 with the WIP peptide was highly soluble compared to the C-terminal fusion (Volkman *et al*., 2002).

The pBHWIP-WASP plasmid was transformed into *E.coli* BL21 cells. Over-produced WIP-WASP was then purified following the method described by Volkman *et al.*, apart from cleavage of the His tag from the fusion protein. The result showed that purified His tagged fusion WIP-WASP was produced in high yield and solubility (Fig. 5.1). Hampton research crystal trails I and II (Hampton Research) were used for WIP-WASP crystallisation trials. 1µl of resovor and 1µl of protein were used for hanging drop method. Protein crystals were not produced.

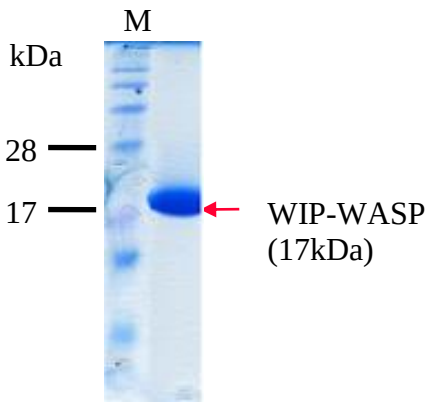



**Figure 5.1 High level overexpression of *wip-wasp* in pBH4 analysed using 12% SDS-PAGE. M:markers(SeeBlue® Pre-stained Protein Standard)**

.

## 5.2 Construction of a plasmid to over-produce the alpha domain of VirG in complex with WASP interaction protein (WIP) and Wiskott-Aldrich syndrome protein (WASP)

Previous experiments showed poor expression of soluble virG-alpha domain and the aggregation of purified VirG. The latter might have been due to the formation of inappropriate disulfide bonds involving the three cysteine residues in the alpha domain of VirG. In order to improve solubility of VirG by promoting disulfide exchange and the correct folding of the protein, expression vector pET32a was used for expression of *virG* as the vector encodes thioredoxin that reduces disulfide bonds to enable their rearrangement and facilitate folding of the protein. In addition, the Neural Wiskott-Aldrich syndrome protein (N-WASP) is known to interact with amino acid residues 53-506 of the alpha domain of VirG to initiate actin-polymerisation (Suzuki *et al*., 1998;, 2002). Co-expression of the *virG* region corresponding to the alpha domain and *wip-wasp* or *wasp* alone may give a soluble complex. Therefore, *wip-wasp* will be cloned downstream of *virG* in pET32a to produce a bicistronic constructs illustrated in Figure. 5.2. Cloning of VirG53 were shown in Chapter 4.

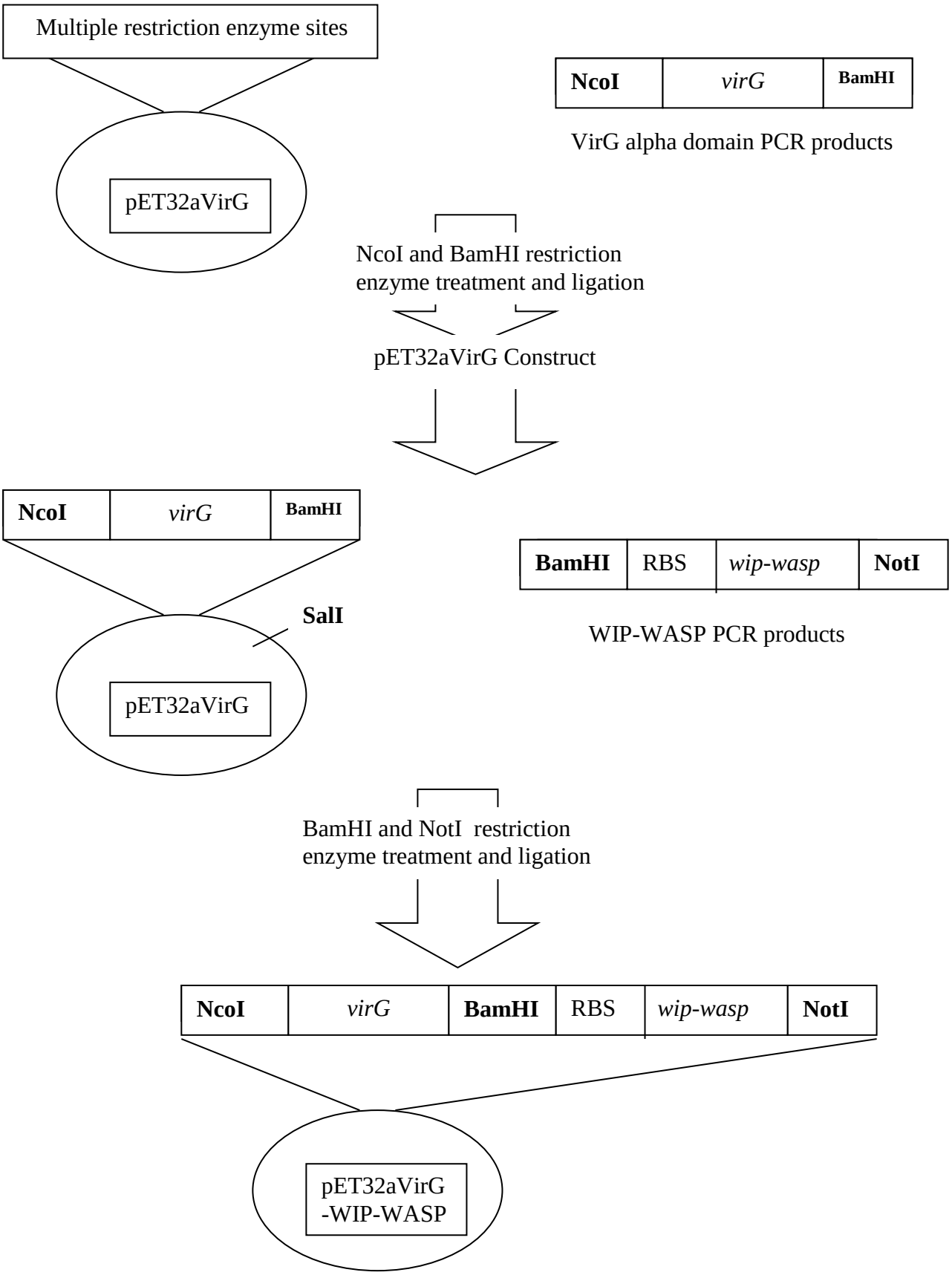



**Figure 5.2 VirG and WIP-WASP in the bicistronic construct**

### 5.2.1 *wip-wasp* PCR amplification

*wip-wasp* were amplified by PCR using following primers.

| **WIP-WASP** | |
|---|---|
| **Primer Forward** | 5'-CATGGATCCAAGGAGATATACATGGATTTG -3' BamHI |
| **primer Reverse** | 5'- ATGGCGGGCGCTCACTCGAGATCTCGTCTT-3' NotI |

**Table 5.1 Primers used for amplification of *wip-wasp***

The reaction mixture was optimized by different primer concentrations (1µM, 0.5µM and 0.1 µM). GoTaq polymerase was used for this amplification. Figure 5.3 (b) shows WIP-WASP was successfully amplified in all conditions. The condition containing high concentration of primer (1µM) had high yield of PCR products. All PCR products were purified by gel extraction kit (QIAGEN). Amplified *wip-wasp* was digested with two enzymes, BamHI and NotI.

**(a)**

| | | |
|---|---|---|
| Initial Denature | 95C° | 2 min |
| Denature | 95C° | 1 min |
| Annealing | 59C° | 30sec |
| Extension | 72C° | 30 sec |
| Number of Cycles | 35 | |
| Extra extension | 72C° | 5 min |

**(b)**

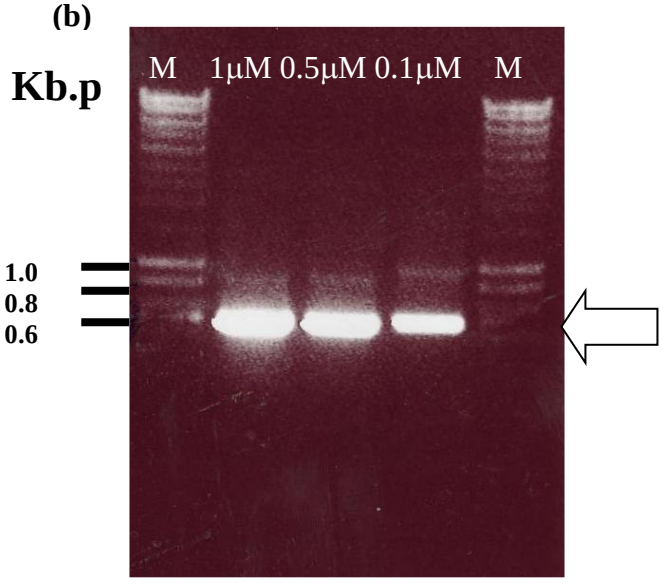


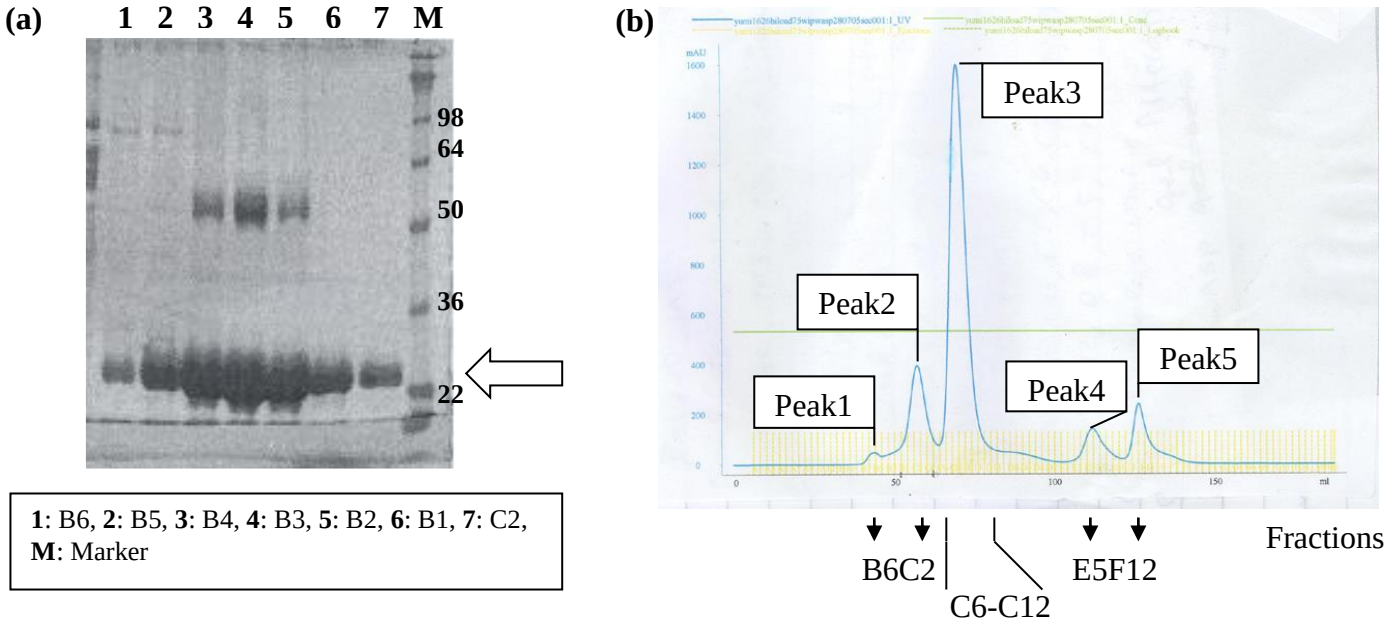

**Figure 5.3 (a) PCR cycle for WIP-WASP using GoTaq polymerase. (b) WIP-WASP PCR products confirmed by gel electrophoresis**

pET32aVirG53 was digested with BamHI and NotI and attempted ligation with WIP-WASP PCR product that was also digested with same restriction enzyme. However, the ligations were not successful.

### 5.3 Purification and crystallization of the WIP-WASP fusion protein

pBH4::*WASP* is the construct which has WIP-WASP insert in a vector derived from pET19b. This plasmid was transferred into BL21plysS (DE3) competent cells and transformants were grown on ampicilin LB-agar plates at 37°C. 1mM final concentration of IPTG was used for protein induction.

The pellet was resuspended in binding buffer (50mM phosphate buffer, 300mM NaCl, 10mM immidazole, pH8.0). The cells were sonicated and centrifuged at 15,000rpm for 45min at 4°C. The supernatant was applied to HiTrap Chelating Ni column (GE Healthcare) and purified WIP-WASP was eluted by using elution buffer (50mM phosphate buffer, 300mM NaCl, 500mM immidazole, pH8.0). Further protein purification was done by size-exclusion chromatography (Figure 5.4(b)). WIP-WASP was eluted by elution buffer (50mM phosphate buffer, 300mM NaCl).



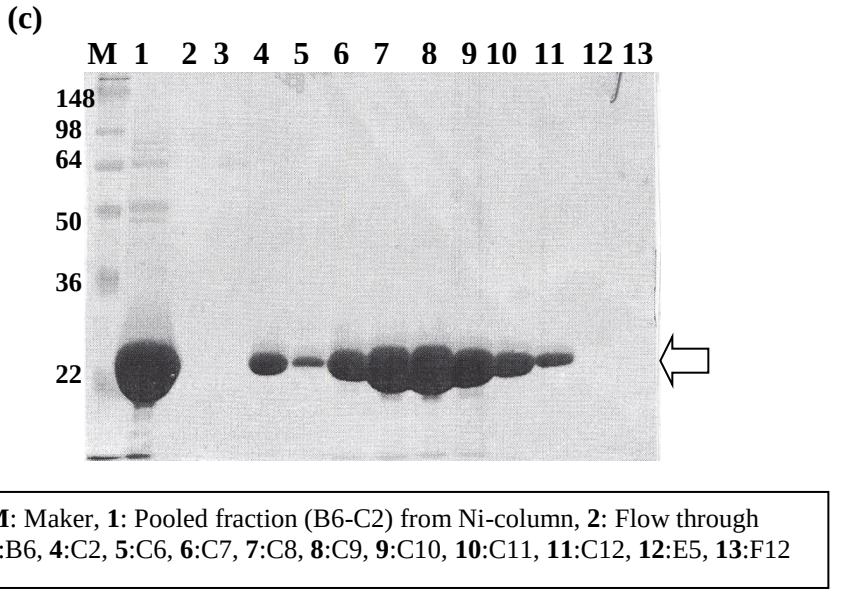




**Figure 5.4 Purification of the WIP-WASP fusion protein. (a) 12% SDS-PAGE of WIP-WASP fusion protein eluted from Ni-Column. Lanes 1 to 7 on the SDS-PAGE gel represents the samples that were taken from each fractions (B6-B1 and C2) at the peak conformed on the elution profile of Ni-column. Arrow shows the WIP-WASP fusion protein (17kDa). The elution profile is not shown. (b) Elution profile from HiLoad 16/600 Superdex 75 pg size-exclusion column (GE healthcare). (c) 12% SDS-PAGE of WIP-WASP fusion protein eluted from size-exclusion column. Samples are WIP-WASP fusion protein from Ni-column (lane1) and from size exclusion column (lane 2-13). Latter samples were taken from each fractions (B6, C2, C6-C12, E5 and F12) at peak 1-5 shown in Figure.5.4(b). Arrow shows the WIP-WASP fusion protein (17kDa).**

There were five distinctive peaks (Figure 5.4 (b)) after WIP-WASP was purified by size exclusions chromatography. Figure 5.4 (c) shows the purified WIP-WASP by size exclusion column analyzed by SDS-PAGE. The SDS-PAGE gel shows that the second and third peaks gives comprise WIP-WASP. However, the second peak was in the void fraction. This suggests that some of the WIP-WASP in the fractions of second peak may have aggregated or formed a high molecular mass complex. Also the result of SDS-PAGE conformed that there were no WIP-WASP in the fractions of first, fourth and fifth peak. Therefore, fractions of third peak were pooled, concentrated and used for crystallization trails using Hampton research crystallization screens I and II. The WIP-WASP fusion can be made in large quantities (4mg from 1L cells) as a soluble protein.

# Chapter 6

# Conclusion and future work

## 6.1 Concluding comments

*Shigella* infection is the cause of a large number of deaths in the developing world. Infection by *Shigella* involves several steps one of which is the movement of *Shigella* between an infected cell and an adjacent uninfected cell. This movement is generated by actin polymerisation at one pole of the *Shigella* bacterium. It is known that VirG is the protein which is anchored on the surface of the *Shigella* bacterium that reclutes the polymerisation of actin.

The alpha domain of VirG, amino acid residues 53-506, which is auto-transported to the surface of the bacterial outer membrane by the beta domain, binds multiple partners including: N-WASP (Suzuki *et al*. 1998; 2002), IcsB and vinculin. VirG also has an IspP cleavage site (Figure 1.7). The binding of N-WASP is known to recruit the APR2/3 complex that causes the polymerisation of actin and the generation of an "actin comet" forcing *Shigella* into the neighbouring cell. In order to understand the molecular basis of the dissemination of *Shigella* infection deeper into tissues, the alpha domain of VirG was studied.

The genes encoded two different domain of VirG alpha (VirG R103-A433 and VirG T53-V506) were cloned in to the several different expression vectors for protein production: pGEX2T, pET32a, and pET20b in order to improve better yield of soluble protein. In total of five constructs were successfully produced in this study. The mutation of one of the construct was succeeded at Cys 130 within the VirGT53-V506 for the purpose of inducing soluble protein. GST-or His-tagged VirG proteins were produced from these constructs. However, VirG protein from *Shigella flexineri* has proven to be remarkably difficult to produce in soluble form in *E. coli* in this study.

The advantage of the pGEX2T, pET14b, pET20b and pET32a expression plasmids are that pGEX2T encodes a glutathione *S*-transferase (GST) tag and pET vector encodes hexahistidine-tag which can be readily purified using a GST and nickel column respectively. Also GST tag is a soluble protein therefore it may aid the folding of the protein to which it is attached. The hexahistidine tag is small compared to the GST-tag and so problems of solubility would not be masked as effectively, this might allow the induction and growth to be investigated unhampered by effects of the GST-tag.

GST-tagged VirG α-domain (T53-V506), produced by the transformation of the pGEX2TVirG53 plasmid into *E. coli* (DE3) cells, was partly in the soluble fraction but mostly in the insoluble fraction. The soluble fraction of α-domain eluted in the void volume of the size exclusion column whether the protein was with or without GST-tag. This protein was confirmed as pure VirG α-domain by SDS-PAGE. The same fraction was analysed by dynamic light scattering the data supported the view that the α-domain. Taken together these results reveal that the α-domain is making a large complex or is aggregated in the solution.

In an attempt to improve the production of soluble VirG α-domain, the pET14bVirG53 construct was used. However, amount of soluble protein produced with this construct were similar to the protein that was produced by pGEX2T vector. These results showed that both proteins that were attached to GST or His-tag could not improve the solubility of VirG53 protein.

From the result above, it is possible that the insoluble protein was due to an incorrect pattern of disulphide bridges within three cysteine residues that were found in the VirG α-domain. Disulfide bonds in proteins are formed between two cysteine residues by oxidation of the thiol (SH) groups of cysteine residues under appropriate oxidizing conditions. To

explore this possibility one of three cysteine (Cys 130) was mutated to serine. Mutated alpha domain VirG C130S was cloned into the pGEX2T vector. The construct produced more soluble protein than did the non-cysteine mutant but the amount of soluble protein was still low.

The T53-V506 and R103-A433 regidues of VirG alpha domains were then cloned into pET32a and pET20b. 103-433 alpha domains were minimum length of alpha domain that contains all known binding site. VirG53-506 was found to be insoluble or aggregated in solution according to the previous experiment with pGEX2T and pET14b constructs. Therefore a shorter alpha domain was also cloned the expression vectors if produced protein can be solubilised better. pET32a encodes thioredoxin which reduces disulphide bonds to enable their rearrangement and facilitate folding of the protein. The pET20b construct has the advantages of targeting VirG to the periplasm by a PelB signal peptide, not the cytoplasm as with the other constructs used, to facilitate disulphide formation and having the affinity-tag at the C-terminus where it is known peptide can be accommodated.

Although the constructs made in pET14b and pGEX2T could produce proteins, the proteins were partially insoluble and aggregated in solution. Protein produced with new constructs with various competent cells were attempted in order to improve the solubility of protein.

However the VirG protein seems very insoluble in solution in *in vitro* experiments. The protein produced by four constructs (pET32aVirG53, pET32aVirG103, pGEX2TVirGC130S and pET20bVirG103) had similar characteristics to those produced by the original constructs. Proteins that were produced by these four constructs were found in inclusion bodies, which werelocated in the insoluble pellet as identified by SDS-PAGE. Precipitation in the soluble fraction was also observed which indicates that the protein is aggregating in solution. A high concentration of NaCl in the buffer seems increase the

solubility of protein as lack of salt in the buffer leads to precipitation of protein. A low concentration of protein in the buffer containing 500mM NaCl had no precipitation but high concentration of protein in the same buffer still precipitated. Adding detergent to the buffer with concentrated protein did not enhance solubility. The study also investigated expression at different temperatures (16 and 30°C), different IPTG concentrations (0.5mM and 1mM) and induction times (3 hours and overnight). Low temperature caused less protein to be produced. Protein production at the low temperature was more soluble than that produced at high temperature presumably because the the protein is being produced less quickly and not immediately put into inclusion bodies. However, the protein produced at the lower temperature may just have insoluble protein in smaller aggregates. Higher temperature, high IPTG concentration and long induction time increases the amount of protein but not the quantity of soluble protein.

The study showed that protein produced by pET20bVirG53 with Rossetta (DE3) pLysS construct eluted from the size exclusion column in the fractionating range, not in the void fraction. This result supported the view that the VirG53 produced was soluble. The pET20bVirG53 construct prodces a protein with a C-terminal tag. In the bacterium VirG is transported to the periplasm where it folds and is transfered to the outer membrane via a transporter. Recombinant VirG produced by pET20b has signal peptide at C-terminus and His tag at the N-terminus. The arrangement of signal peptide and tag at each terminus of the VirG protein is a similar arrangement to that of wild type full length VirG protein. Protein folding in periplasm may have improved and led to an increase in solubility. Also Rosetta2 (DE3) pLysS strains were used as a host of protein production by pET20bVirG53. Rosetta 2 host strains are BL21 derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*. Four rare codons (ATA, AGA, CTA and CCC) were found in VirG53 DNA sequence. Therefore an improvement of solubilisation in this

study may have also been accelerated by tRNAs for 7 rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) that supplied by the host *in vitro* environment.

VirG protein produced from pET20bVirG53 by Rosetta2 (DE3) pLysS was used for protein crystallisation. However, crystals were not produced. Soluble N-WASP was successfully produced in this study. N-WASP was known to bind to VirG alpha domain in nature. The VirG/N-WASP complex might be more soluble and afford a way forward.

**4.2 Future work**

It is surprising that the production of soluble *Shigella flexneri* VirG alpha-domain has proved so troublesome in *E. coli*. It may be that in *Shigella* there is a periplasmic or cytoplasmic chaperone that facilitates the folding of VirG. Chaperones act by preventing inappropriate interactions (i.e. misfolding) and giving the protein a chance to fold in a molecular cavity that facilitates correct folding. Another possibility is that there is an ancillary binding protein that stabilises the folded form of the VirG alpha-domain. It is possible that the β-domain is essential for the folding of the α-domain.

A high throughput cloning strategy where a large number of constructs are generated may be a useful way forward, especially when combined with an expression and solubility screen. Addition of green fluorescent protein to the C-termini of the protein and observation of fluorescence might indicate the correct folding of the attached VirG domain, if the VirG domain folds correctly then the GFP domain might also fold correctly and fluoresce.

A simpler method would be to try expression in a different system. *E. coli* is the simplest system for protein production, next in complexity and timescale is yeast. Expression in *Pitchia pastoris* has recently been used in the laboratory to make active protein that could not be made in active form in *E. coli*. The vector manipulation is done in *E. coli* and then the plasmid is transfected into the yeast. Expression in yeast is what I would like to try next in my search for a method to successfully produce soluble and functional VirG alpha domain.

# Appendix and Reference

# Appendix

### DNA and Protein sequence of VirG 103-433

### DNA 307-1299 (993bp)

```
aga act aag cta cgg att aac tct gat att atg att agc gca ggt cat ggt ggt gct
ggt gat aat aat gat ggt aat agt tgt ggc ggt aat ggt ggt gac tct att acc gga
tct gac ttg tct ata atc aat caa ggc atg att ctt ggt ggt agc ggc ggt agc ggt
gct gac cat aac ggt gat ggt ggt gag gct gtt aca gga gac aat ctg ttt ata ata
aat gga gaa att att tca ggt gga cat ggt ggc gat agt tat agt gat agt gat ggg
ggg aat gga ggt gat gcc gtc aca gga gtc aat cta ccc ata atc aac aaa ggg act
att tcc ggt ggt aat gga ggt aac aat tat ggt gag ggt gat ggc ggt aat gga ggt
gat gcc atc aca gga agc agc ctc tct gta atc aat aag ggc acg ttc gct gga ggc
aac gga ggt gct gct tac ggt tat ggt tat gat ggc tac ggt ggt aat gct atc aca
gga gat aac ctg tct gta atc aac aat gga gct att tta ggc ggt aat ggt gga cat
tgg ggg gat gct ata aat ggt agc aat atg acc att gct aat agc gga tat ata att
tca ggt aaa gaa gat gat gga aca caa aat gta gca ggt aat gct atc cac atc act
ggt gga aac aat tca tta ata ctc cat gaa ggt tct gtc att act ggt gat gta cag
gtt aac aat tca tcc att ctg aaa att atc aac aat gat tac act ggg acc aca cca
act att gaa ggt gat tta tgt gct ggt gat tgt aca act gtt tca cta tca ggt aac
aaa ttc act gtt tca ggt gac gtt tct ttt ggt gag aac agt tct tta aat tta gct
gga atc agt agt ctg gaa gct tct gga aat atg tca ttt ggc aac aat gta aaa gtg
gag gct att ata aat aac tgg gcg
```

### Protein sequence 103-433

```
        10         20         30         40         50         60
RTKLRINSDI MISAGHGGAG DNNDGNSCGG NGGDSITGSD LSIINQGMIL GGSGGSGADH

        70         80         90        100        110        120
NGDGGEAVTG DNLFIINGEI ISGGHGGDSY SDSDGGNGGD AVTGVNLPII NKGTISGGNG

       130        140        150        160        170        180
GNNYGEGDGG NGGDAITGSS LSVINKGTFA GGNGGAAYGY GYDGYGGNAI TGDNLSVINN

       190        200        210        220        230        240
GAILGGNGGH WGDAINGSNM TIANSGYIIS GKEDDGTQNV AGNAIHITGG NNSLILHEGS

       250        260        270        280        290        300
VITGDVQVNN SSILKIINND YTGTTPTIEG DLCAGDCTTV SLSGNKFTVS GDVSFGENSS

       310        320        330
LNLAGISSLE ASGNMSFGNN VKVEAIINNW A
```

Number of amino acids: 330

Molecular weight: 32335.57

Theoretical pI: 4.08

**The rare codon that identified in VirG103-433.**

Rare codon identified by Rare Codon Calculator (RaCC) by NIH MBI Laboratory for Structural Genomics and Proteomics. Underlined alphabets represent the rare codon that identified in VirG103-433.

**Red** = rare Arg codons **AGG, AGA, CGA**
**Green** = rare Leu codon **CTA**
**Blue** = rare Ile codon **ATA**
**Orange** = rare Pro codon **CCC**

for the following input sequence:

aga act aag CTA cgg att aac tct gat att atg att agc gca ggt cat ggt ggt gct
ggt gat aat aat gat ggt aat agt tgt ggc ggt aat ggt ggt gac tct att acc gga
tct gac ttg tct ATA atc aat caa ggc atg att ctt ggt ggt agc ggc ggt agc ggt
gct gac cat aac ggt gat ggt ggt gag gct gtt aca gga gac aat ctg ttt ATA ATA
aat gga gaa att att tca ggt gga cat ggt ggc gat agt tat agt gat agt gat ggg
ggg aat gga ggt gat gcc gtc aca gga gtc aat CTA CCC ATA atc aac aaa ggg act
att tcc ggt ggt aat gga ggt aac aat tat ggt gag ggt gat ggc ggt aat gga ggt
gat gcc atc aca gga agc agc ctc tct gta atc aat aag ggc acg ttc gct gga ggc
aac gga ggt gct gct tac ggt tat ggt tat gat ggc tac ggt ggt aat gct atc aca
gga gat aac ctg tct gta atc aac aat gga gct att tta ggc ggt aat ggt gga cat
tgg ggg gat gct ATA aat ggt agc aat atg acc att gct aat agc gga tat ATA att
tca ggt aaa gaa gat gat gga aca caa aat gta gca ggt aat gct atc cac atc act
ggt gga aac aat tca tta ATA ctc cat gaa ggt tct gtc att act ggt gat gta cag
gtt aac aat tca tcc att ctg aaa att atc aac aat gat tac act ggg acc aca cca
act att gaa ggt gat tta tgt gct ggt gat tgt aca act gtt tca CTA tca ggt aac
aaa ttc act gtt tca ggt gac gtt tct ttt ggt gag aac agt tct tta aat tta gct
gga atc agt agt ctg gaa gct tct gga aat atg tca ttt ggc aac aat gta aaa gtg
gag gct att ATA aat aac tgg gcg

The length is: 993 nucleotides

Rare codon in Rosetta cells : Seven rare codons, **AUA, AGG, AGA, CUA, CCC, GGA, and CGG**

| | Rare codon | | | | | | |
|---|---|---|---|---|---|---|---|
| **VirG - alpha domain** | **ATA (8)*** | **-----** | **AGA (1)*** | **CTA (3)*** | **CCC (1)*** | **-----** | **-----** |
| **Rosetta** | **AUA** | **AGG** | **AGA** | **CUA** | **CCC** | **GGA** | **CGG** |

***The number of codon in the VirG DNA sequence.**

**Double rare codon :** ATA ATA
**Triple rare codon :** CTA CCC ATA

## DNA and Protein sequence of VirG 53-506

### DNA 157-1518(1361bp)

**act cct ctt tcg ggt act caa gaa ctt cat ttt tca gag gac aat tat gaa aaa tta**
**tta aca cct gtt gat gga ctt tct ccc ttg gga gct ggt gaa gat gga atg gat gcg**
**tgg tat ata act tct tcc aac ccc tct cat gca tct aga act aag cta cgg att aac**
**tct gat att atg att agc gca ggt cat ggt ggt gct ggt gat aat aat gat ggt aat**
**agt tgt ggc ggt aat ggt ggt gac tct att acc gga tct gac ttg tct ata atc aat**
**caa ggc atg att ctt ggt ggt agc ggc ggt agc ggt gct gac cat aac ggt gat ggt**
**ggt gag gct gtt aca gga gac aat ctg ttt ata ata aat gga gaa att att tca ggt**
**gga cat ggt ggc gat agt tat agt gat agt gat ggg ggg aat gga ggt gat gcc gtc**
**aca gga gtc aat cta ccc ata atc aac aaa ggg act att tcc ggt ggt aat gga ggt**
**aac aat tat ggt gag ggt gat ggc ggt aat gga ggt gat gcc atc aca gga agc agc**
**ctc tct gta atc aat aag ggc acg ttc gct gga ggc aac gga ggt gct gct tac ggt**
**tat ggt tat gat ggc tac ggt ggt aat gct atc aca gga gat aac ctg tct gta atc**
**aac aat gga gct att tta ggc ggt aat ggt gga cat tgg ggg gat gct ata aat ggt**
**agc aat atg acc att gct aat agc gga tat ata att tca ggt aaa gaa gat gat gga**
**aca caa aat gta gca ggt aat gct atc cac atc act ggt gga aac aat tca tta ata**
**ctc cat gaa ggt tct gtc att act ggt gat gta cag gtt aac aat tca tcc att ctg**
**aaa att atc aac aat gat tac act ggg acc aca cca act att gaa ggt gat tta tgt**
**gct ggt gat tgt aca act gtt tca cta tca ggt aac aaa ttc act gtt tca ggt gac**
**gtt tct ttt ggt gag aac agt tct tta aat tta gct gga atc agt agt ctg gaa gct**
**tct gga aat atg tca ttt ggc aac aat gta aaa gtg gag gct att ata aat aac tgg**
**gcg cag aag gac tat aaa ctg cta agt gca gat aaa ggg ata aca ggt ttc agt gtt**
**tct aat ata tct atc atc aat ccg tta ctc act act ggt gct att gac tat aca aaa**
**agc tat atc agt gac cag aat aaa ttg atc tac ggt ttg agc tgg aat gat aca gat**
**ggc gac agt cat gga gag ttc aat ctg aaa gaa aac gct gaa ctt act gtt**

### Protein Sequence T53-V506

[53]TPLSGTQELH FSEDNYEKLL TPVDGLSPLG AGEDGMDAWY ITSSNPSHAS RTKLRINSDI

MISAGHGGAG DNNDGNSCGG NGGDSITGSD LSIINQGMIL GGSGGSGADH NGDGGEAVTG

DNLFIINGEI ISGGHGGDSY SDSDGGNGGD AVTGVNLPII NKGTISGGNG GNNYGEGDGG

NGGDAITGSS LSVINKGTFA GGNGGAAYGY GYDGYGGNAI TGDNLSVINN GAILGGNGGH

WGDAINGSNM TIANSGYIIS GKEDDGTQNV AGNAIHITGG NNSLILHEGS VITGDVQVNN

SSILKIINND YTGTTPTIEG DLCAGDCTTV SLSGNKFTVS GDVSFGENSS LNLAGISSLE

ASGNMSFGNN VKVEAIINNW AQKDYKLLSA DKGITGFSVS NISIINPLLT TGAIDYTKSY

ISDQNKLIYG LSWNDTDGDS HGEFNLKENA ELTV[506]

Number of amino acids: 454

Molecular weight: 45679.1

Theoretical pI: 4.16

**The rare codon that identified in VirG53-506.**

Rare codon identified by Rare Codon Calculator (RaCC) by NIH MBI Laboratory for Structural Genomics and Proteomics. Underlined alphabets represent the rare codon that identified in VirG53-506 (1362 nucleotides).

**Red** = rare Arg codons **AGG, AGA(1), CGA**
**Green** = rare Leu codon **CTA** (4)
**Blue** = rare Ile codon **ATA** (11)
**Orange** = rare Pro codon **CCC** (3)

for the following input sequence:

act cct ctt tcg ggt act caa gaa ctt cat ttt tca gag gac aat tat gaa aaa tta
tta aca cct gtt gat gga ctt tct CCC ttg gga gct ggt gaa gat gga atg gat gcg
tgg tat ATA act tct tcc aac CCC tct cat gca tct aga act aag CTA cgg att aac
tct gat att atg att agc gca ggt cat ggt ggt gct ggt gat aat aat gat ggt aat
agt tgt ggc ggt aat ggt ggt gac tct att acc gga tct gac ttg tct ATA atc aat
caa ggc atg att ctt ggt ggt agc ggc ggt agc ggt gct gac cat aac ggt gat ggt
ggt gag gct gtt aca gga gac aat ctg ttt ATA ATA aat gga gaa att att tca ggt
gga cat ggt ggc gat agt tat agt gat agt gat ggg ggg aat gga ggt gat gcc gtc
aca gga gtc aat CTA CCC ATA atc aac aaa ggg act att tcc ggt ggt aat gga ggt
aac aat tat ggt gag ggt gat ggc ggt aat gga ggt gat gcc atc aca gga agc agc
ctc tct gta atc aat aag ggc acg ttc gct gga ggc aac gga ggt gct gct tac ggt
tat ggt tat gat ggc tac ggt ggt aat gct atc aca gga gat aac ctg tct gta atc
aac aat gga gct att tta ggc ggt aat ggt gga cat tgg ggg gat gct ATA aat ggt
agc aat atg acc att gct aat agc gga tat ATA att tca ggt aaa gaa gat gat gga
aca caa aat gta gca ggt aat gct atc cac atc act ggt gga aac aat tca tta ATA
ctc cat gaa ggt tct gtc att act ggt gat gta cag gtt aac aat tca tcc att ctg
aaa att atc aac aat gat tac act ggg acc aca cca act att gaa ggt gat tta tgt
gct ggt gat tgt aca act gtt tca CTA tca ggt aac aaa ttc act gtt tca ggt gac
gtt tct ttt ggt gag aac agt tct tta aat tta gct gga atc agt agt ctg gaa gct
tct gga aat atg tca ttt ggc aac aat gta aaa gtg gag gct att ATA aat aac tgg
gcg cag aag gac tat aaa ctg CTA agt gca gat aaa ggg ATA aca ggt ttc agt gtt
tct aat ATA tct atc atc aat ccg tta ctc act act ggt gct att gac tat aca aaa
agc tat atc agt gac cag aat aaa ttg atc tac ggt ttg agc tgg aat gat aca gat
ggc gac agt cat gga gag ttc aat ctg aaa gaa aac gct gaa ctt act gtt

| | Rare codon | | | | | | |
|---|---|---|---|---|---|---|---|
| **VirG - alpha domain** | **ATA (11)*** | **-----** | **AGA (1)*** | **CTA (4)*** | **CCC (3)*** | **-----** | **-----** |
| **Rosetta** | **AUA** | **AGG** | **AGA** | **CUA** | **CCC** | **GGA** | **CGG** |

**Double rare codon :** ATA ATA

**Triple rare codon :** CTA CCC ATA

**VirG alpha domain protein sequence Thr53-Val506** (MW=45750, PI=4.16, a.a =455)

**Gray = Point mutation point (Cysteine (C) → Serine(S) )**

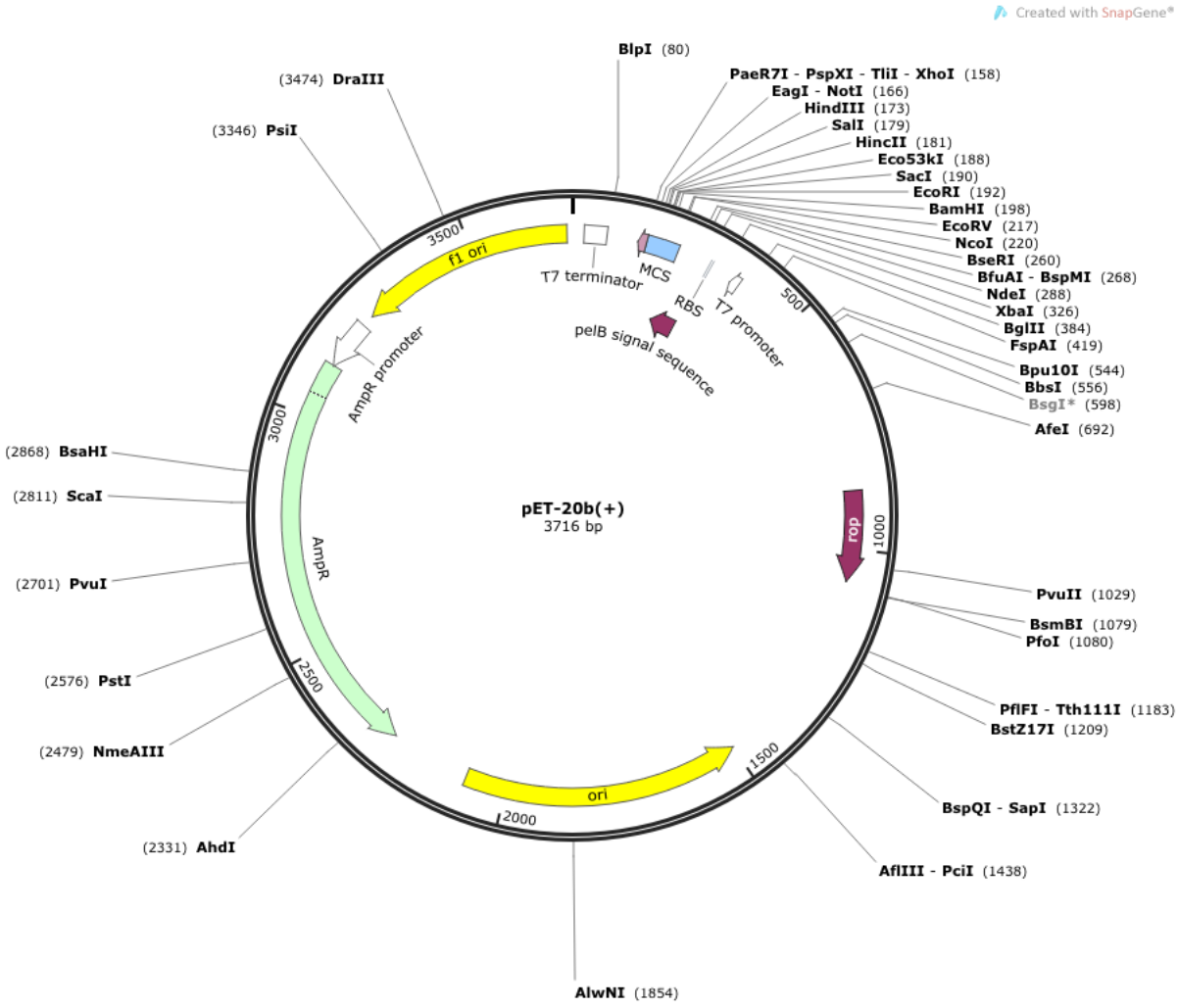
**53T**PLSGTQELH FSEDNYEKLL TPVDGLSPLG AGEDGMDAWY ITSSNPSHAS
RTKLRINSDI MISAGHGGAG DNNDGNSCGG NGGDSITGSD LSIINQGMIL
GGSGGSGADH NGDGGEAVTG DNLFIINGEI ISGGHGGDSY SDSDGGNGGD
AVTGVNLPII NKGTISGGNG GNNYGEGDGG NGGDAITGSS LSVINKGTFA
GGNGGAAYGY GYDGYGGNAI TGDNLSVINN GAILGGNGGH WGDAINGSNM
TIANSGYIIS GKEDDGTQNV AGNAIHITGG NNSLILHEGS VITGDVQVNN
SSILKIINND YTGTTPTIEG DLCAGDCTTV SLSGNKFTVS GDVSFGENSS
LNLAGISSLE ASGNMSFGNN VKVEAIINNW AQKDYKLLSA DKGITGFSVS
NISIINPLLT TGAIDYTKSY ISDQNKLIYG LSWNDTDGDS HGEFNLKENA
ELTVSTILAD NLSHHNINSW DGKSLTKSGE GTLILAEKNT YSGFTNINAG
ILKMGTV**506**

**VirG alpha domain DNA Sequence(1365bp)**

**Gray = Point mutation point TGT (TGT=Cysteine-> AGT=Serine)**

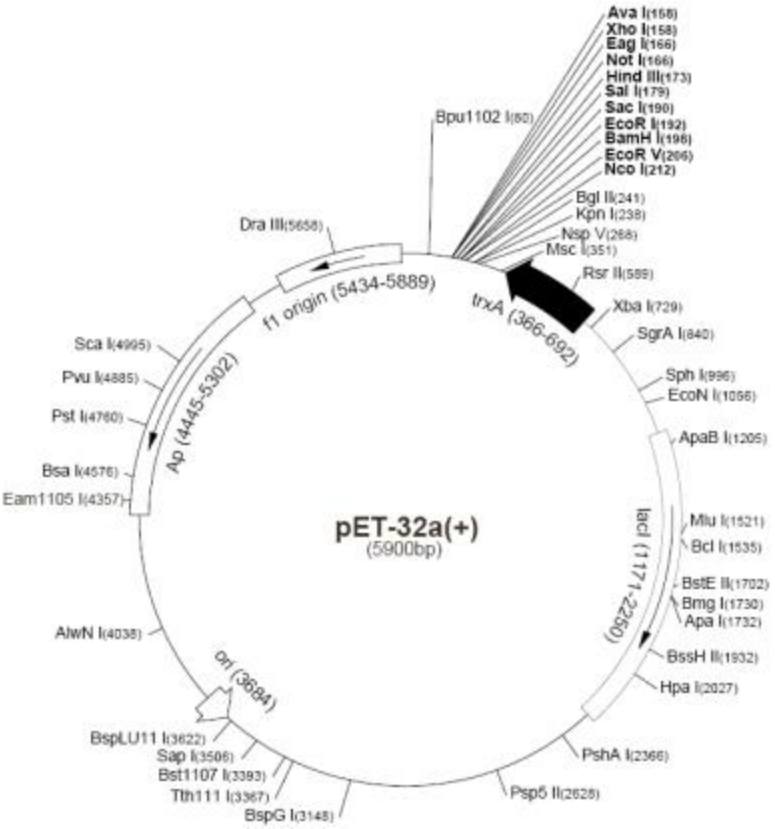
**act cct ctt tcg ggt act caa gaa ctt cat ttt tca gag gac aat tat gaa aaa tta**
**tta aca cct gtt gat gga ctt tct ccc ttg gga gct ggt gaa gat gga atg gat gcg**
**tgg tat ata act tct tcc aac ccc tct cat gca tct aga act aag cta cgg att aac**
**tct gat att atg att agc gca ggt cat ggt ggt gct ggt gat aat aat gat ggt aat**
**agt tgt ggc ggt aat ggt ggt gac tct att acc gga tct gac ttg tct ata atc aat**
**caa ggc atg att ctt ggt ggt agc ggc ggt agc ggt gct gac cat aac ggt gat ggt**
**ggt gag gct gtt aca gga gac aat ctg ttt ata ata aat gga gaa att att tca ggt**
**gga cat ggt ggc gat agt tat agt gat agt gat ggg ggg aat gga ggt gat gcc gtc**
**aca gga gtc aat cta ccc ata atc aac aaa ggg act att tcc ggt ggt aat gga ggt**
**aac aat tat ggt gag ggt gat ggc ggt aat gga ggt gat gcc atc aca gga agc agc**
**ctc tct gta atc aat aag ggc acg ttc gct gga ggc aac gga ggt gct gct tac ggt**
**tat ggt tat gat ggc tac ggt ggt aat gct atc aca gga gat aac ctg tct gta atc**
**aac aat gga gct att tta ggc ggt aat ggt gga cat tgg ggg gat gct ata aat ggt**
**agc aat atg acc att gct aat agc gga tat ata att tca ggt aaa gaa gat gat gga**
**aca caa aat gta gca ggt aat gct atc cac atc act ggt gga aac aat tca tta ata**
**ctc cat gaa ggt tct gtc att act ggt gat gta cag gtt aac aat tca tcc att ctg**
**aaa att atc aac aat gat tac act ggg acc aca cca act att gaa ggt gat tta tgt**
**gct ggt gat tgt aca act gtt tca cta tca ggt aac aaa ttc act gtt tca ggt gac**
**gtt tct ttt ggt gag aac agt tct tta aat tta gct gga atc agt agt ctg gaa gct**
**tct gga aat atg tca ttt ggc aac aat gta aaa gtg gag gct att ata aat aac tgg**
**gcg cag aag gac tat aaa ctg cta agt gca gat aaa ggg ata aca ggt ttc agt gtt**
**tct aat ata tct atc atc aat ccg tta ctc act act ggt gct att gac tat aca aaa**
**agc tat atc agt gac cag aat aaa ttg atc tac ggt ttg agc tgg aat gat aca gat**
**ggc gac agt cat gga gag ttc aat ctg aaa gaa aac gct gaa ctt act gtt**

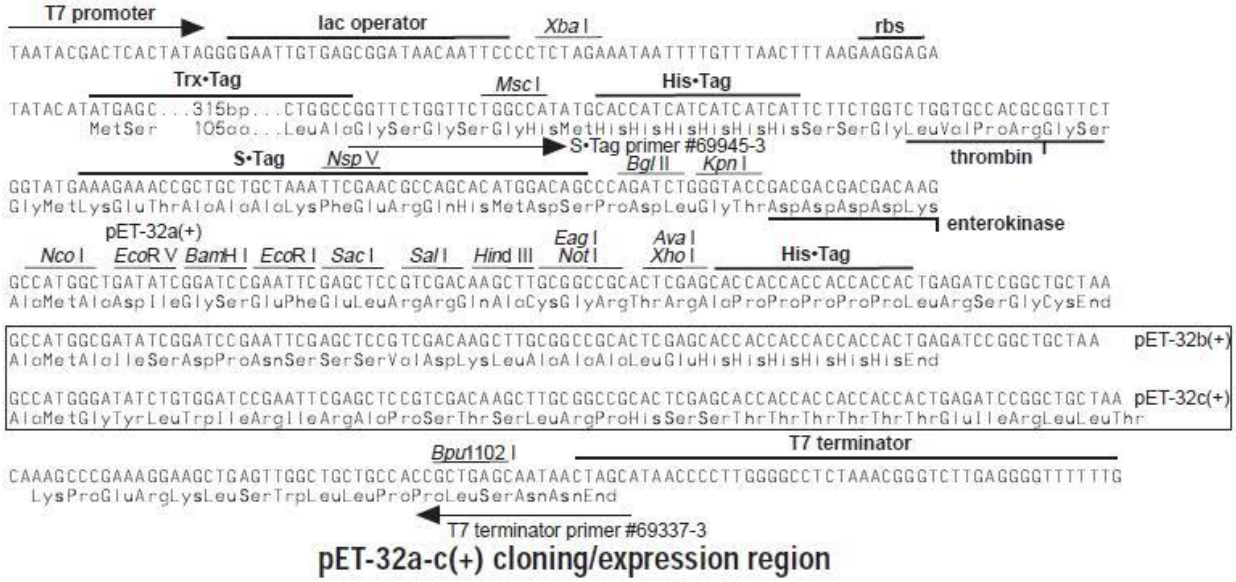
## Sequence and DNA map for pET20b(+)




Sequence and map were extracted from snapgene. http://www.snapgene.com

**Sequence and DNA map for pET32a**

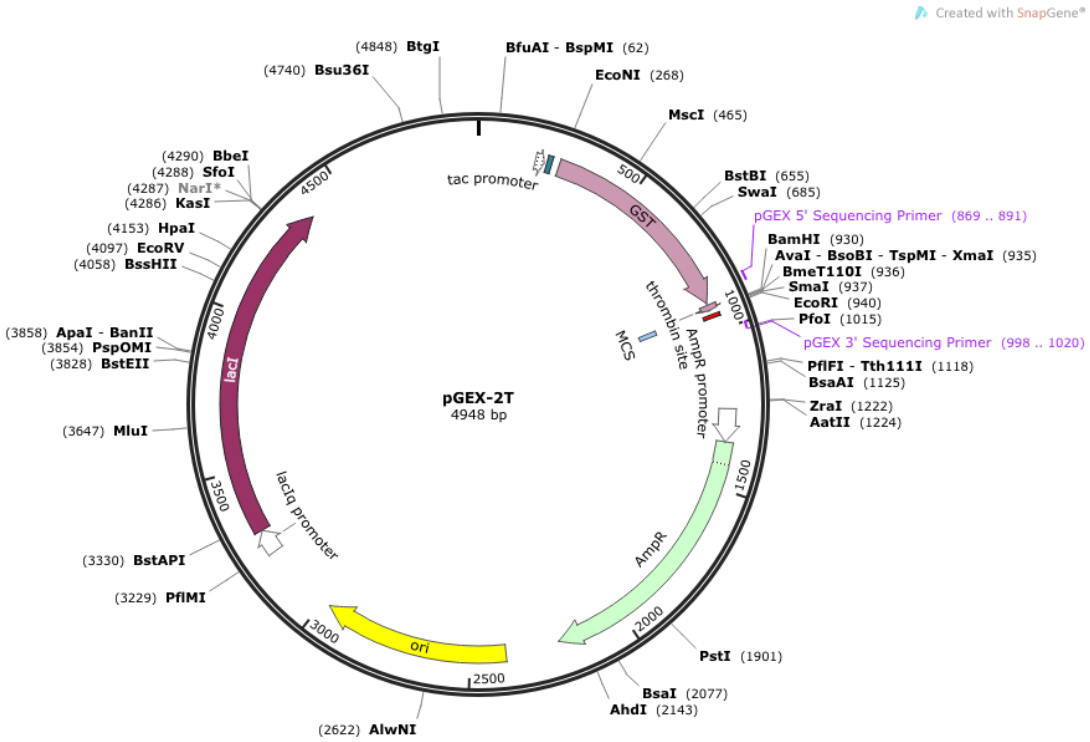


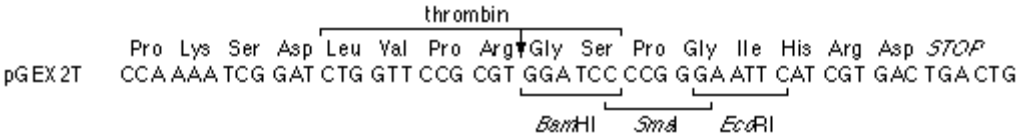




**pET-32a-c(+) cloning/expression region**

Sequence and map were extracted from Novagen.

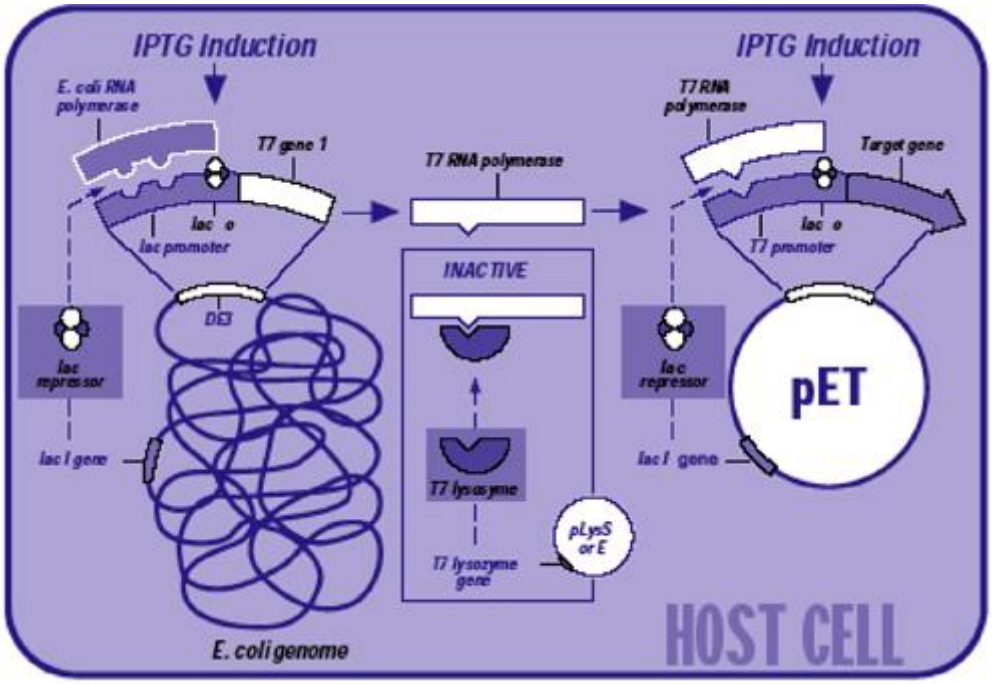
## Sequence and DNA map for pGEX2T




Sequence and map were extracted from snapgene. http://www.snapgene.com

**The schematic overview of pET expression system**




**The figure was extracted from Novagen.**

## Crystal Screen Formulation (Table was taken from Hampton Reserch)

| Reagent # | [Salt] | [Salt] units | Salt | [Buffer] | [Buffer] units | Buffer | pH | [Ppt 1] | [Ppt 1] units | Precipitant 1 | [Ppt 2] | [Ppt 2] units | Precipitant 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.02 | M | Calcium chloride dihydrate | 0.1 | M | Sodium acetate trihydrate | 4.6 | 30 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 2 | | | | | | | | 0.4 | M | Potassium sodium tartrate tetrahydrate | | | |
| 3 | | | | | | | | 0.4 | M | Ammonium phosphate monobasic | | | |
| 4 | | | | 0.1 | M | TRIS hydrochloride | 8.5 | 2.0 | M | Ammonium sulfate | | | |
| 5 | 0.2 | M | Sodium citrate tribasic dihydrate | 0.1 | M | HEPES sodium | 7.5 | 30 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 6 | 0.2 | M | Magnesium chloride hexahydrate | 0.1 | M | TRIS hydrochloride | 8.5 | 30 | % w/v | Polyethylene glycol 4,000 | | | |
| 7 | | | | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 1.4 | M | Sodium acetate trihydrate | | | |
| 8 | 0.2 | M | Sodium citrate tribasic dihydrate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 30 | % v/v | 2-Propanol | | | |
| 9 | 0.2 | M | Ammonium acetate | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 30 | % w/v | Polyethylene glycol 4,000 | | | |
| 10 | 0.2 | M | Ammonium acetate | 0.1 | M | Sodium acetate trihydrate | 4.6 | 30 | % w/v | Polyethylene glycol 4,000 | | | |
| 11 | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 1.0 | M | Ammonium phosphate monobasic | | | |
| 12 | 0.2 | M | Magnesium chloride hexahydrate | 0.1 | M | HEPES sodium | 7.5 | 30 | % v/v | 2-Propanol | | | |
| 13 | 0.2 | M | Sodium citrate tribasic dihydrate | 0.1 | M | TRIS hydrochloride | 8.5 | 30 | % v/v | Polyethylene glycol 400 | | | |
| 14 | 0.2 | M | Calcium chloride dihydrate | 0.1 | M | HEPES sodium | 7.5 | 28 | % v/v | Polyethylene glycol 400 | | | |
| 15 | 0.2 | M | Ammonium sulfate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 30 | % w/v | Polyethylene glycol 8,000 | | | |
| 16 | | | | 0.1 | M | HEPES sodium | 7.5 | 1.5 | M | Lithium sulfate monohydrate | | | |
| 17 | 0.2 | M | Lithium sulfate monohydrate | 0.1 | M | TRIS hydrochloride | 8.5 | 30 | % w/v | Polyethylene glycol 4,000 | | | |
| 18 | 0.2 | M | Magnesium acetate tetrahydrate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 20 | % w/v | Polyethylene glycol 8,000 | | | |
| 19 | 0.2 | M | Ammonium acetate | 0.1 | M | TRIS hydrochloride | 8.5 | 30 | % v/v | 2-Propanol | | | |
| 20 | 0.2 | M | Ammonium sulfate | 0.1 | M | Sodium acetate trihydrate | 4.6 | 25 | % w/v | Polyethylene glycol 4,000 | | | |
| 21 | 0.2 | M | Magnesium acetate tetrahydrate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 30 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 22 | 0.2 | M | Sodium acetate trihydrate | 0.1 | M | TRIS hydrochloride | 8.5 | 30 | % w/v | Polyethylene glycol 4,000 | | | |
| 23 | 0.2 | M | Magnesium chloride hexahydrate | 0.1 | M | HEPES sodium | 7.5 | 30 | % v/v | Polyethylene glycol 400 | | | |
| 24 | 0.2 | M | Calcium chloride dihydrate | 0.1 | M | Sodium acetate trihydrate | 4.6 | 20 | % v/v | 2-Propanol | | | |
| 25 | | | | 0.1 | M | Imidazole | 6.5 | 1.0 | M | Sodium acetate trihydrate | | | |
| 26 | 0.2 | M | Ammonium acetate | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 30 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 27 | 0.2 | M | Sodium citrate tribasic dihydrate | 0.1 | M | HEPES sodium | 7.5 | 20 | % v/v | 2-Propanol | | | |
| 28 | 0.2 | M | Sodium acetate trihydrate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 30 | % w/v | Polyethylene glycol 8,000 | | | |
| 29 | | | | 0.1 | M | HEPES sodium | 7.5 | 0.8 | M | Potassium sodium tartrate tetrahydrate | | | |
| 30 | 0.2 | M | Ammonium sulfate | | | | | 30 | % w/v | Polyethylene glycol 8,000 | | | |
| 31 | 0.2 | M | Ammonium sulfate | | | | | 30 | % w/v | Polyethylene glycol 4,000 | | | |
| 32 | | | | | | | | 2.0 | M | Ammonium sulfate | | | |
| 33 | | | | | | | | 4.0 | M | Sodium formate | | | |
| 34 | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 2.0 | M | Sodium formate | | | |
| 35 | | | | 0.1 | M | HEPES sodium | 7.5 | 0.8 | M | Sodium phosphate monobasic monohydrate | 0.8 | M | Potassium phosphate monobasic |
| 36 | | | | 0.1 | M | TRIS hydrochloride | 8.5 | 8 | % w/v | Polyethylene glycol 8,000 | | | |
| 37 | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 8 | % w/v | Polyethylene glycol 4,000 | | | |
| 38 | | | | 0.1 | M | HEPES sodium | 7.5 | 1.4 | M | Sodium citrate tribasic dihydrate | | | |
| 39 | | | | 0.1 | M | HEPES sodium | 7.5 | 2 | % v/v | Polyethylene glycol 400 | 2.0 | M | Ammonium sulfate |
| 40 | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 20 | % v/v | 2-Propanol | 20 | % w/v | Polyethylene glycol 4,000 |
| 41 | | | | 0.1 | M | HEPES sodium | 7.5 | 10 | % v/v | 2-Propanol | 20 | % w/v | Polyethylene glycol 4,000 |
| 42 | 0.05 | M | Potassium phosphate monobasic | | | | | 20 | % w/v | Polyethylene glycol 8,000 | | | |
| 43 | | | | | | | | 30 | % w/v | Polyethylene glycol 1,500 | | | |
| 44 | | | | | | | | 0.2 | M | Magnesium formate dihydrate | | | |
| 45 | 0.2 | M | Zinc acetate dihydrate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 18 | % w/v | Polyethylene glycol 8,000 | | | |
| 46 | 0.2 | M | Calcium acetate hydrate | 0.1 | M | Sodium cacodylate trihydrate | 6.5 | 18 | % w/v | Polyethylene glycol 8,000 | | | |
| 47 | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 2.0 | M | Ammonium sulfate | | | |
| 48 | | | | 0.1 | M | TRIS hydrochloride | 8.5 | 2.0 | M | Ammonium phosphate monobasic | | | |
| 49 | 1.0 | M | Lithium sulfate monohydrate | | | | | 2 | % w/v | Polyethylene glycol 8,000 | | | |
| 50 | 0.5 | M | Lithium sulfate monohydrate | | | | | 15 | % w/v | Polyethylene glycol 8,000 | | | |

# Crystal Screen 2 Formulation (Table was taken from Hampton Reserch)

| Reagent # | [Salt 1] | [Salt 1] units | Salt 1 | [Salt 2] | [Salt 2] units | Salt 2 | [Buffer] | [Buffer] units | Buffer | pH | [Ppt 1] | [Ppt 1] units | Precipitant 1 | [Ppt 2] | [Ppt 2] units | Precipitant 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.0 | M | Sodium chloride | | | | | | | | 10 | % w/v | Polyethylene glycol 6,000 | | | |
| 2 | 0.5 | M | Sodium chloride | 0.01 | M | Magnesium chloride hexahydrate | | | | | 0.01 | M | Hexadecyltrimethylammonium bromide | | | |
| 3 | | | | | | | | | | | 25 | % v/v | Ethylene glycol | | | |
| 4 | | | | | | | | | | | 35 | % v/v | 1,4-Dioxane | | | |
| 5 | 2.0 | M | Ammonium sulfate | | | | | | | | 5 | % v/v | 2-Propanol | | | |
| 6 | | | | | | | | | | | 1.0 | M | Imidazole pH 7.0 | | | |
| 7 | | | | | | | | | | | 10 | % w/v | Polyethylene glycol 1,000 | 10 | % w/v | Polyethylene glycol 8,000 |
| 8 | 1.5 | M | Sodium chloride | | | | | | | | 10 | % v/v | Ethanol | | | |
| 9 | | | | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 2.0 | M | Sodium chloride | | | |
| 10 | 0.2 | M | Sodium chloride | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 30 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 11 | 0.01 | M | Cobalt(II) chloride hexahydrate | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 1.0 | M | 1,6-Hexanediol | | | |
| 12 | 0.1 | M | Cadmium chloride hydrate | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 30 | % v/v | Polyethylene glycol 400 | | | |
| 13 | 0.2 | M | Ammonium sulfate | | | | 0.1 | M | Sodium acetate trihydrate | 4.6 | 30 | % w/v | Polyethylene glycol monomethyl ether 2,000 | | | |
| 14 | 0.2 | M | Potassium sodium tartrate tetrahydrate | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 2.0 | M | Ammonium sulfate | | | |
| 15 | 0.5 | M | Ammonium sulfate | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 1.0 | M | Lithium sulfate monohydrate | | | |
| 16 | 0.5 | M | Sodium chloride | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 2 | % v/v | Ethylene imine polymer | | | |
| 17 | | | | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 35 | % v/v | tert-Butanol | | | |
| 18 | 0.01 | M | Iron(III) chloride hexahydrate | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 10 | % v/v | Jeffamine M-600 | | | |
| 19 | | | | | | | 0.1 | M | Sodium citrate tribasic dihydrate | 5.6 | 2.5 | M | 1,6-Hexanediol | | | |
| 20 | | | | | | | 0.1 | M | MES monohydrate | 6.5 | 1.6 | M | Magnesium sulfate heptahydrate | | | |
| 21 | 0.1 | M | Sodium phosphate monobasic monohydrate | 0.1 | M | Potassium phosphate monobasic | 0.1 | M | MES monohydrate | 6.5 | 2.0 | M | Sodium chloride | | | |
| 22 | | | | | | | 0.1 | M | MES monohydrate | 6.5 | 12 | % w/v | Polyethylene glycol 20,000 | | | |
| 23 | 1.6 | M | Ammonium sulfate | | | | 0.1 | M | MES monohydrate | 6.5 | 10 | % v/v | 1,4-Dioxane | | | |
| 24 | 0.05 | M | Cesium chloride | | | | 0.1 | M | MES monohydrate | 6.5 | 30 | % v/v | Jeffamine M-600 | | | |
| 25 | 0.01 | M | Cobalt(II) chloride hexahydrate | | | | 0.1 | M | MES monohydrate | 6.5 | 1.8 | M | Ammonium sulfate | | | |
| 26 | 0.2 | M | Ammonium sulfate | | | | 0.1 | M | MES monohydrate | 6.5 | 30 | % w/v | Polyethylene glycol monomethyl ether 5,000 | | | |
| 27 | 0.01 | M | Zinc sulfate heptahydrate | | | | 0.1 | M | MES monohydrate | 6.5 | 25 | % v/v | Polyethylene glycol monomethyl ether 550 | | | |
| 28 | | | | | | | | | | | 1.6 | M | Sodium citrate tribasic dihydrate pH 6.5 | | | |
| 29 | 0.5 | M | Ammonium sulfate | | | | 0.1 | M | HEPES | 7.5 | 30 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 30 | | | | | | | 0.1 | M | HEPES | 7.5 | 10 | % w/v | Polyethylene glycol 6,000 | 5 | % v/v | (+/-)-2-Methyl-2,4-pentanediol |
| 31 | | | | | | | 0.1 | M | HEPES | 7.5 | 20 | % v/v | Jeffamine M-600 | | | |
| 32 | 0.1 | M | Sodium chloride | | | | 0.1 | M | HEPES | 7.5 | 1.6 | M | Ammonium sulfate | | | |
| 33 | | | | | | | 0.1 | M | HEPES | 7.5 | 2.0 | M | Ammonium formate | | | |
| 34 | 0.05 | M | Cadmium sulfate hydrate | | | | 0.1 | M | HEPES | 7.5 | 1.0 | M | Sodium acetate trihydrate | | | |
| 35 | | | | | | | 0.1 | M | HEPES | 7.5 | 70 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 36 | | | | | | | 0.1 | M | HEPES | 7.5 | 4.3 | M | Sodium chloride | | | |
| 37 | | | | | | | 0.1 | M | HEPES | 7.5 | 10 | % w/v | Polyethylene glycol 8,000 | 8 | % v/v | Ethylene glycol |
| 38 | | | | | | | 0.1 | M | HEPES | 7.5 | 20 | % w/v | Polyethylene glycol 10,000 | | | |
| 39 | 0.2 | M | Magnesium chloride hexahydrate | | | | 0.1 | M | Tris | 8.5 | 3.4 | M | 1,6-Hexanediol | | | |
| 40 | | | | | | | 0.1 | M | Tris | 8.5 | 25 | % v/v | tert-Butanol | | | |
| 41 | 0.01 | M | Nickel(II) chloride hexahydrate | | | | 0.1 | M | Tris | 8.5 | 1.0 | M | Lithium sulfate monohydrate | | | |
| 42 | 1.5 | M | Ammonium sulfate | | | | 0.1 | M | Tris | 8.5 | 12 | % v/v | Glycerol | | | |
| 43 | 0.2 | M | Ammonium phosphate monobasic | | | | 0.1 | M | Tris | 8.5 | 50 | % v/v | (+/-)-2-Methyl-2,4-pentanediol | | | |
| 44 | | | | | | | 0.1 | M | Tris | 8.5 | 20 | % v/v | Ethanol | | | |
| 45 | 0.01 | M | Nickel(II) chloride hexahydrate | | | | 0.1 | M | Tris | 8.5 | 20 | % w/v | Polyethylene glycol monomethyl ether 2,000 | | | |
| 46 | 0.1 | M | Sodium chloride | | | | 0.1 | M | BICINE | 9.0 | 20 | % v/v | Polyethylene glycol monomethyl ether 550 | | | |
| 47 | | | | | | | 0.1 | M | BICINE | 9.0 | 2.0 | M | Magnesium chloride hexahydrate | | | |
| 48 | | | | | | | 0.1 | M | BICINE | 9.0 | 2 | % v/v | 1,4-Dioxane | 10 | % w/v | Polyethylene glycol 20,000 |

Morpheus™ Additive OptiMax Kit Formulation (Molecular dimenstions)

| Tube | Additive/precipitant |
|---|---|
| 1 | 1 M Calcium chloride |
| 2 | 1 M Magnesium chloride |
| 3 | 1 M Sodium bromid |
| 4 | 1 M Sodium fluorid |
| 5 | 1 M Sodium iodid |
| 6 | 1 M Ammonium sulfat |
| 7 | 0.5 M Sodium hydrogen phosphate dibasic |
| 8 | 1 M Sodium nitrate |
| 9 | 100% v/v 1-butanol |
| 10 | 100% v/v 1,4-butanediol |
| 11 | 70% v/v 1,6-hexanediol |
| 12 | 100% v/v 1,2-propanediol |
| 13 | 50% v/v 1,3-propanediol |
| 14 | 100% v/v 2-propanol |
| 15 | 50% v/v Di-ethyleneglycol |
| 16 | 0.4 M Penta-ethyleneglycol |
| 17 | 40% v/v Tetra- ethyleneglycol |
| 18 | 100% v/v TEG (Tri- ethyleneglycol) |
| 19 | 1 M N-Acetyl - D-Glucosamine |
| 20 | 1 M L-Fucose |
| 21 | 1 M D-Galactose |
| 22 | 1 M D-Glucose |
| 23 | 1 M D-Mannose |
| 24 | 1 M D-Xylose |
| 25 | 1 M Ammonium acetate |
| 26 | 1 M Potassium sodium tartrate |
| 27 | 1 M Sodium citrate |
| 28 | 1 M Sodium formate |
| 29 | 0.5 M Sodium Oxamate |
| 30 | 1 M DL-Alanine |
| 31 | 0.2 M DL- Glutamic acid |
| 32 | 1 M Glycine |
| 33 | 1 M DL- Lysine HC |
| 34 | 0.2 M DL- Serine |
| 35 | 50% v/v PEG 550 MME |
| 36 | 50% v/v PEG 20,000 |
| 37 | 100% v/v Ethylene glycol |
| 38 | 50% v/v PEG 8000 |
| 39 | 100% v/v Glycerol |
| 40 | 50% v/v PEG 4000 |
| 41 | 100% v/v MPD |
| 42 | 50% v/v PEG 1000 |
| 43 | 50% v/v PEG 3350 |

# References

Abby, S. S. and Rocha, E. P. C. (2012) The Non-Flagellar Type III Secretion System Evolved from the Bacterial Flagellum and Diversified into Host-Cell Adapted Systems. *PLOS. Genetics* **8**(9): 1-15

Bernardini, M. L., Mounier, J., d'Hauteville, H., Coquis-Rondon, M. and Sansonetti, P. J. (1989) Identification of *icsA*, a plasmid locus of *Shigella flexneri* that governs bacterial intra- and intercellular spread through interaction with F-actin. *Proc. Natl. Acad. Sci. (USA)* **86**: 3867-3871

Blocker, A., Jouihri, N., Larquet, E., Gounon, P., Ebel, F., Parsot, C., Sansonetti P. and Allaoui, A. (2001) Structure and composition of the *Shigella flexneri* 'needle complex', a part of its type III secreton. *Mol. Microbiol.* **39**:652-663

Brandon, L.D. and Goldberg, M.B. (2001) Periplasmic Transit and Disulfide Bond Formation of the Autotransported *Shigella* Protein IcsA. *J. Bacteriol*. **183**: 951-958

Brandon, L.D., Goehring, N., Janakiraman, A., Yan, A.W., Wu, T., Beckwith, J. and Goldberg, M. B. (2003) IcsA, a polarly localized autotransporter with an atypical signal peptide, uses the Sec apparatus for secretion, although the Sec apparatus is circumferentially distributed. *Mol. Microbiol.* **50**: 45-60

Charles, M., Perez, M., Kobil, J. H. and Goldberg, M. B. (2001) Polar targeting of *Shigella* virulence factor IcsA in Enterobacteriacae and *Vibiro*. *Proc. Natl. Acad. Sci. (USA)* **98**:9871-9876

Cordes, F. S., Komoriya, K., Larquet, E., Yang, S., Egelman, E. H, Blocker, A. and Lea, S. M. (2003) Helical structure of the needle of the Type III secretion system of *Shigella flexneri*. *J. Biol. Chem*. **278**:17103-17107

Egile, C., Loisel, T. P., Laurent, V., Li, R., Pantaloni, D., Sansonetti, P. J., and Carlier, Marie-France. (1999) Activation of the CDC42 effector N-WASP by the *Shigella flexneri* IcsA protein promotes actin nucleation by Arp2/3 complex and bacterial actin–based motility. *J. Cell Biol.* **146**:1319-1332.

Fantuzzi, G. and Dinarello, C. A. (1999). Interleukin-18 and interleukin-1 beta: two cytokinesubstrates for ICE (caspase-1). *J. Clin. Immunol.* **19**(1): 1-11.

Hilbi, H., Moss, J. E., Hersh, D., Chen, Y., Arondel, J., Banerjee,S., Flavell,R. A., Yuan, J., Sansonetti,P. J. and Zychlinsky, A. (1998). Shigella-induced Apoptosis Is Dependent on Caspase-1 Which Bindsto IpaB. *J. Biol. Chem*. **273**(49): 32895-32900.

Jin, Q., Yuan, Z., Xu, J., Wang, Y., Shen, Y., Lu, W., Wang, J., Liu, H., Yang, J., Yang, F., Zhang, X., Zhang, J., Yang, G., Wu, H., Qu, D., Dong, J., Sun, L., Xue, Y., Zhao, A., Gao, Y., Zhu, J., Kan, B., Ding, K., Chen, S., Cheng, H., Yao, Z., He, B, Chen, R., Ma, D., Qiang, B., Wen, Y., Hou, Y. and Yu, J. (2002) Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. *Nucleic Acids Res.* **30**:4432-4441

Kenjale, R., Wilson, J., Zenk, S. F., Saurya, S., Picking, W. L., Picking, W. D. and Blocker, A. (2005) The needle component of the Type III Secretion of *Shigella* regulates the activity of the secretion apparatus. *J. Biol. Chem*. **280**:42929-42937.

Maurelli, A.T., Baurdry, B., d'Hauteville, H., Hale, T. L. and Sansonetti, P.J. (1985) Cloning of plasmid DNA sequences involved in invasion of HeLa cells by *Shigella flexneri*. *Infect. Immun.* **49**:164-171.

Mimuro, H., Suzuki, T., Suetsugu, S., Miki, H., Takenawa, T. and Sasakawa, C. (2000) Projilin is Required for Sustaining Efficient Intra- and Intercellular Spreading of *Shigella flexneri*. *J. Biol. Chem.* **275**: 28893-28901.

Ogawa, M., Youshimori, T., Suzuki, T., Sagara, H., Mizushima, N. and Sasakawa, C. (2005) Escape of intracellular *Shigella* from autophage. *Science* **307**: 727-731

Ogawa, M., (2006) A functional analysis of IcsB protein secreted via the type III secretion system of *Shigella. Nihon Saikingaku zasshi* 61:229-233

Purdy, G. D., Hong, M. and Payne, S. M. (2002) *Shigella flexineri* DegP Facilitates IcsA Surface Expression and Is Required for Efficient Intercellular Spread. *Infect. Immun.* **70**: 6355-6364

Ramesh, N., Anton, I., M., Hartwig, J.H. and Geha, R.S. (1997) WIP, a protein associated with the Wiskott-Aldrich syndrome protein, induces actin ploymerization and redistribution in lymphoid cells. *Proc. Natl. Acad. Sci. (USA)* **94**:14671-14676

Sansonetti, P. J. (2001) Rupture, invasion and inflammatory destruction of the intestinal barrier by *Shigella*, making sense of prokaryote–eukaryote cross-talks. *FEMS Microbiol. Rev.* **25**:3-14

Suzuki, T., Lett, Marie-Cleaire. and Sasakawa, C. (1995) Extracellular Transport of VirG Protein in *Shigella. J. Biol. Chem*. **270**: 30874-30880

Suzuki, T., Saga, S. and Sasakawa C. (1996) Functional analysis of shigella VirG domains essential for interaction with Vinculin and actin-based motility. *J. Biol. Chem*. **271**:21878-21885

Suzuki, T., Miki, H., Takenawa, T. and Sasakawa, C. (1998) Neural Wiskott-Aldrich syndrome protein is implicated in the actin-based motility of *Shigella flexneri*. *EMBO J.* **17**:2767-2776

Suzuki, T., Miki, H., Takenawa, T. and Sasakawa. C. (2002) Neural Wiskott-Aldrich syndrome protein (N-WASP) is the specific ligand for *Shigella* VirG among the WASP family and Determines the host cell type allowing actin–based spereading. *Cell. Microbiol.* **4**: 223-233

Steinhauer, J., Agha, R., Pham, T., Varga, A.W., and Goldbaerg, M. B. (1999) The unipolar *Shigella* surface protein IcsA is targeted directily to the bacterial old pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Mol. Micorbiol*. **32**:367-377.

Tripathi, P., Beaussart, A., Andre G, Rolain, T., Lebeer, S., Vanderleyden, J., Hols, P. and Dufrêne, Y.F. (2012). Towards a nanoscale view of lactic acid bacteria.*Micron* **43**(12):1323-1330.

Tseng, T.-T., *Tyler*, B. M. and Setubal, J. C. (2009). Protein secretion systems in bacterial-host associations, and theirdescription in the Gene Ontology. *BMC Microbiol.* **9** (Suppl 1): S2.

Volkman, B. F., Prehoda, K. E., Scott, J.A., Peterson, F. C., and Lim, W. A. (2002) Structure of the N-WASP EVH1 Domain-WIP Complex: Insight into the Molecular Basis of Wiskott-Aldrich Syndrome. *Cell*. **111**: 565-576.

Venkatesan, M. M., Goldberg, M. B., Rose, D. J., Grotbeck, E. J., Burland, V., and Blattner, F. R. (2001) Complete DNA Sequence and Analysis of the Large Virulence Plasmid of *Shigella flexneri. Infect. Immun.* **69**: 3271-3285

Wei, J., Goldberg, M. B., Burland, V., Venkatesan, M. M., Deng, W., Fournier, G., Mayhew, G. F., Plunkett III, G., Rose, D. J., Darling, A., Mau, B., Perna, N. T., Payne, S. M., Runyen-Janecky, L. J., Zhou, S., Schwartz, D. C. and Blattner, F. R. (2003) Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect. Immun.* **71**: 2775-2786.

Wing, Helen J., Yan, A.. W., Goldman,S. R., and Goldberg, M. B. (2004). Regulation of IcsP, the Outer Membrane Protease of the Shigella Actin Tail Assembly Protein IcsA, by Virulence Plasmid Regulators VirF and VirB. *J. Bacteriolo.* **186**(No. 3): 699-705,.

Yip, C. K. and Strynadka, N. C. J. (2006). New structural insights into the bacterial type III secretion system. Trends in Biochemical Sciences **31**(4): 223-230.