CHAPTER IV

Searching for virulent *Burkholderia pseudomallei* genes by immunoscreening the ZAPII expressed genomic library

Materials and methods

4.1 Bacterial strains

The *E.coli* strain and genotype used in this study were follows:

E.coli strain	Genotype
1. XL 1- BIUE MRF′ Kan strain	Δ(*mcrA*) 183Δ (*mcr CB- hsdSMR- mrr*) 173 end A1 sup E44 thi-1 rec A1 *gyr A* 96 rel A1 *lac* (F′ pro AB *lac* I^q ZΔM15 Tn 5 (kan^γ)]
2. XLOLR	Δ (*mcr A*) 183 Δ (*mcr CB- hsd SMR- mrr*) 173 end A1 thi-1 rec A1 *gyr A*96 rel A1 *lac* [F′ pro AB *lac* I^q ZΔM15 Tn 10 (Tet^γ)] Su^- (non- suppressing) λ^R (lambda resistant)
3. BL21 (DE3) plysS	F′, *ampT*, *hsd* S_B (r^-_8, m^-_8), *dcm*, *gal*, (DE3), plys S, Cm^R
4. DH 5 α	ϕ 80 d lac ZΔM15, rec A1, end A1 *gyr* A96, thi-1, *hsd* R17 (r_k-, m_k-), sup E44, rel A1, deo R, Δ(*lac zya-arg* F) U169

4.2 Melioidosis sera and normal healthy control sera

Five melioidosis serum (no 1,5 ,9, 10 and 15) were obtained from hospital in-patients with positive blood cultures for *B.pseudomallei*. Five normal serum samples were obtained from healthy university students.The sera were kept at – 80 ^{O}C freezer until use.

4.3 ZAPII expression genomic library construction

4.3.1 DNA preparation

Bacterial chromosomal DNA was isolated from *B. pseudomallei* by using the Wizard genomic DNA purification kit (Promega) as described in Chapter III. The DNA was partially digested with *Sau3A* to optimize production of 1-2 kb fragments as described in the standard protocol (91). The reaction volume of 20 microlitres contained 10 microlitres of *B. pseudomallei* DNA, 2 microlitres of 1 unit of *Sau3A*, 2 microlitres of 10X BSA, 2 microlitres of buffer and 4 microlitres of distilled water. The tube was incubated at 37^{0}C for 30 minutes and then inactivated at 70^{0}C for 10 minutes. The digested DNA was run on 1% agarose gel electrophoresis as described in Chapter III. The bands were visualized on a long wave UV transillumuntor (Gelman), and then the bands of 1-2 kb were excised and put in a microcentrifuge tube. Extraction of DNA fragments from agarose gels was achieved using the Qiaex II DNA Gel extraction kit(Qiagen) , according to the manufacturer's instructions. Three hundred microlitres of solubilisation buffer QX1 per 100 mg of gel were added to the microcentrifuge tube containing the cut gel agarose. Then 10 microlitres of QIAEX II gel matrix was added to the tube and incubated at 50^{0}C for 10 minutes. The mixture was vortexed every 2 minutes. Then the tube was centrifuged at 10,000g for 30 seconds and the supernatant carefully removed with a pipette. The pellet was resuspended in 500 microlitres of buffer QX I and then the tube was centrifuged at 10,000g for 30 seconds. The pellet was collected and

resuspended in 500 microlitres of buffer PE 1, the tube again centrifuged at 10,000g for 30 seconds, and the washing step with PE1 repeated. The pellet was air dried until it turned a white color. Twenty microlitres of distrilled water was added to the pellet and the mixture was incubated at room temperature for 5 minutes. The tube was centrifuged at 10,000g for 30 seconds and the supernatent was collected for the next step of ligation.

4.3.2 Ligating the insert

Five mililitres of the ligation mixture, consisting of 1 microlitre of ZAP Express^{R} vector (Stratagene) (digested with BamH1 and CIA pretreated at 1 microgram/microlitre), 2 microlitres of DNA insert (1-2 kb fragment), 0.5 microlitres of ligase buffer (10X), 0.5 microlitres of 10mM rATP (pH7.5) and distilled water. The reaction was started by adding 2 units of T4 DNA ligase and then the tube was incubated at 4^{0}C overnight.

4.3.3 Packaging protocol

4.3.3.1 Preparing the host bacteria

The isolated colony of *E.coli* XL1BlueMRF' was inoculated in 50 ml of LB broth containing 10mM $MgSO_4$ and 0.2% (w/v) maltose and then the tube was incubated at 30^{0}C with shaking overnight. The pellet was collected by centrifugation and resuspended in sterile 10mM $MgSO_4$ at OD_{600} of 0.5.

4.3.3.2 Packaging

The protocol of ZAP Express^{R} predigested gigapack^{R}III gold cloning kit (Stratagene) was followed. One tube of packaging extract was removed from the –80 freezer and kept on dry ice. The tube was quickly thawed by holding the tube between the fingers until the contents of the tube (25 microlitres) just began to thaw. Two microlitres of the ligated DNA from previous ligation mixture was immediately added to the tube and mixed well by using a pipette. The tube was quickly spun for 3-5 seconds and incubated at 25^{0}C for 2 hours. Then 500 microlitres of SM buffer (appendix A) was added to the tube. Twenty microlitres

of chloroform was added to the tube with gentle agitation, and then the phage solution was ready for the titering and amplification steps.

4.3.4 Titering

The phage in SM buffer was serially diluted to 10^2 and 10^4 dilution and one microlitre of this diluted phage mixture was mixed with 200 microlitres of fresh host cell *E. coli* XL1BlueMRF'. The tube was incubated at 37^0C for 15 minutes. Three ml of 48^0C top agar (appendix A) was added to the tube with gentle mixing and immediately poured on 90mm NZY agar plate. The plate was allowed to harden at room temperature for 10 minutes and placed in a 37^0C incubater overnight. The number of plaques were counted and the concentration of the library (pfu/ml) was determined according to the dilution of the phage.

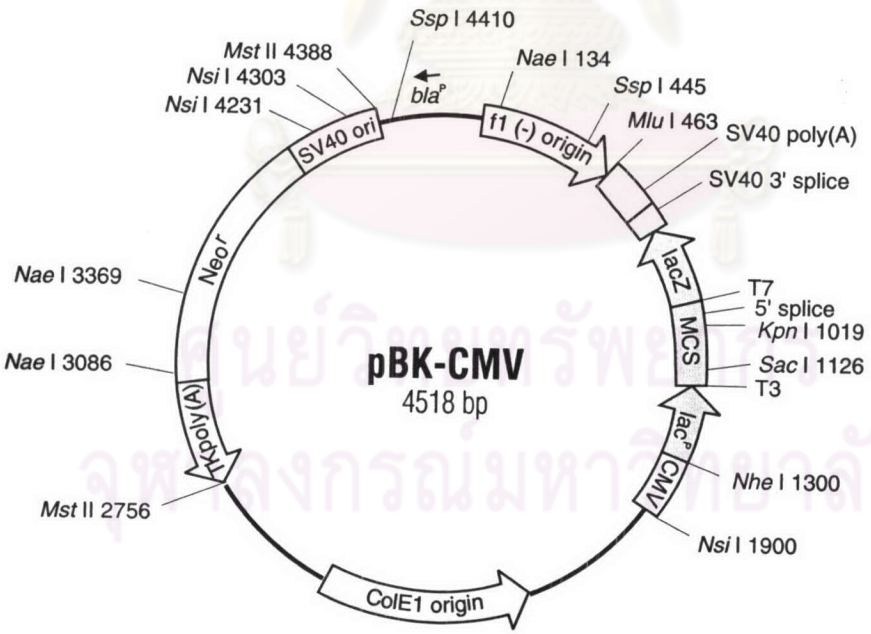


Figure 13 Schematic representation of pBK-CMV

4.3.5 Amplifying the ZAP express library

The phage library suspension containing ~5X10^{4}pfu of bacteriophage was mixed with 600 microlitres of fresh host cell *E. coli* XL1 Blue MRF'. A total of 20 reaction tubes were used to amplify 1X10^6 plaques. The tubes were incubated at 37^{0}C for 15 minutes and then 6.5 ml of 48^{0}C NZY top agar was poured into each tube with gentle inversion. Then the suspensions were spread evenly onto freshly poured 150 mm NZY agar plates. The plates were allowed to harden at room temperature for 10 minutes and incubated at 37^{0}C for 6 hours. Ten ml of SM buffer was poured on each plate and incubated at 4^{0}C overnight. Then the bacteriophage suspensions were pooled on one tube and another 2 ml of SM buffer was poured on the plates and the suspensions were collected. Each suspension was centrifuged at 500g for 10 minutes. The amplified library supernatant was collected with chloroform 0.3% (v/v) final concentration and stored at 4^{0}C. The library was aliquoted in 7% DMSO at –80^{0}C. The phage concentration in the amplified library was determined by titering as previously described .

4.3.6 Mass excision protocol

The isolated colony of *E. coli* XL1 Blue MRF' was inoculated in 20 ml of NZY broth containing 0.2%(w/v) maltose and 10mM $MgSO_4$ while *E. coli* XLOR strain was inoculated in 20 ml of NZY broth. Both tubes were incubated at 30^{0}C with shaking overnight. The pellets were collected and resuspended in 10mM Mg SO_4 to an OD_{600} of 1.0 (8X10^{8}cells/ml). The reaction tube consisted of 10^{7}pfu of the lambda phage, 10^8 *E. coli* XL1 Blue MRF' and 10^9 pfu of ExAssist helper phage. The tube was incubated at 37^{0}C for 15 minutes to allow absorption. Then 2 ml of the NZY broth was added to the tube and incubated at 37^{0}C with shaking for 3 hours. The tube was heated to 65^{0}C for 20 minutes and centrifuged at 4,000g for 10 minutes. The supernatant was transferred to a new

tube. One hundred microlitres of the supernatant was mixed with 200 microlitres of *E. coli* XLOLR and incubated at 37^{0}C for 15 minutes. Then 40 microlitres of the 5X NZY broth was added to the culture tube and incubated at 37^{0}C for 90 minutes. One hundred microlitres of the cell suspension was plated on an LB-kn agar plate and incubated at 37^{0}C overnight. Isolated colonies were randomly selected for phagemid preparation, restriction analysis and gel electrophoresis.

4.4 Pooled melioidosis serum absorption.

The serum was absorbed with *E. coli* XL1 Blue MRF' lysate according to the stratagene protocol with some modifications.

4.4.1 *E. coli* lysate preparation

The isolated fresh colony of *E. coli* XL1 Blue MRF' was inocutated in 200 ml of LB broth and incubated overnight at 37 ^{0}C in a shaking incubator. The pellet was collected by centrifugation at 4000g for 10 minutes. Then the pellet was resuspended in 3 ml of Tris-Cl (TBS) (Appendix A) and kept in a freezer at –70^{0}C for 10 minutes. The freeze and thaw process was repeated 6 times. Then the suspension was sonicated by using a sonicator at high speed for 20 seconds, again repeated 6 times, when the homogeneous lysate of *E. coli* was collected and kept at -70^{0}C until use.

4.4.2 Preparation of *E. coli* coated nitrocellulose membrane

Three hundred microlitres of the *E. coli* lysate was mixed with 1 ml of Tris-Cl-Tween(TBST) (Appendix A). Then two small (90 mm) piecesof nitrocellulose membrane were saturated with the lysate, put on the whatman paper and air dried at room temperature. After washing with TBST 2 times, the membranes were blocked in blocking buffer (TBST+2% casein) for one hour at 25 ^{0}C. Then the membranes were washed with TBST 3 times and immediately transferred to the next step.

4.4.3 Serum absorption step

Forty microlitres of serum from each of 5 melioidosis patients were pooled together in one tube. Then 1.8 microlitres of TBST was mixed with the pooled serum and poured on one of the coated nitrocellulose membranes. The membrane was gently rotated at 37^0C for 10 minutes and discarded from the serum. Then the absorbed pooled serum was transferred to the second coated nitrocellulose membrane, which was gently rotated at 37^0C for 10 minutes. Then the pooled melioidosis serum was collected and diluted to 1:2000 with TBST-0.4% casein and kept in the fridge with 0.05% sodium azide at its final concentration. The pooled normal serum was processed using the same procedure.

4.5 Immunoscreening of the ZAP2-expressed genomic library

The libray was plated on a 150 mm NZY plate at 3000 plaque forming units with XL1 Blue MRF' cells in top agar and incubated in a 42^0C incubator for 5 hours. Then a 10 mM IPTG (Sigma) treated nitrocellulose membrane (Amersham) was put on the plate and incubated at 37^0C overnight. After washing 2 times with TBST, the membrane was blocked in blocking buffer (TBST-2% casein) for one hour at 25^0C. Then the membrane was washed 3 times with TBST and incubated with the pooled absorbed melioidosis serum at 1:2000 for 2 hours at 25^0C. After washing 3 times with TBST, the membrane was incubated with proteinA-alkaline phosphatase conjugated (Sigma, 0.250 microgram/ml) for one hour at 25^0C. Then the membrane was washed with TBST 3 times and rinsed with 1X assay buffer (Appendix A). The substrate CDP-Star (Applied biosystem, Adamantane-1,2-dioxetane phosphate) was added to the membrane and the chemiluminescent reaction was detected by using X-ray film in a dark room. Then the positive plaque was picked up and put on the SM buffer. The released plaques were plated on NZY plate the same as the above

protocol. The immunoscreening step was repeated as previously described. The confirmed positive plaques were kept in SM buffer at 4^0C.

4.6 Plaque dot blot analysis

The positive plaques were diluted with SM buffer to 200-4000 plaque-forming units per ml. Then 5 microlitres of the diluted plaque (1-20 plaque-forming units) were spotted on the surface of an NZY plate with XL1 Blue MRF' cells in top agar. The plate was allowed to stand at 25 ^{0}C for 15 minutes and then incubated at 42^0C incubator for 5 hours. Then a 10 mM IPTG treated nitrocellulose membrane was put on the plate and incubated at 37^0c overnight. The membrane was processed as previously described immunoscreening protocol with both pooled absorbed melioidosis serum and pooled absorbed normal control serum. The positive plaque that reacted only with pooled absorbed melioidosis serum was collected for further characterization.

4.7 Clone characterization

4.7.1 Single-clone excision

An isolated fresh colony of either *E. coli* XL1 Blue MRF' or *E. coli* XLOR was inoculated in 10 ml of NZY broth plus 0.2% (w/v) maltose and 10mM $MgSO_4$ and NZY broth, respectively. The cultures were incubated at 30 ^{0}C with shaking overnight. Then the pellets were collected by gently spinning at 1,000g for 10 minutes and resuspended at an OD_{600} of 1.0 in 10 mM $MgSO_4$. The following components were combined in one tube: 200 microlitres of XL1 Blue MRF' cells, 250 microlitres of positive plaque and 1 microlitre of the ExAssist helper phage (>1X 10^6 pfu/ microlitre). The tube was incubated at 37^0C in a water bath for 15 minutes. Then 3 ml of NZY broth was added to the tube and further incubated at 37^0C with shaking for 3 hours. The tube was heated at 65^0C for 20 minutes and then gently spun at 1000g for 15 minutes. The supernatant

containing the excised pBK-CMV phagemid vector was packaged as filamentous phage particles. One hundred mililitres of the supernatant was mixed with 200 microlitres of *E. coli* XLOR in one tube and incubated at 37^{0}C for15 minutes. Then NZY broth volumed 300 microlitres was added and incubated at 37^{0}C for 45 minutes. One hundred microlitres of the supension was plated on an LB agar plate containing kanamycin (50 micrograms/ml) and incubated at 37^{0}C overnight. The colonies on the plate contained the pBK-CMV double-stranded phagemid vector with the clone DNA insert. The positive clone was restreaked on a new LB-kanamycin agar plate and incubated at 37^{0}C overnight. The positive clones were suspended in 10 % glycerol broth and kept in a freezer at –70^{0}C.

4.7.2 pBK-CMV extraction

The phagemid pBK-CMV was extracted from the host cell (*E. coli* XLOR) by using the alkaline lysis method [74]. The isloated fresh colony of phagemid clone was inoculated in 4.5 ml of LB broth containing kanamycin (50 micrograms/ml) and incubated at 37^{0}C with shaking overnight. The pellet was collected by centrifugation at 2,000g for 10 minutes and resuspended in 200 microlitres of GET buffer pH 8.0 (appendix A) by pipetting up and down. Then 300 mililitres of 3.0 M potassium acetate (appendix A) was added to the suspension and the tube was inverted several times. The tube was microcentrifuged at 10,000 g for 10 minutes at room temperature. The supernatant was carefully transferred to a new clean microcentrifuge and Rnase (Dnase free) was added to the tube at a final concentration 20 micrograms/ml and incubated at 37^{0}C for 20 minutes. The supernatant was extracted with an equal volume of chloroform by inverting the tube for 30 seconds. Then the tube was again centrifuged at 10,000 g for 10 minutes and the upper aqueous phase was transferred to a new microcentrifuge. The supernatant was extracted with an equal volume of 100% isopropanol by inversion for 30 seconds. Then the

tube was centrifuged at 10,000 g for 10 minutes and the supernatant was completely discarded by aspiration. Five hundred microlitres of 70% ethanol was added to the phagemid DNA pellet, the tube was inverted several times and then centrifuged at 10,000g for 10 minutes. The DNA was collected and dried under vacuum for 3 minutes. Then the DNA was dissolved in 100 microlitres of deionized distrilled water and kept in the –80 ^{O}C freezer. This phagemid template preparation was used for restriction digest analysis, transformation to another host XL1blueMRF' for protein expression and as a primary template for the DNA sequencing step.

4.7.3 pBK-CMV transformation to *E. coli* XL1 Blue MRF'

E. coli XL1 Blue MRF' competent cells were prepared by using a chemical treatment [74] with some modifications. An isolated fresh colony of *E. coli* was inoculated in 10 ml of LB broth and incubated at 37^{0}C with shaking for 3 hours (OD_{450} of 0.4). Then the pellet was collected by centrifugation at 4,000g for 5 minutes and resuspended in 5 ml of ice-cold 0.1 M $CaCl_2$. The tube was incubated on ice for 20 minutes and centrifuged at 4,000g for 5 minutes. The pellet was resuspended in 0.1 ml of ice-cold $CaCl_2$ and incubated on ice for 2 hours. The phagemid was transformed to an *E. coli* competent cell by using a heat shock method (91). One microlitre (50 nanograms) of phagemid was added to the competent cell and incubated on ice for 5 minutes. The tube was moved to a water bath (37^{0}C) and incubated for 5 minutes. Then 1 ml of the LB broth was added to the tube and incubated at 37^{0}C for 2 hours. One hundred microlitres of the cell suspension was spread on LB –Kn agar and incubated at 37^{0}C overnight. The isolated colony was picked up and suspended in a 10% glycerol broth. The stock of a positive clone was kept at –80 ^{O}C freezer for the protein expression study.

4.7.4 DNA electrophoresis

The phagemid template was run on 1% agarose gel with the same protocol as the plasmid in Chapter III.

4.7.5 Phagemid DNA sequencing

The phagemid template was prepared as described in (4.6.2). 100 microlitres of DNA was mixed with 55 microlitres of 20% PEG/2.5M NaCl (Appendic A) and incubated on ice for 20 minutes. Then the pellet was collected by centrifugation at 12,000g for 15 minutes at 4^0C and rinsed with 500 microlitres of 70% ethanol. The pellet was air dried and resuspended in 25 microlitres of distrilled water. The purified DNA was transferred to the sequencing step.

4.7.6 Sequencing reaction

Sequencing reaction utilised a Big-Dye-deoxy terminator cycle sequence kit (PRISM). Reactions were set up in 0.5 ml thin soft walled microfuge tubes. Each sequencing reaction contained 8 microlitres premix, 200-400 nanograms of purified plasmid template, 3.2 pmol of primer T3 or T7, and the addition of sterile water to a final reaction volume of 20 microlitres. The reaction mixtures were then placed in a thermal cycler (PCR sprint, Hybaid) and cycled under the following conditions: preheat at 94^0C 1 minute, 28 cycles of 95^0C 30 seconds: 50^0C 15 seconds: 60^0C 4 minutes.

4.7.7 Extraction of sequencing reactions

On completion of the sequencing reaction, the reaction mixture was removed and placed in a sterile 1.5 ml microfuge tube containing 2.5 microlitres of 3M NaOAc pH5.5 and 50 microlitres of 100% ethanol. This was then vortexed briefly and incubated at 25^0C for 15 minutes and centrifuged at 14,000g for 20 minutes at 4^0C. The fluorescent labeled extension products were collected and washed with 70% ethanol and then dried under vacuum.

4.7.8 Sequence determination

Following completion of the sequencing reactions, the extension products were resolved on an applied Biosystems (ABI)373A Automated sequencer (Perkin Elmer). The nucleotide sequence was edited by using SeqEd-version1.03.

4.8 Protein expression study

The isolated colony of the positive clone was inoculated in 2 ml of 2YT broth containing kn 50 micrograms/ml (Appendix A) and incubated at 37^0C with shaking overnight. Then one ml of the overnight broth was transferred to one ml of new 2YT broth containing kn 50 micrograms/ml and IPTG 10 mM. The tube was incubated at 37^0C with shaking for 2 hours, then the pellet was collected and resuspended in 50 microlitres of 5X sample buffer. The tube was boiled for 3 minutes and centrifuged at 13,000 g for 15 minutes at 4^0c. The protein in the supernatant was analyzed on 12.5% SDS-PAGE as described in Chapter III.

4.9 Western blot analysis

The expressed proteins of positive clones in acrylamide gel were transferred to the nitrocellulose membrane by using a semi-dry electrophoretic transfer cell (Trans-blot^R SD, Biorad) as described in Chapter III.

4.10 DNA analysis using a bioinformatics program

The nucleotide sequence of inserted DNA of positive clones obtained from both T3 and T7 primer sequence reactions were searched for similarity with GenBank data by using the BLAST program (and blasted against the *B. pseudomallei* genome at the Sanger web site (http://www.sanger.ac.uk). The whole sequence of the inserted DNA of each clone was obtained and then searched for similarity with genbank data by using BLAST. The hit position on

the *B. pseudomallei* genome was extended in both directions (right and left) for around 2000 bps and cut for further DNA analysis. The whole insert DNA of each clone was analyzed for inframed open reading framed (ORF) and genes arrangment by using Macvector V7.0.1 (Accelrys company) and GeneMark programn[80]. The antigenic index of each gene was analyzed by using Hopp and Woods scales (Macvector V7.0.1).The secondary protein structure was predicted by using Chou Fasman(Macvector V7.0.1) and PHD program[88]. The transmembrane protein was predicted by using PHDhtm[88] and Kyte/Doolittle scale (Macvector V7.0.1). The protein analysis of each gene was analyzed by using Macvector V7.01.

4.11 Strategies to isolate the *in vivo* expressed genes

One ORF found in the inserted DNA basically indicated one *in vivo* expressed gene. If there was more than one ORFs in the inserted DNA, the estimated molecular weight of expressed protein in each clone was compared with the predicted molecular weight of inframed ORF in the inserted DNA sequence. One expressed protein band may contain 2 ORFs according to the molecular weight of the expressed protein. If the inframed ORF was an incomplete gene, the full sequence of the gene was searched for from the extended sequence in both directions (right and left) of each positive clone in the *B.pseudomallei* genome by using BLAST and GeneMark programn.

4.12 Multiple sequence alignment and phylogenetic tree construction

The aligments were generated with the Clustal W program (Macvector V7.0.1). The phylogenetic tree was constructed by distance based clustering algorithm, specifically neighbor joining method (Macvector V7.0.1)

4.13 Amplification of the gmhA and *wcbM* genes from the *B. pseudomallei* chromosome by using polymerase chain reaction (PCR)

The full gene sequences of *gmhA* (594 bp) and *wcbM* (693bp) were amplified from the whole genome of *B. pseudomallei* by using PCR [74]. The primers used in this PCR reaction were described in table 6. The *B. pseudomallei* DNA was isolated as described in 4.2.1. 50 microlitres of the PCR reaction mixture consisted of 1 microliter of *B. pseudomallei* DNA template (10-100 nanograms), 5 microliters of PCR buffer(10X) , 3 microliters of $MgCl_2$(25mM) , 2 microliters of primer each 50pmol , 2 microliters of dNTP(5 mM each), 1 microliter of Taq polymerase(2.5 unit) and distrilled water. The reaction mixture tubes were then placed in a thermal cycle (PCR Sprint, Hybaid) and cycled under the following conditions: 95^0C for 15 seconds, 45^0C for 30 seconds, and finally 74^0C for 90 seconds.The machine was operated for 30 cycles. The reaction mixtures were loaded on 1% agarose gel and electrophoresed as described in Chapter III. The DNA band was visualized by using long wave UV in the dark room and cut to a new clean microfuge tube. The DNA was extracted from the gel by using QIAEX II (Qiagen) as described in 4.2.1. The DNA was digested with two restriction endonuclease enzymes, *BamH1* and *EcoR1*, as described in Chapter III. The digested DNA of both genes was loaded on 1% agarose gel and electrophoresed. The single band at the position 594 bp and 694 bp were cut under long wave UV in the dark room and put on a new microfuge tube. The DNA was extracted from the gel by using QIAEXII (Qiagen). The purified digested DNA genes of *gmhA* and *wcbM* were kept at 4^0C for the ligation step. The expression plasmid vector pRSETA (Invitrogen) was isolated from *E. coli* BL21 DE3 as described in Chapter III. The vector was digested with two restriction endonulease enzyme *BamH1* and *EcoR1* as described in chapter III. The digested vector was loaded on 1% agarose gel and electrophoresed.

The single band was cut under long wave UV in the dark room. Then the DNA was extracted from the gel by using QIAEX II (Qiagen). The vector was ligated with the digested DNA of both genes in 20 microlitres of the ligation mixture as described in the standard protocol in Chapter III. The reaction tubes were incubated at 16^0C overnight. Ten microliters of the ligated DNA was mixed with 100 microliters of *E. coli* DH5 α competent cells and then incubated on ice for 30 minutes. The tubes were moved to a 42^0C incubator for 90 seconds and incubated on an ice box for another 2 minutes. Then one ml of SOC broth was added to the tube and incubated at 37^0C with shaking for 45 minutes. One hundred microliters of the cell suspension was plated on LB agar containg 50 micrograms/ml of ampicillin and incubated at 37^0C overnight. An isolated colony was randomly picked up for plasmid isolation and restriction digestion analysis. Colonies that contained the insert DNA of similar size with the *gmhA* and *wcbM* genes were moved to further steps. The plasmid from the positive clone was transformed to a new host competent cell *E. coli* BL21DE3. The *gmhA* and *wcbM* clones in vector pRSET A were used for protein expression study and western blot analysis against both pooled absorbed melioidosis serum and pooled absorbed normal serum as described in 4.9

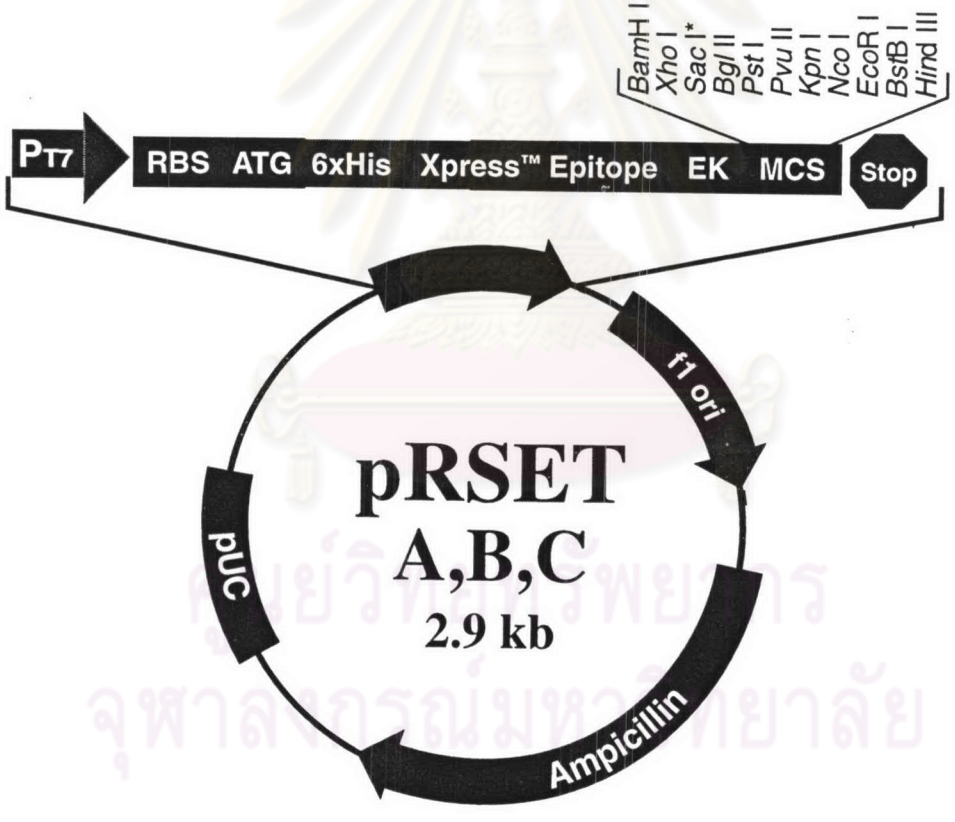


Figure14 Schematic representation of pRSET A

Table 5 Sequence of oligonucleotide primer used to amplify the *gmhA* and *wcbM* from *Burkholderia pseudomallei* genome

gene	Primer	sequence (5′ – 3′)
gmhA	gmhA-forward	CGG GAT CCA TGG AGA ATC GCG AAT TGA
	gmhA-reverse	GGA ATT CTC ACT GCT TCC CGA AAAT
wcbM	wcbM-forward	CGG GAT CCA TGA GAG AAG CGA TCA TCTT
	wcbM-reverse	GGA ATT CTC ACT TGC AGA TTC CGGA

Nucleotides underlined in each primer sequence show the positions of the restriction endonuclease sites incorporated to facillitate cloning

Results

1. ZAP II expressed genomic library construction

The titre of the amplified ZAP II library was ~$2.5x10^9$ pfu/ml.The phagemids obtained from the mass excision were demonstrated in figure15. The phagemids were cut by using endorestriction enzyme *SacI* and *KpnI* as illustrated in figure 16. The phagemids contained different size of DNA insert of *B.pseudomallei* ranged from 1-2 kb.

2. Immunoscreening of the ZAP II expression library

The pooled absorbed melioidosis serum has been demonstrated to react with only the *B.pseudomallei* whole lysate not E.coli XL1Blue MRF' lysate by using dot blot analysis . Fourteen positive clones were isolated from immunoscreening of 30,000 plaques . All of them reacted only with pooled absorbed melioidosis serum and not pooled absorbed normal serum by using plaque dot blot assay as in figure 17. The phagemid of the positive clone were digested with restriction endonuclease *SacI* and *KpnI* as shown on figure 18. The clone insert size ranged from 1-2.5 kb. DNA sequence analysis revealed some clones that had a similar DNA sequence with others;these could be divided into 6 groups as Bp1 (1,2,8,13), Bp6 (6,18), Bp9, Bp3, Bp7 (7,19), Bp5 (4,5,20,10).The list of suspected *in vivo* expressed genes isolated from this immunoscreening experiment is illustrated in table 7.

Table 7 List of suspected *in vivo* expressed genes isolated from the immunoscreening experiment

clone	insert size (bp)	average % G+C	in vivo expressed genes	size (bp)	putative function
Bp 1	1378	59.29 %	*gmh A*	594	phosphoheptose isomerase
Bp 3	1738	69.68 %	Conserved hypothetical protein	462	hypothetical protein
			hypothetical signal peptide protein	738	hypothetical protein
Bp 6	2500	70.16 %	*bip D*	933	cell invasion protein
Bp 7	2303	67.54 %	transmembrane protein	813	membrane protein
Bp 9	1303	63.69 %	*gro EL*	1496	chaperonine protein

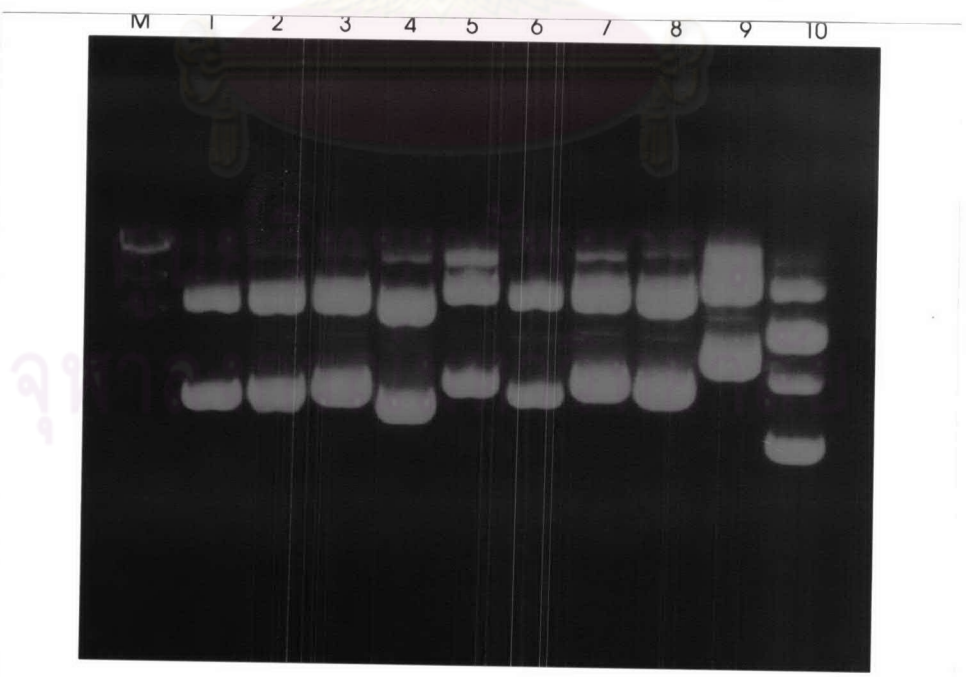


Figure 15 Recombinant phagemid pBK-CMV isolated from the mass excision

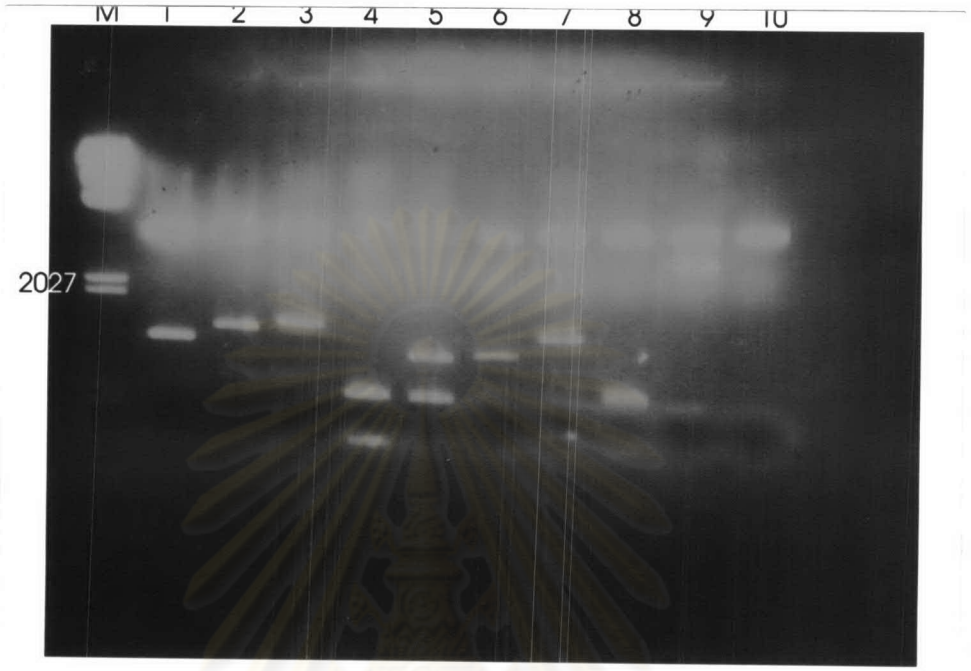


Figure 16 Restriction analysis of recombinant phagemid pBK-CMV isolated from the mass excision. Agarose gel(1%w/v) electrophoresis of recombinant phagemid digested with *kpnI* and *sacI*.

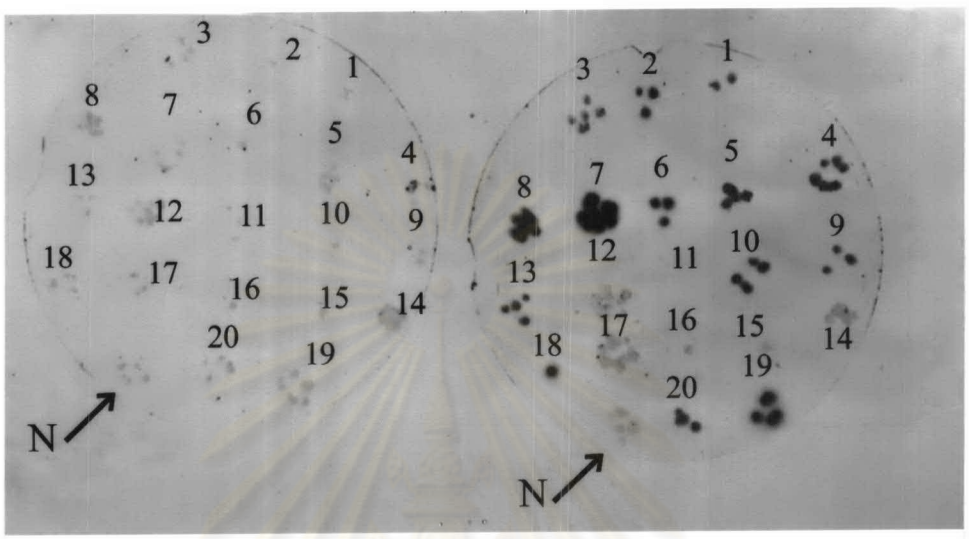


Figure 17 Plaque dot blot analysis.

Fourteen positive clones were confirmed to react only with pooled absorbed melioidosis serum(B) and not pooled absorbed normal serum(A).

The position of each clone was put in the plate.

N = negative control clone that did not react with serum in the immunoscreenin.

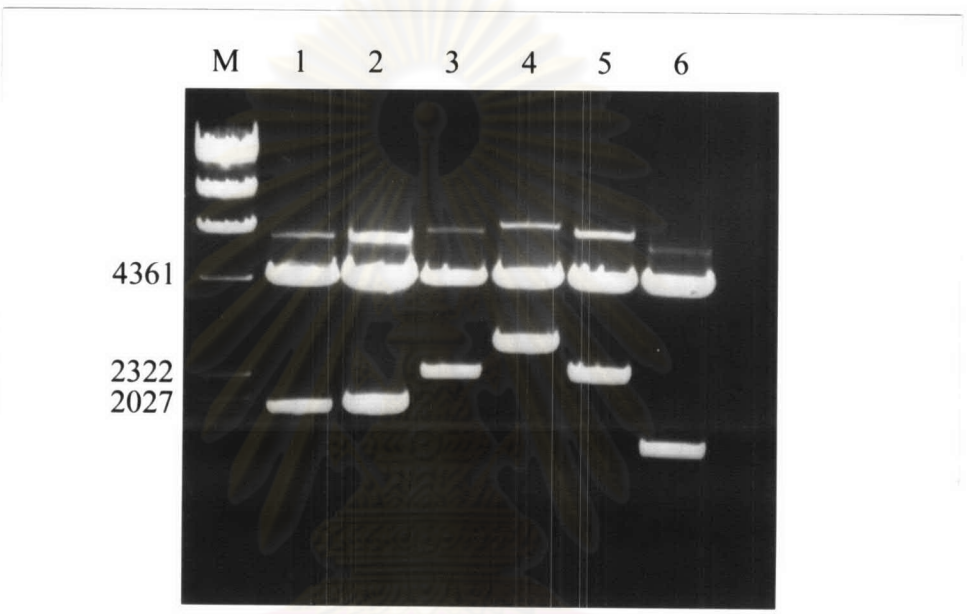


Figure 18 Restriction endonuclease analysis of Bp3,5,6 and 7 clone. Agarose gel (1% w/v) electrophoresis of recombinant phagemid pBK-CMV digested with *kpnI* and *SacI*. M:DNA marker λ DNA digested *HindIII*(fragment size in bp are shown on the left). Lanes 1-2:Bp3 clone. Lanes 3,5:Bp7 clone.Lane4: Bp6 clone. Lane 6: bp5 clone

2.1 Clone Bp1(2,8,13), *gmhA* gene,Capsule gene

The inserted DNA size of clone Bp1 was 1378 bp as demonstrated in appendix B.The restriction endonuclease analysis of insert DNA of this clone is shown in figure 19.The average percentage of G+C base was 59.29% .The signal of immunoscreening of this clone is demonstrated in figure 20. It expressed one protein band with approximately MW of 25 kDa. When the insert DNA was blasted against *B.pseudomallei* genome, it hit the chromosome 1 at the position 3,341,332 –3,342,709(positive strand/ positive strand).With the McVector analysis, there was one inframed open reading frame(ORF 6, 537 bp) as shown in figure 21 . GeneMark and BLAST analysis revealed two genes in the insert DNA, of which one was nearly full gene of *gmhA*(ORF 6, 537 bps) and the other was full gene of *wcbM*(ORF4, 693 bps). The direction of gene expression of both ORFs was the same as T3 promoter of the pBK-CMV.The arrangement of genes in this clone was demonstrated in figure 22. Following a BLAST analysis, the insert sequence was located in a capsule operon(around 20 genes, 38,280 bp). Both can be defined as biosynthetic capsule genes. The hit area on the the chromosome 1 was extended to cover the *gmhA* gene . Then both *gmhA* and *wcbM* gene were cut from the first chromosome of *B.pseudomallei* and analyzed by using various bioinformatic programns.

The full sequence of *gmhA* gene is demonstrated in appendix B. This gene encodes a protein of 197 amino acid with molecular weight 20.8 Kd. Blast analysis revealed high scores with protein of *gmhA* gene from various bacterial species such as *gmhA* of *B.mallei* (S=382, E=e-105, I=98%) :putative phosphoheptose isomerase(*Campylobacter jejuni*, S=221,E=5e-57, I=57%): phosphoheptose isomerase(*Clostridium acetobutylicum*,S=204, E=7e-52, I=51%): putative 6-deoxy-D-manoheptose pathway protein (*Yersinea pseudotuberculosis*, S=197, E=1e-49, I=53%) . There were many conserved domain hits to this protein such as COG 0279, pfam 01380(SIS domain),COG

1737, COG 2103,COG 2222, COG 0794, COG 4821 as demonstrated in figure 23. The main function of the hit domain were carbohydrate transportation and metabolism. Hopp/Woods hydropathy scale was applied to this protein as illustrated in figure 24. There were many minima peaks(hydrophilic region or antigenic site). The picture was supported by the secondary structure prediction by using choua Fasman approach as demonstrated in figure 25. The protein contained mostly alpha helix and beta sheet residues .By PHD approach, this gmhA protein contained alpha helix, beta sheet and others as 52.3%, 14.2% and 33.5% respectively.

The full sequence of *wcbM* gene is shown in appendix B. This gene encodes protein of 230 amino acid with molecular weight 24.92.Blast analysis revealed a high score with *wcbM* (*B.mallei*,S=463, E=e-129, I=100%): D-glycero-D-manno-heptose1-phosphate guanosyltransferase(*Aneurinibacillus thermoaerophilus*, S=168, E=8e-41, I=39%): putative sugar –phosphate nucleotidyltransferase(*Campylobacter jejuni*, S=160, E=2e-38, I=41%): putative 6-deoxy-D-mannoheptose pathway(*Yersinea pseudotuberculosis*, S=159, E=4e-38, I=40%).There were many conserved domain hits to this protein , such as COG 1208, pfam 00483,COG 1207, COG 1209, and COG 1210.By PHD approach, this protein consisted of alpha helix, beta sheet and other as 23.9% , 28.7% and 47.4% respectively.

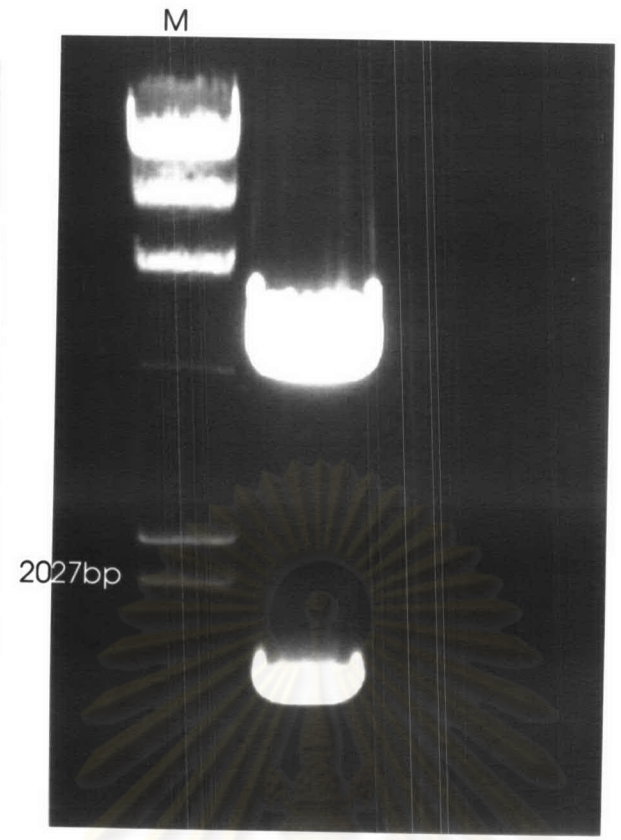


Figure 19 Restriction endonuclease analysis of clone Bp1.Agarose gel (1% w/v) electrophoresis of recombinant phagemid pBK-CMV of Bp 1 clone was digested with *KpnI* and *sacI*. M: DNA marker λ DNA digested *HindIII*(fragment size in bp are shown on the left).

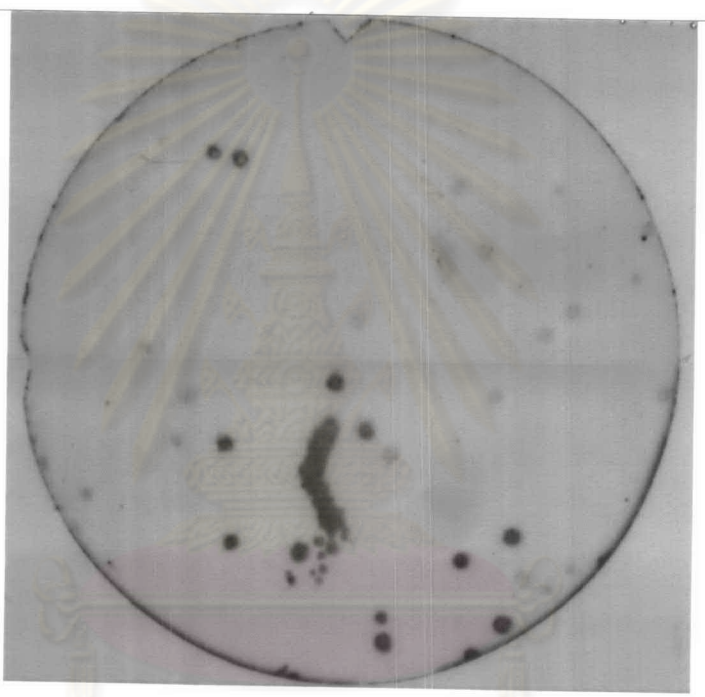


Figure 20 Immunorescreening of clone Bp1.

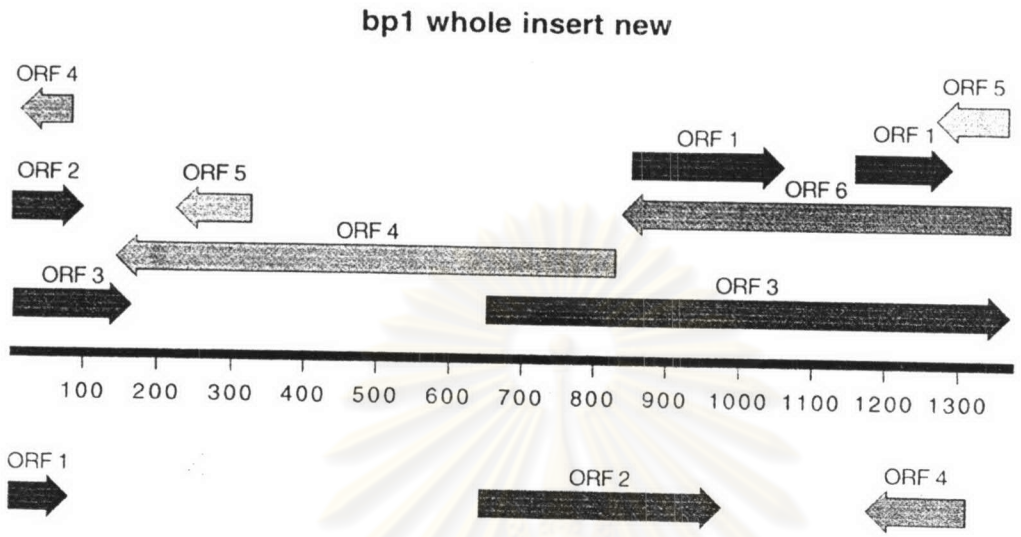


Figure 21 ORF analysis of Bp1.

ORF analysis using Macvector programn V 7.0.1. ORF 6(537) was identified as inframed ORF.

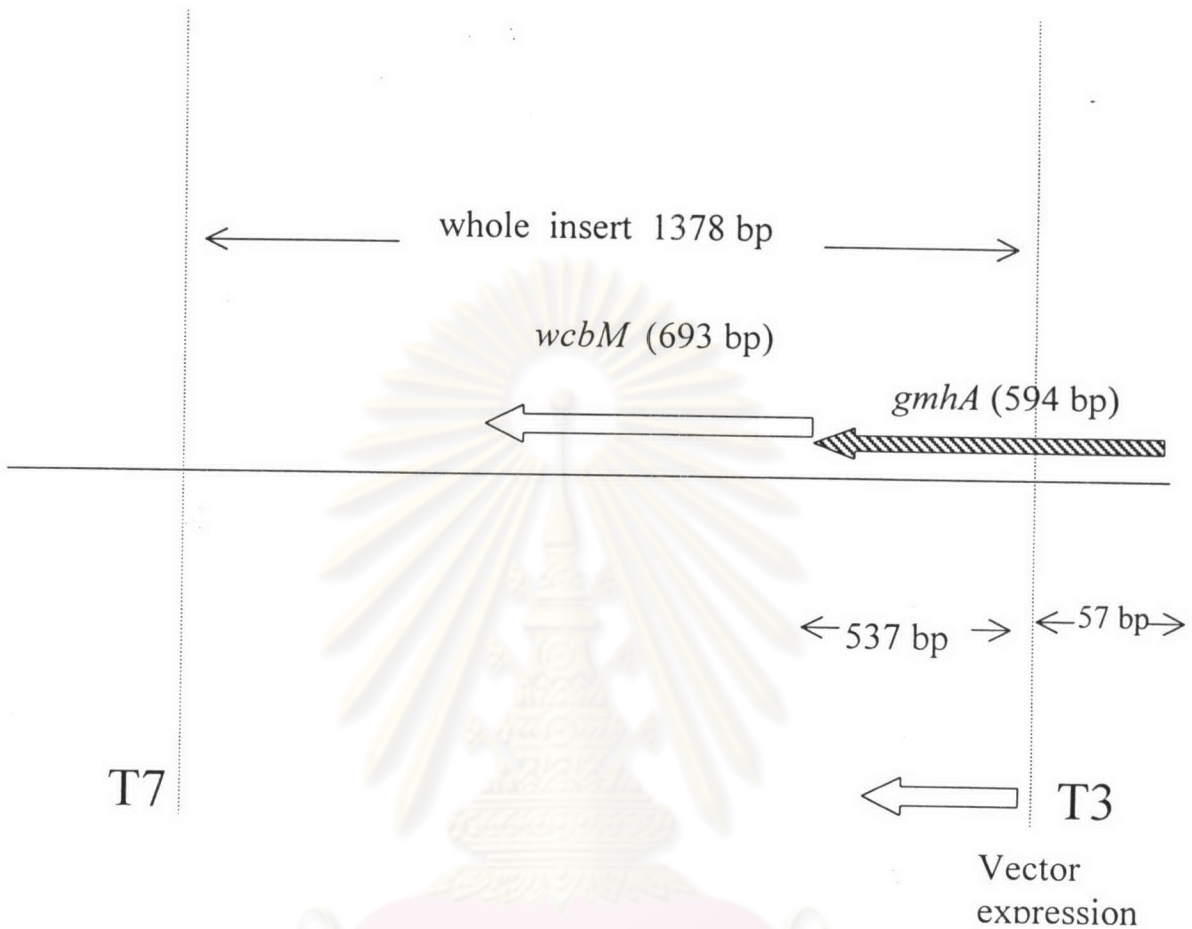


Figure 22 Arrangement of genes in Bp 1 clone using GeneMark programn. Inframed ORF(537 bp) of *gmhA* gene correlated with one band of expressed protein with approximately 25 kDa molecular weight.

NCBI

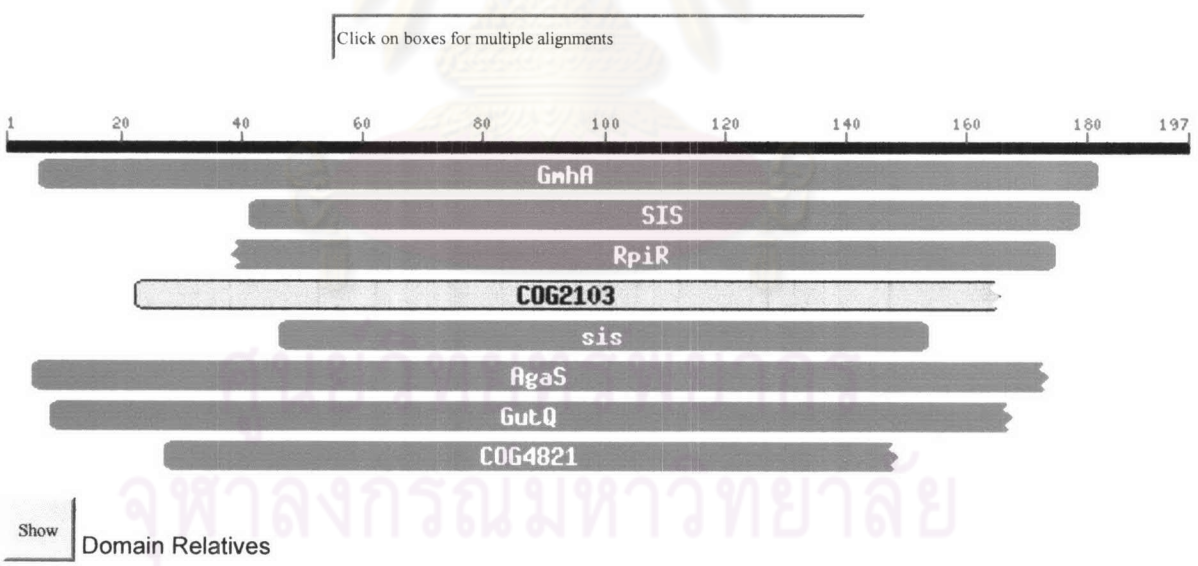
NCBI Conserved Domain Search

New Search PubMed Nucleotide Protein Structure CDD Taxonomy Help?

RPS-BLAST 2.2.6 [Apr-09-2003]

Query= local sequence:
(197 letters)

Database: #cdd.v1.62
11,088 PSSMs; 2,717,223 total columns



- .. This CD alignment includes 3D structure. To display structure, download **Cn3D**!

Figure 23 Conserved protein domains of gmhA protein

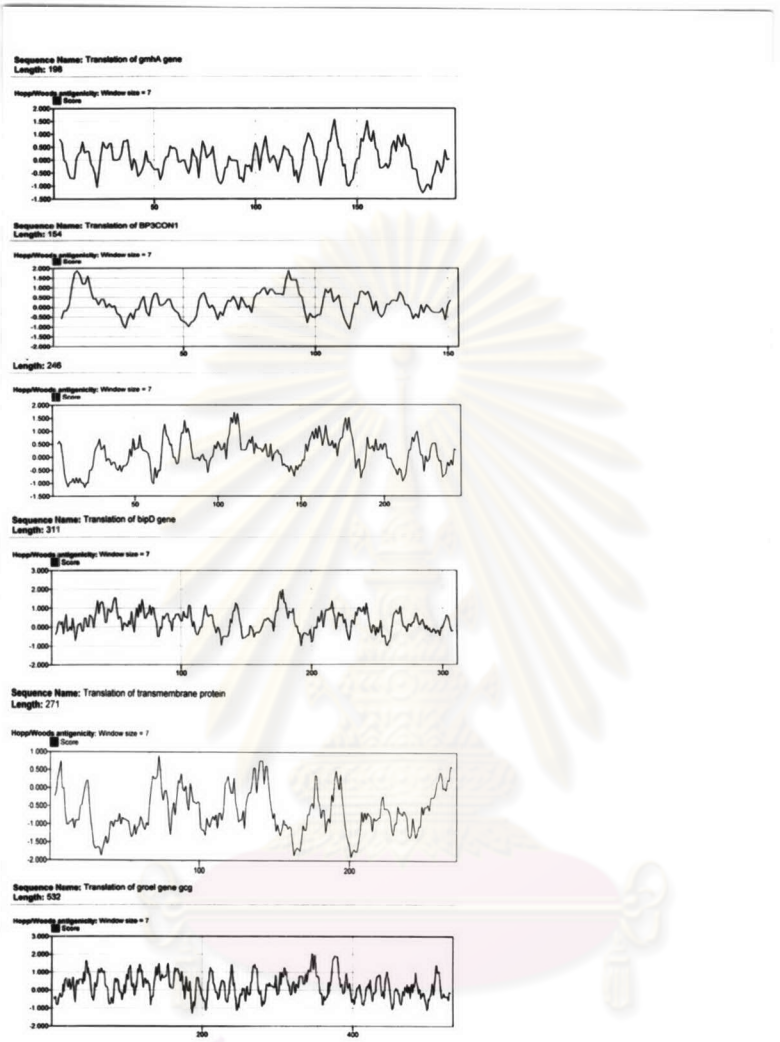


Figure 24 The Hopp and Wood hydropathy profiles for 6 suspected in vivo expressed genes using a scanning window size 7 amino acids. Transmembrane gene shows highest minima peak (hydrophilic region)

Sequence Name: gmhA gene
Length: 198

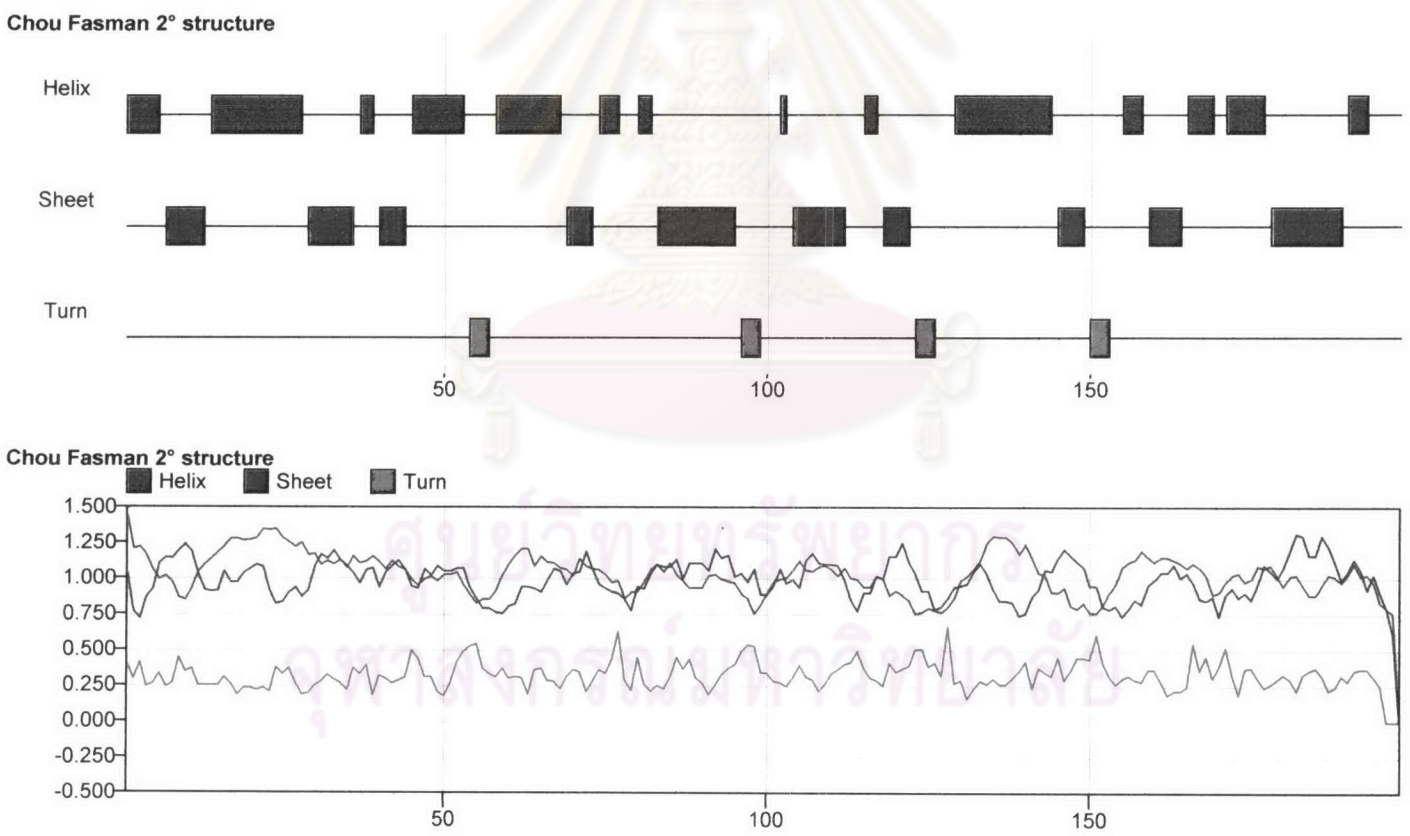


Figure 25 Chou Fasman secondary structure of gmhA protein

2.2 Clone Bp6 (18) , (SPI-1 like clusters)

The insert size of this clone was approximately 2.5 Kb as demonstrated in appendix B. The average percentage of G+C base content was 70.16% .The DNA insert hit the *B.pseudomallei* chromosome 2 at position 2,081,650-2,084,150.(negative strand/positive strand). The positive signal from immunoscreening experiment is shown in figure 26.The expressed protein size was ~ 45 Kd. The expressed protein band reacted only with pooled absorbed melioidosis serum and not pooled absorbed normal serum as demonstrated in figure 27.There was no original insert in the sequence because of an incompleted sequencing experiment. From BLAST and GeneMark anaysis, the insert sequence at T3 site hit at the middle of *bpH3* ORF. The inframed ORF may be included both of bpH3 ORF(291 bp) and *bip*(*SipD*)(933bp)ORF.The direction of the vector expression was the same as *bpH3* and *bipD*.The genes arrangement in this whole insert clone are demonstrated in the figure 28. From the BLAST study, this inserted DNA was located in the pathogenicity islands (PAIs) of SPI-1 like cluster that encoded type three secretion protein(~25 genes).The full size of 291 Bp gene and 933 Bp gene were cut from the second chromosome of *B.pseudomallei* and analyzed by using various bioinformatic programns.

The full sequence of 291 Bp gene is shown in appendix B.Blast analysis of 291 bp gene revealed low score with putative HNS-like transcription regulator protein(*Ralstonia solanacearum*, S=78.2, E=3e-14, I=48%).The protein encoded by this gene has one conserved domain HNS, DNA-binding protein .This gene encodes a protein of 96 amino acid with molecular weight 10.9 Kd.The protein secondary structure prediction was demonstated by PHD approach.This protein contained alpha helix, beta sheet and other as 61.5%, 3.1% and 35.4% respectively.This gene is called *bpH3* gene.

The 933 Bp gene sequence is shown in appendix B.This gene encodes protein of 310 amino acid with molecular weight 33.98 kd. Blast analysis revealed low scores with some well known virulence factors such as IpaD,secreted by the Mxi-Spa machinery, required for entry of bacteria into epithelial cells (*Shigella flexneri*, S=97.4, E=3e-19, I=36%):SipD,pathogenicity island 1 effector protein (*Salmonella enterica* subsp.enterica serovar Typhi, S=97.4, E=3e-19, I=33%): ipaD protein(*Shigella dysenteriae*, S=93.6, E=4e-18, I=34%). There were no conserved domain hits to this protein.The hydropathy scales of this protein is illustrated in figure 24. There were many antigenic sites in this Hopp and Woods approaches.The secondary structure prediction of this protein is demonstrated in figure 29. The protein consisted mainly of a mixture of both alpha helix and beta sheet structure.By PHD approach, this protein consisted of alpha helix, beta sheet and others as 56.8%, 5.2% and 38.1% respectively.This gene is named *bipD* .

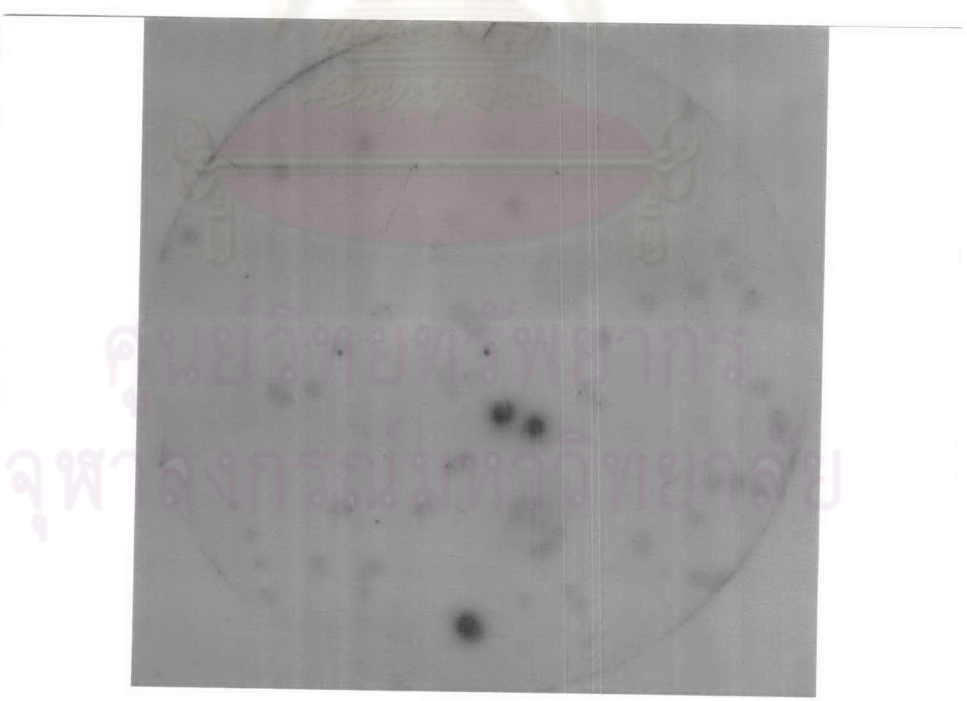


Figure 26 Immunorescreening of clone Bp6

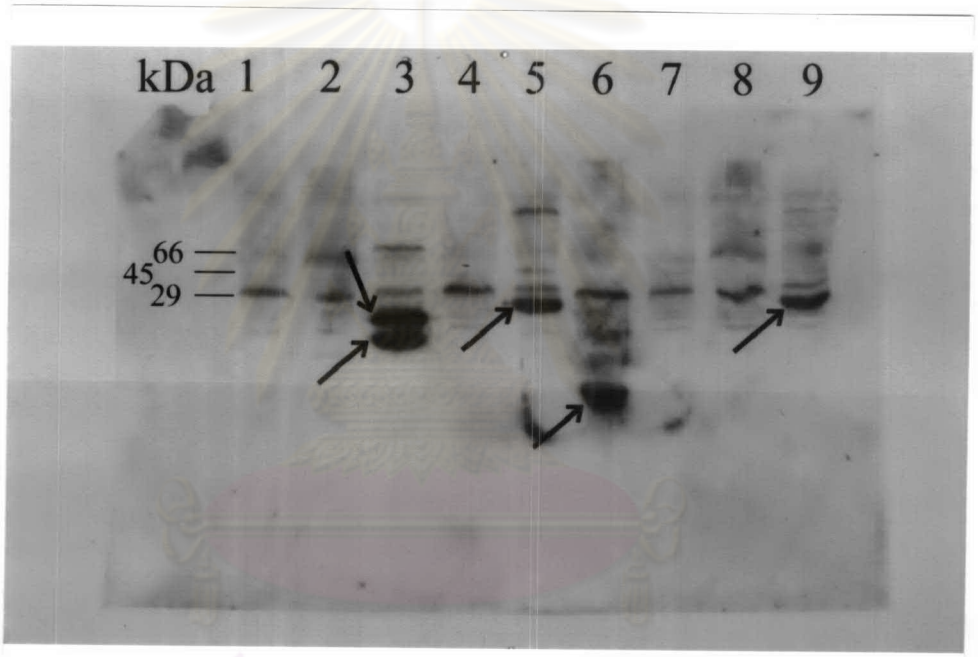


Figure 27 Expression of recombinant phagemid pBK-CMV in *E.coli XL1Blue MRF'*:after induction with I mM IPTG, whole cell lysates of *E.coli* were subjected to SDS-PAGE, transferred to nitrocellulose, and blotted with 1;1000 dilution of pooled absorbed melioidosis serum. Lane 1-2:*E.coli XL1 Blue MRF'*. Lane 3: Bp3 clone. Lane4: *E.coli XL1 Blue MRF'* with vector pBK-CMV only. Lane 5,9: Bp6 clone. Lane6:Bp 7 clone.Lane 7:Bp1 clone. Lane8: Bp 9 clone.

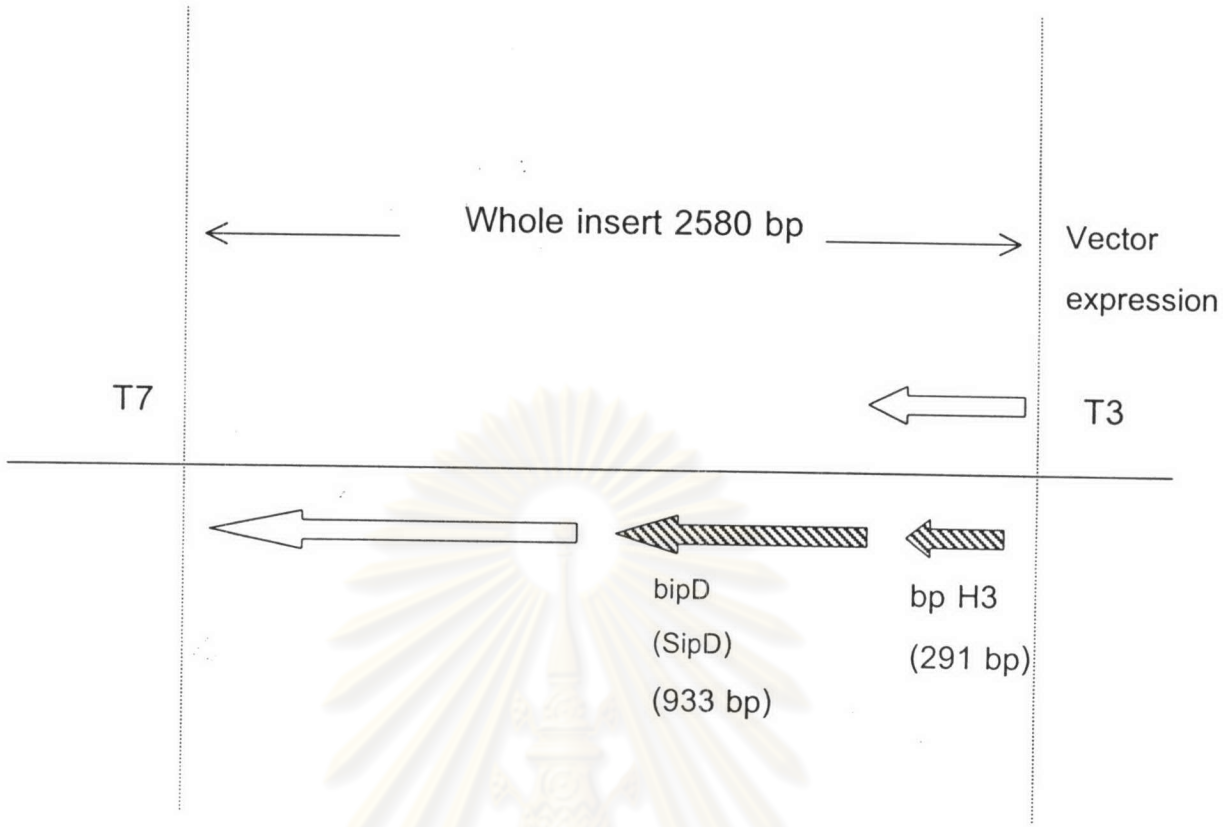


Figure 28 Arrangement of genes in Bp 6 clone by using GeneMark program

Sequence Name: bipD (sipD)gene
Length: 311

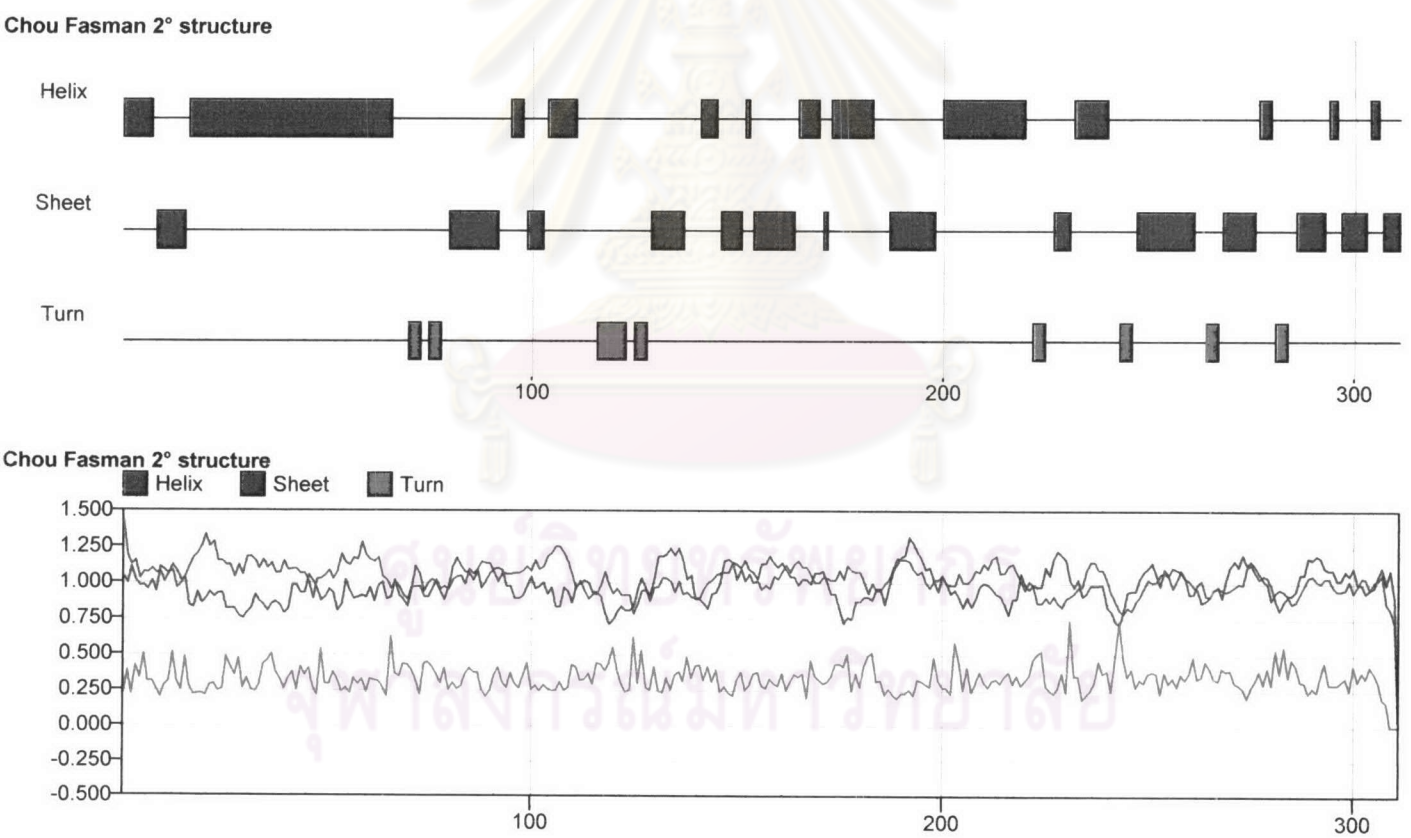


Figure 29 Chou Fasman secondary structure of bipD protein

2.3 Clone Bp7(19) Transmembrane protein

The insert DNA size was 2303 bp as demonstrated in appendix B. The percentage of G+C content was 67.54 % .The clone hit with the second chromosome of *B.pseudomallei* at position 2,437,533-2,439,835(negative strand/positive strand). This clone generated the stongest signal in the screening reaction as shown in figure 30. It expressed one band of protein ~10 Kd.The western blot experiment resulted in one band that reacted with only pooled absorbed melioidosis serum as demonstrated in figure 27. From MacVector analysis, one inframed open reading frame ORF2 (571 bp) was found . This inframed ORF consisted of 2 incomplete genes(246 bp and 334 bp) by analysing with GeneMark program. The full gene at the T3 site was 813bp. The other gene that was located near the first gene had 996 bp.The organization of genes in this clone is demonstrated in figure31. The size of inframed ORF(571 bp) did not correlate with the size of low molecular weight of the expressed protein (~10 Kd).According to the the size of the expressed protein of this clone(10 Kd) and the the size of some part of the first gene(246 bp) , the first gene had the highest probability to be expressed. The full gene of both 813 bp , 996 bp and 399 bp were cut from the second chromosome of *B.pseudomallei* and analyzed by using various bioinformatic programns.

BLAST analysis of 813 bp gene revealed high scores with probable transmembrane protein of *Ralstonia solanacearum*(S=205, E=4e-53,I=41%) and putative membrane protein of *Bordetella bronchiseptica* and *Bordetella parapertussis*(S=125,E=8e-31,I=30%) .There was no conserved domain hit to this protein.The full gene sequence is demonstrated in figure appendix B.The gene encodes a protein of 270 amino acid with molecular weigth 28.8 kd.This protein is consisted of 76 % of hydrophobic amino acid group such as glycine, alanine, isoleucine, leucine, methionine, phenylalanine, tyrosine, tryptophan, valine, cysteine or proline. The hydropathy scale of both Hopp/

Woodsand Kyte/ Doolittle were applied to this protein as demonstrated in figure 32. There were many of high broad maximum peaks of hydrophobic region as revealed by Kyte and doolittle scale.The high minima peaks of the antigenic site were also observed with Hopp and Woods antigenicity scale.The secondary structure of the protein was predicted by using chou and Fasman approach as illustrated in figure 33. The beta –sheet structure was covered nearly all of the protein sequence.The percentage of each secondary protein structure in the protein of alpha-helix, beta-sheet and others as revealed by PHD programn were 52.59%, 10.74% and 36.67 % respectively . The transmembrane regions in protein was predicted by using Argos Helix transmembrane and Von Heijne transmembrane approach as demonstrated in figure 34.There were many broad high maxima peaks that extended more over 20 amino acids lenght in both methods.The PHDhtm also detected 7 membrane helices at the position of amino acid 28-52, 57-74, 106-126,148-167,172-189,198-217 and225-242.This gene is named transmembrane protein gene.

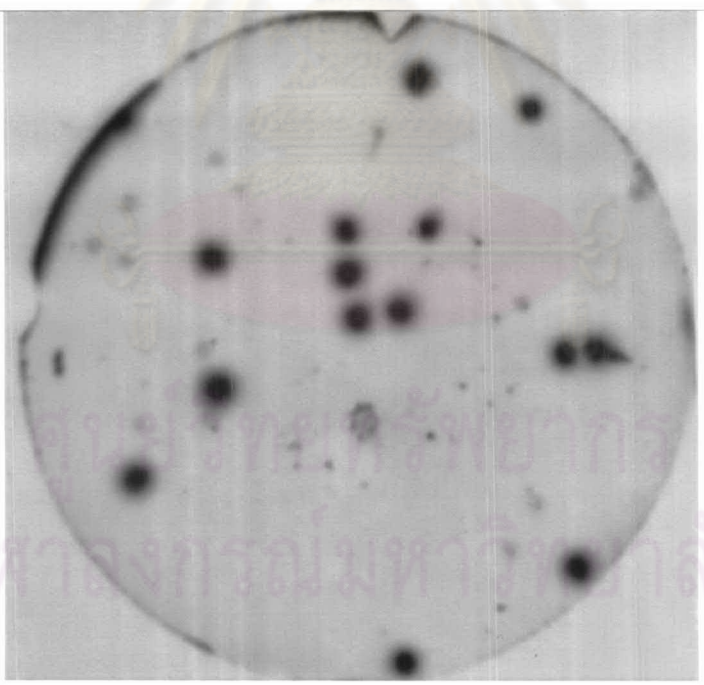
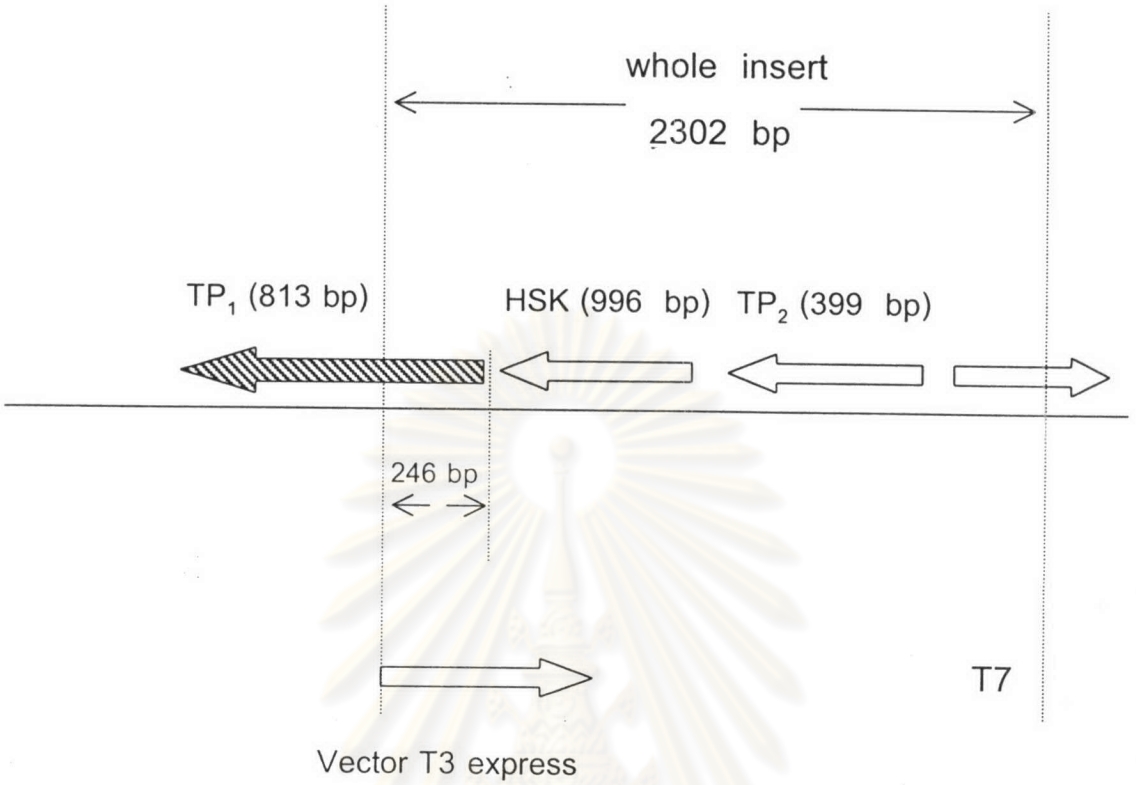


Figure 30 Immunorescreening of clone Bp7



TP_1 = Transmembrane protein 1

HSK = Homoserine kinase

TP_2 = Transmembrane protein 2

Figure 31 Arrangement of genes in Bp 7 clone by GeneMark program

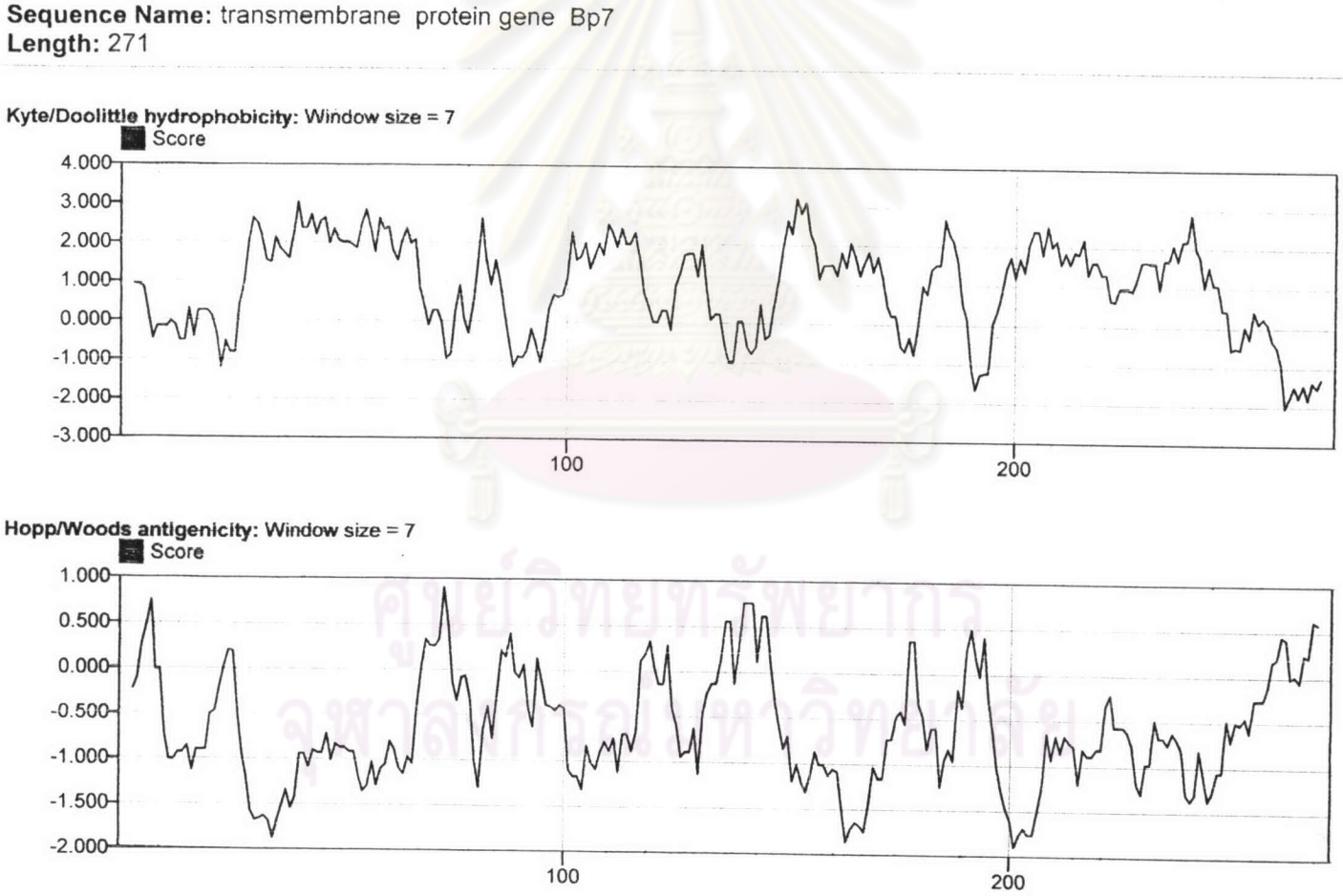


Figure 32 Kyte/Doolittle; Hopp/Woods antigenicity analysis of clone Bp7

Sequence Name: transmembrane protein gene Bp7
Length: 271

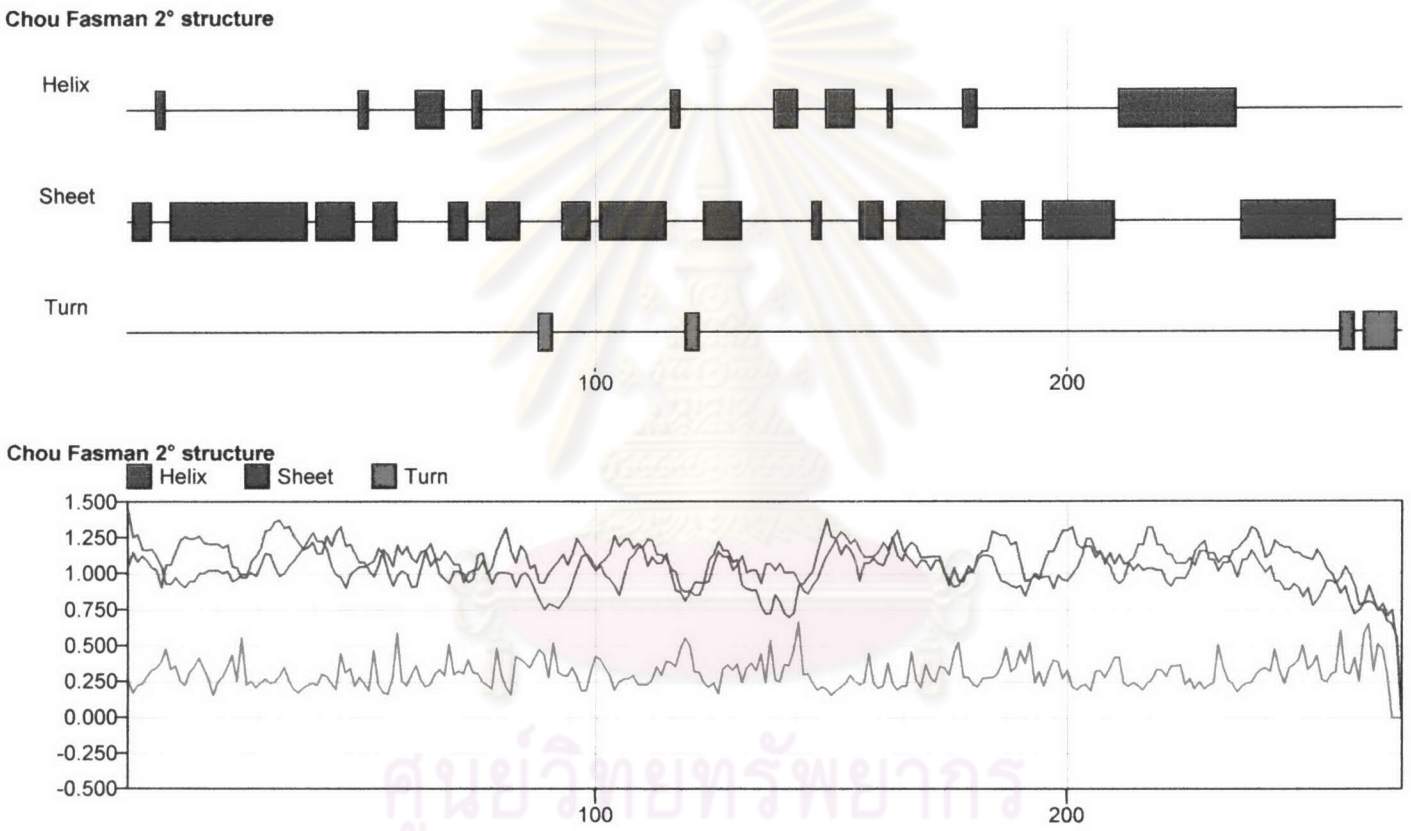


Figure 33 Chou Fasman secondary structure of transmembrane protein

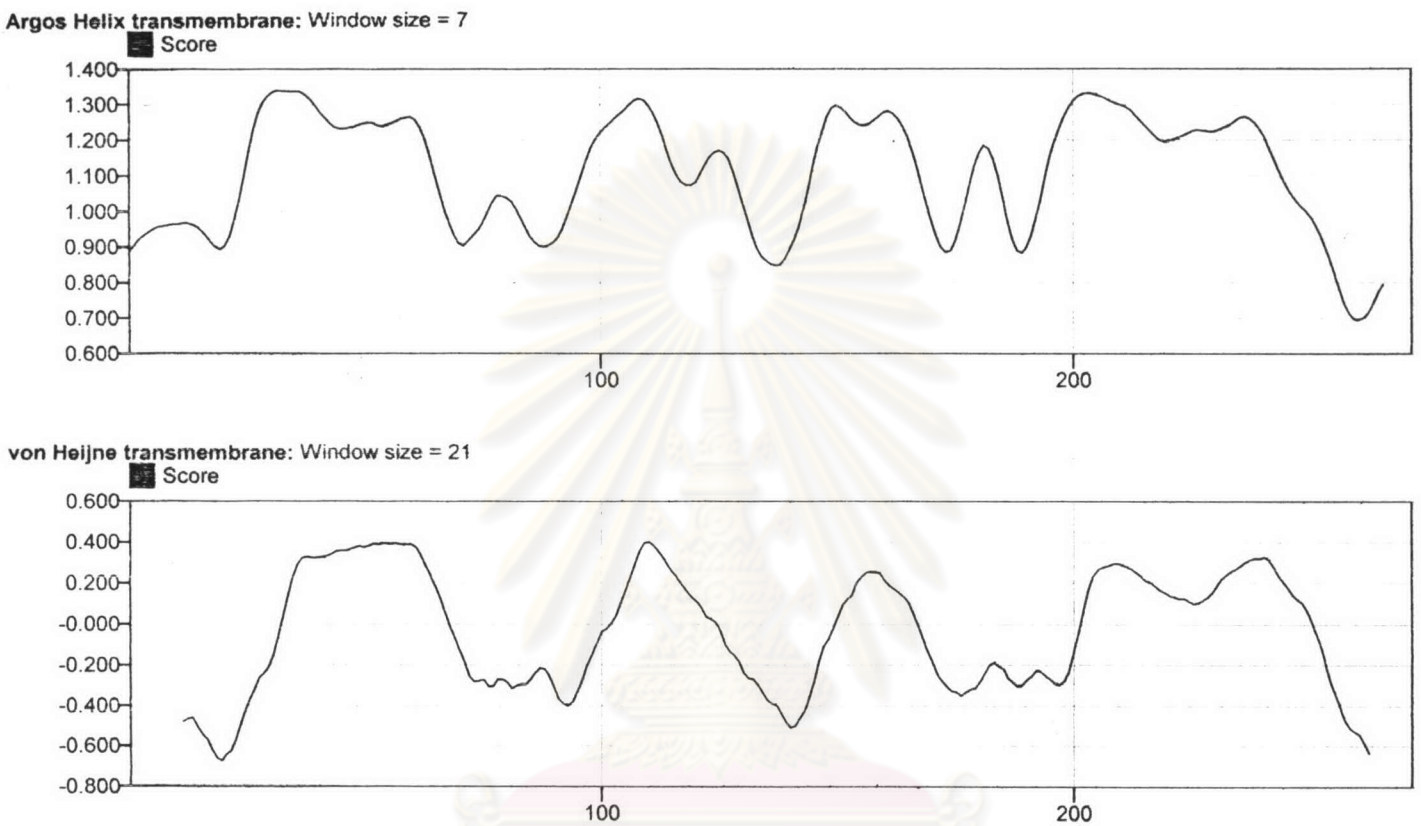


Figure 34 Argos Helix transmembrane and Von Heijne transmembrane analysis of transmembrane protein

2.4 Clone Bp3(Conserved hypothetical protein and hypothetical signal peptide protein)

The insert DNA size was approximately 1738bp as shown in appendix B. The average percentage of G+C base content was 69.68% . The clone DNA hit the *B.pseudomallei* chromosome 1 at position 1,131,998-1,130,260.(positive strand/ negative strand). The original insert could not be identified according to incomplete sequencing. The signal of immunoscreening of this clone was illustrated in figure35. This clone expressed 2 bands of protein with approximately 17 kDa and 25 kDa. The western blot also resulted with 2 bands with only pooled absorbed melioidosis serum as demonstrated in figure 27. From sequence analysis, the expressed ORF may be ORF2 (321 bp) andORF2 (731 bp) as shown in figure 36 . The direction of both ORFs was the same as T3 promoter of pBK-CMV. The whole area ~5 Kb around the inserted DNA sequence was analysed withGeneMark progrman,showing the first gene at T3 promoter was 462 bp and the second gene as 738 bp.The arrangement of genes in this clone is demonstrated in figure 37.The 462 bp gene and 738 bp gene were cut from the first chromosome of *B.pseudomallei* and analyzed by using bioinformatic programs.

The full sequence of 462 bp gene was illustrated in appendix B.This gene codes a protein of 153 amino acid with molecular weight 17.49 kDa.Blast analysis demonstrated that this gene codes for conserved hypothetical protein (*Ralstonia solanacearum*, S=232, E=1e-60, I=72%). There were 2 conserved protein domains hits to this protein ; COG2947 and DUF 589 .These conserved protein function were unknown. The hydropathy scales of this protein revealed many high peaks antigenic sites as demonstrated in figure 24.The secondary structure prediction of this protein was shown in figure 38. The protein consisted of mainly both alpha helix and beta – sheet structure.However, this

protein contained alpha helix , beta sheet and other as 16.2%, 13% and 70.8% respectively.This gene is named conserved hypothetical protein gene.

The 738 Bp gene sequence is shown in appendix B. This gene codes a protein of 245 amino acid with molecular weight 25.63 kDa.Blast analysis revealed high scores with conserved hypothetical protein; hypothetical protein(*Burkholderia fungorum*, S = 272, E = 3e-72, I = 66%) ; hypothetical protein (*Ralstonia metallidurans*, S=204,E=1e-51,I=52%); hypothetical signal peptide protein (*Ralstonia solanacearum*,S=145,E=3e-34, I=56%). There were three conserved domains protein hits to this protein;COG 3471;pfam 0442, Duf541 and COG 2968 . The function of this protein remains unknown. It is postulated to be located in the periplasm or outer membrane.There were many minima peaks of antigenic sites as illustrated in figure 24. The secondary structure prediction of this protein is shown in figure 39. The alpha helix structure is located in nearly all the protein sequence.From PHD approach, this protein contained alpha helix, beta sheet and others as 42.9%, 23.7% and 33.5% respectively.This gene is named hypothetical signal peptide protein gene.

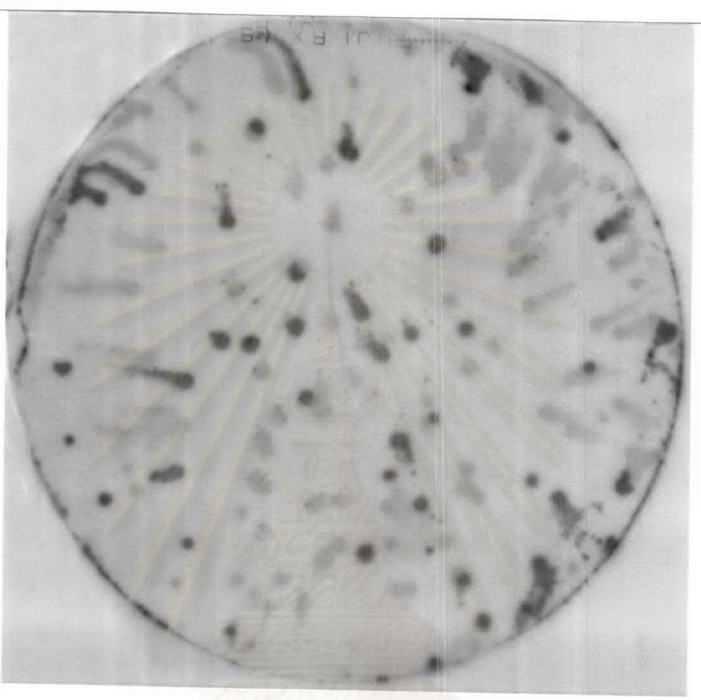


Figure 35 Immunorescreening of clone Bp3

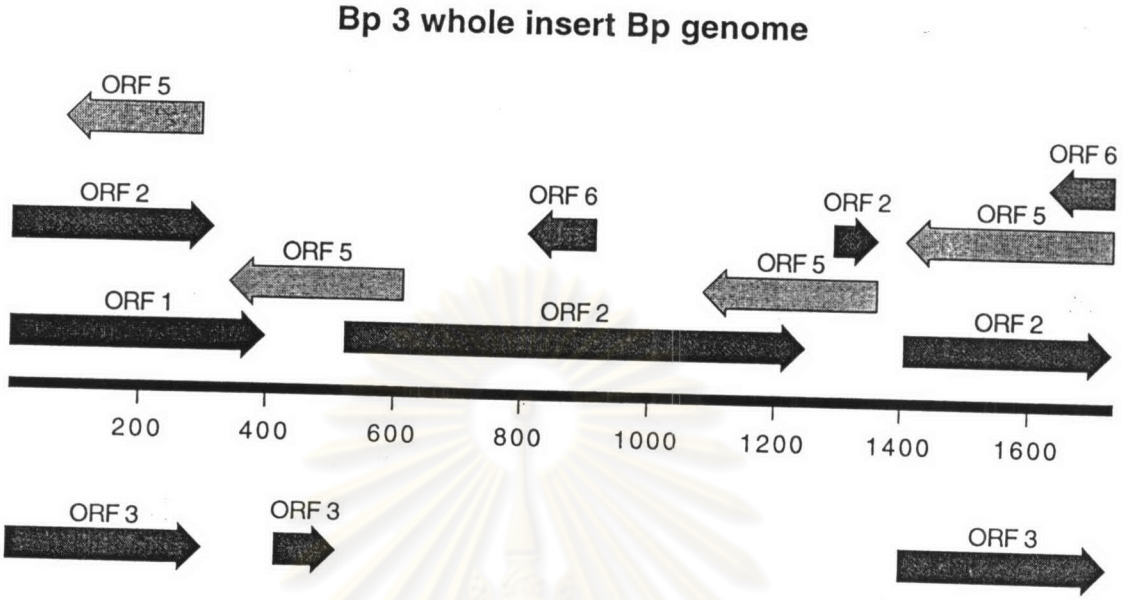
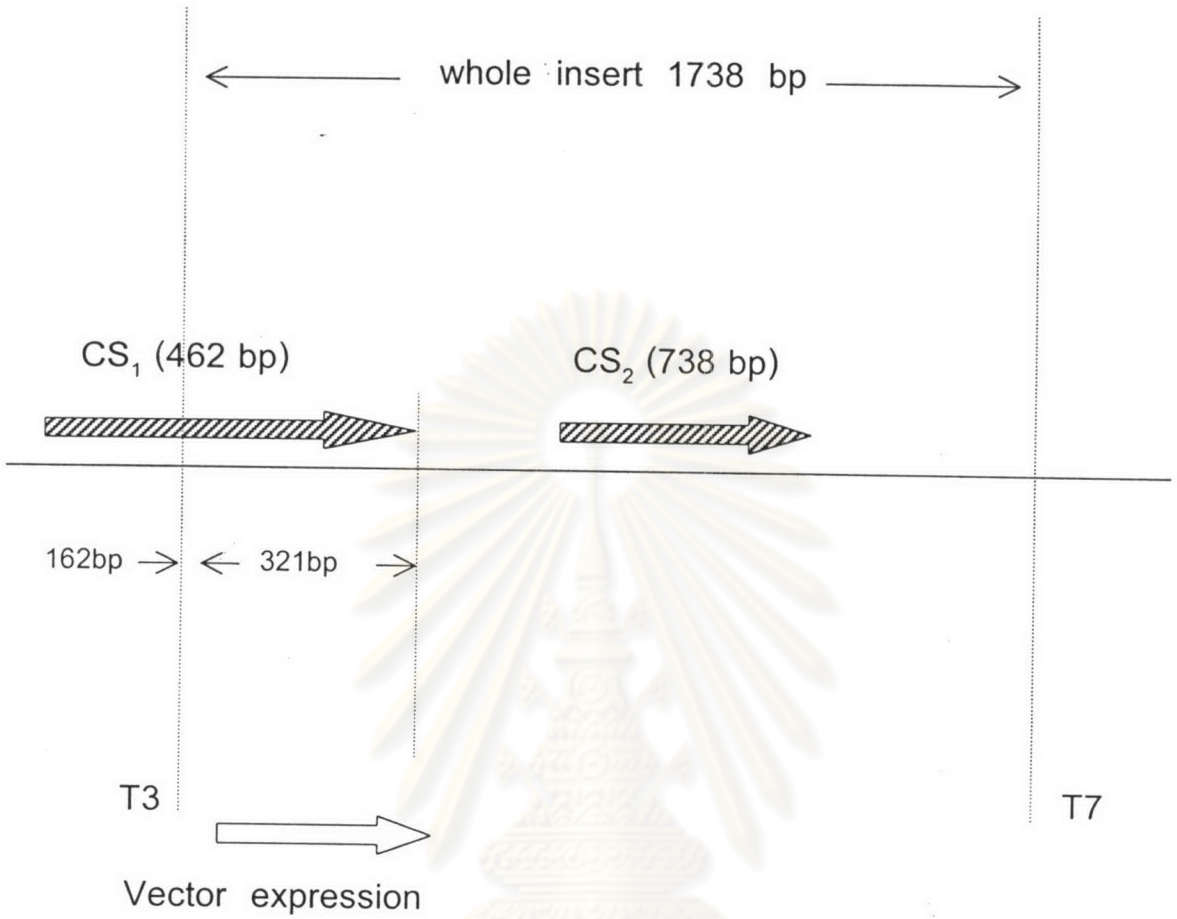


Figure 36 ORF analysis of clone Bp3 inserted DNA using Macvector program V7.0.1



CS_1 = Conserved hypothetical protein

CS_2 = hypothetical signal peptide protein

Figure 37 Arrangement of genes in Bp 3 clone by using GeneMark program

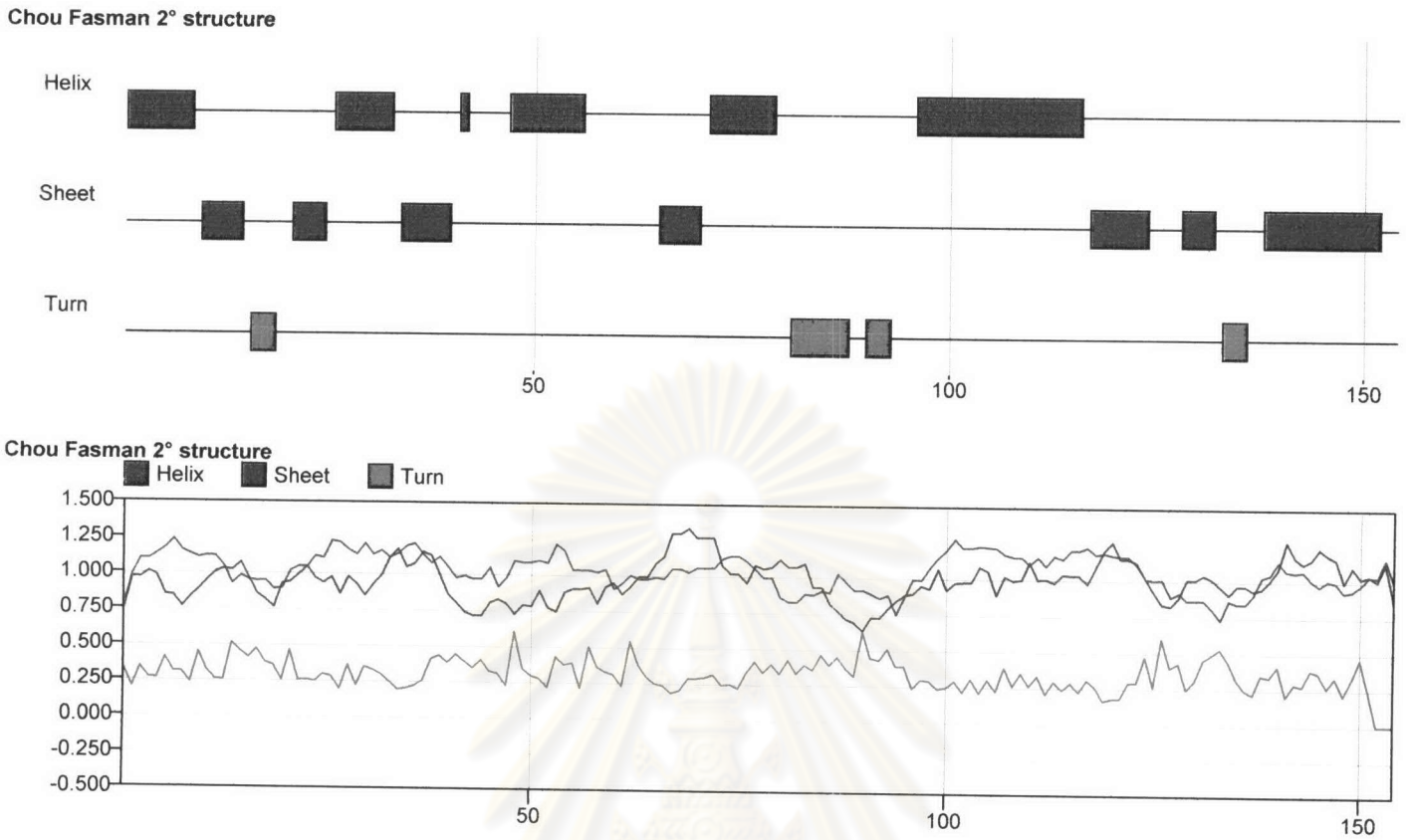


Figure 38 Chou Fasman secondary structure of conserved hypothetical protein 1 of clone Bp3

Sequence Name: hypothetical signal peptide gene
Length: 246

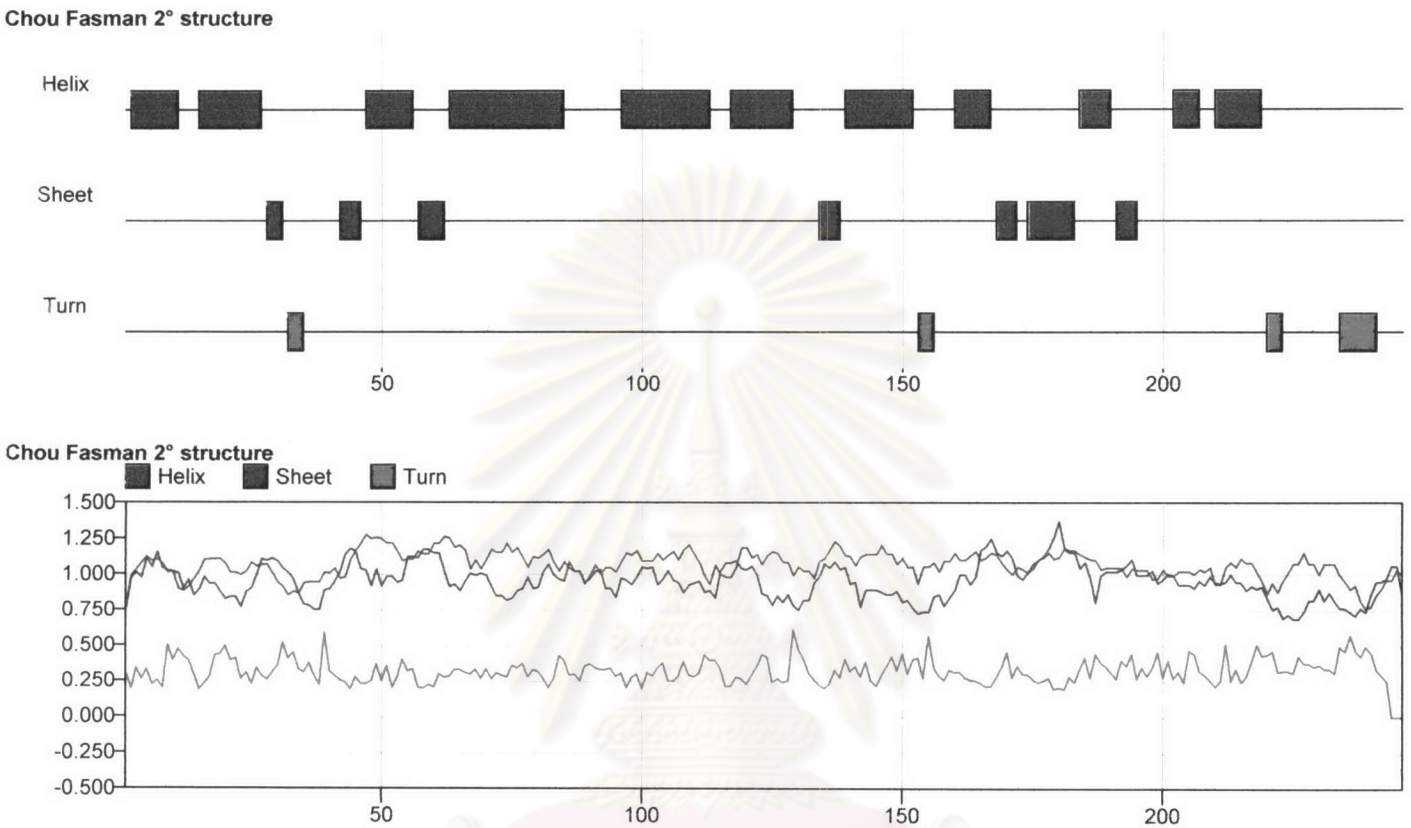


Figure 39 Chou Fasman secondary structure of conserved hypothetical signal peptide protein of clone Bp3

2.5 Clone Bp9 groEL, Chaperonine protein

The inserted DNA size was 1300 bp as illustrated in appendix B.The average percentage of G+C base content was 63.69% .The signal of immunoscreening of this clone is illustrated in figure 40. The original insert could not be found. The clone hit the first chromosome of *B.pseudomallei* genome at position 3,223,718 –3,225,021(negative strand/ positive strand). The expressed protein could not be demonstrated via SDS-PAGE. From sequence analysis, this insert DNA was belonged to a nearly complete *groEL* gene. The full gene size of groEL (1596 bp) was cut from the first chromosome of *B.pseudomallei* genome by using BLAST and GeneMark programn as shown in appendix B.

The *groEL* gene codes a protein of 531 amino acid with molecular weight 55.57 kDa.Blast analysis revealed high score with chaperonin protein from many bacterial species; *groEL* (*Burkholderia thailandensis*, S=868,E=0,I=90%); 57 kDa heat shock protein *groEL* (*B.vietnamiensis*, S=860,E=0,I=89%); 57 kDa heat shock protein *groEL* (*B.cepacia* ,S=858 ,E=0,I=89%). There were two conserved protein domain hits to this protein ;GroI and pfam 00118 .The general function of the heat shock protein 60 family are posttranslational modification, protein turnover and chaperones. There were many minima peaks of antigenic site along with the protein sequence of Hopp/Woods approches as illustrated in figure 24.The secondary protein structure prediction of the protein is illustrated in figure 41. The protein contained mostly both alpha helix and beta sheet structures. However, by PHD htm, this protein contained alpha helix, beta sheet and others as 49%, 16.4%, 34.7% , respectively.

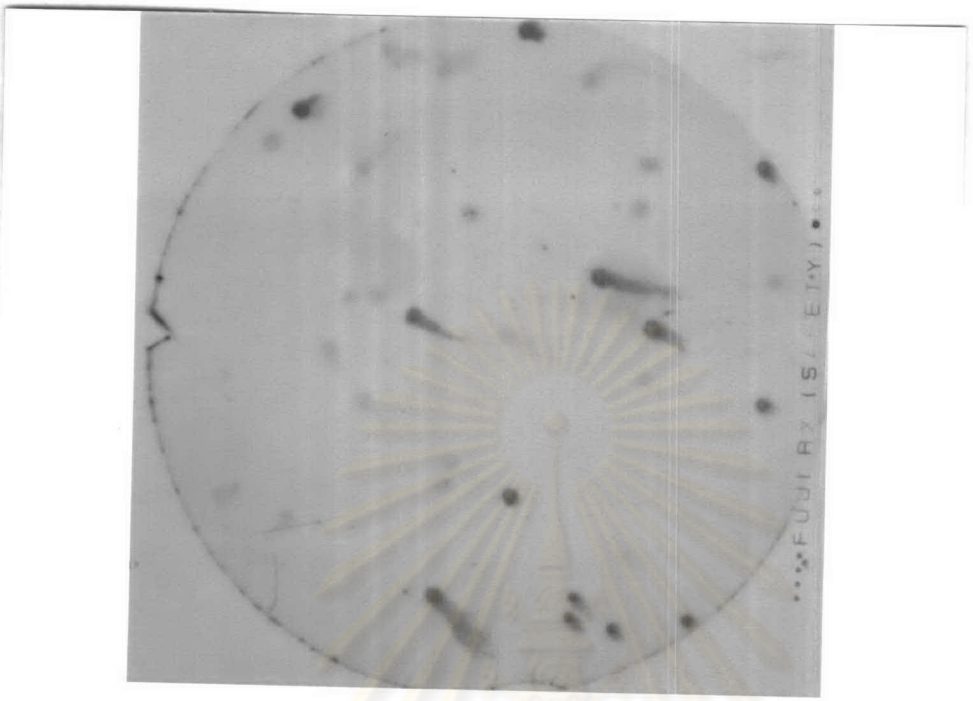


Figure 40 Immunorescreening of clone Bp9

Sequence Name: Translation of /usr/tmp/aaaa13291
Length: 532

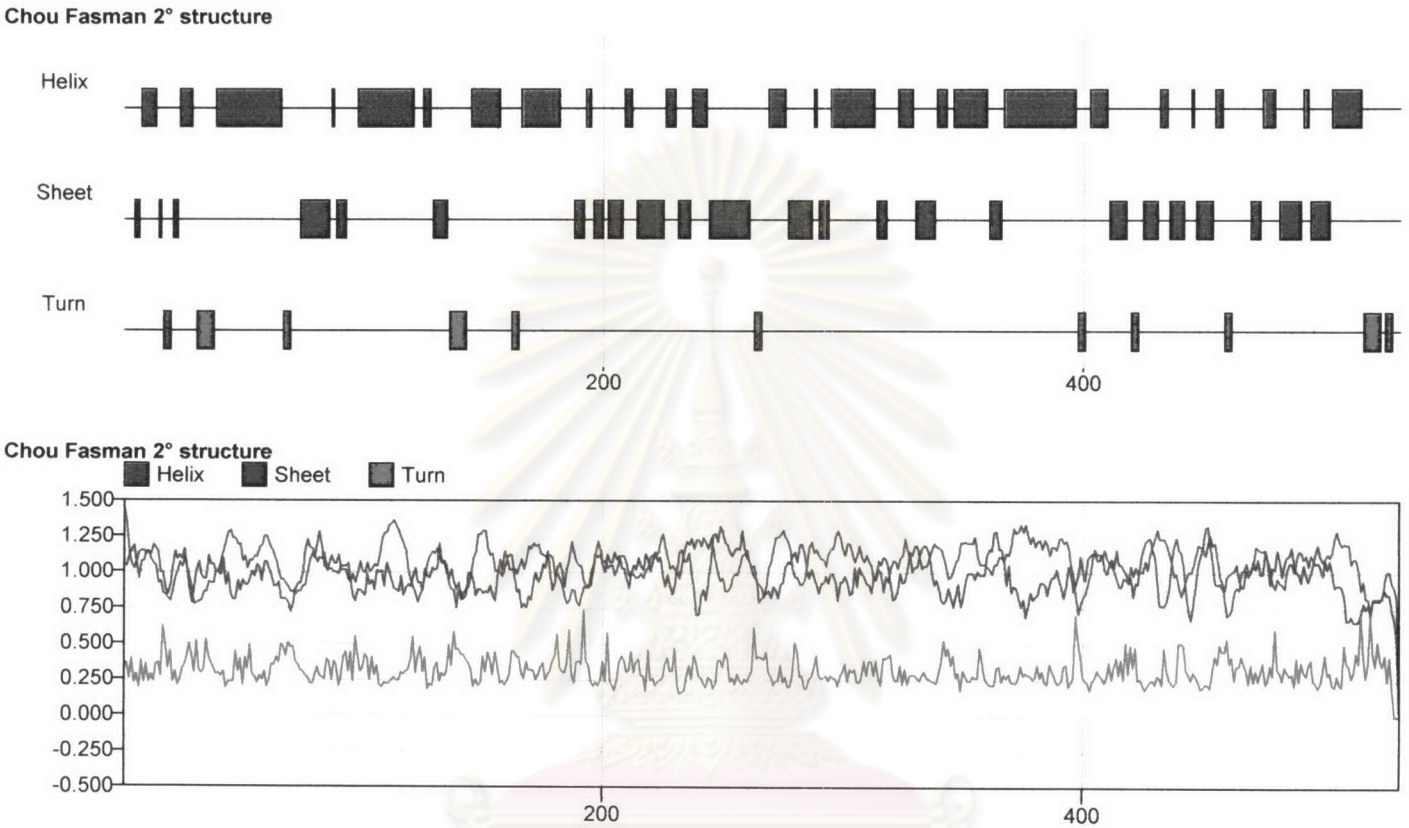


Figure 41 Chou Fasman secondary structure of groEL protein of clone Bp9

2.6 Clone Bp 4(5, 20), ompA

The size of this DNA insert was 1427bp as illustrated in appendix B The average percentage of G+C base content was 60.06 %.The signal of immunoscreening of this clone is showed in figure 42. The clone hit the first chromosome of *B.pseudomallei* at position 3,039,125- 3,040,551(positive strand/positive strand) .The original insert and inframed ORF could not be identified. The expressed protein also could not be demonstrated via SDS-PAGE. From sequence analysis, there was one gene ORF 1(675 bp) that located near the T3 site.(as shown in figure 43). The direction of expression of this gene was opposite of The vecter.The arrangement of genes in this clone is shown in figure 44. The full size of 675 bp gene was cut from the first chromosome of *B.pseudomallei* and analyzed by using bioinformatic programs.

The full sequence of 675 bp gene is illustrated in appendix B. This gene codes a protein of 224 amino acid with molecular weight 24.07. Blast analysis revealed high score with hypothetical protein (*Burkholderia fungorum*, S=279, E=2e-74, I=77%); hypothetical protein(*Ralstonia metallidurans*, S=225, E=5e-58, I=53%); probable signal peptide protein(*Ralstonia solanacearum*, S=220, E=1e-56, I=53%); outer membrane protein A precursor(*Bordetella pertussis*, S=160, E=2e-38, I=48%), outer membrane protein A precursor(*Bordetella parapertussis*, *B.bronchiseptica*, S=159, E=4e-38, I=48%).There were 3 conserved protein domain hits to this protein including COG2885 (ompA);pfam00691(ompA family);COG 1360(MotB, Flagella motor protein) .This protein structure is a porin-like integral membrane protein , with half of its C-terminal a conserve domain of *ompA*..This gene is called *ompA* gene.

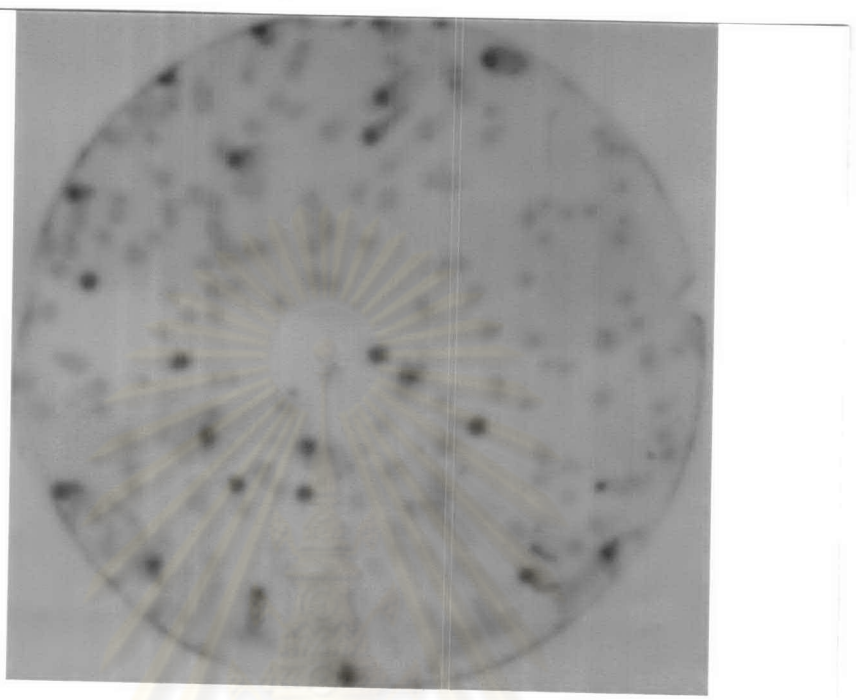


Figure 42 Immunorescreening of clone Bp5

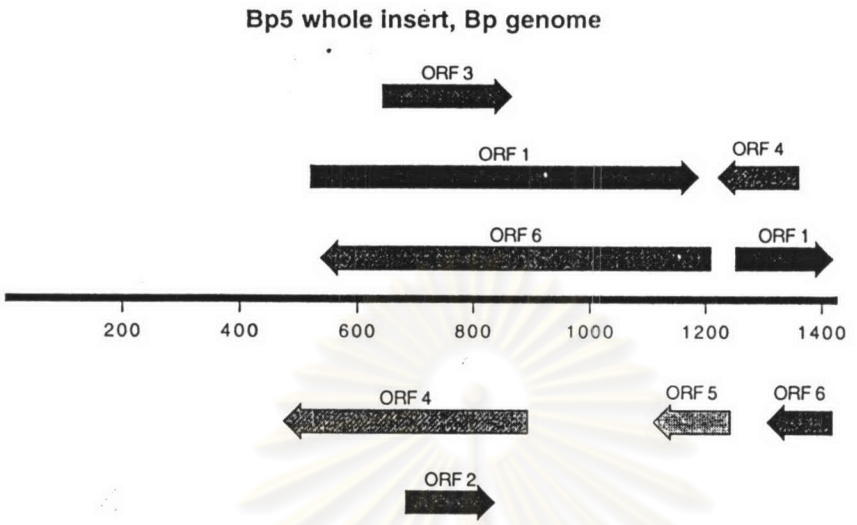


Figure 43 ORF analysis of clone Bp5 using Macvector V 7.0.1

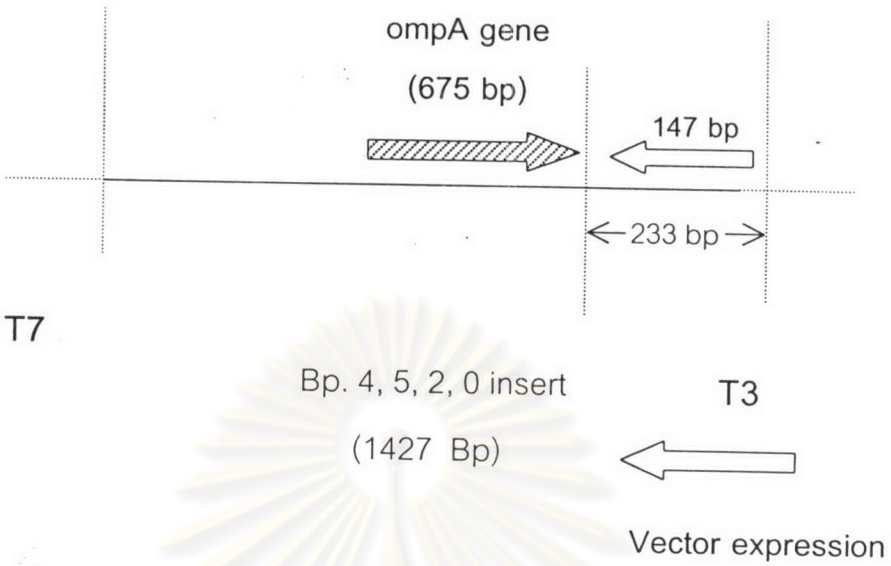


Figure 44 Arrangement of genes in Bp 5 clone by usingGeneMark program

3. Antigenic index analysis

The Hopp and Woods hydropathy scales were applied to 6 suspected in vivo expressed genes as illustrated in figure 24. All of the genes have been demonstrated to contain many antigenicity peaks. The transmembrane protein has highest minima peaks of hydrophilicity.

4. Multiple aligment analysis and phylogenetic tree construction

Multiple aligment of amino acid translated of *gmhA* is shown in figure 45. The aligment was generated with the program Clustal W (MacVector V 7.0.1) using a gap penalty of 10.0 and extend gap penalty of 0.05. The gmhA showed high similarity to other gmhA protein, with the most similarity to *B.pseudomallei* 1026 b strain (98% identity, 100% positive), *B.mallei* (98% identity, 100% positive), *A.thermoaerophilus* (63% identity, 74% positive). The gmhA protein had a conserved domain through a diversed bacterial population. The phylogenetic tree of *gmhA* is illustrated in figure 46. The tree revealed the cluster of *Burkholderia* groups that were closely related to an environmental bacteria, *Aneurinibacillus thermoaerophilus*.

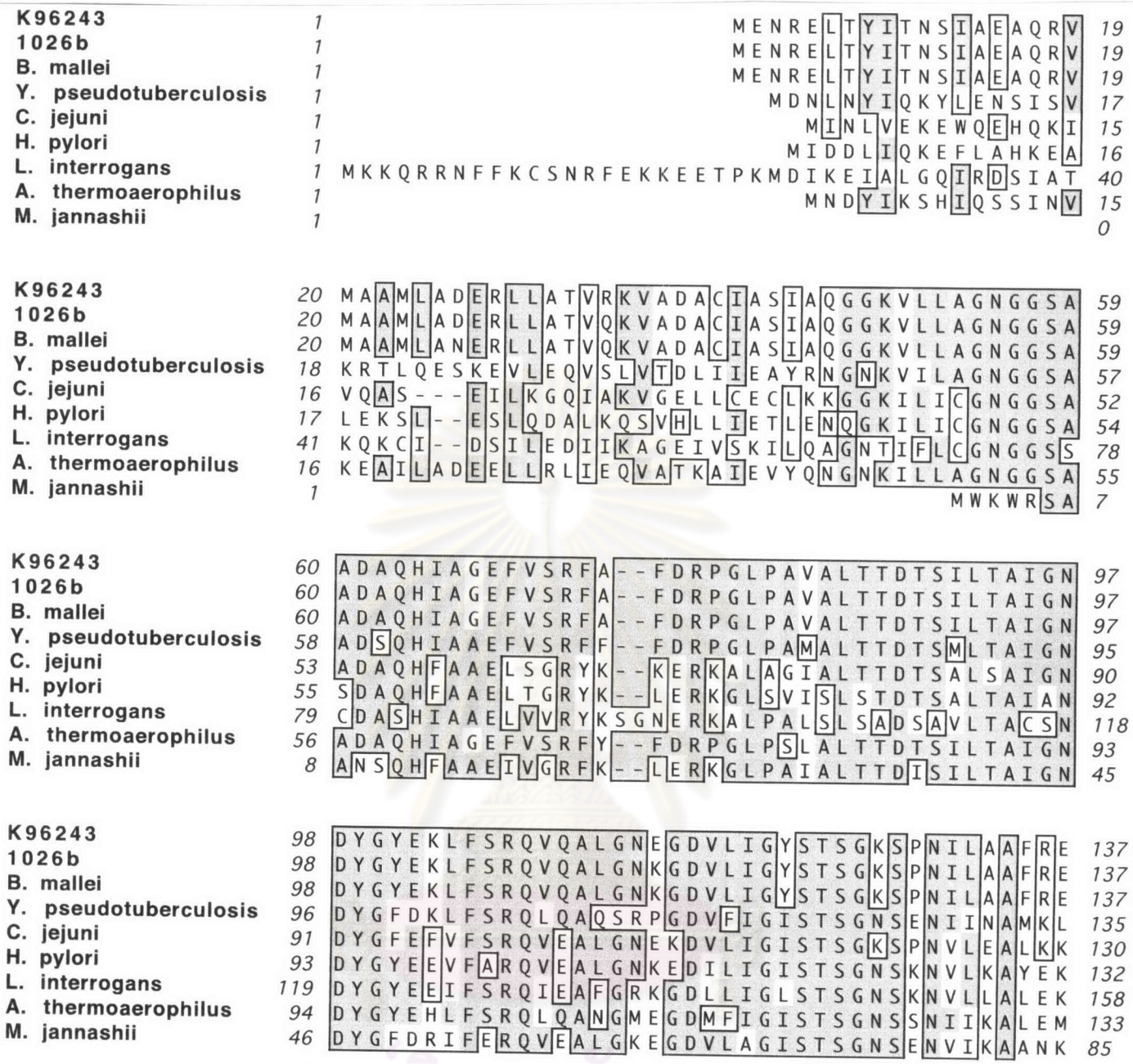


Figure 45. Multiple aligment of the amino acid sequence of gmhA of *B.pseudomallei* (K96243), *B.pseudomallei* gmhA(GenBank accession number.AAK 49808.1), *B.mallei* gmhA(GenBank accession number .AF 28563620),*Yersinia pseudotuberculosis* gmhA (GenBank accession number.AF 46176915),*Campylobacter jejuni* gmhA(GenBank accession number.H81287), *Helicobacter pylori* gmhA(GenBank accession number.AAD 06367.1), *Leptospira interrogans* gmhA(GenBank accession number.NP 712273.1), *Aneurinibacillus thermoaerophilus* gmhA(GenBank accession number. AF 3248364), *Methanocaldococcus jannaschii* gmhA(GenBank accession number.Np248337.1).Where four or more amino acids are identical, they are boxed in shade.Amino acid numbers for each protein are on both sides.

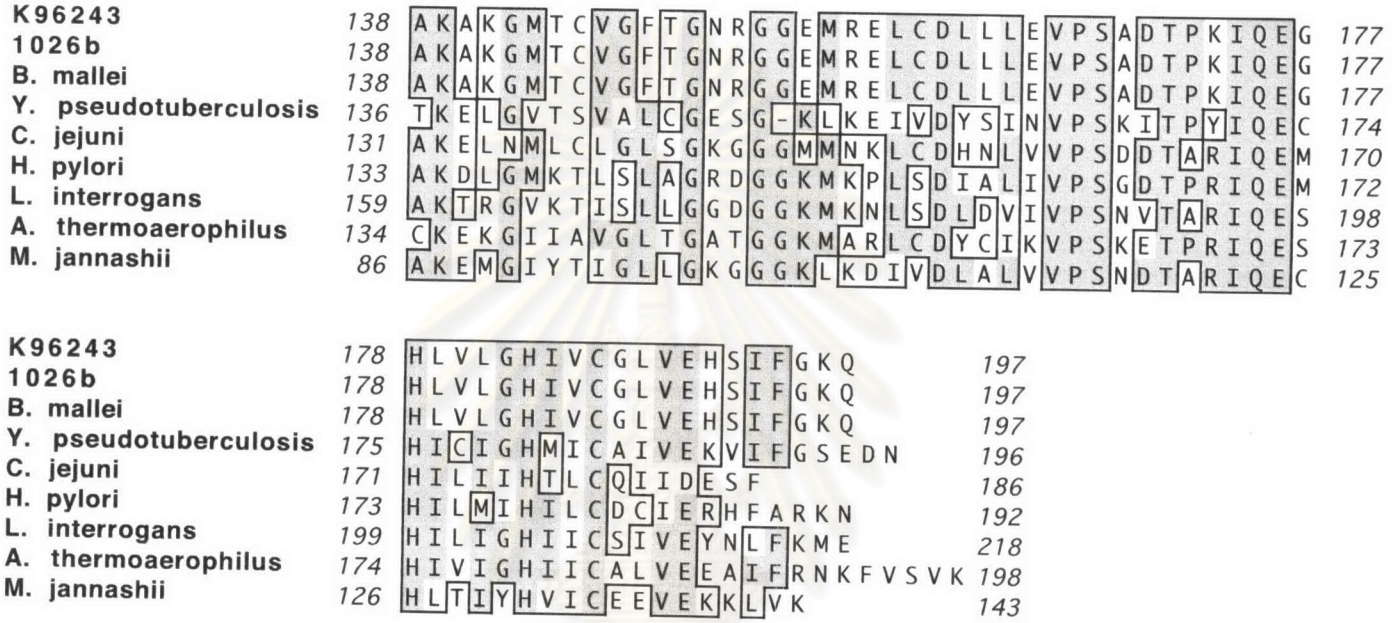


Figure 45 (continue)

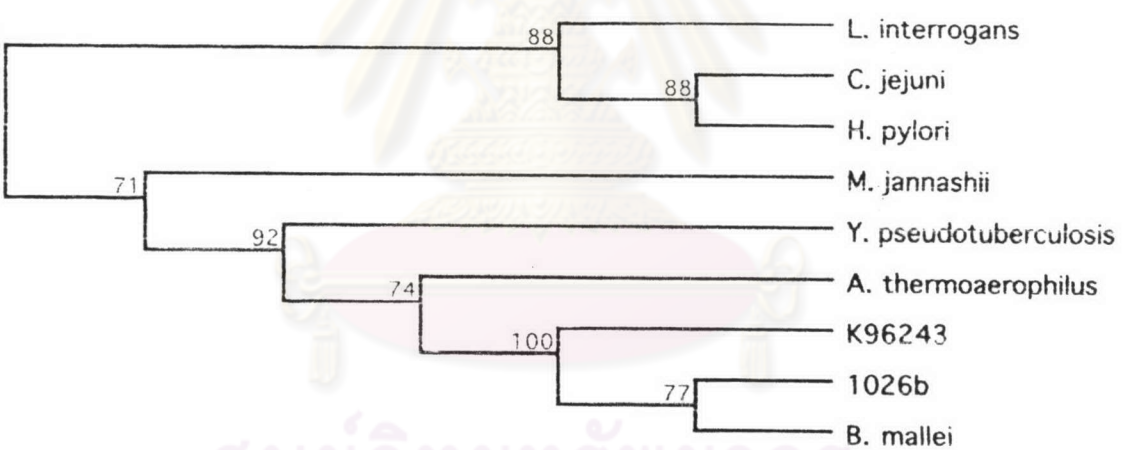


Figure 46 Phylogenetic tree construction of gmhA protein.Phylogram rooted tree based on amino acid sequence of gmhA using phylogenetic inference (distance methods). The tree was constructed from distance data by using neighbor joining method. The number of each node indicated the percentage of bootstrap replicates (of 100 total)

5. PCR cloning of *gmhA* and *wcbM* gene from the *B.pseudomallei* genome.

The *gmhA* gene (594bp) and *wcbM* gene (693 bp) were amplified from the whole *B.pseudomallei* by using PCR as demonstrated in figure47. The amplified genes and vecter pRSET A were digested with endorestriction enzyme *BamH1* and *EcoR1* as demonstrated in figure 48. The recombinant clones of gmhA and wcbM in host cell E.coli DH5 were demonstrated in figure 49 . Both gmhA and wcbM clone were contained the inserted gene size of 594 bp and 693 bp as shown in figure 50. The plasmid isolated from both clones were transformed to a new host cell, E.coli BL21DE3 pLYS, the expressed protein were then demonstrated via SDS-PAGE as in figure 51.The western blot of expressed protein from *gmhA* and *wcbM* clone with both pooled absorbed melioidosis serum and pooled absorbed normal serum could not be demonstrated .

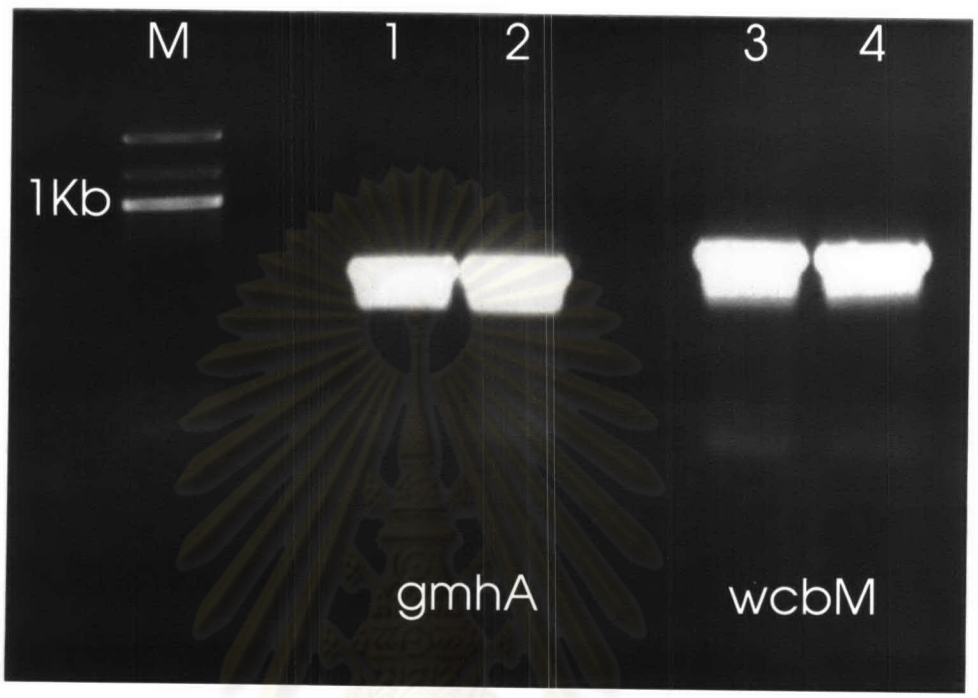


Figure 47 PCR amplification of *gmhA* and *wcbM* genes from the *B.pseudomallei* genome.Agarose gel (1% w/v) electrophoresis of PCR products.M: DNA marker 1 Kb.Lanes 1-2:*gmhA* gene. Lanes 3-4: *wcbM* gene.

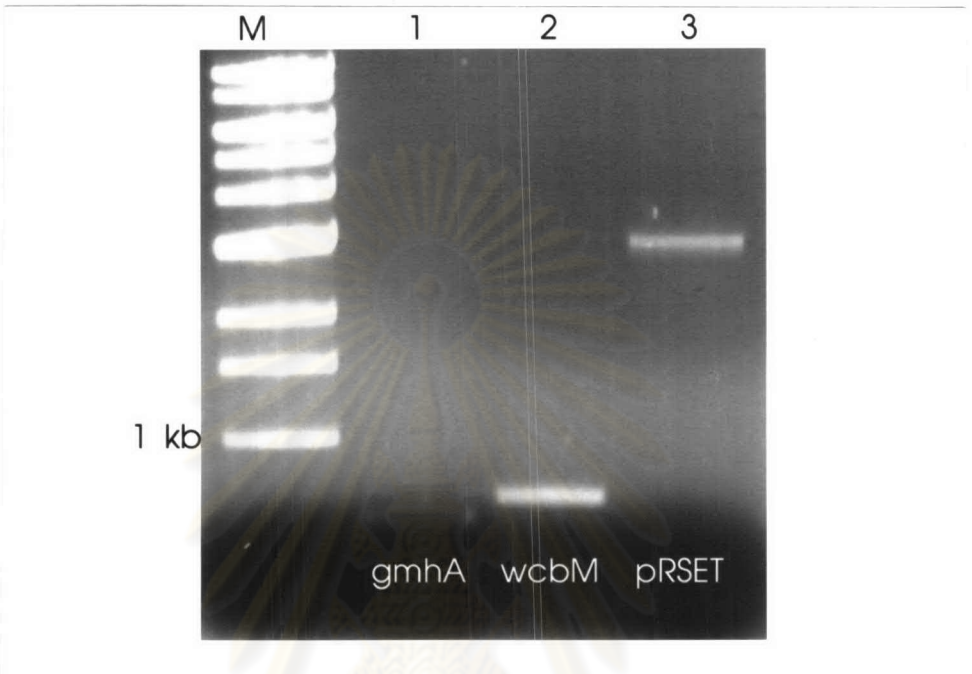


Figure 48 Restriction endonuclease analysis of PCR products of *gmhA* and *wcbM* genes;pRSETa. Agarose gel (1% w/v) electrophoresis of PCR products and pRSETa digested with BamH1 and EcoR1. M: DNA marker 1Kb.Lane 1: *gmhA* gene(594 bp). Lane 2: *wcbM* gene (693 bp). Lane 3: pRSETa

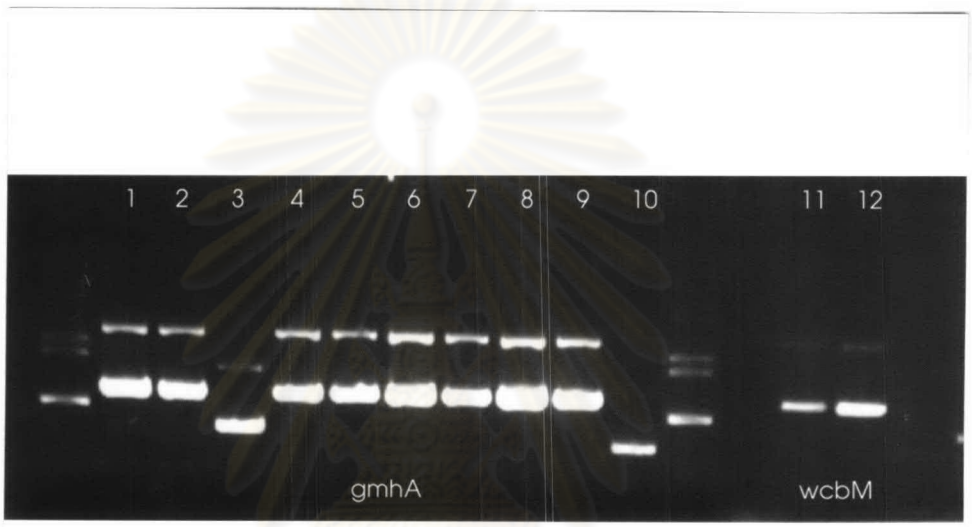


Figure 49 The recombinant clones of *gmhA* and *wcbM* in host cell E.coli DH5 α. Agarose gel (1% w/v) electrophoresis of recombinant pRSETa . Lanes 1,12: pRSETa with out insert DNA. Lanes 2-11: gmhA recombinant clones. Lanes 13-14: wcbM recombinant clones

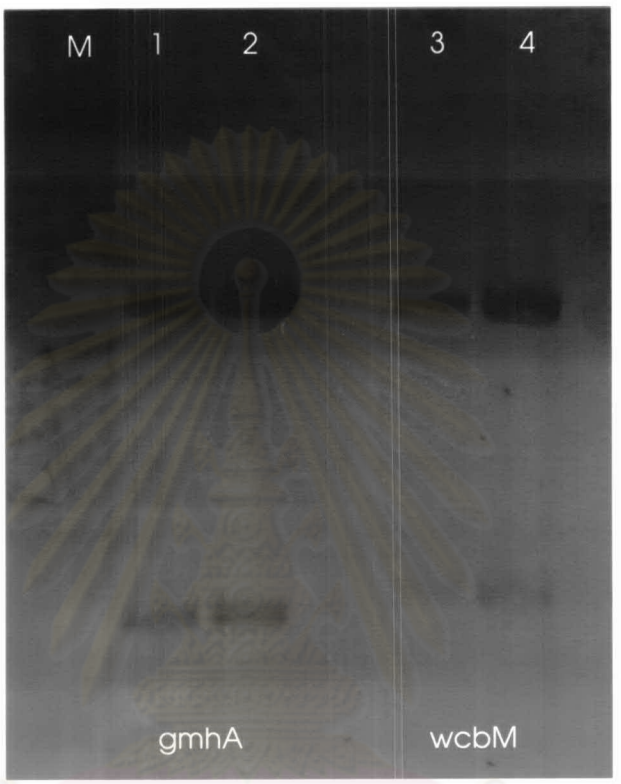


Figure 50 Restriction endonuclease analysis of recombinant clones of gmhA and wcbM.Agarose gel(1% w/v) electrophoresis of recombinant pRSETa digested with BamH1 and EcoR1 (fragment size in bp are shown on the left). M: DNA marker 1 Kb. Lanes 1-2: gmhA clone. Lanes 3-4: wcbM clone.

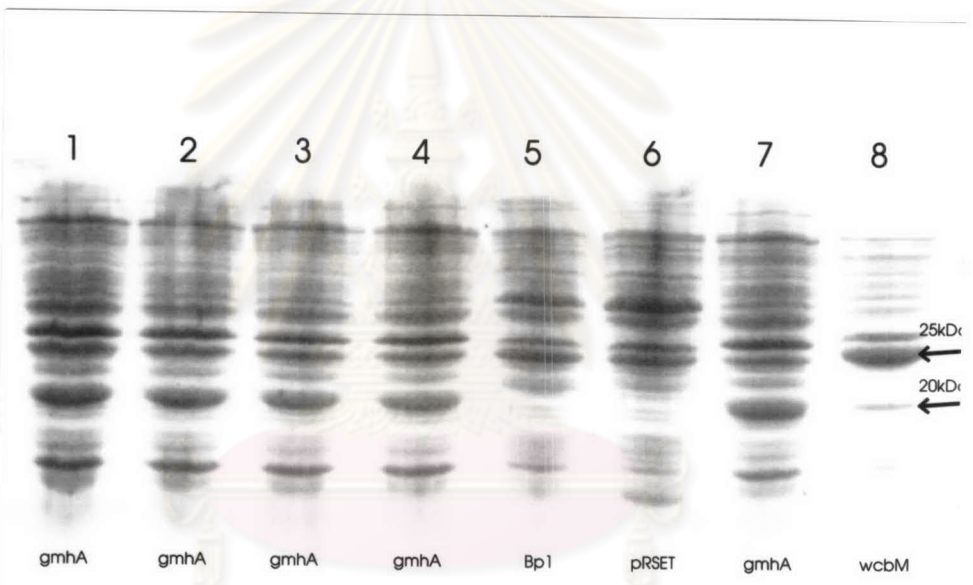


Figure 51 SDS-PAGE analysis of *gmhA*, wcbM and bp1 clone.Lanes 1-4,7: *gmhA* clone. Lane5: Bp1 clone (in E.coli XL1 Blue MRF'). Lane 6; pRSETa without insert in E.coli BL21 DE3 pLYS. Lane 8: *wcbM* clone

Discussion

The results from this immunoscreening approach demonstrate that the developed protocols work very well, with at least 6 genes expressed *in vivo* of the bacteremic melioidosis patient. All of the isolated genes are potentially good diagnostic antigens for melioidosis. Two genes belonged to a well known virulent gene- the capsule gene (*gmhA*, biosynthetic capsule gene) and type three secretion protein gene (PAIs of SPI-1 like cluster, *bipD*). There were also two other genes that are classified as hypothetical conserved proteins, the conserved hypothetical protein gene and the hypothetical signal peptide protein gene. There was one gene (*groEL)* that encodes a chaperonine protein. More interesting, the transmembrane protein gene was a novel gene that generated the strongest signal from the immunoscreening reaction. This transmembrane protein should be the most candidate diagnostic antigen. None of the isolated clones reacted with the pooled absorbed normal serum in plaque dot blot analysis. These findings support our hypothesis that the bacterial genes that are expressed only in a patient and not in a normal person would be virulent genes. Although some isolated genes are novel genes with unknown functions, further experiments are needed to examine their role(s) in melioidosis pathogenesis.

The inserted DNA sequence of positive clone number one hit the first chromosome of *B. pseudomallei* at the capsule operon. This capsule cluster contains approximately 40 Kb with 20 genes [47]. The genes are arranged in a similar way to the group 3 capsule gene found in other gram negative bacterial pathogens such as *Neisseria menigitidis*, *Hemophilus influenzae* and *E. coli*. The common features of this group of genes are as follows: always coexpressed with O serogroup, not thermoregulated, transferred by an abc-2 exporter system, and do not contain KpsU and/or kpsF genes. The genes are organized

in to 3 regions and divergently transcribed. Regions 1 and 3 are generally conserved and contain genes involved in the export of the polysaccharide. The biosynthesis genes are not conserved between serotypes and are located between regions 1 and 3[89]. These capsule genes encode an exopolysaccharide that is an unbranched polymer of a repeating tetrasaccharide unit with the structure –3)-2-O-acetyl-β-D-Galp-(1-4)-alpha-D-Galp-(1-3)-β-D-Galp-(1-5)-β-D-KDOp-(2- [90]. The bacterial capsule has long been recognized as a major virulent determinant of many both gram- negative and gram- positive bacterial pathogens. The capsule enables bacteria to evade adaptive and specific immune defense mechanisms. The most well known effect of the capsule is to prevent phagocytosis by preventing complement deposition and hiding the immunogenic proteins and lipids present in their outer cell walls, including lipopolysaccharides [89]. The role of the *B. pseudomallei* capsule in virulence is unknown [47]. It has been suggested that the polysaccharide capsule of *B. pseudomallei* plays a role in environmental protection, immune system evasion and attachment to epithelial cells that enable the organism to survive within the host. There is some evidence to support this view - a non-capsulated strain of *B. pseudomallei* by insertational mutagenesis of some genes such as *wcbA*, *wzt2*, *wcbE*, *wcbB*, *wcbH*, *wcbP* or *wcbO* was markedly less virulent than the parent strain[47], and a non-capsular mutant was subsequently shown to be unable to induce a protective immune response in mice[90] ; passive immunization with the antibody to this capsule reduces the lethality of infection in mice ; and an antibody response develops to the capsule during an acute infection [90]. The capsule may help the bacterial *B. pseudomallei* to disseminate to other sites of infection and develop bacteremia *in vivo*. There are some solid evidences to support this possibility, as *H. influenzae type b*, *N. menigitidis* and *S. pneumoniae* bacteria are all highly encapsulated, and are the major cause of

bacterial meningitis in children [91]. When the bacteria *N. meningitis* is isolated in carriers, it usually does not produce a capsule [91].

The *gmhA* and *wcbM* genes were amplified from the whole *B. pseudomallei* genome. Both of expressed protein from gmhA and wcbM did not reacted with the pooled melioidosis serum in a western blot experiment.The nature of antigenicity of *gmhA* may be conformational antigenicity that lost during the SDS-PAGE denature process. This preliminary data indicates that *gmhA* was expressed in vivo of bacteremic melioidosis patient. From blast analysis, *gmhA* shows a strong similarity to the phosphoheptose isomerase genes from many bacterial species. The phylogenetic tree of *gmhA* demonstrated a close relationship with *A.thermoaerophilus* (formerly Bacillus), an environmental bacteria. The *gmhA* gene in *A.thermoaerophilus* has been demonstrated to catalyze the coversion of D-sedoheptose-7-phosphate to D-glycero-D-manno-heptose7-phosphate, the substrate for D-glycero-D-manno-heptose production that has been described as constituent of lipopoly-saccharide cores of gram negative bacteria[92]. The *gmhA* gene in *B.pseudomallei* may function in a similar way with *A.thermoaerophilus*. This gene should be expressed inside the bacterial cell for sugar production during active growth in vivo to fight against the body immune system. There are some bacterial cells that are ingested and killed by the normal adaptive defense of the host during early infection that give the gmhA protein to expose with the immune system . The antibodies to a gmhA protein may be indirect evidence of heavy capsule production in the early phase of melioidosis infection. The capsule operon of group A streptococcus has been reported to be induced when the organism was introduced to the pharynx or in the blood stream [93]. In addition, the *gmhA* gene was not found in *Burkhoderia thailandensis* under a southern blot hybridization study , which supports the importance of this gene in virulence determination. The role played by the *gmhA* biosynthetic capsule gene in the

pathological mechanism associated with melioidosis infection requires further investigation.

The Bp 6 clone contained *bpH3* and *bipD* genes that belonged to part of the type three secretion system (TTSS) operon. This TTSS cluster consisted of 30 genes [53]. The organization of the genes in this cluster are similar to TTSS-SpI-1 (inv/spa/prg) of *Salmonella typhimurium* and TTSS (ipa/mxi/spa) of shigella [94]. There are another two TTSS operons in the *B. pseudomallei* genome; TTSS- hrp (hyper sensitive response and pathogenicity) and the TTSS 3 operon which is similar to other plant TTSS operons. The TTSS-hrp and TTSS3 are similar in gene organization, but their nucleotide sequences are sufficiently divergent to suggest that the two TTSS systems may have different roles [95]. TTSS-hrp is present only in *B. pseudomallei* and not in the avirulent *B. thailandensis* [96]. TTSS 3 is found in both *B. pseudomallei* and *B. mallei*. Some parts of TTSS 3 are also present in *B. thailandensis* [96]. TTSS is a highly specialized secretion system and well known virulent gene in more than 12 genera of gram negative bacteria of both animals and plants. The TTSS functions as a machine to deliver the effector molecules into the cytoplasm of the host cell or at the cell wall of bacteria. The effector molecules may have different functions, such as interfering with or subverting normal cells by changing the cytoskeleton structure of the target cell leading to membrane ruffling and apoptosis of infected macrophages; mediating cytotoxicity; or interfering with phagosomal fusion enabling the bacterium to reside in the phagosome unharmed by host cell defenses [94]. The TTSS is usually associated with pathogenicity islands (PAIs) that have a distinct boundary locus [94]. The TTSS could be a very useful tool in at least three ways - the targeted delivery of engineered proteins to influence cellular signal transduction and other processes; they may prove especially suited for delivery of highly toxic or labile

compounds; and they may provide targets for new drugs which specifically attenuate bacterial pathogens without affecting the commensal flora [94].

The role of the TTSS-SpI-1 like operon in melioidosis pathogenesis is unknown. This cluster contains the same average percentage of G+C content as the whole *B. pseudomallei* genome. According to the highest molecular weight of protein expressed band of the Bp6 clone(~45 kDa) , both bpH3 and bipD should be expressed. The protein that reacts with the antibodies in a western blot experiment should be bpH3 or bipD or both. The antibodies in the pooled absorbed melioidosis were polyclonal antibodies that may react with more than one antigenic site in the expressed protein. However, the bipD gene has been reported to be expressed in melioidosis patients [53]. The PCR cloning of bpH3 from the whole genome of *B. pseudomallei* and western blot experiment would support the role of this bp3 gene. The bipD is a secreted protein that may function as a translocator of this TTSS cluster. Mutation in the bipD gene not only disrupts the ability of *B. pseudomallei* to induce membrane protrusion and an actin tail but also renders it unable to escape from endocytic vesicles. The mutant bipD strain of *B. pseudomallei* were almost (95.9%) exclusively observed in association with LAMP-1 (lysosomal marker, lysosome-associated-membrane glycoprotein-1) [53]. *B. pseudomallei* has been demonstrated to be able to invade eukaryotic cells and replicate within the phagocytic cells, and then escape from the phagosome [53]. *B. pseudomallei* may use the same strategy as TTSS-ipa/mxi/spa of shigella to escape from the phagosome. The absence of intracellular replication of the bipD mutant of *B. pseudomallei* in murine macrophage-like cells may be caused by either the confinement of these strains to endosomes and later endosome-lysosome fusion or inability to inhibit the normal killing mechanism of the phagocyte. *Salmonella typhimurium* sipB and sipC and sip D and sopE2 proteins have been demonstrated to regulate inducible nitric oxide synthase (iNOS) [53]. So it is possible that homologues of

these proteins perform a similar function in *B. pseudomallei*. The role of this bipD gene in melioidosis pathogenesis requires further study.

The Bp7 clone has been shown to be the strongest positive signal clone in immunoscreening. Gene analysis indicated that the protein encoded by 246 bp of transmembrane protein gene reacted to the antibodies in the immunoscreening. The strong signal is supported by antigenic index analysis of transmembrane proteins (figure 32). This protein has both the highest hydrophobic maxima peak (Kyte/Doolittle) and highest minima peak of the antigenic site (Hopp/Woods). The nature of this protein is transmembrane not globular, as it contains more than 75% hydrophobic amino acids. This integral protein spans its helix parts (7 regions) in the bacterial membrane and leaves all of the 4 outside regions exposed to the immune system . This membrane protein structure may be the cause of its high signal in immunoscreening. In addition, there is another similar gene with a transmembrane called transmembrane 2 that is located beside the homoserine kinase gene (Fig 31). This arrangement of the genes in this clone may enhance the immune response to a transmembrane protein. The role of this gene in melioidosis pathogenesis requires further investigation. However, according to its high immunogenicity and a novel gene property, this transmembrane protein should be a good candidate diagnostic antigen. The full gene length of the transmembrane protein (813bp) should express a protein that reacts more strongly with antibodies in the melioidosis than the original clone of Bp7.

The Bp3 clone contains two expressed genes named the conserved hypothetical protein gene and hypothetical signal peptide protein gene. The conserved hypothetical protein has two conserved protein domains covering nearly all the protein length. The nature of this protein is not a transmembrane protein. Members of this protein of unknown function have been found in many bacterial species. The role of this gene in melioidosis pathogenesis is unknown.

The hypothetical signal peptide protein has three conserved domains of proteins of unknown function. More interesting with the DUF541 domain, members of this family have been found in bacteria, mouse (GenBank:144450) and humans (GenBank:AI478629). A mouse family member has been named SIMPL (signaling molecule associated with mouse pelle-like kinase). This SIMPL appears to facilitate and/or regulate complex formations between IRAK/mPLK (IL-1 receptor-associated kinase) and IKK (inhibitor of kappa-B kinase) containing complexes, and thus regulate NF-kappa-B activity [97]. Separate experiments demonstrated that a mouse family member (named LaXp189) binds the *Listeria monocytogenes* surface protein ActA, which is a virulent factor that induces actin polymerization. It may also bind stathmin, a protein involved in signal transduction and in the regulation of microtubule dynamic [98]. In bacteria, its function is unknown, but it is thought to be located in the periplasm or outer membrane. Both the conserved hypothetical protein and the hypothetical signal peptide protein have been shown to have a good antigenic index, and both should be good diagnostic antigens for melioidosis diagnosis. The role in pathogenesis of melioidosis requires further investigation.

Bp9 contains nearly the full gene of *groEL*, a chaperonine protein gene. The full gene size of *groEL* gene is 1,596 bp, which is shorter than previously reported (1,641 bp) [99] which may be due to a different genetic strain. This protein contains 49% alpha helix structure and 16.4% beta-sheet with a good antigenic index. *GroEL*, or heat shock proteins, are evolutionarily highly conserved stress-inducible or constitutive proteins that maintain homeostasis in eukaryotic and prokaryotic cells [99]. In *Clostridium difficile* infection, *groEL* encodes a protein of 58 kDa surface exposed adhesin mediating adherence to culture cells [100]. The chaperonine protein may function as a protein folding support in stressful conditions [99]. The role of this *groLI* gene in melioidosis is unknown. This *groEL* gene has been demonstrated to be a highly antigenic

protein in *B. pseudomallei*. Its recombinant protein and monoclonal antibody may be useful for melioidosis serodiagnosis [99]. The *groEL* gene sequence may potentially be useful for a hierarchical identification of medically important Burkholderia at the genus and species level [101]. Our experimental data indicate that the transmembrane protein in the Bp7 clone generates a stronger magnitude signal than this groEL clone.

The Bp5 clone contains only one gene in the inserted DNA. Even though the expressed protein band could not be demonstrated via SDS-PAGE, the *ompA* gene has a high possibility of expression as this gene is located only 233 bp from the T3 expressed vector. The *ompA* has a good antigenic index that makes this protein a good diagnostic antigen for melioidosis. This is not a transmembrane protein. It contains *ompA* conserved domain at the C-terminal. The role of this gene in melioidosis pathogenesis is unknown. The protein may be located at the outer membrane, which represents a good cell surface for host humoral immunity.

Many experimental trials in immunoscreening were performed to improve the sensitivity and specificity of the test. The genomic library was constructed in lambda ZAPII expressed vector to increase the efficiency of clone production [74]. The protein A conjugated alkaline phosphatase was used instead of using antihuman gamma globulin. The protein A has 4 potential binding sites for antibodies, however, only 2 of them can be used at one time. There are at least 2 binding sites on any antibody molecule for protein A. One important disadvantage of protein A is its failure to bind effectively to the subclass from human IgG3 [102]. CDP-star was used as a chemiluminescent substrate to increase the sensitivity of signal detection. CDP-star has MW 496 g/mol with structure disodium 2-chloro-5-(4-methoxyspiro(1,2-dioxetane-3,2'-(5-chloro)-tricyclo(3.3.1.1 ^{3,7})decan)-4-yl)-1-phenyl phosphate [103]. The signal was strong and stable for a long time, so that the x-ray film could be exposed many times.

The reagent substrate was ready to be used and the used reagent could be repeated more than one time. It is usually recommend that the immunoscreening experiment is reacted with the pooled absorbed melioidosis serum as the first step, then the positive plaques will be picked up for reacting with the pooled absorbed normal serum in plaque dot blot assay. This experimental was designed to decrease labor and reagent costs in the immunoscreening. Our experimental data indicates that some protocols need to be reviewed to get more *in vivo* expressed genes efficiently through a process such as follows:

1.The convalescence serum of melioidosis should be an ideal sample for this immunoscreening experiment. The serum should contain all types of antibodies to the expressed antigen of the bacteria *B. pseudomallei*. The main reason to collect the serum from bacteremia melioidosis patients who were admitted to the hospital was the ease in collection from known melioidosis cases - most recovered melioidosis patients who live in rural areas do not return to the hospital even when asked with a letter.

2.Before moving to a step of PCR cloning of each suspected in vivo expressed gene, the dot slot blot experiment may be performed against each melioidosis patient serum at different stages of disease such as bacteremia, localized infection, or disseminated acute septicemic patient. In addition, the serum from bacteremia patients caused by other gram negative bacterial infections such as *E. coli*, Pseudomonas or nonfermentative organisms, should be included in this experiment.

3.The ZAP II expressed genomic library need to be reconstructed according to many isolated positive clones that contained the same genes . There would be meaningless if the more library clones continue to be screened by immunoscreening. The first step of amplification of the ZAP II express library is a critical step that may lead some groups of recombinant

phage dominant than others. Although the phagemids obtained from the mass excision were demonstrated to contain different size of DNA insert.

Bacterial infections are complex and dynamic processes .There are only a few number of bacterial species that can cause disease in human being.The unique features of the bacterial pathogen are the ability to compete with the normal bacterial flora and proliferate in the host enviromment. The factors that support the bacteria to become a pathogen or virulence are called virulent factors. The virulent factors are broadly defined as the gene products responsible for infection(invasion, toxin production, colonization) and manipulation of the host cell functions ultimately lead to host damage or death [61]. The Robert Koch postulate and revision molecular Koch postulate became the gold standard to define bacterial virulence with some limitations [61]. The classic characterization and identification of bacterial virulent genes are carried out by the generation of random mutants then search for avirulent phenotypes. This approach is helpful but a limit number of genes are discovered [65,66]. The growing data indicated that bacteria behave in a different gene expression between *in vivo* and *in vitro* condition [63,64]. The *in vivo* genes expression are regulated by many changeable environment factors that could not be simulated *in vitro* condition [63]. Cloning and identification and characterization of in vivo expressed genes is important for understanding the molecular bacterial pathogenesis. Many new molecular biological techniques are developed to detect *in vivo* expressed genes of bacterial pathogen. Although the new insights of pathogenesis are demonstrated in some bacterial pathogens, there are some limitations to be noted [65,66].

The protocol developed in this study has been demonstrated to work very well. Two well known virulent genes operon are recovered by this approach. In addition all of the isolated genes are potentially good diagnostic antigens for melioidosis. These findings support our hypothesis that the genes

that are expressed only in a patient and not in a normal person would be virulent genes. Although some isolated genes are novel genes with unknown function, further experiments are needed to examine their role(s) in melioidosis pathogenesis. At the same time as this study, there were two reports of immunoscreening of ZAPII expressed genomic libraries of *B.pseudomallei*, in which only the groEL gene could be isolated by using ECL chemiluminescence detection [99] and one novel 18.7 kDa recombinant protein was recovered by using enzyme immunoassay [104]. Some advantages of the current study could be demonstrated as follow : no animal model required: no genetic manipulation in the pathogen required: direct relevance to natural host: potential to detect stage stage and/or route of infection-specific genes; technically simple, fast and inexpensive. However, some disadvantage also should be considered as: not all virulent genes detected , and antibody response required.

The number of isolated *in vivo* expressed genes of melioidosis would be highly expected to be increased if the new constructed ZAPII expressed genomic library are screened with larger number of convalescence bacteremia patient serum. The genes isolated from this approach could be confirmed to be expressed *in vivo* with mRNA transcript detection by RT-PCR in both clinical specimens and animal model. In comparison with other molecular technique, signature tagged mutagenesis (STM) approach, only one biosynthetic operon is detected [105]. With drastically increase in genomic datas in GenBank and continuing development in both bioinformatic and molecular biology, this current study could be efficiently isolate *in vivo* expressed genes in other bacterial pathogens. The isolated genes profile would be useful for pathogenesis study, vaccine development , treatment and diagnosis.